

Organização e Representação do Conhecimento em diferentes contextos:

Desafios e perspectivas na era da datificação

Organizadoras:

Natália Bolfarini Tognoli

Ana Cristina de Albuquerque

Brígida Maria Nogueira Cervantes



ISKO BRASIL
International Society for
Knowledge Organization



Natália Bolfarini Tognoli
Ana Cristina de Albuquerque
Brígida Maria Nogueira Cervantes
Organizadoras

**ORGANIZAÇÃO E REPRESENTAÇÃO DO
CONHECIMENTO EM DIFERENTES CONTEXTOS:
DESAFIOS E PERSPECTIVAS NA ERA DA DATIFICAÇÃO**



© 2023 A reprodução desse livro na íntegra ou em parte é permitida, desde que citados os créditos.
Proibida a venda

COMISSÃO ORGANIZADORA

Diretoria ISKO-Brasil (2022-2023)

Presidente: Natalia Bolfarini Tognoli - Universidade Federal Fluminense

Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina

Coordenadora: Ana Cristina de Albuquerque - Universidade Estadual de Londrina

COMISSÃO CIENTÍFICA

Presidente: Brígida Maria Nogueira Cervantes - UEL

Ana Carolina Simionato Arakaki (UFSCar)

Ana Célia Rodrigues (UFF)

André Vieira de Freitas Araújo (UFPR)

Camila Monteiro de Barros (UFSC)

Carlos Guardado da Silva (Universidade de Lisboa), Portugal

Célia da Consolação Dias (UFMG)

Clarissa Moreira dos Santos Schmidt (UFF)

Daniel Martinez-Ávila (Universidad de León), Espanha

Fabio Assis Pinho (UFPE)

Gercina Ângela de Lima (UFMG)

Gilberto Gomes Cândido (UFPA)

Icléia Thiesen (Unirio)

Jorge Manuel Rias Revez (Universidade de Lisboa), Portugal

José Augusto Chaves Guimarães (Unesp), Campus de Marília

José Carlos Francisco dos Santos (Faculdades Londrina)

Juliana de Assis (UFRJ)

Leilah Santiago Bufrem (UFPR)

Luciana de Souza Gracioso (UFSCar)

Maíra Fernandes Alencar (UEL)

Marcos Luiz Cavalcanti de Miranda (Unirio)

Mariângela Spotti Lopes Fujita (Unesp), Campus de Marília

Mario Barité (Universidad de la Republica), Uruguay

Marivalde Moacir Francelin (USP)

Miriam Gontijo de Moraes (Unirio)

Nanci Elizabeth Oddone (Unirio)

Paula Carina de Araujo (UFPR)

Rodrigo de Sales (UFSC)

Rogério Aparecido Sá Ramalho (UFSCar)

Rosane Suely Álvares Lunardelli (UEL)

Suellen Oliveira Milani (UFF)

Thiago Henrique Bragato Barros (UFRGS)

Vânia Mara Alves Lima (USP)

Zaira Regina Zafalon (UFSCar)

Coordenação Editorial e colaboradores na Comissão Científica

Produção gráfica: Antonio Lucio Barizon Filho

Leitora de Prova: Maíra Fernandes Alencar

Normalização: Patrícia Ofélia Pereira de Almeida

Editoração: Miguel Ivan Magarzo Arias

Colaboradores: Ana Paula Pereira, Aurea Celeste Pires de Souza, Danieli Pinto,

Eduarda Adiemla Ferreira, Júlio César Silveira Tauil, Liliane Cristina Soares Sousa,

Maria Ligia Triques, Pedro Cremonez Rosa, Raíssa Yuri Hamanaka.

Divulgação e Mídias sociais: Raquel Torrão Valentim e Antonio Lucio Barizon Filho

Revisão textual: os autores

Dados Internacionais de Catalogação-na-Publicação (CIP)

068 Organização e representação do conhecimento em diferentes contextos: desafios e perspectivas na era da datificação / Organizadoras Natália Bolfarini Tognoli, Ana Cristina de Albuquerque, Brígida Maria Nogueira Cervantes; produção gráfica: Antonio Lucio Barizon Filho. – Londrina: ISKO-Brasil : PPGCI-UEL, 2023.
818 p. (Estudos Avançados em Organização do Conhecimento, 6)

Livro digital.
ISBN: 978-65-980448-0-0
Conteúdo: **Eixo 1:** Dimensão epistemológica da organização do conhecimento. **Eixo 2:** Dimensão aplicada da organização do conhecimento. **Eixo 3:** Dimensão sociocultural da organização do conhecimento. Inclui os resumos II Fórum de Organização do conhecimento no Brasil na perspectiva dos pós-graduandos: pesquisas em andamento.

1. Organização do Conhecimento. 2. Representação do Conhecimento. 3. Ciência da Informação. I. Tognoli, Natália Bolfarini. II. Albuquerque, Ana Cristina de. III. Cervantes, Brígida Maria Nogueira. IV. Barizon Filho, Antonio Lucio. V. Série.

CDD:020

PREFÁCIO

O Capítulo Brasileiro da *International Society of Knowledge Organization* (ISKO - Brasil) reúne, neste volume, os estudos apresentados e discutidos no VII Congresso Brasileiro em Organização e Representação do Conhecimento, realizado em Londrina, em junho de 2023, o primeiro evento presencial da ISKO-Brasil pós-pandemia. O evento congregou pesquisadores, acadêmicos e profissionais dedicados aos estudos da Organização do Conhecimento e áreas afins, proporcionando um espaço enriquecedor para a troca de ideias, discussões e avanços científicos acerca do tema central “**Organização e Representação do Conhecimento em Diferentes Contextos: Desafios e Perspectivas na Era da Datificação**”.

Especialmente na era da rápida evolução tecnológica e crescente disponibilidade de informações, a organização e representação adequadas do conhecimento se tornam ainda mais relevantes e desafiadoras. Nesse contexto, os participantes do Congresso apresentaram trabalhos de pesquisa que exploram diferentes perspectivas nas três dimensões fundamentais da Organização do Conhecimento: epistemológica, aplicada e sociocultural.

Na dimensão epistemológica, os trabalhos investigam as bases teóricas e conceituais da Organização do Conhecimento, buscando a compreensão mais profunda das estruturas, classificações e formas de representação que fundamentam a organização e a recuperação eficaz do conhecimento em diferentes domínios. Na dimensão aplicada, são destacados os desafios e as soluções encontradas na prática da Organização do Conhecimento em diversos contextos, com especial ênfase às Inteligências Artificiais. Na dimensão sociocultural, os estudos abordaram a interação entre a Organização do Conhecimento e a sociedade, considerando aspectos culturais, políticos e éticos. Exploraram-se questões relacionadas à diversidade cultural, inclusão digital, privacidade, acesso à informação e aspectos sociais da Organização do Conhecimento, reconhecendo a importância de uma abordagem sensível e consciente para promover a equidade e a justiça no acesso ao conhecimento.

Além dos trabalhos completos, reunimos aqui os resumos das pesquisas de Mestrado e Doutorado apresentadas no evento no âmbito do segundo Fórum de *Organização do conhecimento no Brasil na perspectiva dos pós-graduandos: pesquisas em andamento*. Trata-se de iniciativa inédita e que recomendamos que seja continuada nas futuras edições.

Este livro é um testemunho do comprometimento dos participantes do VII Congresso Brasileiro em Organização e Representação do Conhecimento em compartilhar suas pesquisas e contribuições para a área. Os capítulos apresentam uma ampla gama de tópicos, reflexões e descobertas que refletem a riqueza e a diversidade dos debates ocorridos no evento. Cada autor e autora trouxe uma perspectiva única, enriquecendo o campo da Organização do Conhecimento e abrindo novos horizontes para pesquisas futuras.

Esperamos que este livro seja uma prova duradoura do nosso compromisso coletivo em impulsionar a pesquisa e o desenvolvimento da Organização do Conhecimento no Brasil, incentivando a reflexão em estudantes, pesquisadores, profissionais e todos aqueles que têm interesse nos estudos da área. Que as ideias aqui compartilhadas contribuam para uma sociedade mais informada, conectada e preparada para enfrentar os desafios que a era da datificação nos apresenta.

Natália Tognoli

Presidente do Capítulo Brasileiro da ISKO (2023-2024)
Programa de Pós-graduação em Ciência da Informação
Universidade Federal Fluminense
Junho/2023

SUMÁRIO

Dimensão Epistemológica da Organização do Conhecimento

Análise de Conteúdo na Ciência da Informação: Limites e Possibilidades no Uso da Disciplina Foucaultiana	1
Aspectos culturais e a Teoria dos Objetos de Fronteira: Aproximações em Busca de Caminhos mais Inclusivos na Organização do Conhecimento	8
Categorias Fundamentais da Organização da Informação: do Conceitual, Metodológico e Pragmático ao Normativo e Tecnológico	18
De que Serve o Autor em Ciência da Informação? Um Estudo Sobre a Relação Entre a Noção de Autor e o Indivíduo Que Escreve	29
Indexação e Desinformação: Possíveis Relações	37
John Dewey e Birger Hjørland: Pragmatismo e Lógica nos Fundamentos da Organização do Conhecimento	44
A Modelagem Conceitual, Organização do Conhecimento e Organização da Informação: Aproximações Epistemológicas no Universo Bibliográfico	50
<i>Ordo Librorum Hviivs Operis</i> : Estrutura e Fundamentos da Classificação Bibliográfica Gesneriana	57
Organização do Conhecimento e Epistemologia Social: Relações Teóricas, Epistemológicas e Aplicadas	65
A Organização do Conhecimento na Biblioteconomia: uma Perspectiva por Meio da Análise da Série Estudos Avançados em Organização do Conhecimento da ISKO	77
<i>The Purpose of Epistemology of Knowledge Organization</i>	84
Referentes Teóricos em Indexação Arquivística: Estudo em Artigos Científicos Brasileiros	92
Representação do Conhecimento Arquivístico: uma análise comparativa entre entidades de <i>Records in Contexts Ontology</i> e entidades derivadas de <i>Basic Formal Ontology</i>	101
Sistemas de Organização do Conhecimento no Contexto da ISKO: Uma Análise de Domínio dos Capítulos Brasileiro e Norte-Americano	112

Dimensão Aplicada da Organização do Conhecimento

Análise da Representação da Música em Bibliotecas Digitais	126
Análise de Domínio em Biodiversidade	134
Aspectos Constitutivos e Atributos da Fotografia Documental	143
Base de Registros de Tesouros: um Mapeamento de <i>Software</i> Gestor de Tesouros e Tecnologias de Acesso	152
Big Data e Dados Pessoais: Questões Legais Emergentes na Perspectiva da Ciência da Informação	159

A Catalogação de Pinturas no Cenário Linked Data	169
Ciência Cidadã em Botânica: um Estudo Sobre a Colaboração de Cidadãos em Herbários e Jardins Botânicos	179
<i>Computational Archival Science</i> : Uma Nova Abordagem Para a Organização do Conhecimento Arquivístico.....	187
As Comunidades Virtuais de Aprendizagem e o Uso das Tecnologias Digitais na Formação do Gestor de Compras e Licitações no IFPB	194
O Conceito <i>Linked Data</i> no Brasil.....	203
Construção de Modelo para Enriquecimento Semântico de Imagem Biomédica: Metodologia para Extração de Conceitos.....	213
A Convergência da Organização do Conhecimento e do Design para Apresentação de Informação Acessível: a Criação de um Modelo de Tesouros Acessível.....	221
Correlações entre a Intencionalidade dos Sujeitos Informativos e o Conceito de Polirrepresentação	231
Dados de Pesquisa em Enfermagem: um Estudo Sobre a Organização Para Transformação em Dados FAIR (<i>FAIRification Process</i>).....	240
O <i>Data Quality Vocabulary</i> (DQV): um Sistema de Organização do Conhecimento?.....	250
Datificação, Inteligência Artificial e Imagens: o Paradigma Dominante na Representação do Conhecimento em Imagens	260
A Dimensão Aplicada da Organização do Conhecimento no Campo da Inteligência Artificial	270
O Domínio Científico da Arquivologia na Organização do Conhecimento: uma Análise a Partir do Acoplamento Bibliográfico no Periódico <i>Knowledge Organization</i> (KO).....	279
Epistemologia do Mapa Conceitual em Sistema de Organização do Conhecimento por Meio da Historicidade dos Conceitos.....	288
Estratégias de Pesquisa e Análise de Dados no Domínio da Organização do Conhecimento: Análise de Abordagens Metodológicas Utilizadas nos Trabalhos do Eixo Dimensão Aplicada da ISKO-Brasil.....	299
Ética da Informação nos Compiladores de Tendências das Plataformas de Mídia Digital.....	308
<i>Fake News</i> e a Organização do Conhecimento: Aproximações Possíveis.....	318
A Importância da Organização E Representação de Informação para a <i>Search Engine Optimization</i>	327
Indexação Social como Estratégia para Visibilidade de Conteúdos Científicos no Serviço de Rede Social Online Tiktok.....	335
Introdução a Práticas de Mineração de Dados: Considerações Sobre os Dados Abertos Governamentais em LONDRINA	347
Investigações Acerca da Representação da Informação em Repositórios de Dados de Pesquisa na Ciência da Informação Brasileira.....	353

<i>Knowledge Organization</i> no Processo de FAIRificação de <i>Datasets</i> : Estruturando a Semântica e Interligando as Notícias do Banco de Dados de Periódicos Musicais Oitocentistas	363
Levantamento e Tratamento dos Dados de Pesquisa da Ciência da Informação à Luz dos Dados Abertos da CAPES	373
O Lugar da Normalização na Organização do Conhecimento: Análise da Coleta de Referências em Teses e Dissertação de um Programa de Pós-Graduação em Ciência da Informação	383
Metodologia para Categorização das Concepções Científicas em Ciência da Informação.....	391
<i>Mineral Exploration Knowledge Organization System (MEKOS)</i> : um Modelo Conceitual para o Domínio da Exploração Mineral	402
Organização do Conhecimento e os Desafios Tecnológicos da Era Big Data	410
Percepção Bibliotecária Sobre Política de Indexação: um Estudo em Bibliotecas Especializadas	420
Perspectivas do Conhecimento Arquivístico no Cenário de Acesso à Informação e Proteção de Dados Pessoais	430
Possibilidades para a Indexação em Museus: Identificação e Análise de Orientações para Leitura Técnica de Objetos Museológicos.....	439
Potencial da Aplicação de Técnicas de Processamento de Linguagem Natural (PLN) e de <i>Machine Learning</i> (ML) no processo de categorização para a Análise de Domínio (AD).....	447
A Presença Online dos Autores Mais Produtivos da ISKO Brasil: Indicadores de Atividade, Visibilidade e Impacto nas Plataformas Sociais.....	457
Preservação de Documentos Digitais: Estratégias de Preservação	467
Procedimentos Para Leitura Documentária de Objetos Museológicos de Coleções Zoológicas: Contribuições Para Análise de Assunto em Museus	474
Proposta de Classificação para Promoção da Competência em Dados de Pesquisa.....	481
Proposta de Ontologia para Gestão Documental Arquivística de Acervo Acadêmico.....	493
Prospecção de Termos Candidatos no Domínio do Transtorno do Espectro Autista.....	503
A Relevância da Análise de Domínio: um Estudo da Classificação das Atividades-fim nas Instituições Federais de Ensino Superior -IFES.....	511
A Representação da Informação Arquivística nos Arquivos Estaduais Brasileiros.....	520
Representação de Figurino: a Moda como Informação.....	527
Representação e Organização do Conhecimento dos Dados Legislativos na Câmara dos Deputados do Brasil: Análise e Classificação em Direção à Agenda 2030	538
Os Sistemas de Organização do Conhecimento nos Currículos dos Programas de Pós-Graduação em Ciência da Informação nas Regiões Sul e Sudeste do País	545
Thesa 2.0: a Interação dos Tesouros em Tempos de Dados Lincados	554
O Uso de Taxonomias para o Enriquecimento Semântico de Objetos	562

Dimensão Política e Social da Organização do Conhecimento

Aspectos éticos da organização e representação do conhecimento no contexto das políticas de desenvolvimento em coleções de obras raras	573
As Bases Epistemológicas do Álbum Fotográfico e Suas Relações com a Organização do Conhecimento	580
O Capacitismo e a Ética na Representação do Conhecimento: Um Estudo Quanto à Ética na Representação do Conhecimento Sobre a Temática da Neuroatipicidade	591
<i>Chatbots</i> em Bibliotecas: Aplicações, Desafios e Perspectivas	600
O Desafio da Representação da Etnoliteratura: o Caso do Glossário Decolonial de Macunaima	608
Desafios da Organização do Conhecimento Para a Representação Temática dos Saberes Originários em Sistemas de Informação: Reflexões Preliminares	615
Em Busca de uma Terminologia: o Circo no Brasil de 1820 a 2000.....	622
Espoliação de Artefatos Africanos: o Caso dos Bronzes de Benin e Suas Representações no <i>British Museum</i>	632
O Fenômeno do Juridiquês e seu Impacto na Representação do Conhecimento	641
<i>La Garantía Indígena: Aportes para su Aplicación al Patrimonio Cultural de los Pueblos Aborígenes Latinoamericanos</i>	650
Intersecção de Três Lentes da Análise de Domínio para a Caracterização de Comunidades Discursivas com Focos Decoloniais/Críticos-Sociais.....	661
Linguagem Documentária e Terminologia: Fundamentos Linguísticos para a Organização do Conhecimento	671
Mediação Cultural da Informação e Organização do Conhecimento: Uma Abordagem Crítica de Enfretamento ao Epistemicídio.....	682
Meninas e Mulheres na Ciência: o Caso das Pesquisadoras Brasileiras e m Ciência da Informação no Norte e Nordeste do Brasil.....	693
Organizar Para Quem? Uma Reflexão Sobre a Mediação da Informação	700
Representação Arquivística: Percepção dos Arquivistas das Instituições Federais de Ensino Superior sobre a neutralidade	709
A Representação Colaborativa do Conhecimento no WikiFavelas - Dicionário de Favelas Marielle Franco	717
Representação do Conhecimento na Prisão: Palavras e Significados em Jogo	727
Sistemas de Organização do Conhecimento Inclusivos: a importância das Notas de Escopo, Notas de Uso e Definições	734
Terminologia da Documentação Sanitária na Literatura Científica	745
A Utilização de Dados Como Moedas Digitais Através da Coleta Por Cookies	753

Resumos dos trabalhos apresentados no Fórum de discentes

Bibliografias em Museus: um estudo em museus brasileiros do início da república a meados do século XX (1893-1949)	760
Enfoques sobre Sistemas de Organização do Conhecimento (SOC) nos currículos dos programas de pós-graduação em Ciência da Informação do Brasil	763
Organização e Representação do Conhecimento aplicado na área de “controle externo” sob a luz de Foskett: mapeamento dos processos para aperfeiçoamento do tesouro de contas nacional	766
Proposta de formalização de política de indexação de biblioteca universitária	769
Mapeamento de padrões de metadados para dados de pesquisa alinhados aos princípios FAIR	771
Análise do domínio <i>Computacional Archival Science</i> : contribuições para a Arquivologia e a Organização do Conhecimento Arquivístico	773
Aparecimento e descontinuidade de termos: um resgate histórico em periódicos de Ciência da Informação	776
O contexto social de criação dos álbuns fotográficos: uma reflexão no âmbito da Organização do Conhecimento	779
Contribuições da Linguística Sistêmico-Funcional (SFL) para o aprimoramento da construção de Sistemas de Organização do Conhecimento (SOCs)	781
As concepções de Michel Foucault para a compreensão da informação como poder na Ciência da Informação	784
A decolonialidade na Organização e Representação do Conhecimento para a construção de Sistemas de Organização do Conhecimento Decoloniais	786
O domínio dos patrimônios culturais na Ciência da Informação pela perspectiva da Organização e Representação do Conhecimento em plataformas digitais	788
A importância da privacidade para a Organização do Conhecimento	790
Organização do Conhecimento e inteligência artificial: interseções teórico-conceituais	792
Organização e Representação do Conhecimento em Arquivos de associações político-partidárias: mapeamento de acervos documentais do Partido dos Trabalhadores no Estado de São Paulo	794
Paratextos: elementos que acompanham o texto	797
O período de escravização no Brasil e a Organização e Representação da Informação e do Conhecimento: estudos nas instituições arquivísticas	799
A produção científica dos líderes dos grupos de pesquisa em Organização e Representação da Informação e do Conhecimento: uma análise nos anais dos eventos da ISKO-Brasil	801
<i>Scientia Oblitus</i> : o esquecer e o resgate dos saberes dentro da Organização e da Representação do Conhecimento	804
Uma história da informação e seus efeitos éticos e epistemológicos	807



**DIMENSÃO EPISTEMOLÓGICA DA
ORGANIZAÇÃO DO CONHECIMENTO**

*The Epistemological Dimension
of the Knowledge Organization*

Análise de Conteúdo na Ciência da Informação: limites e possibilidades no uso da disciplina foucaultiana

Content Analysis in Information Science: limits and possibilities in the use of the foucauldian discipline

Liliane Cristina Soares Sousa (1), Fábio Parra Furlanete (2)

(1) UEL, Rodovia Celso Garcia Cid, PR-445, Km 380, lilianeli.sousa@uel.br, (2) ffurlanete@uel.br

Resumo

Objetivo: Este artigo apresenta uma análise crítica sobre os parâmetros conceituais adotados pelos pesquisadores do Programa de Pós-Graduação em Ciência da Informação da UEL ao utilizar a Análise de Conteúdo como metodologia de análise em seus objetos de pesquisa. Para isso, foi utilizado procedimento de controle do discurso chamado por Michel Foucault de “disciplina”. Compreendemos que esse esboço tem grande contribuição para o campo da Organização do Conhecimento, visto que, os múltiplos debates teóricos na área, contribui para novas perspectivas conceituais, ampliando e problematizando as diversas formas de organizar e representar o conhecimento na Ciência da Informação. **Metodologia:** A pesquisa foi realizada com base em uma abordagem bibliográfica, utilizando tanto métodos quantitativos quanto qualitativos. **Resultados:** Foi observado que a maioria dos pesquisadores optou pela Análise de Conteúdo como metodologia, com uma taxa de utilização de 43,16% das 95 dissertações defendidas no período de 2014 a 2022. A análise também revelou que muitos dos pesquisadores utilizaram apenas uma referência bibliográfica para discutir a Análise de Conteúdo, o que pode ser interpretado como um exemplo de controle do discurso denominado disciplina. **Conclusão:** Conclui-se que a falta de uma abordagem mais ampla, incluindo múltiplas referências bibliográficas, pode restringir a compreensão e o desenvolvimento da Análise de Conteúdo como metodologia de análise em Ciência da Informação. Portanto, é necessário promover a ampliação do repertório teórico e conceitual dos pesquisadores para que possam escolher suas metodologias com maior clareza e fundamentação.

Palavras-chave: Metodologia de pesquisa. Análise de conteúdo. Disciplina. Ciência da Informação.

Abstract

Objective: This article presents a critical analysis of the conceptual parameters adopted by researchers from the Graduate Program in Information Science at UEL when using Content Analysis as an analysis methodology in their research objects. For this, a discourse control procedure called “discipline” by Michel Foucault was used. We understand that this sketch has a great contribution to the field of Knowledge Organization, since the multiple theoretical debates in the area contribute to new conceptual perspectives, expanding and problematizing the different ways of organizing and representing knowledge in Information Science. **Methodology:** The research was carried out based on a bibliographical approach, using both quantitative and qualitative methods. **Results:** It was observed that most researchers opted for Content Analysis as a methodology, with a utilization rate of 43.16% of the 95 dissertations defended in the period from 2014 to 2022. The analysis also revealed that many of the researchers used only one bibliographic reference. ca to discuss Content Analysis, which can be interpreted as an example of discourse control called discipline. **Conclusion:** It is concluded that the lack of a broader approach, including multiple bibliographic references, can restrict the understanding and development of Content Analysis as an analysis methodology in Information Science. Therefore, it is necessary to promote the expansion of the theoretical and conceptual repertoire of researchers so that they can choose their methodologies with greater clarity and foundation.

Keywords: Research Methodology. Content analysis. Discipline. Information Science.

Resumen

Objetivo: Este artículo presenta un análisis crítico de los parámetros conceptuales adoptados por investigadores del Programa de Posgrado en Ciencias de la Información de la UEL al utilizar el Análisis de Contenido como metodología de análisis en sus objetos de investigación. Para ello se utilizó un procedimiento de control del discurso denominado “disciplina” por Michel Foucault. Entendemos que este bosquejo tiene un gran aporte al campo de la Organización del Conocimiento, ya que los múltiples debates teóricos en el área contribuyen a nuevas perspectivas conceptuales, ampliando y problematizando las diferentes formas de organizar y representar el conocimiento en las Ciencias de la Información. **Metodología:** La investigación se llevó a cabo con base en un enfoque bibliográfico, utilizando métodos tanto cuantitativos como cualitativos. **Resultados:** Se observó que la mayoría de los investigadores optó por el Análisis de Contenido como metodología, con una tasa de utilización del 43,16% de las 95 disertaciones defendidas en el período de 2014 a 2022. El análisis también reveló que muchos de los investigadores utilizaron una sola referencia bibliográfica. para discutir el Análisis de Contenido, que puede ser interpretado como un ejemplo de control del discurso llamado

disciplina. **Conclusión:** Se concluye que la falta de un enfoque más amplio, que incluya múltiples referencias bibliográficas, puede restringir la comprensión y el desarrollo del Análisis de Contenido como metodología de análisis en Ciencias de la Información. Por lo tanto, es necesario promover la ampliación del repertorio teórico y conceptual de los investigadores para que puedan elegir sus metodologías con mayor claridad y fundamento.

Palabras clave: Metodología de investigación. Análisis de contenido. Disciplina. Ciencias de la Información.

1 Introdução

O Programa de Pós-Graduação em Ciência da Informação da UEL (PPGCI/UEL, 2023) iniciou suas atividades em 2012, com o curso de Mestrado Acadêmico. No ano de seu primeiro quadriênio, o PPGCI/UEL obteve conceito 4 da CAPES, o que possibilitou a abertura do Doutorado em 2018, com sua primeira turma ingressante no ano de 2019. Diante desse cenário, há um ciclo contínuo de produções acadêmicas, que fortalecem o comprometimento com as necessidades sociais, nacionais e regionais, sendo o único curso *strictu sensu* em Ciência da Informação na área de Comunicação e Informação da CAPES, do Estado do Paraná.¹

No período de vigência do programa, foram desenvolvidas, entre os anos de 2014 a 2022, 98 dissertações, sendo que, ao consultar o catálogo do Sistema de Bibliotecas da UEL, 3 dissertações não estavam disponíveis para acesso, possibilitando a pesquisa em 95 trabalhos acadêmicos produzidos neste período. Diante desse contexto, analisamos as metodologias de análise dos dados nas dissertações, obtendo um repertório de métodos utilizados nas pesquisas.

Identificamos que a técnica preponderantemente utilizada nas investigações foi a Análise de Conteúdo. O percentual de pesquisadores que utilizaram o método citado foi de 43,16%, totalizando 41 dissertações. Perante o exposto, direcionamos nosso esboço com a seguinte questão: quais são os parâmetros conceituais que os pesquisadores do Programa de Pós-Graduação em Ciência da Informação da UEL utilizaram ao escolher a Análise de Conteúdo como metodologia de análise para seus objetos de pesquisa? Nosso objetivo é compreender as proposições de utilização da técnica de Análise de Conteúdo pelos pesquisadores, na produção das dissertações no campo da Ciência da Informação, especificamente, no PPGCI/UEL, consequentemente compreender o impacto na organização do conhecimento.

A análise de conteúdo é uma técnica de pesquisa qualitativa que visa identificar padrões de significado dentro de um conjunto de dados, como entrevistas, textos, imagens e vídeos. A técnica se baseia na identificação de categorias e temas presentes nos dados coletados, permitindo a compreensão de fenômenos complexos e a identificação de padrões e tendências. Segundo a observação de Moraes (1999, p. 8) a análise de conteúdo “[...] ajuda a reinterpretar as mensagens e

a atingir uma compreensão de seus significados num nível que vai além de uma leitura comum.” Para o autor, a análise de conteúdo é uma técnica metodológica capaz de aprofundar o entendimento dos significados presentes nas mensagens. Sendo assim, permite uma leitura mais crítica e reflexiva das mensagens, contribuindo para uma compreensão mais ampla e profunda do conteúdo analisado.

A Ciência da Informação é uma área multidisciplinar que se dedica ao estudo da produção, disseminação, organização e uso da informação em diversos contextos. No entanto, essa área do conhecimento é muito mais do que um conjunto de técnicas e práticas para lidar com a informação. Em sua essência, a Ciência da Informação é discurso que produz conhecimento e está permeado de relações de poder. Na observação de Capurro (2003, sem paginação) “[...] se queremos identificar o papel de uma ciência da informação autônoma, devemos transportá-la a nível mais abstrato.” Para o autor, é preciso ir além das questões práticas e técnicas e pensar nos conceitos fundamentais e nas questões mais abrangentes que norteiam a ciência da informação. Ele ainda defende que a ciência da informação é uma disciplina que lida com questões filosóficas, sociológicas, psicológicas e tecnológicas relacionadas ao uso e ao significado da informação na sociedade.

Propomos uma análise crítica, a partir da leitura da obra de Michel Foucault, pois consideramos uma fonte importante na estruturação de uma teoria crítica, em particular, para a Ciência da Informação. Essa teoria crítica busca compreender os mecanismos de poder e controle presentes na sociedade, e como eles afetam a produção, organização e disseminação da informação. Na observação de Foucault (1996) o poder não é algo individual ou institucional, mas uma coisa difusa, presente nas relações sociais e nas práticas discursivas. Ele ressalta que o poder não é algo que se possui, mas algo que se exerce. E está presente em todas as relações sociais, e é usado para controlar, moldar e dirigir a conduta das pessoas.

Nesta pesquisa, a abordagem da “disciplina” considera o procedimento de controle do discurso, conforme explícito na obra “A Ordem do Discurso”. Foucault (1996) destaca que o discurso é moldado por práticas discursivas que estabelecem regras, normas e formas de exclusão. Esse sentido para o termo disciplina é diferente daquele definido em “Vigiar e Punir”, onde a disciplina é descrita como uma forma específica de

Sousa, Liliane Cristina Soares; Furlanete, Fábio Parra. Análise de Conteúdo na Ciência da Informação: limites e possibilidades no uso da disciplina foucaultiana. *Organização e Representação do Conhecimento em diferentes contextos: desafios e perspectivas na era da datificação*. 2023.

poder disciplinar, baseada em técnicas de controle, vigilância e normalização do comportamento.

Quanto a metodologia empregada neste texto, o estudo é do tipo bibliográfico, visto que, busca em materiais já publicados, o objeto de estudo proposto, neste caso: os parâmetros que os pesquisadores do PPGCI/UUEL, utilizaram para a seleção do método de Análise de Conteúdo, para a análise dos dados das respectivas pesquisas. A abordagem empregada é a quantitativa.

A pesquisa iniciou com a delimitação e recorte das fontes que seriam utilizadas para a coleta e análise dos dados. Diante disso, selecionamos as 41 dissertações, que tinham como método principal a Análise de Conteúdo. Posterior à seleção dos trabalhos, optamos por analisar a introdução, os objetivos e a metodologia de cada dissertação, para compreender as referências dos pesquisadores, ao utilizar o método de análise.

No campo científico, esta investigação justifica-se por disponibilizar aos pesquisadores da área de Ciência da Informação, um parâmetro a respeito das metodologias utilizadas nas pesquisas acadêmicas, mesmo que um recorte de um determinado local, esse mapeamento é significativo pela relevância da aplicação de metodologia científica para o fazer ciência. Goldenberg (2009) “[...] infere que a metodologia ensina o pesquisador a ter um olhar crítico e científico, desenvolvendo nele o pensamento científico, criativo, organizado, claro e crítico”. Ao seguir os processos de metodologia científica, somos incentivados a avaliar nossos próprios dados e conclusões.

2 Desenvolvimento

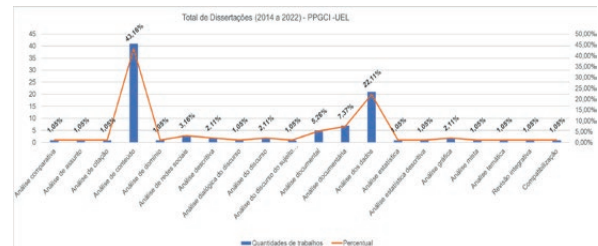
2.1 Mapeamento das técnicas de análise em dissertações do PPGCI/UUEL

O presente artigo traz uma análise dos dados obtidos a partir da revisão de 95 dissertações, conforme Gráfico 1, que objetivou identificar as principais técnicas de análise utilizadas pelos autores. Os resultados indicam que a técnica de análise de conteúdo foi a mais utilizada pelos pesquisadores, representando 43,16% das dissertações analisadas.

A análise documental e a análise dos dados também se destacaram, representando respectivamente, 5,26% e 22,11% das dissertações. Já as técnicas de análise de assunto, análise comparativa, análise de citação, análise de domínio, análise descritiva, análise dialógica do discurso, análise do discurso, análise do discurso do sujeito coletivo, análise estatística, análise estatística descritiva, análise gráfica, análise mista, análise temática, compatibilização e revisão integrativa foram menos utilizadas, representando menos de 3% cada uma. A técnica de análise de redes sociais apresentou

um percentual de 3,16%, o que indica uma utilização moderada pelos pesquisadores.

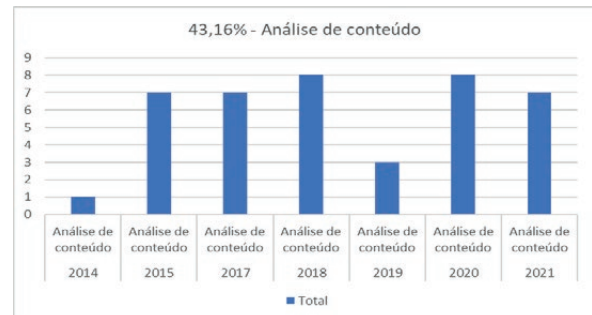
Gráfico 1. *Dissertações publicadas no PPGCI/UUEL – 2014 a 2022*



Fonte: Elaborado pelos autores (2023).

O estudo apontou que um total de 41 dissertações, representado pela porcentagem de 43,16%, utilizaram a análise de conteúdo em suas pesquisas, conforme visualizamos no Gráfico 2.

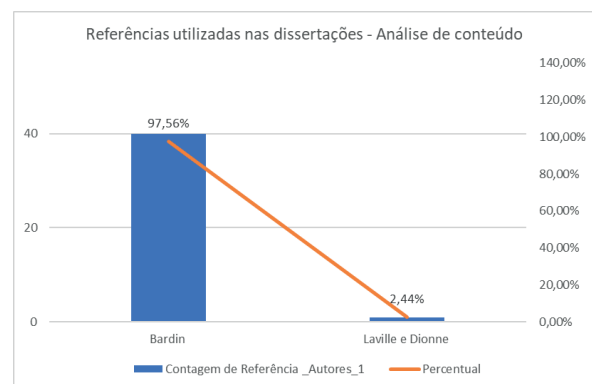
Gráfico 2. *Dissertações publicadas no PPGCI/UUEL – Análise de Conteúdo*



Fonte: Elaborado pelos autores (2023).

Dos trabalhos que se basearam em análise de conteúdo para construir suas análises, 97,56% utilizaram a referência de Bardin para compor a análise da pesquisa, enquanto apenas 2,44% se basearam na teoria de Laville e Dionne.

Gráfico 3. *Referências utilizadas nas dissertações – Análise de Conteúdo*



Fonte: Elaborado pelos autores (2023).

Dentre essas dissertações, 23 trabalhos, representando 56,09%, utilizaram referências de dois ou mais autores

que teorizam a análise de conteúdo para compor a análise da pesquisa. Ao todo, foram utilizadas referências de 14 diferentes autores, com destaque para Minayo e Moraes, que foram citados em 13,04% e 17,39% das dissertações respectivamente. Outros autores foram citados, como Alves, Bortolin, Alcará; Chizzotti; Flick e Valentim, cada um com 8,70% das citações. Já Campos; Cavalcante; Franco; Gil; Marconi, Lakatos; Mozzato, Grzybovski; Santos e Triviñus foram citados em 4,35% das dissertações cada. Conforme visualizamos no Gráfico 4.

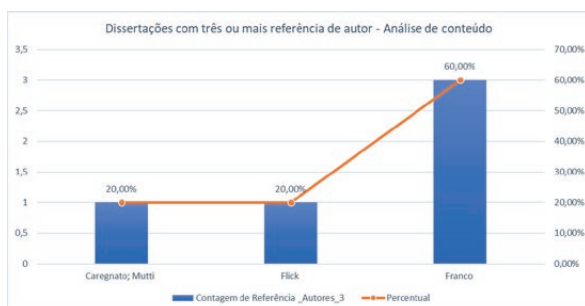
Gráfico 4. *Dissertações com mais de uma referência de autor – Análise de conteúdo*



Fonte: Elaborado pelos autores (2023).

Observamos que 5 trabalhos, o equivalente a 12,19%, utilizaram referências de três ou mais autores que teorizam a análise de conteúdo para compor a análise da pesquisa. Dentre esses trabalhos, verificou-se que 20% utilizaram as referências de Caregnato e Mutti, outros 20% utilizaram as referências de Flick, enquanto 60% utilizaram as referências de Franco. Conforme visualizamos no Gráfico 5.

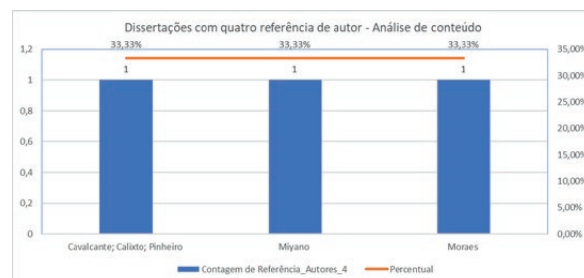
Gráfico 5. *Dissertações com três ou mais referências de autor – Análise de conteúdo*



Fonte: Elaborado pelos autores (2023).

Por fim, após análise das dissertações, constatamos que apenas 3 trabalhos, o equivalente a 7,31%, utilizaram quatro referências de autores que teorizam a análise de conteúdo para compor a análise da pesquisa. Nesses trabalhos, foi verificado que 33,33% utilizaram as referências de Cavalcante, Calixto e Pinheiro, outros 33,33% utilizaram as referências de Miyano, enquanto 33,33% utilizaram as referências de Moraes. Conforme visualizamos no Gráfico 6.

Gráfico 6. *Dissertações com quatro referências de autor – Análise de conteúdo*



Fonte: Elaborado pelos autores (2023).

A partir do processamento dos dados coletados, é possível observar que há a predominância de uma referência de autor prevalece. A presença de uma única referência bibliográfica nos trabalhos analisados pode ser considerada um limitante para a construção de um diálogo científico mais amplo e profundo.

2.2 Contribuições dos múltiplos autores na validade e confiabilidade da pesquisa científica

Segundo Lakatos e Marconi (2003, p. 80) a ciência se constitui como “[...] uma sistematização de conhecimentos, um conjunto de proposições logicamente correlacionadas sobre o comportamento de certos fenômenos que se deseja estudar [...]”. A ciência se constitui por um conjunto de proposições baseadas em observações empíricas, análises críticas e reflexões teóricas, e buscam descrever e explicar os comportamentos dos fenômenos estudados.

Köche (2003, p. 132) ressalta que “O importante é notar os pontos positivos e negativos nas teorias analisadas, inter-relacionando-as umas com as outras.” Ao analisar diferentes teorias, o objetivo é identificar os pontos fortes e fracos de cada uma, bem como suas limitações e contribuições para o entendimento do fenômeno em questão.

Na observação de Gil (2008), é importante referenciar diferentes autores sobre o mesmo tema pode ser útil para contextualizar e embasar teoricamente uma pesquisa, além de oferecer um panorama mais abrangente das questões envolvidas.

Godoy (1995) destaca a importância de se ter referências bibliográficas múltiplas em uma mesma produção científica, pois isso permite a abertura de diálogos entre diferentes autores e correntes teóricas, sem prejudicar a coerência e a coesão do trabalho.

Seguindo a mesma premissa, Chalmers (1995) destaca que, ao escrever um trabalho científico, é importante referenciar as fontes mais relevantes e significativas para a sua pesquisa, mas também é importante considerar outras fontes que possam fornecer uma perspectiva mais ampla sobre o tema.

Em síntese, utilizar múltiplos autores em um trabalho científico, em particular na metodologia, é fundamental para enriquecer a discussão do tema, apresentar diferentes perspectivas teóricas e metodológicas, aumentar a validade e a confiabilidade do trabalho. Por isso, é recomendado que o referencial teórico do trabalho científico apresente mais de um autor para compor a discussão, inclusive no elemento “Metodologia” da pesquisa. Nesse sentido, desejamos compreender de que forma a utilização de múltiplos autores na metodologia de um trabalho científico pode contribuir para a validade e confiabilidade da pesquisa e como isso pode enriquecer a discussão do tema em diferentes perspectivas teóricas e metodológicas.

2.3 Discursos e relação de poder na área da Ciência da Informação: uma perspectiva crítica

Para compreendermos a questão anterior, precisamos ter em mente que a Ciência da Informação é uma área do conhecimento que abrange a produção, aquisição, organização, disseminação e uso da informação em um contexto social, político e cultural específico, que está intrinsecamente relacionado aos sistemas de poder e controle presentes em nossa sociedade. Nesse sentido, ela é permeada de discursos e relações de poder. E para entendermos estas questões, vamos fazer uma leitura das concepções de Foucault (1996, p. 9), a respeito da construção dos discursos:

[...] suponho que em toda sociedade a produção do discurso é ao mesmo tempo controlada, selecionada, organizada e redistribuída por certo número de procedimentos que têm por função conjurar seus poderes e perigos, dominar seu acontecimento aleatório, esquivar sua pesada e temível materialidade.

Para ele, a produção do discurso em uma sociedade é submetida a uma série de procedimentos que visam controlar e selecionar a maneira como as ideias são expressas e difundidas. Ele argumenta que o discurso é a linguagem entendida a partir das relações de poder nas quais estão inseridas, e que podem ser usados para influenciar e transformar a sociedade, mas também pode representar um perigo, uma vez que pode ser usado para questionar e desestabilizar o *status quo*. Sendo assim, o discurso pode ser utilizado como uma ferramenta para influenciar e transformar uma sociedade. E essas relações de compõem o discurso também pode ser considerado perigoso, visto que, pode ser usado para questionar a maneira como as coisas são estabelecidas e mantidas na sociedade.

O conceito de poder para Foucault é fundamental para entendermos a questão que desejamos analisar neste texto. Segundo Castro (2009, p.326), para Foucault “[...] o poder não é uma substância ou uma qualidade, algo que se possui ou se tem; é, antes, uma forma de relação.” Ou seja, o poder não é uma coisa que as

pessoas ou instituições possuem, mas sim uma relação que existe entre elas. Ele ressalta que o poder é algo que está presente em todas as relações sociais, sejam elas políticas, econômicas ou culturais. Dessa forma, é importante entender que o poder não é algo que possa ser tomado ou entregue, mas é algo que está em constante negociação e disputa entre as pessoas e grupos que o exercem.

Como observa Foucault (2010, p. 185): “Temos que deixar de descrever sempre os efeitos de poder em termos negativos: ele ‘exclui’, ‘prime’, ‘recalca’, ‘censura’, ‘abstrai’, ‘mascara’, ‘esconde’. Na verdade, o poder produz; ele produz realidade; produz campos de objetos e rituais da verdade.” Para ele, o poder não é apenas um elemento repressor, mas também capaz de produzir realidades, verdades e sujeitos. Ele defende que é necessário abandonar a descrição dos efeitos do poder em termos negativos, como exclusão, repressão, censura, entre outros. Ao invés disso, é preciso compreender que o poder é produtivo, criando campos de objetos e rituais da verdade que moldam a forma como entendemos o mundo ao nosso redor. Dessa forma, o poder é uma força presente em todas as relações sociais e não deve ser visto apenas como algo negativo, mas como algo que tem a capacidade de moldar e criar realidades.

2.4 Disciplina: procedimento de controle de discurso

Ao nos encontrarmos diante dessa situação de análise, compreendemos ser instigante refletir sobre o procedimento de controle do discurso chamado por Foucault de “Disciplina”. Segundo o autor (1996, p. 30), para caracterizar-se disciplina como controle do discurso “[...] é preciso, pois, que haja possibilidade de formular, e de formular indefinidamente, proposições novas[...]”. Ele, refere-se à importância da produção de proposições novas dentro de um determinado campo do conhecimento. Observando a necessidade de que haja sempre a possibilidade de se formular novas ideias e conceitos, que possam enriquecer e expandir o conhecimento já existente. Foucault (1996, p.30) reflete ainda que “uma disciplina não é a soma de tudo o que pode ser dito de verdadeiro sobre alguma coisa; não é nem mesmo o conjunto de tudo o que pode ser aceito, [...] em virtude de um princípio de coerência[...]”. Ressalta-se assim, a importância da produção do conhecimento como um processo contínuo e dinâmico, que envolve a formulação constante de novas ideias e conceitos. Isso implica em uma ruptura com a ideia de que o conhecimento é algo acabado e estático, que pode ser transmitido de forma definitiva e inquestionável.

Sob este prisma, Foucault (1996, p. 30), observa que a disciplina “[...] é aquilo que é requerido para a construção de novos enunciados.” Para o autor, é importante compreender de que maneira a disciplina se relaciona com a produção de conhecimento e com a

formação de subjetividades. Ao mencionar “novos enunciados”, ele está se referindo à possibilidade de produção de novas ideias, conceitos e teorias. Visto que, para que esses novos enunciados possam surgir, é necessário que haja um campo disciplinar que estabeleça as regras, as normas e os critérios que vão permitir que essas ideias sejam formuladas de maneira coerente e sistemática. Sendo assim, a disciplina funciona como uma espécie de moldura que delimita o campo do discurso e que orienta a produção de conhecimento. Ela estabelece os limites do que é aceitável e do que não é, o que pode ser dito e o que não pode ser dito, o que é verdadeiro e o que é falso.

Deste modo, a disciplina não é apenas procedimentos de controle, mas sim uma prática social que está presente em diversas esferas da vida, como na educação, na ciência, na religião, na política, na mídia, entre outras. Além disso, a disciplina não é algo estático e imutável, mas sim um processo histórico e contingente, que está em constante transformação e adaptação às condições sociais e culturais.

Compreendemos que a escolha de apenas uma referência teórica, no elemento “Metodologia” do trabalho acadêmico, é um fator que reflete uma limitação na forma como o discurso e o conhecimento são estabelecidos em torno de um determinado tema ou objeto de estudo. Nesse sentido, ocorre uma exclusão de outras perspectivas, discursos e maneiras de conhecimento que poderiam enriquecer e ampliar o entendimento do objeto em questão.

Diante disso, entendemos que essa escolha limitada pode estar relacionada a um controle que molda a forma como pensamos e experimentamos o objeto de estudo. Destacamos que essa limitação não se trata apenas de uma censura ou proibição direta, mas sim de um mecanismo que influencia e estabelece as maneiras como lidamos com o conhecimento e a produção acadêmica.

É importante salientar que a escolha do procedimento de controle do discurso “disciplina”, visa explorar uma faceta, diante da complexidade dos procedimentos de controle do discurso nas relações de poder, seguindo a leitura de Foucault. O procedimento de controle do discurso, denominado “disciplina”, foi escolhido como foco da investigação em conformidade com a concepção de Foucault (1996, p. 32) no qual afirma que “A disciplina determina as condições que uma determinada proposição deve cumprir para entrar no campo do verdadeiro: estabelece de que objetos se deve falar, que instrumentos conceituais ou técnicas há que utilizar, em que horizonte teórico deve inscrever-se.” Essa escolha visa explorar um aspecto, dentre outras camadas possíveis e complexas dos procedimentos de controle do discurso nas relações de poder, a partir do estudo das práticas disciplinares e seus efeitos no campo do conhecimento.

Na observação de Castro (2003, p. 111), Foucault evidencia que, no final do século XVIII, com o surgimento de novas formas de produção e de exigências econômicas, tornou-se necessário ordenar esse campo (disciplina), e isso instaurou uma luta econômico-política em torno dos saberes, diante disso, conforme o autor, “[...] É nessa e por essa luta também que surge a universidade moderna: seleção de saberes, institucionalização do conhecimento e, conseqüentemente, o desaparecimento do *sábio-amateus*.”¹ Segundo Foucault, com o surgimento da universidade moderna, houve uma institucionalização do conhecimento, o que resultou na seleção e organização dos saberes em disciplinas. Essa prática, por sua vez, levou a um controle do discurso e do conhecimento. Destacamos que o procedimento de controle do discurso “disciplina”, na academia, permite uma organização do conhecimento, por meio da seleção e classificação de saberes em diferentes áreas, uma maior especialização dos profissionais, bem como uma facilidade na transmissão de conhecimentos de maneira estruturada. Por outro lado, esse procedimento de controle do discurso, incluem o risco de se estabelecer um controle excessivo do discurso e do conhecimento, o que poderia limitar a produção de ideias inovadoras e críticas, pode criar hierarquias e exclusões, restringir o acesso ao conhecimento a determinados grupos e limitar a diversidade de perspectivas e abordagens.

Ao reconhecer a existência desse procedimento de controle do discurso “disciplina” e trabalhar para desafiar-los, é possível ampliar a compreensão dos discursos e enriquecer a produção acadêmica de forma mais plural e crítica. A inclusão de diferentes perspectivas e autores pode trazer novos olhares sobre um mesmo tema, e assim, contribuir para um conhecimento mais amplo e complexo.

2.5 A importância dos múltiplos autores na organização do conhecimento

A organização do conhecimento, na Ciência da Informação, dedica-se ao estudo e à prática dos sistemas e metodologias para representar e recuperar o conhecimento em diferentes domínios e contextos. O conhecimento, na observação de Bräscher (2012, p. 10), é entendido como “o resultado da interação entre a informação e o indivíduo,” ou seja, a informação que foi assimilada e transformada pelo sujeito. Para que o conhecimento possa ser compartilhado e reutilizado, é preciso organizá-lo e representá-lo de forma sistemática e padronizada. Para isso, a organização do conhecimento se utiliza de, segundo Carlan e Medeiros (2011, p. 54), “estruturas conceituais que permitem a organização e a representação do conhecimento de forma sistemática e padronizada,” denominados de sistemas de organização do conhecimento. A eficiência e a multiplicidade das concepções da organização do

conhecimento dependem da diversidade da produção desse conhecimento, especialmente no âmbito acadêmico. Portanto, é imprescindível que esse conhecimento seja construído com base em uma abordagem multidisciplinar e em um diálogo crítico com a comunidade científica.

3 Conclusão

A investigação realizada teve como objetivo central o mapeamento dos métodos de análise utilizados nas pesquisas produzidas no Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Universidade Estadual de Londrina (UEL). Através da análise das dissertações produzidas, foi possível identificar que a técnica de análise mais utilizada pelos pesquisadores foi a análise de conteúdo. Nesse sentido, compreendemos a relevância da análise de conteúdo como um método de pesquisa utilizado na Ciência da Informação.

Ao abordarmos a concepção de disciplina na pesquisa, desejamos compreender como o discurso e o conhecimento estabelecem limites em torno do que pode ser dito e pensado sobre um determinado objeto. Esse controle do discurso pode levar a uma homogeneização do pensamento acadêmico e à exclusão de perspectivas menos conhecidas ou marginais. Tendo como consequência a restrição de novas proposições e impedir a emergência de perspectivas críticas ou alternativas.

Diante da análise, concluímos que a não utilização de múltiplos autores em trabalhos científicos, especificamente na “Metodologia”, podem ser metodologicamente entendida como um procedimento de controle do discurso “disciplina”, não exclusivamente, visto que, como explicamos brevemente no texto, na análise do discurso há diversas camadas de possibilidade de investigação, pois os procedimentos de controle do discurso não estão desconectados de relações diversas de poder. Sendo assim, ao limitar a escolha de autores em uma metodologia, o pesquisador pode estar criando barreiras que impedem a análise e o entendimento mais amplo e plural do objeto estudado. Ao escolher trabalhar apenas com um autor, o pesquisador limita o espectro de abordagens teóricas possíveis e impõe uma visão específica ao seu trabalho, deixando de lado outras perspectivas que poderiam enriquecer a análise.

Notas

[1] No passado, o *sábio-amateur* era aquele que produzia e divulgava o conhecimento de maneira livre e independente

Referências

- BRÄSCHER, M. Organização do conhecimento. **Ciência da Informação**, v. 43, n. 1, p. 9-20, 2012.
- CAPURRO, Rafael; HJORLAND, Birger. O Conceito de informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.12, n.1, p. 148-207, jan./ abril. 2007.
- CARLAN, E.; MEDEIROS, M. B. B. Sistemas de organização do conhecimento na visão da ciência da informação. **Revista Ibero-Americana de Ciência da Informação**, v. 4, n. 2, p. 53-73, 2011.
- CASTRO, Edgardo. **O vocabulário de Michel Foucault**. Belo Horizonte: Autêntica, 2009.
- CHALMERS, Alan F. **O que é ciência afinal?** São Paulo: Brasiliense, 1995.
- DEMO, Pedro. **Metodologia do Conhecimento Científico**. São Paulo: Atlas, 2009.
- FOUCAULT, Michel. **A ordem do discurso**. São Paulo: Ed. Loyola, 1996.
- FOUCAULT, M. **Microfísica do poder**. 23. ed. São Paulo: Graal, 2004.
- FOUCAULT, Michel. **Vigiar e punir**. Nascimento da prisão, trad. Raquel Ramallete. Petrópolis: Vozes, 2010.
- GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 7. ed. São Paulo: Atlas, 2022.
- GODOY, Arilda Schmidt. **Introdução à pesquisa qualitativa e suas possibilidades**. Revista de Administração de Empresas, São Paulo, v. 35, n. 2, p. 57-63, mar./abr. 1995.
- GOLDENBERG, Mirian. **A Arte de Pesquisar**. 11ª Ed. Rio de Janeiro: Record, 2009.
- KÖCHE, José Carlos. **Fundamentos de Metodologia Científica**: Teoria da Ciência e Iniciação à pesquisa. 21ª ed. Rio de Janeiro, 2003.
- LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.
- MORAES, R. Análise de conteúdo. **Revista Educação**, Porto Alegre, v. 22, n. 37, p.7-32, 1999. Disponível em: <https://edisciplinas.usp.br/pluginfile.php/4125089/mod_resource/content/1/Roque-Moraes_Analise%20de%20conteudo-1999.pdf>. Acesso em 03 de maç. de 2023.
- PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO – PPGCI/UEL. **Apresentação**. Disponível em: <http://www.uel.br/pos/ppgci/portal>. Acesso em: 27 maio 2023.

Aspectos culturais e a teoria dos objetos de fronteira: aproximações em busca de caminhos mais inclusivos na Organização do Conhecimento

Cultural aspects and the theory of boundary objects: approaches searching for more inclusive paths in the Knowledge Organization

Nina G. S. Barcellos D'Almeida (1), Linair Maria Campos (2), Suellen Oliveira Milani (3)

Universidade Federal Fluminense (1) mninabarcellos@id.uff.br (2) lmcamos@id.uff.br (3) suellenmilani@id.uff.br

Resumo

Introdução: Os sistemas de organização do conhecimento são largamente utilizados como instrumentos para recuperação da informação contida nos documentos. Porém, na literatura da Ciência da Informação encontramos diversos relatos e pesquisas que expressam a potencialidade desses instrumentos em reproduzir preconceitos sociais ou a predominância de grupos específicos sobre os demais. Muitas das referências encontradas debatem esses problemas os relacionando com a ética envolvida no fazer profissional. Nos últimos anos assistimos o crescimento de estudos cooperativos, onde integrantes de diferentes áreas do conhecimento se debruçam sobre o mesmo tema ou objeto com a finalidade de resolver problemas ou apresentar inovações. Esse contexto traz novas características para a Ciência da Informação e novos desafios para a organização do conhecimento. **Objetivo:** Indicar aspectos teórico-metodológicos que destaquem a Teoria dos Objetos de Fronteira como aliada para elaboração dos sistemas de organização do conhecimento, garantindo a construção de diferentes caminhos de acesso à informação para diferentes comunidades de usuários de forma equitativa. **Metodologia:** O presente trabalho foi desenvolvido por pesquisa exploratória, em relação aos objetivos, pesquisa qualitativa em relação à abordagem do problema e por pesquisa bibliográfica pelo ponto de vista do método. **Resultados:** São sistematizados aspectos da Teoria dos Objetos de Fronteira que colaboram com a elaboração dos sistemas de organização do conhecimento. **Conclusão:** A Teoria indicada pode ser utilizada como ferramenta teórica-metodológica para garantir a construção de diferentes caminhos de acesso à informação para diferentes comunidades de usuários de forma equitativa.

Palavras-chave: Sistemas de organização do conhecimento. Teoria dos Objetos de Fronteira. Comunidades de usuários. Perspectiva cultural.

Abstract

Introduction: Knowledge organization systems are widely used as tools for retrieving information contained in documents. However, in the literature of Information Science, we find several reports and research that express the potential of these tools to reproduce social prejudices or the predominance of specific groups over others. Many of the references found discuss these problems by relating them to the ethics involved in professional practice. In recent years, we have seen the growth of cooperative studies, where members from different areas of knowledge come together to solve problems or present innovations. This context brings new characteristics to Information Science and new challenges for knowledge organization. **Objective:** Indicate theoretical-methodological aspects that highlight the Boundary Objects Theory as an ally for the elaboration of knowledge organization systems, guaranteeing the construction of different paths of access to information for different user communities in an equitable way. **Methodology:** This work was developed through exploratory research in relation to objectives, qualitative in relation to problem approach and bibliographic research from the method point of view. **Results:** Aspects of Boundary Objects Theory that collaborate with the elaboration of knowledge organization systems are systematized. **Conclusion:** The Theory can be used as a theoretical-methodological tool to guarantee the construction of different paths of access to information for different user communities in an equitable way.

Keywords: Knowledge organization systems. Boundary Objects Theory. User communities. Cultural perspective.

Resumen

Introducción: Los sistemas de organización del conocimiento se utilizan ampliamente como herramientas para recuperar información contenida en documentos. Sin embargo, en la literatura de la Ciencia de la Información encontramos varios informes e investigaciones que expresan el potencial de estas herramientas para reproducir prejuicios sociales o la predominancia de grupos específicos sobre otros. Muchas de las referencias encontradas discuten estos problemas relacionándolos con la ética involucrada en la práctica profesional. En los últimos años, hemos visto el crecimiento de estudios cooperativos, donde miembros de diferentes áreas del conocimiento se unen para resolver problemas o presentar innovaciones. Este contexto trae nuevas características a la Ciencia de la Información y nuevos desafíos para la organización del conocimiento. **Objetivo:** Indicar aspectos teórico-metodológicos que destaquen la Teoría de los Objetos Frontera como aliada para la elaboración de sistemas de organización del conocimiento,

garantizando la construcción de diferentes caminos de acceso a la información para diferentes comunidades de usuarios de manera equitativa. **Metodología:** Este trabajo se desarrolló a través de investigación exploratoria en relación con los objetivos, cualitativa en relación con el enfoque del problema y bibliográfica desde el punto de vista del método. **Resultados:** Se sistematizan aspectos de la Teoría de los Objetos Frontera que colaboran con la elaboración de sistemas de organización del conocimiento. **Conclusiones:** La Teoría puede ser utilizada como herramienta teórico-metodológica para garantizar la construcción de diferentes caminos de acceso a información para diferentes comunidades de usuarios de manera equitativa.

Palabras clave: Sistemas de organización del conocimiento. Teoría de los Objetos Frontera. Comunidades de usuarios. Perspectiva cultural.

1 Introdução

Os sistemas de organização do conhecimento podem ser entendidos como estruturas sistemáticas que visam à construção de modelos abstratos do mundo real (CARLAN, 2010, p. 29) e são elaborados a partir da representação do conhecimento, por um processo de modelagem conceitual, que Mylopoulos (1992, p. 20) define como “a atividade de descrever formalmente o mundo físico e social ao nosso redor com o propósito de compreensão e comunicação” (MYLOPOULOS, 1992, p. 20).

Os sistemas de organização do conhecimento são utilizados nas atividades de representação temática de documentos e objetos informacionais, auxiliando os processos de indexação e classificação realizados pelos profissionais. Porém, vários relatos encontrados na literatura da Ciência da Informação e, mais especificamente, da Organização do Conhecimento expressam que esses instrumentos podem dificultar ou inviabilizar o acesso a conhecimentos específicos, ou então reproduzir preconceitos e marginalizações sociais em sua estrutura classificatória, o que Milani (2014) chamam de *biases*, ou seja, quando o instrumento reflete preconceitos e tendenciosidades.

Um exemplo é citado no relato de Wiebke Keim (2008) acerca das bases de dados internacionais para recuperação de trabalhos em sociologia. Keim (2008) identifica, através da análise da visibilidade da produção sociológica nacional em três bases de dados internacionais, que “os textos norte-americanos ficam com a maior parte das publicações, enquanto o resto do mundo é relegado a uma visibilidade extremamente limitada” (KEIM, 2008, p. 2). Ainda em sua pesquisa sobre as bases de dados, Keim (2008) revela, através de uma verificação cruzada com o banco de dados alternativo da UNESCO, que esta falta de visibilidade dos dados de outros países não reflete graus objetivos sobre os índices de publicação internacional, como por exemplo, sobre os periódicos africanos de ciências sociais ou então, livros chineses sobre o assunto. Segundo o autor:

[...] os dados da UNESCO para publicação de livros revelam que a China é, mundialmente, o primeiro produtor de literatura de ciências sociais, uma realidade que não se reflete em nenhum dos três índices (KEIM, 2008, p. 2).

Por fim, o autor conclui que: “essas bases de dados revelam-se, portanto, não apenas um indicador de

marginalidade, mas também um instrumento de marginalização” (KEIM, 2008, p. 2).

No exemplo de Keim (2008) fica nítido que um sistema de organização do conhecimento pode, não só reproduzir uma marginalização social, como também ser utilizado como meio de reprodução desta marginalização, dificultando e inviabilizando o acesso a determinados conhecimentos. Outro exemplo é citado em Milani (2010, p. 103), onde é observado que o Vocabulário Controlado USP “disponibilizou apenas o descritor *História da Mulher* inserido abaixo do descritor *Historiografia*”. Neste exemplo, um enfoque é dado à “[...] história especificamente da mulher, a linguagem não considera a participação das mulheres na história como um todo, o que, atualmente, não é mais aceito” (MILANI, 2010, p. 103).

Podemos dizer que nos dois exemplos os sistemas de organização do conhecimento veiculam tendenciosidades. Enquanto no exemplo de Keim (2008) fica nítido a marginalização de grupos sociais relacionados aos temas de raça e etnicidade, no exemplo de Milani (2010) fica nítida a influência de preconceitos sociais relacionados às questões de gênero no processo de elaboração da estrutura conceitual do Vocabulário Controlado USP, onde consideraram a ‘História da Mulher’ como categoria a parte da ‘Historiografia’. Assim, o papel da mulher como parte da historiografia geral fica às margens, não viabilizando, inclusive, a representação da historiografia geral por seus agentes e participantes totais, comprovado pela ausência de um descritor intitulado “História do Homem”.

Os problemas relacionados a veiculação de tendenciosidades nos sistemas de organização do conhecimento envolvem tanto o fazer profissional em forma de dilemas éticos, quanto a elaboração dos instrumentos utilizados pelos profissionais. Nos últimos anos, os estudos relacionados a gênero, sexualidade, raça, etnicidade, religião e linguagem aumentaram exponencialmente, tangenciando diversas áreas do conhecimento e fomentando a cooperação entre pessoas envolvidas em diferentes disciplinas.

Destaca-se também outros temas que vêm envolvendo a contínua cooperação entre diversas áreas do conhecimento, como estudos relacionados ao combate das *fake news*, desenvolvimento tecnológico e algumas tendências sobre as pesquisas relacionadas a ética. Dessa forma, o crescente envolvimento de diferentes

comunidades para a produção do conhecimento implica na colaboração de pessoas com diferentes formações e compreensões terminológicas. Esse contexto traz novos desafios para a organização do conhecimento. Nesse sentido, vale-se o questionamento: como a Teoria dos Objetos de Fronteira pode ser aliada no combate às tendenciosidades e preconceitos na elaboração e manutenção de sistemas organização do conhecimento?

Este trabalho é motivado pela necessidade de se elaborar sistemas de organização do conhecimento que colaborem com a pesquisa e trabalho cooperativo e pela preocupação com a construção desses instrumentos de forma inclusiva, não reproduzindo preconceitos e/ou marginalizações sociais. O objetivo deste trabalho é indicar aspectos teórico-metodológicos que indiquem como a Teoria dos Objetos de Fronteira pode ser uma aliada para a elaboração de sistemas de organização do conhecimento, garantindo a construção de diferentes caminhos de acesso à informação para diferentes comunidades de usuários de forma equitativa.

O presente trabalho foi desenvolvido a partir de uma pesquisa exploratória, em relação aos objetivos, pesquisa qualitativa em relação à abordagem do problema e por pesquisa bibliográfica pelo ponto de vista do método. A pesquisa bibliográfica foi realizada na base de dados Brapci e no site da isko.org, utilizando o termo “objetos de fronteira” e “sistemas de organização do conhecimento”. A partir dos artigos recuperados foi efetuada a leitura do resumo, título e palavras-chaves e foram selecionados os que atenderam ao propósito delineado pelo objetivo geral deste trabalho. Também foram especialmente considerados aqueles trabalhos voltados para elaboração e utilização de sistemas de organização do conhecimento por diferentes comunidades de usuários ou que envolviam contextos cooperativos. O trabalho também utiliza a estratégia de ‘cultivo de pérolas’ ou ‘mineração de citação’, de Hawkins e Wagers (1982) e Hadfield (2019), que indica a identificação de conceitos-chaves de pesquisa nos documentos iniciais e das citações e referências utilizadas em relação ao tema, no intuito de ampliar a quantidade de documentos e informações, aprofundando o conhecimento sobre o tema. Segundo Hadfield (2019), a abordagem é eficaz para a pesquisa sistemática da literatura, ajudando a garantir que toda a literatura relevante sobre um determinado tema tenha sido identificada.

2 Aspectos culturais na Organização do Conhecimento

Os processos e os instrumentos de organização e representação do conhecimento utilizados para a

criação de produtos e serviços que visam promover o acesso e a apropriação da informação das comunidades de usuários são elaborados e avaliados teórica e metodologicamente no âmbito da Organização do Conhecimento. Ao acessar e se apropriar de informações e documentos, esses indivíduos terão elementos para exercer plenamente a sua cidadania, fomentando sua participação nos fluxos de informação e, conseqüentemente, de outros fluxos como os de direitos sociais, riqueza, poder, entre outros.

Por lidar com informações, sistemas de representação e pessoas, elementos do conceito de cultura perpassam a Organização do Conhecimento de forma transversal.

Vinda do verbo latino *colere*, na origem cultura significa o cultivo, o cuidado. Inicialmente, era o cultivo e o cuidado com a terra, donde agricultura, com as crianças, donde puericultura, e com os deuses e o sagrado, donde culto. Como cultivo, a cultura era concebida como uma ação que conduz a plena realização das potencialidades de alguma coisa ou de alguém; era fazer brotar, frutificar, florescer e cobrir de benefícios. No correr da história do ocidente, esse sentido foi-se perdendo até que, no século XVIII, com a Filosofia da Ilustração, a palavra cultura ressurgiu, mas como sinônimo de um outro conceito, torna-se sinônimo de civilização. Sabemos que civilização deriva da ideia de vida civil, portanto, de vida política e de regime político. Com o Iluminismo, a cultura é o padrão ou o critério que mede o grau de civilização de uma sociedade. Assim, a cultura passa a ser encarada como um conjunto de práticas (artes, ciências, técnicas, filosofia, os ofícios) que permite avaliar e hierarquizar o valor dos regimes políticos, segundo um critério de evolução. No conceito de cultura introduz-se a ideia de tempo, mas de um tempo muito preciso, isto é, contínuo, linear e evolutivo, de tal modo que, pouco a pouco, cultura torna-se sinônimo de progresso. Avalia-se o progresso de uma civilização pela sua cultura e avalia-se a cultura pelo progresso que traz a uma civilização (CHAUÍ, 2008, p. 55).

O termo cultura é polissêmico e, atualmente, conforme explica Lee (2015), pode expressar: nacionalidade ou região geográfica; contexto; fenômeno coletivo; ou criações humanas no ambiente. Quando se reflete sobre a perspectiva cultural na Organização do Conhecimento, um enfoque pode ser dado a cada uma das facetas do conceito de cultura, que pode ser pensada na perspectiva das informações e documentos que serão representados, isto é, a partir da coleção, dos sistemas de organização do conhecimento, enquanto manifestações de cultura que podem ser influenciados e influenciar visões de mundo, bem como dos atores da Organização do Conhecimento, incluindo os bibliotecários, criadores e mantenedores de sistemas de organização do conhecimento, entre outros.

A literatura, principalmente em fontes de informação advindas da ISKO Internacional e de periódicos como *The Indexer e Cataloging and Classification Quarterly*, vem sinalizando que ao se refletir sobre uma dimensão cultural da Organização do Conhecimento está se

falando, também, sobre a estruturação de sistemas de organização do conhecimento, que podem causar danos aos usuários e prejudicar o seu acesso à informação. Esses danos podem ser causados pelo reforço ou marginalização de determinadas perspectivas ou culturas, conforme ensina Olson (1999).

Assim, é fundamental garantir reflexões atualizadas sobre os danos ou impactos que os sistemas de organização do conhecimento podem fomentar às comunidades de usuários das bibliotecas e sistemas de informação em busca de soluções contextualizadas em um domínio, tempo e espaço.

Por coexistirem muitos aspectos culturais dentro de um indivíduo (etnia, gênero, religião etc.), os quais não são estáticos, haverá opiniões conflitantes sobre as maneiras de organizar e representar informações na prática o que exige dos criadores e mantenedores dos esquemas de classificação, por exemplo, a seleção de um conjunto de aspectos culturais como basilares para sua tomada de decisão, como lembra Lee (2015).

Nesta reflexão, destacamos a dimensão cultural que envolve tanto as atitudes do catalogador, indexador ou classificador frente à informação ou ao documento durante a etapa de análise conceitual ou de assunto, quanto as suas atitudes frente à tarefa de criação ou manutenção de um sistema de organização do conhecimento.

Em ambas as tarefas, enquanto um conjunto de práticas também influenciadas pelos componentes tecnológicos, os aspectos culturais são estruturantes. Nesse contexto, a cultura consiste em

[...] uma estrutura cognitiva construída por uma comunidade. A estrutura influencia como as pessoas dentro de uma comunidade percebem o mundo. Para nós, a cultura é aprendida, dinâmica e coexistente (LEE, 2015, p. 305).

Para realizarmos discussões a partir da perspectiva cultural na Organização do Conhecimento, algumas abordagens são possíveis, inclusive no âmbito da denominada Organização do Conhecimento Crítica, a partir de: teorias e métodos advindos das teorias críticas; estudos éticos a partir de diferentes posturas epistêmicas; estudos terminológicos; análise dos processos de organização e representação do conhecimento a partir de abordagens sociológicas, entre outros.

Um conceito que se destaca nesta seara é o de garantia cultural. A garantia cultural consiste em princípio largamente debatido, a partir de sua primeira utilização, em 1976 por J. M. E. Lee (BARITÉ, 2011). De acordo com Begthol (2002, p. 511), “a suposição de que indivíduos em diferentes culturas precisam de diferentes tipos de (e caminhos para) informação tem sido chamada de ‘garantia cultural’”.

Segundo Barité (2011), a garantia cultural é assegurada quando se reconhece que as classificações e as relações semânticas são dependentes do contexto cultural. Dessa forma, a introdução da garantia cultural como um dos princípios para a apreensão e seleção dos termos que serão incluídos nos sistemas de organização do conhecimento, se coloca como um caminho para um posicionamento ético, frente à necessidade de se levar em consideração o interesse de diferentes grupos de usuários.

Como forma de complementar os princípios apresentados que asseguram a garantia de relevância dos termos selecionados para compor o sistema de organização do conhecimento elaborado, a partir de uma perspectiva cultural, busca-se, nesse trabalho, ressaltar outras orientações epistêmico-metodológicas que podem auxiliar o processo de elaboração desses instrumentos a fim de se levar em consideração a suposição que indivíduos em diferentes culturas precisam de diferentes tipos de (e caminhos para) informação, como já alertava Clare Beghtol (2002, 2005).

Dessa forma, na próxima seção, iremos considerar a Teoria dos Objetos de Fronteira como mais um eixo de pesquisa quando o objetivo é debater a perspectiva cultural na organização do conhecimento, utilizando a visão ecológica do conhecimento.

3 Tendenciosidades e preconceitos nos sistemas de organização do conhecimento

Os sistemas de organização do conhecimento são elaborados para a organização do conhecimento gerado (tanto em forma de documentos quanto em abstração teórica) e se desdobram nas taxonomias, ontologias, tesouros, vocabulários controlados, listas de cabeçalho de assunto, arquivos de autoridade, redes semânticas, dicionários, glossários e nos sistemas de classificação, responsáveis pelas notações dos livros, como, por exemplo, a Classificação Decimal de Dewey e Classificação Decimal Universal (BRASCHER; CAFÉ, 2008).

Em relação aos objetivos dos sistemas de organização do conhecimento, Soergel (1999) enumera: fornecer um mapa semântico para domínios individuais e para os relacionamentos entre domínios oferecendo orientação e servindo como ferramenta de referência; melhorar a comunicação, ensino e aprendizagem; fornecer base conceitual para execução e implementação de pesquisas e fornecer uma classificação para a ação – uso prático dos SOC, apoiando e permitindo a recuperação da informação e de documentos.

Os sistemas de organização do conhecimento assim, podem ser utilizados para indexação de documentos

em diversos suportes, tanto convencionais, quanto digitais, visando à recuperação das informações contidas nesses documentos, e podem ser utilizados como facilitadores da produção do conhecimento (D'ALMEIDA, 2022). No âmbito das bibliotecas por exemplo, os sistemas de organização do conhecimento podem ser elaborados, representando o assunto do livro ou documento, gerando seu substituto documental, que pode ser um cabeçalho de assunto, catálogo etc. onde estará indicado a localização do material físico, sua descrição física ou representação temática.

Segundo Milani e Guimarães (2017, p. 73):

Quando este processo [representação de assunto] não é realizado de forma defensável, pode-se promover, censurar, omitir e distorcer informações. Assim, pode-se privar um usuário de obter informações ou de conhecer o ponto de vista de um autor sobre um assunto.

Ainda segundo os autores,

Quando os substitutos documentais refletem preconceitos, tendenciosidades, vieses por meio dos descritores de indexação, notações de classificação, diz-se que esses substitutos veiculam *biases* (MILANI; GUIMARÃES, 2017, p. 74).

Assim, um sistema de organização do conhecimento veicularia inclinações ou preconceitos quando possui a tendência em evocar os valores e crenças de quem os criou, causando danos e/ou benefícios a comunidades específicas de usuários. Dessa forma, compreende-se que os interesses das comunidades de usuários não foram levados em consideração para a elaboração do instrumento, ou então, que somente os interesses de uma ou algumas comunidades de usuários foram levados em consideração.

Nesse sentido, a atuação profissional que atua na organização e representação do conhecimento ao se deparar com essas questões, deve promover espaços de negociação entre os sistemas de organização do conhecimento e o acesso de diferentes comunidades de usuário que dele farão uso.

4 A importância da interdisciplinaridade e da Teoria dos Objetos de Fronteira

De acordo com Yeganeh, Su e Chrysostome (2004), as dificuldades encontradas em pesquisas ou trabalhos cooperativos interculturais são de natureza variada e podem ir além das limitações metodológicas e práticas. Dessa forma, as diferenças culturais podem se apresentar de forma mais sutil, como diferenças no significado de termos, construção de diferentes relações com objetos a partir da prática e a elaboração de diferentes arranjos sociais a partir dos mesmos objetos de conhecimento.

Knorr-Cetina (1999) utiliza a noção de culturas epistêmicas no contexto dos estudos dos laboratórios

científicos, deslocando o foco da pesquisa sobre os processos sociais da construção do conhecimento científico dos “objetos mentais”, como interesse e estratégia, para se concentrar no estudo sobre a organização, as estruturas e dinâmicas do contexto de produção.

As culturas epistêmicas são consideradas por Knorr-Cetina (1999), como:

[...] amálgamas de arranjos e mecanismos ligados por afinidade, necessidade e histórico coincidente, que em determinado campo, determinam como sabemos o que conhecemos. São culturas que criam e garantem o conhecimento e a principal instituição de conhecimento no mundo ainda é a ciência (KNORR-CETINA, 1999, p.1).

A disciplinaridade inculcada aos sistemas de organização do conhecimento por vezes pode dificultar a compreensão e a inclusão de diferentes comunidades de prática, ou então, de diferentes culturas epistêmicas nos instrumentos que se visa construir. Quando se colocam as presunções na Organização do Conhecimento, comumente se cita a universalidade, o bibliocentrismo, a neutralidade e a imparcialidade das categorias mutuamente exclusivas. As presunções, de acordo com Milani e Guimarães (2017, p. 78), “vêm sendo tomadas como verdades inquestionáveis” e podem causar danos. Elas são citadas na literatura especializada de maneira implícita, no tocante às críticas relacionadas aos sistemas de organização do conhecimento (MILANI; GUIMARÃES, 2017).

Uma questão que não fica nítida é que nem sempre, entre nossas comunidades de usuários, teremos grupos que seguem a mesma orientação epistemológica ou mesmo, disciplinar, para as buscas e recuperação das informações, dado principalmente os novos contextos de ambientes cooperativos para a produção e validação do conhecimento.

Teremos provavelmente, nos grupos de usuário prioritários, certa similaridade entre a orientação teórica, sistema de valores e utilização de termos, porém da mesma forma que podemos ter grupos sociais distintos, que precisam ser respeitados e representados adequadamente, como no caso das mulheres, podemos também ter diferentes configurações epistemológicas. As diferentes configurações epistemológicas podem ser encaradas como culturas epistêmicas, que são formadas a partir da contínua interação entre indivíduos com certo interesse em comum e do estabelecimento de padrões e valores em relação ao conhecimento, sua produção e validação (KNORR-CETINA, 1999). Vale destacar que podemos encontrar diferentes configurações epistemológicas tanto entre diferentes disciplinas, quanto em uma mesma disciplina.

Quando estamos falando de diferentes comunidades de usuários e destacamos diferenças em relação à marginalização e exclusão de grupos específicos, pode

ser válido considerar diferentes pontos de vista entre as comunidades envolvidas para que, por exemplo, não se marginalizem grupos epistêmicos específicos. Podemos, nesse sentido, pensar em um ambiente de trabalho e pesquisa que envolva, por exemplo, criadores amadores de pássaros, acadêmicos e cientistas especializados em pássaros, ambos interessados em saber tanto informações a respeito da biologia que envolve as espécies quanto seus significados sociais e culturais ao longo dos anos. Ou então, em setores de uma empresa envolvendo, por exemplo, equipes responsáveis pela engenharia do produto e de marketing e propaganda, que precisam ter acesso a diferentes informações sobre o processo de produção e impacto social de um produto específico. Exemplos como esses nem sempre vão se beneficiar da divisão disciplinar *a priori* do conhecimento.

Em outras palavras, quando pensamos a cooperação entre sujeitos que possuem diferentes formações e que por vezes serão oriundos de diferentes culturas epistêmicas, destacamos que cada comunidade utiliza *seu próprio sistema de valores*, que precisa ser respeitado e representado quando se faz necessária a elaboração de um sistema de organização do conhecimento. Nesse cenário se faz pertinente a Teoria dos Objetos de Fronteira e a visão ecológica que a teoria se propõe a seguir (D'ALMEIDA; CAMPOS, 2021, D'ALMEIDA, 2022, CAMPOS, 2018).

A Teoria de Objetos de Fronteiras foi proposta inicialmente em 1989, por Leigh Star e Griesemer (1989) em um projeto de estudo sobre práticas informacionais no Museu de Zoologia de Vertebrados de Berkeley, utilizando aportes teóricos da Sociologia (CAMPOS, 2018, p.5). A Teoria propõe a existência de objetos que são pontos de ligação entre comunidades, desde que construídos cooperativamente e respeitando os pontos de vista e necessidades informacionais de cada mundo social (CAMPOS, 2018).

Adotando uma abordagem ecológica, a Teoria admite diversos pontos de vista, sem privilegiar um viés em detrimento dos demais. Baseia-se na teoria ator-rede (CALLON, 1980, LATOUR, 1997, LAW, 1991), no contexto da Sociologia da Ciência e Tecnologia, e se diferencia desta ao adotar uma análise ecológica, onde se admite diferentes pontos de vista, sem que se busque apresentar uma visão específica de um viés sob as demais visões. Dessa forma, sua unidade de análise é tomada como todo, não um ponto de vista específico, permitindo mapeamento diversos (de muitos para muitos). Nesse sentido, não se trata de estabelecer uma visão consensual, que sirva de intermediação para diversos pontos de vista, mas sim sobre pontes entre as representações expressas pelas diferentes linguagens envolvidas em determinado contexto cooperativo.

Os objetos de fronteira servem para mediar interações, ao mesmo tempo em que essas interações ajudam a

construir em conjunto o significado conceitual desses objetos (TROMPETTE; VINCK, 2009). Segundo Star (2010), os objetos de fronteiras surgem em infraestruturas orgânicas com base nas percepções locais de grupos diferentes de usuários que desejam trabalhar de forma cooperativa, das necessidades de informação e de requisitos de trabalho.

Ao evidenciar os objetos, a Teoria dos Objetos de Fronteira sugere que se atente para as possíveis forças invisíveis que atuam no contexto, ou seja, possíveis disputas políticas, ideológicas, jogos de poder, níveis de hierarquia, ou outras forças que possam influenciar o processo de significação e troca de conhecimento.

A Teoria dos Objetos de Fronteira deixa em evidência a existência de comunidades de prática, onde a prática é composta de atividades, tarefas, funções e entendimentos que não existem isoladamente, mas são parte de um sistema mais amplo de relações, no qual possuem significados (AKRICH, 2010). As relações são construídas a partir das interações que circunscrevem aquele contexto.

As comunidades de prática podem ser vistas como sistemas sociais, que podem ser mais simples ou mais complexos, dependendo da sua constituição em termos de interações entre comunidades que abrangem, ou então, como sistemas de aprendizagem social (WENGER, 2010). Para Wenger (2010), o conceito de comunidade de prática não existe por si só, mas:

É parte de uma estrutura conceitual mais ampla para pensar sobre a aprendizagem em suas dimensões sociais. É uma perspectiva que localiza aprender, não na cabeça ou fora dela, mas na relação entre a pessoa e o mundo, que para os seres humanos, se dá com uma pessoa social em um mundo social. Nessa relação de participação, o social e o individual constituem-se mutuamente (WENGER, 2010).

Nesse sentido, as comunidades de prática partilham também um tipo de conhecimento que é incorporado aos seus membros, que permite compartilhar entendimentos sobre o que estão fazendo, o que isso significa, valores, orientações e interpretações, em outras palavras, as comunidades de prática podem envolver e/ou gerar culturas epistêmicas distintas, que criam e garantem o conhecimento, moldando a forma pela qual a comunidade 'sabe' o que sabe (AKRICH, 2010).

Quando consideramos a tendência atual de cooperação entre diferentes áreas do conhecimento, ou entre pesquisadores oriundos de comunidades de prática diferentes, com diferentes formações e vocabulários, encontramos propostas interessantes na literatura de Organização do Conhecimento que parecem despontar aportes teóricos-metodológicos que buscam abarcar diferentes caminhos de acesso à informação para diferentes comunidades de usuários de forma

equitativa. O Manifesto de Leon (2007) elaborado após encontro da ISKO, já destacava que:

A tendência atual para uma crescente interdisciplinaridade do conhecimento exige essencialmente novos sistemas de organização do conhecimento (SOC), baseados em uma revisão substantiva dos princípios subjacentes ao SOC tradicional baseado em disciplina; [...] em vez de disciplinas, as unidades básicas do novo KOS devem ser fenômenos do mundo real tal como é representado no conhecimento humano; o novo SOC deve permitir que os usuários mudem de uma perspectiva ou ponto de vista para outro [...]. Em particular, deve permitir que eles busquem independentemente fenômenos particulares, teorias particulares sobre fenômenos (e sobre relações entre fenômenos) e métodos particulares de investigação; as conexões entre os fenômenos, entre os fenômenos e as teorias que os estudam, e entre os fenômenos e os métodos para investigá-los, podem ser expressas e gerenciadas por técnicas analítico-sintéticas já desenvolvidas na classificação facetada.

Assim, frisamos que o Manifesto já aponta alguns aportes teórico-metodológicos que podem ser utilizados para a elaboração de sistema de organização do conhecimento a fim de contemplar diferentes comunidades, como a utilização de fenômenos ao invés de disciplinas como primeiras divisões nos instrumentos e das técnicas analítico-sintéticas oferecidas pela classificação facetada.

Segundo Beghtol (1998), uma resposta adequada à pesquisa multidisciplinar pode ser explorada através da elaboração de instrumentos construídos pela combinação de estruturas hierárquicas com o princípio analítico-sintético. Para o desenvolvimento deste instrumento, de acordo com Beghtol (1998), seria necessário o uso da análise de facetas para o trabalho sintático e de notação, como vem sendo estudado por diversos grupos desde o trabalho de Ranganathan (1967). Dessa forma, os sistemas de organização do conhecimento abarcariam a representação de múltiplas perspectivas em um mesmo contexto (BEGHTOL, 1998).

Acreditamos que a Teoria dos Objetos de Fronteira e a visão ecológica que ela propõe podem colaborar positivamente para a elaboração de sistemas de organização do conhecimento que contemplem diferenças culturais e sociais, em outras palavras, a criação de instrumentos que sejam verdadeiramente inclusivos[1] e que considerem diferentes comunidades de usuários e a necessidade de construção de diferentes caminhos para que eles acessem o conhecimento e a informação.

Como forma de ilustrar o debatido até agora neste trabalho, utilizamos a lista de autoridades disponível pela Biblioteca Nacional, onde encontramos o termo tópico ‘Etnicismo’, apresentado conforme ilustrado na Figura 1, extraída do sistema Sophia, da Biblioteca Nacional.

Figura 1. *Etnicismo como assunto na Biblioteca Nacional*

Etnicismo (Subdividido geograficamente)
Termo tópico

Remissiva Complexa Ver Também - Assunto
Também as subdivisões - Identidade racial e -Identidade étnica sob nome de raça ou grupos étnicos específicos. Ex. Negros - Identidade racial; Sírios - Rio Grande do Sul - Identidade étnica.

Remissiva Ver (US/UF)
Ethnic identity
Group identity, Ethnic

Remissiva Ver Também (TR)
Identidade social TG
Etnicismo em crianças TE
Negros - Identidade racial TE
Fusão cultural TR
Multiculturalismo TR
Pluralismo cultural TR

Fonte positiva dos dados
LCSH

Nota geral pública
Usado para obras que tratam do senso subjetivo de participação de um grupo étnico específico. Obras sobre o aspecto teórico do conceito de grupos de pessoas que têm seus limites comuns unidos por laços de ancestralidade e de cultura, entram em Grupos étnicos. Obras sobre a disciplina e sobre a...
Ver mais

Fonte: Biblioteca Nacional.

O termo ‘Etnicismo’ está presente na indexação de 51 registros, a maior parte deles relacionados às questões da população negra, conflitos étnico-raciais e identidade étnica. Destaca-se que na seção “Outros vocabulários” consta a *Library of Congress* com o termo ‘*Ethnicity*’, porém em uma tradução oficial o termo em português seria normalmente traduzido para ‘Etnia’ apenas. Sendo ou não uma questão de tradução, encontramos entre suas remissivas Ver Também (TR) o termo ‘Negros – identidade racial’, porém o termo ‘Branços’ só está relacionado ao termo ‘Etnologia’, que no instrumento analisado, no campo de “Outros Vocabulários”, estaria relacionado ao termo ‘*Ethnology*’, na *Library of Congress*. Além disso, o sufixo -ismo é largamente utilizado pelo vocabulário especializado para a denominação de doenças, patologias ou vícios, ficando confusa sua utilização no campo de assuntos, principalmente relacionado à negritude.

Ainda, utilizando a lista de autoridades, destacamos que, no sistema Sophia da Biblioteca Nacional, o termo de entrada ‘Branços’, não possui uma especificação como identidade racial, enquanto a entrada relacionada a ‘Negros’, especifica com ‘Negros – identidade racial’. Fato esse que pode levar o usuário leigo ou desatento a compreensão que enquanto ‘negro’ é uma identidade racial, ‘branco’ não o é.

O campo de estudos que tange os estudos étnico-raciais, conforme já expresso neste trabalho, é um ponto de encontro entre diferentes comunidades como Educação, Psicologia, Medicina, História, entre tantas outras. Logo, a classificação das diferentes etnias é fundamental para orientar a prática profissional de

diferentes formações. Assim, a classificação e denominação das diferentes etnias, de forma a respeitar as características e diferenças de cada uma, sendo construída conjuntamente, pode ser utilizada como objeto de fronteira entre diferentes áreas do conhecimento ou mesmo entre diferentes culturas epistêmicas, colaborando com a produção do conhecimento cooperativo e envolvendo sujeitos com diferentes formações.

Ao enfatizar a existência das comunidades de prática, a Teoria dos Objetos de Fronteira sugere a possível existência de diferentes comunidades, inclusive em contextos de aparente consenso, trazendo a necessidade de investigação do contexto utilizado na elaboração do sistema de organização do conhecimento para a identificação das comunidades envolvidas ou para comprovação de que apenas uma comunidade está presente no contexto.

Tomando como exemplo a elaboração de um cabeçalho de assunto, a utilização da Teoria dos Objetos de Fronteira como teoria complementar aquelas já comumente utilizadas, indicaria a necessidade de investigação sobre o contexto de utilização dos termos associados aos objetos. No caso do cabeçalho de assunto da Biblioteca Nacional, a Teoria dos Objetos de Fronteira, ao aproximar diferentes comunidades, poderia facilitar a identificação da inconsistência em relação aos termos 'Etnicismo' e 'Etnologia', ou então sobre as remissivas 'Branços' e 'Negros – identidade racial'. Essa facilitação se daria por conta da identificação dos objetos e pela validação dos termos e relações pelos representantes das diferentes comunidades envolvida com o contexto, assim, a utilização dos termos 'Etnicismo' e 'Etnologia' passariam pelo crivo de representantes de comunidades, por exemplo, da Psicologia ou Medicina, que teriam maior facilidade em identificar a inconsistência. Ou então, em relação a 'Branços' e 'Negros – identidade racial', que, provavelmente, passariam pelo crivo de representantes das Ciências Sociais e dos Estudos Étnico-raciais e poderiam identificar a inconsistência no que tange à representação de 'identidade racial'.

No último exemplo utilizado, ainda, se destaca a presença de forças invisíveis, no caso, a difundida compreensão social que a população negra possui especificidades em relação a raça e etnia, enquanto a população branca não as possuiria.

A Teoria dos Objetos de Fronteira, conforme exposto neste trabalho, coloca ênfase nos objetos, uma vez que para identificar os objetos de fronteira é necessário conhecer os objetos envolvidos no contexto. Dessa forma, acreditamos que ao identificar os objetos e por sua vez, os objetos de fronteira, um caminho teórico-metodológico para se identificar os objetos que são utilizados por diferentes comunidades é traçado,

facilitando a compreensão de possíveis diferenças em relação a compreensão, significado ou linguagem utilizados para a denominação ou uso desses objetos. Assim, possíveis diferenças na utilização dos conceitos relacionados aos Estudos Étnico-raciais seriam mais facilmente identificadas, diminuindo a possibilidade de reprodução de incoerências e invisibilidades na elaboração dos sistemas de organização do conhecimento.

Com a Teoria dos Objetos de Fronteira, provavelmente, também seria possível a identificação de diferentes e novas relações entre assuntos, uma vez que os objetos estão relacionados a diferentes comunidades. Dessa forma, ao se identificar os objetos de fronteiras, com objetivo de construir sistemas de organização do conhecimento verdadeiramente representativos, estariam representadas no instrumento as diferentes comunidades, facilitando e viabilizando a descoberta do conhecimento em contexto diverso.

A investigação sobre o contexto, em relação à elaboração do sistema de organização do conhecimento, está de certa forma presente quando da etapa de seleção dos termos que farão parte do instrumento. Uma vez que, como visto anteriormente na seção 3, pode ser utilizada, por exemplo, a garantia cultural para atestar a relevância dos termos selecionados. Porém, com a Teoria dos Objetos de Fronteiras, a investigação sobre quantas comunidades existem e quais são elas, provavelmente anteciparia a seleção de termos, facilitando a identificação de diferentes grupos sociais ou políticos que possam estar envolvidos com o contexto.

A seguir, na Tabela 1, serão sistematizados alguns pontos expostos neste trabalho como forma de explicitar aspectos teórico-metodológicos pelos quais a Teoria dos Objetos de Fronteira pode colaborar com a elaboração de sistemas de organização do conhecimento orientados a diferentes caminhos para informação direcionados a diferentes comunidades de usuários.

Em síntese, a Teoria dos Objetos de Fronteira:

- a. focaliza os objetos presentes no encontro de comunidades distintas;
- b. destaca a existência de diferentes comunidades;
- c. ressalta a existência de forças invisíveis no contexto e pode complementar o arcabouço teórico-metodológico já utilizado para a elaboração do sistema de organização do conhecimento.

Tabela 1. *Sistematização de aspectos da Teoria dos Objetos de Fronteira e sua influência nos sistemas de organização do conhecimento visando diferentes caminhos para a informação e o conhecimento*

<i>Aspecto</i>	<i>Influência nos sistemas de organização do conhecimento</i>
Destaque para os objetos	Identificação dos objetos em uso no contexto e daqueles que são utilizados, denominados e significados por diferentes comunidades
Existência de diferentes comunidades	Inserção de diferentes comunidades nas etapas de elaboração dos sistemas de organização do conhecimento, levando em consideração os valores e compreensões de cada uma
Visão ecológica do conhecimento	Participação de representantes de cada uma das comunidades envolvidas, destacando a importância de o instrumento estar de acordo com a visão de cada uma, dificultando possíveis invisibilidades e/ou sub-representações
Atenção às forças invisíveis que podem estar presente no contexto	Levar em consideração que possíveis forças ou disputas políticas, ideológicas, sociais, jogos de poder, níveis de hierarquia, entre outros, podem influenciar o processo e organização do conhecimento no contexto, marginalizando ou sobrepondo pontos de vista específicos

Fonte: *Autores*.

5 Conclusão

A partir do exposto, destacamos os estudos sobre os aspectos éticos do fazer profissional em informação e biblioteconomia como meio para analisar, compreender e combater a marginalização de grupos específicos, as tendências e preconceitos na organização e representação do conhecimento.

Sugerimos, a partir da literatura da Organização do Conhecimento, aspectos teórico-metodológicos a serem considerados para elaboração dos sistemas de organização do conhecimento, como a classificação facetada, técnicas analítico-sintéticas e a utilização da garantia cultural e de teorias que evidenciem os fenômenos, por exemplo, a Teoria dos Níveis Integrativos, como forma de construir diferentes caminhos de acesso à informação para diferentes comunidades de usuários de forma equitativa.

Sugerimos que a Teoria dos Objetos de Fronteira seja considerada como um dos aportes teórico-metodológicos que podem colaborar com os objetivos

elencados, por entender que a partir da visão ecológica que a teoria propõe, se pode materializar a importância de considerar diferentes pontos de vista entre a comunidade de usuários na elaboração dos sistemas de organização do conhecimento. Assim, destacamos que como forma de combater as inclinações, tendências ou preconceitos nos sistemas de organização do conhecimento, a identificação de objetos compartilhados entre diferentes comunidades e a identificação das próprias comunidades pode ser uma etapa metodológica fundamental para garantir que todos os pontos de vista estão devidamente representados, sem marginalizações ou reprodução de preconceitos.

Acreditamos ainda que direcionar a atenção para os objetos utilizados na produção e partilha do conhecimento pode colaborar para a representação do conhecimento a partir dos fenômenos estudados, que por sua vez, pode ajudar no processo de desenvolvimento de trabalhos e pesquisas cooperativos, conforme explicitado no presente trabalho, e na elaboração de diferentes caminhos de acesso à informação para diferentes comunidades de usuários de forma inclusiva e equitativa.

Notas

[1] Neste trabalho, o termo 'inclusivo' foi utilizado com base na Teoria dos Objetos de Fronteira, a fim de demarcar a importância da inclusão de diferentes perspectivas que podem estar envolvidas no contexto dos sistemas de organização do conhecimento e evitar, assim, a sub-representação ou marginalização de outras perspectivas ou culturas, como nos exemplos apresentados.

Referências

- AKRICH, M. From Communities of Practice to Epistemic Communities: Health Mobilizations on the Internet. *Pesquisa Sociológica Online*, v. 15, n. 2, 2010.
- BARITÉ, M. La garantía cultural como justificación en sistemas de organización del conocimiento: aproximación crítica. *Palabra Clave (La Plata)*, Buenos Aires, v. 1, n. 1, p.2-11, 2011
- BEGHTOL, C. A proposed ethical warrant for global knowledge representation and organization systems. *Journal of Documentation*, London, v. 58, n. 5, p. 507-532, 2002.
- BEGHTOL, C. Ethical decision-making for knowledge representation and organization systems for global use. *Journal of the American Society for Information Science and Technology*, New York, v. 56, n. 9, p. 903-912, 2005.
- BRÄSCHER, M.; CAFÉ, L. Organização da informação ou organização do conhecimento? *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 9, 2008, São Paulo. *Anais eletrônicos*. São Paulo: USP, 2008.

- CALLON, M. Struggles and negotiations to define what is problematic and what is not: the sociology of translation. In: KNORR-CETINA, K.; KROHN, R.; WHITLEY, R. (Ed.). **The social process of scientific investigation**. Dordrecht [Holanda]: Ridel, p.197-220, 1980.
- CAMPOS, L. M. Classificação de objetos de fronteira na organização do conhecimento e o papel das ontologias. **Liinc em Revista**, [S. l.], v. 14, n. 2, 2018.
- CARLAN, E. **Sistemas de Organização do Conhecimento: uma reflexão no contexto da Ciência da Informação**. 2010. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília, Brasília, 2010.
- CHAUÍ, Marilena, Cultura e democracia. Crítica y emancipación: **Revista latino-americana de Ciências Sociales**, Buenos Aires, v. 1, n. 1, 2008.
- D'ALMEIDA, N. G. S. B.; CAMPOS, L. M. Influência das culturas epistêmicas nos objetos de fronteira. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 21, 2021, Rio de Janeiro. **Anais** [...], 2021.
- D'ALMEIDA, N. G. S. B. **Diretrizes para a etapa de captura do conhecimento visando a elaboração de sistemas de organização do conhecimento para atuarem como objetos de fronteira**. 2022. 125 f. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Instituto de Arte e Comunicação Social, Universidade Federal Fluminense, Niterói, 2022.
- ETNICISMO. In: Sistema de Gestão de Bibliotecas Sophia. Rio de Janeiro: Biblioteca Nacional. Disponível em: https://acervo.bn.gov.br/sophia_web/autoridade/detalhe/44513?i=1&guid=1675254546358&returnUrl=%2Fsophia_web%2Fautoridade%2Findex%3Fguid%3D1675254546358%26p%3D1. Acesso em: 20 abr. 2023.
- HADFIELD, R. Pearl growing as a strategy in systematic literature searches. **Mediwrite**. 2019. Disponível em: <https://www.mediwrite.com.au/medical-writing/pearlrowing/>. Acesso em: 15/01/2022.
- HAWKINS, D. T.; WAGERS, R. Online bibliographic search strategy development. **Online...**, v. 6, n. 3, p. 10-15, maio, 1982.
- KEIM, W. Social Sciences Internationally: The Problem of Marginalisation and Its Consequences for the Discipline of Sociology. **African Sociological Review / Revue Africaine de Sociologie**, v. 12, n. 2, pp. 22–48, 2008.
- KNORR-CETINA, K. Epistemic cultures: how the sciences make knowledge. Harvard. Harvard University Press, 1999.
- LATOUR, B. **Ciência em ação: como seguir cientistas e engenheiros sociedade afora**. São Paulo: Ed. Unesp, 1997.
- LAW, J. **A sociology of monsters, essays on power, technology and domination**. London: Routledge, 1991.
- LEE, Wan-Chen. Culture and Classification: An Introduction to Thinking about Ethical Issues of Adopting Global Classification Standards to Local Environments. **Knowledge Organization**, Würzburg, v. 42, n. 5, p. 302-307, 2015.
- MANIFESTO de Léon. La interdisciplinariedad y la transdisciplinariedad en la organización del conocimiento científico = Interdisciplinarity and transdisciplinarity in the organization of scientific knowledge. CONGRESO ISKO-ESPAÑA, 8, León, 18, 2007. **Actas** [...]. León, Universidad de León. Secretariado de Publicaciones, 2007.
- MILANI, S. O.; GUIMARÃES, J. A. C. Problemas relacionados a biases em sistemas de organização do conhecimento: perspectivas para a representação de assunto. **IRIS - Revista de Informação, Memória e Tecnologia**, Recife, v. 3, n. esp., p. 72-92, 2017.
- MILANI, S. O. **Estudos éticos em representação do conhecimento: uma análise da questão feminina em linguagens documentais brasileiras**. 2010. 141 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- MILANI, S. O. **Bias na Representação de Assunto: uma discussão de posições binárias nos Functional Requirements for Subject Authority Data (FRSAD)**. 2014. 134 f. Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2014.
- MYLOPOULOS, J. Conceptual Modelling and Telos. In: LOUCOPOULOS, P.; ZICARI, R (eds.). **Conceptual modeling, databases, and case an integrated view of information systems development**. New York: Wiley. p. 49–68, 1992.
- OLSON, H. A. Exclusivity, teleology and hierarchy: our Aristotelian legacy. **Knowledge Organization**, Würzburg, v. 26, n. 2, p. 65-73, 1999.
- SOEGEL, D. The rise of ontologies or the reinvention of classification. **Journal of the American Society for Information Science**, v. 50, p. 12, p. 1119 – 1120, out. 1999.
- STAR, S. L. This is not a boundary object: reflections on the origin of a concept science. **Technology & Human Values**, v.35, p. 601-617, 2010.
- STAR, S. L.; GRIESEMER, J. R. Institutional Ecology, “Translations”, and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. **Social Studies of Science**, [S.l.], v.19, n.3, p.387-420, 1989.
- TROMPETTE, P.; VINCK, D. Revisiting the notion of boundary object. **Revue d'anthropologie des connaissances**, v. 3, n. 1, p. 3-25, 2009.
- WENGER, E. Communities of practice and social learning systems: the career of a concept. In: BLACKMORE, C. (ed.). **Social learning systems and communities of practice**. London: Springer, p. 179 – 198, 2010.
- YEEGANEH, H; Su, Z.; CHRYSOSTOME, E. V. M. A Critical Review of Epistemological and Methodological Issues in Cross-Cultural Research. **Journal of Comparative International Management**, v. 7, n. 2, p. 66-86. 2004.

Categorias fundamentais da Organização da Informação: do conceitual, metodológico e pragmático ao normativo e tecnológico

Fundamental categories of Information Organization: from conceptual, methodological and pragmatic to normative and technological

Cristina Dotta Ortega

Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627 – Belo Horizonte, MG;
Universidade de São Paulo, Av. Prof. Lúcio Martins Rodrigues, 443 – São Paulo, SP, ortega@eci.ufmg.br

Resumo

A Organização da Informação compõe o conjunto das ações de mediação documental em abordagem bibliográfica, cuja função social é a de promover a apropriação da informação por um público. **Objetivo:** Explorar categorias fundamentais para a Organização da Informação, buscando contribuir para explicitá-la de maneira integrada, coerente e consistente. **Metodologia:** Discussão do tema em três tópicos: terminologia da Organização da Informação, haja vista que a compreensão de um campo de conhecimento depende de sua terminologia especializada e esta, por sua vez, permite observar a consolidação científica do campo; categorias de Organização da Informação abordadas como níveis de manifestação de seus processos, quais sejam: conceitual, metodológico, pragmático, normativo e tecnológico; e, questão da primazia atribuída às categorias normativa e tecnológica da Organização da Informação, a qual pode ser relacionada ao entendimento de que a concretude dos instrumentos normativos e tecnológicos suscita segurança e rapidez na pesquisa, no ensino e na prática profissional. **Considerações finais:** A maturidade científica de um campo de conhecimento depende de propostas de categorias que o fundamentem e explicitem. Esse tipo de investigação permite identificar lacunas e desvios, mas também acúmulos e avanços.

Palavras-chave: Organização da Informação. Terminologia da Organização da Informação. Categorias de Organização da Informação.

Abstract

Information Organization comprises the set of documentary mediation actions in a bibliographical approach, whose social function is to promote information appropriation by an audience. **Objective:** To explore fundamental categories for Information Organization to contribute to make it explicit in an integrated, coherent and consistent way. **Methodology:** Discussion of the subject in three parts: Information Organization terminology, given that the understanding of a field of knowledge depends on its specialized terminology and this, in turn, allows observing the scientific consolidation of the field; Information Organization categories, addressed as manifestation levels of its processes, namely: conceptual, methodological, pragmatic, normative and technological; and, question of the primacy attributed to the normative and technological categories of Information Organization, which can be related to the understanding that the concreteness of normative and technological instruments generates confidence and speed in research, in teaching and in professional practice. **Final considerations:** The scientific maturity of a field of knowledge depends on proposals for categories that support and explain it. Among others, this type of investigation not only allows identifying gaps and biases, but also aggregation and advances.

Keywords: Information Organization. Information Organization Terminology. Information Organization Categories.

Resumen

La Organización de la Información comprende el conjunto de acciones de mediación documental en un enfoque bibliográfico, cuya función social es promover la apropiación de la información por un público. **Objetivo:** Explorar categorías fundamentales para la Organización de la Información, buscando contribuir a explicitarla de manera integrada, coherente y consistente. **Metodología:** Discusión del tema en tres partes: terminología de la Organización de la Información, dado que la comprensión de un campo de conocimiento depende de su terminología especializada y esta, a su vez, permite observar la consolidación científica del campo; categorías de Organización de la Información, abordadas como niveles de manifestación de sus procesos, a saber: conceptual, metodológica, pragmática, normativa y tecnológica; y, cuestión de la primacía atribuida a las categorías normativa y tecnológica de la Organización de la Información, la cual puede relacionarse con la comprensión de que la concreción de los instrumentos normativos y tecnológicos eleva la seguridad y la rapidez en la investigación, la docencia y la práctica profesional. **Consideraciones finales:** La madurez científica de un campo de conocimiento depende de propuestas de categorías que lo sustenten y expliquen. Este tipo de investigación permite identificar vacíos y desviaciones, pero también acumulaciones y avances.

Palabras clave: Organización de la información. Terminología de la Organización de la Información. Categorias de Organización de la Información.

1 Introdução

Ações de mediação documentária são ações materiais e simbólicas sobre documentos que têm por função social a apropriação da informação por um público. Neste trabalho, tratamos das ações de mediação documentária em abordagem bibliográfica.

As ações de mediação se realizam por meio de processos documentários, da seleção à busca, passando pelos serviços e produtos.

A Organização da Informação é balizada pela previsão de demandas de públicos, enquanto serviços e produtos são a sistematização de informações de documentos a partir de demandas concretas. Em ambos os casos, trata-se de processos realizados sobre documentos para um público, de maneira sistêmica e integrada e que se articulam entre si em sequência lógica.

A função social de apropriação da informação depende de todos os processos documentários.

No entanto, os parâmetros adotados nos processos de Organização da Informação se refletem significativamente em todos os processos documentários, de tal maneira que seu estudo apresenta relevância no que tange à totalidade dos processos.

Dessa maneira, para uma fundamentação da Organização da Informação, é produtivo discorrer sobre categorias que a configurem como tal.

A motivação atual para o desenvolvimento de categorias de Organização da Informação decorre da observação sobre a reiterada ênfase à prescrição normativa e às tecnologias mais recentes, as quais, supostamente, delinham e caracterizam o campo. Essa ênfase, ao privilegiar alguns dos elementos constituintes da Organização da Informação, secundariza os demais, dificultando que esse subcampo seja compreendido e operacionalizado. Ademais, essa visão do campo subestima e negligencia a literatura e as práticas profissionais.

O objetivo desse trabalho é o de explorar categorias fundamentais para a Organização da Informação, buscando contribuir para explicitá-la de maneira integrada, coerente e consistente.

Como metodologia, discutimos o tema em três tópicos.

Após a Introdução, o primeiro tópico é voltado a uma discussão terminológica a respeito da Organização da Informação, haja vista a variação que cerca este nome e outros, de alguma maneira relacionados a ele. A compreensão de um campo de conhecimento depende de sua terminologia especializada e esta, por sua vez, permite observar a consolidação científica do campo. Os autores que sustentam a discussão desse tópico são: Otlet (1903), Taylor (1999), Svenonius (2000) e Sales e Murguía (2015).

No segundo tópico, são apresentadas categorias de Organização da Informação, abordadas como níveis de manifestação de seus processos, quais sejam: conceitual, metodológico, pragmático, normativo e tecnológico. Uma explicação do percurso de proposição e de desenvolvimento das categorias é apresentada. Vários autores fizeram propostas deste tipo, sendo citado neste trabalho Otlet (1934), segundo elaboração de García Gutiérrez (1999).

No terceiro e último tópico, discutimos a questão da primazia das categorias normativa e tecnológica da Organização da Informação. Apresentamos ideias de autores que, desde o início do século XX, trataram do papel da normatividade para a funcionalidade dos sistemas. São eles: Briet (1951), Gardin (1966), García Gutiérrez (1984, 1985), Blanquet (1993), Calenge (1998), Fayet-Scribe (2001) e Béguin-Verbrugge (2002). Ressaltamos a precedência dos processos, problematizando a ênfase atribuída a alguns instrumentos normativos, tomados como universais, em grande medida em referência ao modelo estadunidense de gestão de bibliotecas. Apontamos a concretude dos instrumentos normativos e tecnológicos como característica que suscita segurança e rapidez na pesquisa, no ensino e na prática profissional.

2 Entre terminologia e categorias da Organização da Informação

2.1 Organização da Informação: a variação terminológica

Um campo de conhecimento não se constitui de maneira linear. Várias concepções manifestam-se simultaneamente, com a predominância de umas sobre as outras. Esse feixe de pontos de vista evidencia-se em sua terminologia na forma de variações. As variações terminológicas decorrem da existência de proposições diferentes, que são identificadas e confrontadas em processos de interlocução contínua, conduzindo o campo em direção à maturidade científica.

No entanto, se essas variações não forem discutidas, as diferentes proposições não serão suficientemente conhecidas, dificultando o acúmulo que seria derivado desse movimento.

Em capítulo de livro produzido com duas colegas, tratamos da variação terminológica em Organização da Informação presente no uso dos termos Análise de Assunto e Análise Documentária. Devido ao uso recorrente de termos em Organização da Informação sob a ausência de princípios epistemológicos explicitados, afirmamos que, subjacente a esta variação terminológica, encontram-se construções culturais, sociais e técnicas que resultaram em abordagens próprias sobre o mesmo processo (ORTEGA; FUJITA; SIMÕES, 2020, p. 148).

A terminologia de um campo, mesmo que fragmentada, e por esta razão mesma, deve ser investigada.

No campo que nos ocupa, como em muitos outros, o idioma inglês é tomado como referência para a produção e comunicação científica. Há avanços significativos realizados no campo em outros idiomas e fora do eixo dos Estados Unidos, mas importa que tratemos da forte presença da terminologia em inglês.

É usual a antiga expressão *Cataloguing and Classification* em parcela significativa da literatura técnico-científica de gestão de bibliotecas, assim como em cursos, eventos e produtos e serviços.

Em uma outra vertente, bastante disseminada hoje, podemos mencionar o termo *Knowledge Organization*, adotado em 1989, com a criação da *International Society for Knowledge Organization* (ISKO), por Ingetraut Dahlberg. Não se pode afirmar que *Organization of Knowledge*, proposto por Henry Bliss, em 1929, tenha relação direta com o conceito expresso por *Knowledge Organization*, segundo Dahlberg. A força política da ISKO deve ser considerada nos estudos sobre a terminologia do campo. Rodrigo de Sales e Eduardo Murguia (2015) discorreram sobre a proposição deste termo, relacionando aspectos epistemológicos e político-ideológicos, e tratando de continuidades e reorientações.

Temos ainda a vertente representada pelos termos: *Organization of Information*, adotado por Arlene Taylor (1999), e *Information Organization*, como usado por Elaine Svenonius (2000). De alguma forma, Taylor e Svenonius, nestas publicações e nas posteriores, buscaram atualizar a expressão Catalogação e Classificação, promovendo uma reorganização sob nova denominação.

As vertentes desenvolvidas em idioma inglês remetem a aspectos evidenciados por: catalogação e classificação, tanto quanto por organização, seja da informação, seja do conhecimento. Comentamos cada vertente a seguir.

Foi no contexto da produção de catálogos de acervos locais de bibliotecas que a expressão Catalogação e Classificação se desenvolveu. Catalogação contempla usualmente a Catalogação Descritiva e a Catalogação de Assunto. Sob estes nomes não são consideradas certas formas de representação dos documentos, como é o caso dos resumos documentários. Ao mesmo tempo, é comum que o uso de Catalogação contemple como elemento obrigatório o número de chamada, modelo de base classificatória adotado para a produção de arranjos de documentos. Classificação, por sua vez, é tomada tanto como a base do processo de organizar informação quanto como o método adotado na produção de arranjos de documentos, como no caso do número de chamada.

A expressão Catalogação e Classificação, sendo abordagem regida por um modelo particular de gestão

de coleções locais de bibliotecas, não favorece a organização de tipologias e suportes documentais diversificados e a oferta de recursos amplos para busca à informação. Essa afirmação pode ser confirmada pelo estudo de Philip Hider (2020) sobre os currículos produzidos nos Estados Unidos e Inglaterra há várias décadas. O estudo nos permite inferir que há necessidade de revisão desse modelo, tornando-o mais abrangente, de modo a contemplar as diversas bases de dados construídas até hoje e a produção de arranjos de documentos diversos, incluindo os eletrônicos, segundo características institucionais e de públicos. A Catalogação/Classificação, também denominada *cat and class*, manteve-se como matriz curricular, eventualmente apresentando alguns outros conteúdos, porém não articulados a ela, como a produção de resumos documentários e a documentação audiovisual.

Quanto à Organização do Conhecimento, como adotado na ISKO, apresentamos as seguintes reflexões. O termo se refere à operação cognitiva realizada pelo indivíduo ou à organização do conhecimento coletivo. Neste segundo caso, trata-se de organização do conhecimento como aproximação a uma configuração ideal de um determinado conhecimento, seja ele científico ou não. Partimos da consideração de que é necessário distinguir a organização do conhecimento coletivo, como ocorre na ciência, e a organização que tem por fim a apropriação da informação por um público.

Otlet, entre outros autores, entendia que não se tratava de organizar o conhecimento, mas os documentos que dispunham sobre ele, ou seja, o conhecimento registrado. Dessa maneira, Otlet (1903, p. 129, 132) afirmava que se deve separar claramente a ciência, de um lado, e a documentação, de outro, e que a documentação deve ser distinguida – com não menos cuidado – da organização da ciência. Para ele, tudo que se refere à documentação científica pode ser considerado como um ramo da organização da ciência, contudo, em razão de sua importância, este ramo deve ser uma disciplina (*matière*, no original) independente e autônoma. No artigo escrito por Sales e Murguia, citado anteriormente, a posição de Otlet é sintetizada como segue:

existe o documento e suas unidades informativas, e cabe à documentação extrair e disponibilizar tais informações para o eficaz acesso à informação. Sendo o conhecimento uma esfera dentro da qual os documentos são necessários para a sua perpetuação, cabe às associações organizar o conhecimento e à documentação agir dentro dos campos dessa organização prévia (SALES; MURGUIA, 2015, p. 414).

Seguindo essa linha de raciocínio, desde as últimas décadas do século XX, teorias e métodos da Terminologia foram apropriados para a construção de sistemas de conceitos relativos ao domínio de que tratam os documentos selecionados, tendo em vista públicos determinados. Este trabalho de sistematização de termos

e conceitos é realizado como parte da metodologia de construção de tesouros. Trata-se de atividade que não tem por fim a elaboração de terminologias especializadas.

Posteriormente, Gail Hodge (2000) propôs o termo Sistemas de Organização do Conhecimento (SOC) incluindo, sob essa denominação, de dicionários e glossários a sistemas de classificação bibliográfica, tesouros e ontologias. Ainda que Hodge considere os níveis de complexidade estrutural destes instrumentos, partindo de listas de termos e suas definições (como os dicionários) até estruturas terminológicas complexas (como os tesouros), ele não releva a especificidade documentária de alguns desses instrumentos. Se Hodge não enfatiza a função dos instrumentos documentários, distintamente daquela exercida pelos instrumentos terminológicos, a apropriação do seu texto no Brasil também não priorizou esta distinção, adotando-se o termo SOC de maneira genérica. Esse movimento dispersivo tem fomentado o apagamento dos estudos sobre os fundamentos das linguagens documentárias e suas metodologias de construção e atualização.

O uso genérico do termo Organização do Conhecimento no Brasil tem implicações semelhantes. No caso das bases de dados (incluindo catálogos), a estrutura central dos registros que as compõem responde por campos descritivos e campos temáticos. Por muitos anos, adotou-se os termos Catalogação, Catalogação Descritiva ou Representação Descritiva para indicar a produção dos campos descritivos, e, Representação Temática para indicar a produção dos campos temáticos. Mais recentemente, tornou-se comum a opção pela expressão Representação Descritiva e Organização do Conhecimento. Essa expressão evidencia deslocamento conceitual do segundo termo em direção à produção de sistemas de conceitos de um domínio, destoando do primeiro termo, no que tange às relações que ambos exercem na produção de bases de dados, como concretizado na prática profissional e consolidado na literatura.

Além dos termos advindos do idioma inglês, no Brasil, foi se tornando cada vez mais presente o termo Organização e Representação. Nesse caso, seria preciso observar que ‘representação’ remete tanto à etapa final da Organização da Informação (etapas essas estabelecidas como leitura, análise, síntese e representação) quanto ao seu produto. Fica a dúvida, portanto, sobre o que se pretende comunicar quando se adota o termo Organização e Representação. Além disso, é comum que Organização e Representação seja seguido de ‘conhecimento’ ou ‘informação e conhecimento’, o que promove outras dificuldades de interpretação.

O termo Tratamento da Informação também é usado no Brasil. Ele é adotado ao menos desde os anos 1970 em publicações brasileiras, incluindo traduções, como é o

caso de: Simpson (1975), Dusoulier (1976) e Cavalcanti (1978). Os modos de uso deste termo são questão a investigar. Termo semelhante – Tratamento Documentário – é usado em francês e espanhol (como nos textos de Béguin-Verbrugge e García Gutiérrez, citados no último tópico), o que pode indicar sua origem no Brasil.

A exploração da relação entre termos e conceitos é tarefa necessária à prática científica, pois ela subsidia o processo de escolhas terminológicas pertinentes e orienta a discussão sobre as implicações das escolhas realizadas.

Para este trabalho, dada a pertinência do significado atribuído ao termo Organização da Informação, mantivemos o seu uso em continuidade à nossa produção científica anterior, somando-nos à produção de outros autores. Consideramos também a solidez e a estabilidade que a literatura sobre Organização da Informação oferece.

Dessa maneira, anteriormente definimos Organização da Informação como:

o conjunto de procedimentos sobre documentos voltados a propiciar seu uso por públicos específicos, segundo necessidades de informação de ordem científica, educacional, profissional, estética, de entretenimento, utilitária. O objetivo é o de promover aderência entre os produtos documentários – produzidos segundo características dos documentos e previsões de busca – e as buscas efetivamente realizadas pelos usuários (ORTEGA, 2013b, p. 185).

Reconhecemos que a primeira parte da definição não discrimina a Organização da Informação, por contemplar todos os processos documentários realizados como ações de mediação documentária. Na segunda parte da definição, a menção às previsões de busca indica a especificidade dos processos de Organização da Informação.

Como discorreremos inicialmente, há sistematização de informações desde a produção de sistemas até a elaboração de produtos e serviços. Trata-se sempre de processos de representação de documentos, realizados uns sobre os outros, a favor de um certo público. No entanto, a produção de sistemas de informação de que se ocupa a Organização da Informação influencia significativamente os demais processos.

É a esse subcampo da mediação documentária em abordagem bibliográfica – a Organização da Informação –, que nos debruçamos a seguir.

2.2 *Categorias da Organização da Informação: os níveis de manifestação dos processos*

No decorrer do século XX, vários autores empreenderam esforços na elaboração de categorias da Organização da Informação.

Mencionando a obra emblemática de Otlet, o Tratado de Documentação, publicado em 1934, García Gutiérrez

(1999, p. 27) afirma que a Documentologia é a ciência documentária, enquanto a Documentotecnia volta-se às técnicas documentárias e a Documentonomia se ocupa da normativa documentária.

Por sua vez, os aspectos pragmáticos, assim como a noção de usuário, embora considerados por Otlet, apenas mais tarde foram explicitamente tratados. Quanto à tecnologia, Otlet é reconhecido como visionário de suas possibilidades de aplicação à Organização da Informação, além de considerar os novos suportes documentais à época, como é o caso dos audiovisuais.

Há pouco mais de 10 anos, apresentamos uma proposta de categorias de Organização da Informação. Desde então, a proposta foi discutida e reformulada em comunicações em eventos, palestras e cursos, e teses e dissertações orientadas.

A proposição inicial foi apresentada no I Encontro Nacional de Catalogadores (ENACAT) e III Encontro de Estudos e Pesquisa em Catalogação (EEPC), realizado na Biblioteca Nacional, no Rio de Janeiro, em outubro de 2012. A apresentação discorreu sobre a Catalogação nas atividades acadêmicas da UFMG. Após uma definição de Organização da Informação, afirmamos que ela envolve: pesquisa teórica-conceitual, pesquisa teórica-procedimental e pesquisa aplicada ou empírica.

As duas publicações da proposta ocorreram no ano seguinte, em 2013, explorando o tema de maneira diversa.

A primeira publicação foi enunciada como aspectos teóricos, procedimentais, normativos e pragmáticos, tomados como categorias para uma epistemologia da Organização da Informação (ORTEGA, 2013a). A publicação foi produto da apresentação realizada na segunda edição do evento da ISKO Brasil, que ocorreu no Rio de Janeiro, em 2013.

Na segunda publicação, a proposta foi explicitada como níveis de manifestação dos processos de Organização da Informação, que foram elencados como: conceitual, procedimental, tecnológico e aplicado (ORTEGA, 2013b). A proposta foi publicada na revista *Perspectivas em Ciência da Informação*, em artigo a respeito da Organização da Informação, com enfoque para a Catalogação, na Escola de Ciência da Informação, da UFMG. O artigo foi resultado da sistematização apresentada no ENACAT/EEPC, de 2012, que tratou da Catalogação nas atividades acadêmicas da UFMG, como dissemos.

No ano seguinte, em seu relatório para exame para qualificação de mestrado, Tolentino (2014) adotou a proposta da segunda publicação (ORTEGA, 2013b), adaptando-a ao processo de descrição. Ele segmentou os níveis de manifestação do processo de descrição em dois planos: o abstrato e o concreto. O primeiro plano foi

denominado abstrato, contemplando os níveis conceitual, procedimental e tecnológico, e o segundo plano foi denominado concreto para caracterizar o nível aplicado. Tolentino agregou à proposta, fornecendo mais sentido aos níveis de manifestação, ao propor o nível aplicado como elemento do plano concreto, em distinção aos outros níveis, que seriam constituintes do plano abstrato.

Dois anos depois, em 2016, apresentamos a proposta no III Seminário Internacional A Arte da Bibliografia, em Belo Horizonte, discorrendo sobre o papel da Bibliografia em Ciência da Informação, a partir do caso da Escola de Ciência da Informação da UFMG. Nesta apresentação, os níveis de manifestação dos processos de Organização da Informação foram ampliados, como adotamos hoje, para: conceitual, metodológico, pragmático, normativo e tecnológico. Os planos abstrato e concreto, elaborados por Tolentino, foram incorporados à proposta.

No início desta década, em suas teses de doutorado, Tolentino (2021) e Silva (2022) desenvolveram o que chamaram, respectivamente, de níveis de manifestação e de dimensões dos processos de Organização da Informação. As definições propostas por eles para as categorias foram resultantes da discussão que realizaram entre si, a partir da proposição descrita anteriormente, depois discutidas no âmbito das atividades de orientação de doutorado. Das propostas apresentadas nas duas teses, foram inseridos os modelos conceituais, pois apenas os instrumentos documentários estavam indicados, e acrescidos os documentos da coleção aos elementos constituintes do nível pragmático, o qual estava restrito aos públicos e às instituições. Também foram aproveitadas as ideias desenvolvidas pelos dois autores.

Desse modo, o percurso de desenvolvimento da proposta contou com a explicitação de aspectos como categorias para uma epistemologia da Organização da Informação e, de outra maneira, como níveis de manifestação dos processos como categorias para a Organização da Informação.

O aspecto procedimental e o normativo foram tratados ora em conjunto, ora em separado. Inicialmente, em 2012, entendemos que as normas se pautam por metodologias, mas reconhecemos depois que o método se distingue da prescrição que cada norma imprime. Desse modo, em uma das publicações de 2013, passamos a considerar distintamente as metodologias propriamente ditas e as prescrições adotadas em cada sistema. Além disso, buscamos destrinchar os usos atribuídos às normas, tratando das normas elaboradas como referência para implantação adaptada em sistemas de informação específicos e das normas elaboradas para cada sistema de informação, eventualmente pautadas nas primeiras. Os dois usos identificados permitem dizer que as normas referenciais podem ser classificadas

como metodologias, embora muitas vezes não o sejam, como quando há insuficiente elaboração conceitual e generalização.

O aspecto pragmático não foi considerado na primeira proposta, mas avaliamos depois que ele deve ser explicitado, compondo o plano abstrato, em contraposição à ideia preconizada de que os aspectos relacionados a usuários são pensados apenas no momento da aplicação, de maneira meramente empírica.

O aspecto tecnológico não foi contemplado inicialmente por não ser central sobre o tema. Depois foi incluído no plano abstrato, de modo a ser elaborado conceitualmente, e não tomado apenas como uma contingência de cada aplicação, a semelhança da ideia preconizada sobre o aspecto pragmático.

A proposta atual é a que segue:

Quadro 1. *Níveis de manifestação dos processos de Organização da Informação*

Plano	Níveis de manifestação dos processos	Explicitação dos níveis de manifestação dos processos
plano abstrato	conceitual (teórico / fundamental)	Conceitos, e os termos que lhes correspondem, que fundamentam a Organização da Informação, no que tange a seus objetos, processos, modelos/ instrumentos e produtos, orientados à função de apropriação da informação por públicos destinatários.
	metodológico (procedimental / técnico)	Metodologias científicas adotadas para realizar os processos e produzir modelos e instrumentos correspondentes.
	pragmático (contextual)	Parâmetros que permitem elaborar processos e instrumentos aderentes à linguagem dos documentos das coleções e das comunidades discursivas que conformam o público destinatário, ambos condicionados institucionalmente, no contexto de fatores sociais e culturais.
	normativo (prescritivo)	Orientações normativas adotadas para a realização dos processos e para a construção dos instrumentos, de modo a dar regularidade e funcionalidade a cada sistema de informação.
	tecnológico	Tecnologias existentes a cada época que realizam a concretização material dos processos e instrumentos adotados na produção e gestão do sistema.
plano concreto	aplicado (empírico)	Aplicações concretas, como as da prática profissional, mas também as das pesquisas experimentais e dos estudos de caso.

Fonte: Desenvolvido a partir de Ortega (2013a, p. 36) e (2013b, p. 185-186), com as contribuições de Tolentino (2014, p. 10), Tolentino (2021, p. 27) e Silva (2022, p. 23).

Como consta no Quadro 1, abordamos categorias de Organização da Informação como níveis de manifestação de seus processos, compondo dois planos. O plano abstrato contempla os níveis conceitual (teórico/fundamental), metodológico (procedimental/técnico), pragmático (contextual), normativo (prescritivo) e tecnológico, todos eles congregados no plano concreto, que se explica pelo nível de manifestação aplicado ou empírico.

No plano abstrato, o nível conceitual é o mais abrangente e verticalizado e os níveis normativo e tecnológico são os mais contingenciais e datados. O plano concreto representa cada uma das aplicações realizadas, as quais, no seu conjunto, devem ser passíveis de explicação fundamentada pelos níveis que compõem o plano abstrato.

A segmentação em níveis tem por fim possibilitar melhor compreensão dos processos. No entanto, cada nível responde por um aspecto do todo, e todos os níveis são integrados entre si. Cada um dos níveis envolve os outros e só pode ser compreendido a partir de sua relação com eles.

Fazendo uso da proposta em discussão para estabelecer categorias de análise para o processo de ordenação, Silva (2022, p. 23) afirma em sua tese que o plano abstrato e o plano aplicado não são estanques, mas permeáveis e plásticos em função do mutualismo das alterações e de sua codependência.

Em adaptação de enunciado da tese de Tolentino (2021, p. 28), na qual o autor deu continuidade aos estudos sobre o processo de descrição, elaboramos que: os processos de Organização da Informação são apreendidos a partir de um conjunto de conceitos, realizados por métodos, balizados por aspectos pragmáticos, e regulados em cada sistema por normas próprias, que são implementadas concretamente por tecnologia avaliada como adequada, funcional e viável.

A despeito da complexidade da Organização da Informação, alguns de seus aspectos foram privilegiados como parâmetros para a pesquisa, o ensino e a prática profissional. Foi o caso dos aspectos normativos e dos aspectos tecnológicos. Enquanto os primeiros são equivocadamente abordados como universais, falta aos segundos que sejam abordados de maneira abstrata. A discussão da abordagem que enfatiza esses aspectos é feita a seguir.

2.3 A primazia da norma e da tecnologia: entre permanências e mudanças

As categorias fundamentais de um campo dependem de algum nível de permanência. As categorias normativa e tecnológica da Organização da Informação apresentem a maior carga contingencial dentre as demais, mas seus elementos fundamentais se preservam no tempo, já que eles são intrínsecos ao campo.

O problema posto é o da primazia atribuída à normatização e à tecnologia, tendo em conta os aspectos conceituais, metodológicos e pragmáticos que os precedem e sustentam. Considerando que os níveis de manifestação são instâncias diferentes do mesmo processo, atribuir mais peso a um nível que a outro diminui a dimensão das atividades documentárias, limitando seu significado e funcionalidade.

Blanquet (2018, p. 228-229, publicação original de 1993) afirma que toda profissão pode ser representada a partir de dois eixos: um eixo essencial e permanente, representado por suas funções e sustentado por conhecimentos fundamentais compartilhados, e um outro eixo, tributário de variações, pois decorrente de técnicas circunstanciais e instrumentos adotadas para tal. Ela diz que se deve perguntar sobre a solidariedade entre os dois eixos, haja vista que a profissão estruturada na ferramenta desaparece quando esta se torna obsoleta. Por isso, Blanquet entende que os instrumentos mudam, mas a função documentária permanece.

Quanto aos instrumentos, Gardin, já em 1966, ao elaborar tipologias para as classificações, por meio de sua caracterização distintiva, nos permitiu inferir que instrumentos concretos não devem ser avaliados por si mesmos. Ele recomendou descartar as observações relativas aos códigos associados às classificações, como sistema decimal, símbolos alfanuméricos e afixos, afirmando que seria mais razoável definir os principais tipos de classificação por meio de seus traços estruturais (GARDIN, 2022, p. 197, publicação original de 1966).

Béguin-Verbrugge (2002, p. 329), décadas depois de Gardin, apresenta as mesmas orientações e questionamentos: ela trata do lugar da linguagem no tratamento documentário, por considerá-lo um ato complexo de comunicação realizado via enunciação. Apesar disso, ela diz que o tratamento documentário é tomado predominantemente como uma prática empírica, a partir da qual a aplicação rigorosa de regras pelos profissionais garantiria ao usuário o resultado de sua busca de informação. Béguin-Verbrugge afirma que é necessário retomar a forma de ver as normas, passando a considerar os usuários a partir de uma perspectiva interativa e pragmática que integra os conjuntos de signos que formam o enunciado do tratamento da informação.

No mesmo período, Calenge e Fayet-Scribe também elaboram reflexões sobre as normas documentárias.

Para Calenge (1998, p. 17), todo modelo ou parte dele, ao ser utilizado como norma, é adaptado, respeitando seu processo e tendo consciência dos pontos abandonados. Segundo ele, o modelo fornece um quadro de análise formalizada, enquanto a norma impõe uma prática sem distinção possível.

Fayet-Scribe (2001, p. 77) considera que as normas são respostas pragmáticas e concretas a novas situações, em vez de camisas de força para as quais os profissionais e os usuários devem se adaptar.

Afinal, muitos anos antes, Briet (2016, p. 35, publicação original de 1951) já falava que as formas que as atividades documentárias podem assumir são tão numerosas quanto as necessidades para as quais elas nascem.

García Gutiérrez (1984, 1985) está entre os autores que discorreram especificamente sobre a normatividade, o que lhe permitiu explicitar seu papel nas atividades documentárias. Na primeira publicação, o autor (1984, p. 19-21, 31) afirma que a normalização deve ser aplicada a todos os âmbitos da atividade humana que necessitam de regulação. Ele define normalização documentária como a organização racional dos conhecimentos e seus suportes e o tratamento e dinamização do conjunto acumulado deles. Assim, haveria três níveis de incidência normativa na cadeia documentária: seleção, tratamento e difusão documentária. García Gutiérrez explica que a normalização documentária implica a regulação das distintas fases do processo sobre documentos, os quais afloram como a causa (documento original), o efeito (documento referencial) e o fim (informação científica) do trabalho realizado nos centros e sistemas documentários.

Para Tolentino (2021, p. 26), a norma, enquanto modelo operativo de caráter prescritivo, advém da construção contínua de procedimentos que buscam ser suficientes para a finalidade do processo.

Ainda que olhares literais de modelos e normas representem fortemente o campo, em detrimento da identificação de seus princípios subjacentes, há reflexões significativas sobre o tema. Segundo Calenge, podemos inferir que as normas devem ser interpretadas e contextualizadas, de forma que se possa operar com elas reconhecendo o modelo do qual partiram.

A ênfase às normas, com a secundarização de seus princípios, é construção histórica. Dois eventos ao menos reforçaram os aspectos normativos da Organização da Informação, em relação estreita com aspectos tecnológicos.

O primeiro evento foi a criação e coordenação de um sistema, pela *Library of Congress*, dos Estados Unidos, desde os primeiros anos do século XX, em bibliotecas desse país. Nesse sistema, era realizada catalogação cooperativa, por meio da venda de fichas catalográficas, que eram reproduzidas por estereotípia. O segundo evento decorre do primeiro: na década de 1960, no mesmo país, foram adotados computadores para a operação dessa rede de catalogação cooperativa. O instrumento documentário então criado foi o formato MARC, não à toa denominado formato de intercâmbio.

Ambos os projetos tiveram dois pilares que permitiram sua viabilização, além de disseminação para além dos limites nacionais. Esses pilares foram: a adoção de normas para a produção dos registros dos documentos e a adoção de tecnologias adequadas a esses sistemas. Por sua vez, ambos os projetos abandonaram significativamente as propostas dos teóricos da Catalogação que lhes antecederam: Panizzi e Cutter, e, depois, os mesmos teóricos, além de Lubetzky.

No que tange a estes eventos, podemos dizer que houve um recrudescimento da normatividade como parâmetro para a realização da atividade documentária, em detrimento das teorias que a sustentam, e uma amplificação da percepção da tecnologia como recurso. Dito de outro modo, esses movimentos levaram à proeminência dos aspectos normativos, colocando em destaque também os aspectos tecnológicos.

Tratando em especial das normas de catalogação adotadas nesses dois projetos, podemos observar que elas ganharam forte presença na compreensão e produção de respostas às demandas por informação. Como qualquer instrumento de Organização da Informação, normas de catalogação são datadas por seu momento de concepção e uso, ou seja, são condicionadas pela maturidade metodológica, pelo estágio de desenvolvimento tecnológico e pelas variantes econômicas e políticas do momento e do local em que são propostas e implementadas.

A despeito dessas contingências, os esforços realizados em torno de propostas normativas de catalogação são dignos de nota. Algumas das propostas mais significativas partiram dos primeiros manuais de bibliotecas e bibliografias, seguidos dos códigos de catalogação, formatos de registro bibliográfico e padrões de metadados, tendo como referência a ISBD (*International Standard Bibliographic Description*) e os modelos conceituais de registros bibliográficos da família FRBR (*Functional Requirements for Bibliographic Records*). A partir desses instrumentos, podemos identificar parâmetros metodológicos, elaborados de modo racional e objetivo, relativos a estruturas de representação de documentos e recursos para acesso a elementos dessas estruturas. Trata-se de um percurso intelectualmente relevante, se considerarmos a demanda cognitiva exigida para contemplar aspectos conceituais e metodológicos, segundo as possibilidades tecnológicas de cada época.

Como pontuamos no tópico anterior, é preciso distinguir as normas adotadas em cada sistema, que são particulares, únicas, e as normas referenciais produzidas na forma de metodologias, ainda que pretensamente. Estas normas referenciais devem ser capazes de subsidiar a produção de sistemas de informação, segundo documentos, instituição e públicos específicos.

A observação dos instrumentos documentários amplamente adotados como normas referenciais nos remete ao modelo estadunidense de gestão de bibliotecas. Esse modelo, que se tornou marcante em praticamente todo o mundo, é baseado em um conjunto de instrumentos, cujos principais são: AACR2 (*Anglo-American Cataloging Rules*) e MARC (*Machine Readable Cataloging*), ambos em processo de substituição pelo RDA (*Resource Description and Access*), CDD (no original: DDC-*Dewey Decimal Classification*) e LCSH (*Library of Congress Subject Headings*).

Os modos de produção, de uso e de disseminação do modelo estadunidense de gestão de bibliotecas foram fundamentais para a cristalização da concepção universal dos instrumentos. A ideia de padrão ou norma foi deslocada da prescrição necessária à entrada de informações em um sistema para um conjunto de orientações tomadas como universais. Resulta daí a recorrente menção a ‘o padrão’ ou ‘a norma’, em referência aos instrumentos desse modelo, ao invés da menção a ‘um padrão’ ou ‘uma norma’. O problema colocado fica evidente na perspectiva da mediação documentária: a pretensão universal se opõe à perspectiva comunicacional da atividade documentária, pois esta é baseada nas possibilidades de uso de informação por um público. A universalidade atribuída aos instrumentos documentários é uma impossibilidade conceitual e operacional.

No Brasil, o modelo estadunidense de gestão de bibliotecas é, em geral, estruturante dos currículos, compondo disciplinas de Organização da Informação, que são muitas vezes centradas na execução de exercícios, à falta de abstração e generalização. A escolha por esse modelo é comumente feita sob a alegação de que ele é o mais usado nas instituições bibliotecárias no Brasil e no mundo. Desse modo, temos um círculo vicioso: o modelo é adotado no ensino por acreditar-se que ele é grandemente adotado nas instituições bibliotecárias e muitas instituições bibliotecárias adotam esse modelo porque ele é, usualmente, a base da formação profissional. Nessa cultura formativa, esses instrumentos são implicitamente indicados para serem adotados em qualquer prática profissional e sem adaptações. No entanto, como os contextos documentários são diversos, o modelo estadunidense é por vezes usado com algum tipo de adaptação ou soluções variadas são desenvolvidas.

Visando superar a orientação normativa do ensino, baseada principalmente no modelo estadunidense de gestão de bibliotecas, seria pertinente e produtivo elaborar conteúdos programáticos pautados pelas categorias de Organização da Informação apresentadas. Segundo essas categorias, os aspectos conceituais e metodológicos são centrais e os aspectos pragmáticos

devem balizar o estudo de todos os tipos de processos e instrumentos, além dos sistemas concretos. Como podemos constatar em antigos e atuais programas de ensino, é interessante contemplar a historicidade de vários modelos, seus percursos históricos e características, assim como suas potencialidades e fragilidades. Juntamente, deve-se discutir a prevalência do modelo estadunidense, abordando as condições econômicas e ideológicas a partir das quais ele se desenvolveu e ganhou força e legitimidade. Em tempos de estudos sobre decolonialismo, está dada a oportunidade para retomar a questão da crença na universalidade, que é sustentada por modelos hegemônicos, e realizar esse enfrentamento.

De modo semelhante às normas, as tecnologias são em geral abordadas à falta de profundidade. A tecnologia, e não apenas a tecnologia eletrônica, é elemento constituinte do campo que nos ocupa. A dimensão histórico-conceitual da atividade documentária nos permite compreender o papel exercido pelas tecnologias para além do apelo do discurso sobre o novo. Por este motivo, os recursos tecnológicos, em função do estágio de desenvolvimento em que se encontram a cada época, são sempre limitantes das aplicações possíveis.

Desse modo, importa a pergunta sobre qual seria a compreensão do campo das ações de mediação documentária, se limitada às contingências da tecnologia eletrônica de cada época e lugar. Outro ponto é o das implicações do entendimento da tecnologia como instância de aplicação apenas, esvaziada do significado dos processos e das funções que esses processos devem cumprir. Uma dessas implicações é a da ruptura entre estruturas de representação da informação: segundo esse pensamento, haveria um tipo de representação da informação realizado de modo analógico, os quais seriam do escopo da Organização da Informação, e um tipo de representação da informação realizado em meio eletrônico, que seria do âmbito da tecnologia. O questionamento a essa divisão, artificialmente produzida, faz-se urgente.

Parte do cenário aqui apresentado pode ser explicado pela concretude dos instrumentos normativos e tecnológicos, pois eles suscitam segurança e rapidez na pesquisa, no ensino e na prática profissional.

Alfaro López discorre sobre essa questão de maneira interessante. Segundo ele (ALFARO LÓPEZ, 2010, p. 4), a elaboração abstrata, que permite a construção teórica sobre a diversidade de práticas e de objetos próprios e definitórios do campo, é vista como alheia à dinâmica do dia a dia da atividade bibliotecária. Para o autor, a atividade bibliotecária se estabeleceu a partir de fazeres orientados pragmaticamente, cuja sustentação funda-se na funcionalidade técnica de tradição empirista, ou, sob variáveis mais favoráveis, na vertente positivista. Alfaro López afirma que essa visão de mundo forneceu segurança psíquica e cognoscitiva à

maioria das pessoas que exerce a atividade bibliotecária. Ele diz que a construção de abstrações é rechaçada de diversas maneiras, mas que é imperioso reparar essa situação para evitar a estagnação ou mesmo a regressão do campo.

A concretude a que se refere Alfaro López está instalada no que podemos chamar de pensamento instrumental-mecânico. Segundo esse pensamento, o instrumento – normativo ou tecnológico – se antepõe ao processo a que deve responder, restringindo-o a suas contingências. Nessa visão de mundo, o bibliotecário deixa de ser um produtor de informações documentárias e é cada vez mais caracterizado como um facilitador do uso da informação, que atua formando (ou treinando) pessoas para que elas se tornem competentes para esse uso.

A perspectiva mediacional, no entanto, exige que a totalidade dos processos documentários seja contemplada, de maneira articulada e integrada. Segundo essa concepção, o profissional concebe, constrói e gerencia dispositivos documentários que funcionem como recursos de aprendizagem aos diversos públicos, fornecendo percursos a eles pelo mundo da informação. Trata-se de sistemas e serviços de informação, elaborados de maneira a produzir vínculos com as pessoas, possibilitando a elas a navegação na complexidade informacional que caracteriza a contemporaneidade.

3 Considerações finais

Os fundamentos da Organização da Informação são sua base intelectual. Fundamentos possuem certa estabilidade e permanência, ao mesmo tempo em que são elaborados a partir de exercícios de aproximação contínua.

A maturidade científica de um campo de conhecimento depende da realização e discussão de propostas de categorias fundadoras. Essas propostas permitem identificar lacunas e distorções, que devem ser superadas em direção à construção epistemológica do campo.

A sistematização de categorias de Organização da Informação evidenciou desequilíbrio entre elas. O aspecto conceitual nem sempre é priorizado e as metodologias tendem a não ser desenvolvidas, já que a norma é considerada o método a seguir e uma tecnologia já traria os elementos estruturantes e funcionais necessários.

Instrumentos normativos e tecnológicos costumam ser tomados como caminho imediato para a realização das operações. No entanto, esses instrumentos não expressam por si mesmos a atividade intelectual documentária que responde por sua função social. É necessário considerar procedimentos e instrumentos a partir dos princípios em que se baseiam, portanto,

independente de uma norma ou outra, ou de uma tecnologia ou outra.

A categoria pragmática merece pesquisa própria. As características a considerar sobre as pessoas em estudos informacionais, em geral, prescindem de sua articulação com as ações de mediação documentária que visam a apropriação da informação. Seria pertinente considerar as atividades que congregam pessoas – por este motivo chamadas de membros de uma comunidade, e tomadas como públicos –, a partir das quais o uso de informação pode ser interessante.

O desafio colocado é o de uma epistemologia para uma ciência social aplicada: ela deve fornecer, simultaneamente, condições cognitivas de compreensão do campo e fundamentação às suas aplicações.

Referências

- ALFARO LÓPEZ, Hector Guillermo. **Estudios epistemológicos de Bibliotecología**. México: UNAM/CUIB, 2010. 133 p. (Teoría y métodos).
- BÉGUIN-VERBRUGGE, Annette. Le traitement documentaire est-il une énonciation? In: SOCIÉTÉ FRANÇAISE DES SCIENCES DE L'INFORMATION ET DE LA COMMUNICATION. **Les recherches en information et communication et leurs perspectives**: Histoire, objet, pouvoir et méthode. Marseille: SFSIC, 2002. p. 329-335.
- BLANQUET, Maria-France. A função documentária: estudo em uma perspectiva histórica. Trad. Camila M. A. da Silva, Marcelio de Brito e Cristina Dotta Ortega. **Perspectivas em Ciência da Informação**, v. 23, n. 4, p. 221-232, out./dez. 2018. Trad.: La fonction documentaire: etude dans un perspective historique, publicado em 1993.
- BRIET, Suzanne. **O que é a documentação?** Trad. M. Nazareth Rocha Furtado. Brasília: Briquet de Lemos/Livros, 2016. Trad. de: Qu'est-ce que la documentation?, publicado em 1951.
- CALENGE, Bertrand. Peut-on définir la Bibliothéconomie?: essai théorique. **Bulletin de Bibliothèques de France**, n. 2, p. 8-20, 1998.
- CAVALCANTI, M. das Neves Niederauer Tavares. Interdependência entre a aquisição planejada e a disseminação seletiva da informação: uma experiência do Centro de Informação Científica para Saúde da Secretaria de Estado de Saúde do Rio de Janeiro. **Revista de Biblioteconomia de Brasília**, v. 6, n. 2, 1978.
- DUSOULIER, Nathalie. Orientação dos novos serviços de documentação; novas técnicas, problemas futuros. **Ciência da Informação**, v. 5, n. 1, 1976.
- FAYET-SCRIBE, Sylvie. **Histoire de la documentation en France**: culture, science et technologie de l'information: 1895-1937. Paris: CNRS Ed., 2001. (CNRS Histoire).
- GARCÍA GUTIÉRREZ, Antonio Luis. Aproximación al concepto y al objeto de la Información/Documentación. In: (Ed.). **Introducción a la documentación informativa y periodística**. Alcalá de Guadaíra (Sevilla): Ed. Mad, 1999. Cap. 1: p. 23-61.
- GARCÍA GUTIÉRREZ, Antonio Luis. **Lingüística Documental**: aplicación a la documentación de la comunicación social. Barcelona: Mitre, 1984. 279 p.
- GARCÍA GUTIÉRREZ, Antonio Luis. Normalización general y documental: concepto, historia e instituciones. **Documentación de las Ciencias de la Información**, v. 9, n. 55, p. 55-96, 1985.
- GARDIN, Jean-Claude. Elementos de um modelo para a descrição de léxicos documentários. Trad. Marilda Lopes Ginez de Lara e Nair Yumiko Kobashi, **Perspectivas em Ciência da Informação**, v. 27, n. 1, p. 190-202, 2022. Trad. de Éléments d'un modèle pour la description des lexiques documentaires, publicado em 1966.
- GUIMARÃES, José Augusto Chaves. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-Americana de Ciência da Informação (RICI)**, v. 1 n. 1, p. 77-99, jan./jun. 2008.
- HIDER, Philip. Origins of the Knowledge Organization field. In: HJØRLAND, Birger; GNOLI, Claudio (Eds.). **ISKO Encyclopedia of Knowledge Organization**. International Society for Knowledge Organization, 2020.
- HODGE, Gail. **Systems of Knowledge Organization for digital libraries**: beyond traditional authority files. Washington D.C.: The Digital Library Federation/Council on Library and Information Resources, 2000.
- ORTEGA, Cristina Dotta. Aspectos teóricos, procedimentais, normativos e pragmáticos como categorias para uma epistemologia da organização da informação. In: COMPLEXIDADE e organização do conhecimento: desafios de nosso século. Organizado por Vera Dobedei e José Augusto Chaves Guimarães. Rio de Janeiro: ISKO-Brasil ; Marília: FUNDEPE, 2013a. p. 34-38.
- ORTEGA, Cristina Dotta. Contexto de desenvolvimento da Organização da Informação, com enfoque para a Catalogação, na Escola de Ciência da Informação da UFMG. **Perspectivas em Ciência da Informação**, v. 18, n. 2, p. 182-215, 2013b.
- ORTEGA, Cristina Dotta; CARVALHO, Maria da Conceição; TOLENTINO, Vinicius de Souza. **O papel da Bibliografia na construção do conhecimento em Ciência da Informação**: o caso da Escola de Ciência da Informação da UFMG. Trabalho apresentado no III Seminário Internacional A Arte da Bibliografia: as faces culturais do gesto bibliográfico. Belo Horizonte, 2016.
- ORTEGA, Cristina Dotta; FUJITA, Mariângela Spotti Lopes; SIMÕES, Maria da Graça. A variação terminológica em Organização da Informação: Análise de Assunto e Análise Documentária. In: SIMÕES, Maria da Graça; LIMA, Gercina Ângela de (Coords.); MACULAN, Benildes Coura Moreira dos Santos; DIAS, Célia da Consolação (Orgs.). **Do tratamento à Organização da Informação**: reflexões sobre concepções, perspectivas e tendências. Coimbra: Imprensa da UC, 2020. Cap. 5: p. 145-192.
- OTLET, Paul. Les sciences bibliographiques et la documentation. **Bulletin de l'Institut International de**

Bibliographie, Bruxelles/Paris/Zurich, ano 8, n. I-3 (n. 44-46 da coleção), p. 125-147, 1903.

SALES, Rodrigo de; MURGUIA, Eduardo Ismael. Instaurações discursivas da organização do conhecimento: H. E. Bliss e a International Society for Knowledge Organization (ISKO). **Ciência da Informação**, v. 44, p. 396-415, 2015.

SILVA, Camila Mariana Aparecida da. **A ordenação como processo de Organização da Informação**: uma discussão (necessária) sobre classificação bibliográfica. Belo Horizonte, 2022. Tese (Doutorado) – Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Ciência da Informação.

SIMPSON, D. J. Antes que as máquinas cheguem. **Revista da Escola de Biblioteconomia da UFMG**, v. 4, n. 2, 1975.

SVENONIUS, Elaine. **The intellectual foundation of information organization**. Cambridge: MIT Press, 2000.

TAYLOR, Arlene. **The Organization of Information**. Englewood, Colo.: Libraries Unlimited, 1999. 280 p.

TOLENTINO, Vinicius de Souza. **A construção da técnica da descrição em catálogos e bibliografias**: contribuição aos fundamentos da Catalogação. Belo Horizonte, 2014. Relatório (Exame de Qualificação de Mestrado) – Universidade Federal de Minas Gerais/Programa de Pós-Graduação em Ciência da Informação.

TOLENTINO, Vinicius de Souza. **Aspectos conceituais do processo de descrição**: uma abordagem sob a ótica da Bibliografia e da Catalogação a partir do livro impresso. Belo Horizonte, 2021. Tese (Doutorado) – Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Ciência da Informação.

De que serve o autor em Ciência da Informação? um estudo sobre a relação entre a noção de autor e o indivíduo que escreve

What is the author for in Information Science? A study on the relationship between the notion of author and the individual who writes

Fábio Furlanete

Universidade Estadual de Londrina, Londrina-PR, ffurlanete@uel.br

Resumo

Este artigo apresenta o resultado de uma pesquisa a respeito da representação das práticas artísticas em Ciência da Informação. O recorte, neste trabalho, se concentra nas discussões sobre a noção de autor na literatura da área. Nosso objetivo é fazer o levantamento do tratamento dado à noção de autor pela literatura da CI que toma o trabalho de Michel Foucault como ponto de partida. Usamos o método da análise crítica para elaborar nossa síntese da noção de função-autor e de seu papel na teoria foucaultiana a partir dos três textos nos quais ele trata do assunto e, junto dessa síntese, apresentamos nossa leitura dos textos da CI que participam desse debate. Concluímos que há um viés na literatura de CI que dificulta a separação, proposta por Foucault, entre a função-autor e o indivíduo que escreve. Em seguida, elaboramos sugestões, tomando como modelo a abordagem metodológica foucaultiana, para o debate dessa questão.

Palavras-chave: Função-autor. Discurso. Procedimento de controle

Abstract

This article presents the result of a research about the representation of artistic practices in Information Science. The focus, in this work, is on the discussions about the notion of author in the literature of the area. Our objective is to survey the treatment given to the notion of author by the IS literature that takes Michel Foucault's work as a reference. We use the method of critical analysis to elaborate our synthesis of the notion of author-function and its role in Foucault's theory from the three texts in which he deals with the subject and, together with this synthesis, we present our reading of the IS texts that participate in this debate. We conclude that there is a bias in the IS literature that hinders the separation, proposed by Foucault, between the author-function and the individual who writes. Then, we make suggestions, taking Foucault's methodological approach as a model, for the debate on this issue.

Keywords: Author-role. Speech. Control procedure

Resumen

Este artículo presenta el resultado de una investigación sobre la representación de las prácticas artísticas en las Ciencias de la Información. El recorte, en este trabajo, se centra en las discusiones sobre la noción de autor en la literatura del área. Nuestro objetivo es relevar el tratamiento dado a la noción de autor por la literatura del CI que toma como punto de partida la obra de Michel Foucault. Utilizamos el método del análisis crítico para elaborar nuestra síntesis de la noción de autor-función y su papel en la teoría de Foucault a partir de los tres textos en los que trata el tema y, junto a esta síntesis, presentamos nuestra lectura de los textos del IC que participan en este debate. Concluimos que existe un sesgo en la literatura del CI que dificulta la separación, propuesta por Foucault, entre el autor-función y el individuo que escribe. Luego, hacemos sugerencias, tomando como modelo el enfoque metodológico de Foucault, para el debate sobre este tema.

Palabras clave: Autor-rol. Discurso. procedimiento de control

1 Introdução

Este artigo apresenta o resultado de uma pesquisa, realizada junto ao Programa de Pós-graduação em Ciência da Informação da Universidade Estadual de Londrina, a respeito da representação das práticas artísticas em Ciência da Informação. O recorte, neste trabalho, se concentra nas discussões sobre a noção de autor na literatura da área.

É difícil exagerar a importância dessa noção para a CI. Segundo Smiraglia,

[t]he fundamental principle of order in the library catalogue at present is the authorship principle. The notion of authorship serves as the organizing node of an alphabetic-classed system, in which "texts" of "works" are organized first alphabetically by uniform title of the progenitor work and then are subarranged using titles for variant instantiations, under the heading for an "author." (2012, p. 35)

Uma importância tão fundamental pode facilmente colocar uma noção na posição de dado inquestionável e levá-la a uma condição de quase invisibilidade. Essa posição, entretanto, vem sendo desafiada há pelo menos 10 anos em um debate feito por autores que tomam como ponto de partida para a discussão o trabalho de Michel Foucault.

O filósofo se torna uma referência reconhecida no debate sobre o tema a partir de três textos do final da década de 1960: o livro “A arqueologia do saber”, de 1969, onde ele define a abordagem metodológica usada em seus livros anteriores (em particular, “As palavras e as coisas”); a conferência “O que é um autor?”, também de 1969, na qual ele se dedica à ideia, já presente no debate anterior, da “morte do autor”; e a aula inaugural de sua cadeira no *Collège de France* em 1970, intitulada “A ordem do discurso”, que faz a síntese da abordagem apresentada em “A arqueologia do saber” adicionada de sua proposta para as investigações futuras. Nesses textos, Foucault traça uma imagem bastante clara do sentido atribuído por ele à noção de autor e o papel que ela desempenha no quadro geral de sua teoria.

Em nossa pesquisa, percebemos que, apesar de os trabalhos da CI sistematicamente dialogarem com o texto foucaultiano no tratamento dessa questão, há uma discrepância consistente entre suas abordagens e a teoria foucaultiana sobre a noção de autor. Assim, trazemos o levantamento do tratamento dado à noção de autor pela literatura da CI que toma o trabalho de Michel Foucault como ponto de partida, fazemos nossa síntese da noção de função-autor e de seu papel na teoria foucaultiana a partir dos três textos nos quais ele trata do assunto e, junto dessa síntese, apresentamos nossa leitura crítica dos textos da CI que participam desse debate, seguida de sugestões, retiradas da própria abordagem metodológica foucaultiana para o debate dessa questão.

2 Foucault e a noção de autor na literatura de CI

O potencial da teoria foucaultiana para a análise das práticas da Ciência da Informação é reconhecido praticamente no mesmo momento em que se percebe a necessidade de se trazer o olhar da teoria crítica para a área. Já no começo da década de 1990, Radford (1992, p. 417) estabelece o paralelo entre o trabalho de Thomas Kuhn e a teoria foucaultiana, e defende sua leitura como modo de superar o positivismo do modelo tradicional de biblioteca. Modelo a partir do qual a biblioteconomia propõe-se como uma instância tão neutra quanto a ciência que ela pretende representar, e a partir do qual clama para si o título de ciência (RADFORD, 1992, p. 412). O trabalho, no entanto, ao tratar da teoria foucaultiana, tende a se concentrar mais

no binômio saber/poder, e não aborda a noção de autor para além do senso comum.

Na década seguinte, a continuação dos esforços para superar esse modelo positivista a partir de um olhar crítico diversificado continuará incluindo Foucault como uma referência importante para a análise, como podemos ver em Minguers e Willcocks (2004) e em Leckie et al (2010). Gordon e Grant (2004, p. 27) apontam que o tratamento limitado da relação saber/poder na literatura da KO torna boa parte dessa literatura problemática. Eles argumentam que, segundo Foucault, apesar de a linguagem ter um papel importante na transferência de significado, ela não constrói significado por si só, e sugerem a associação entre sistemas de significados e produção de verdade como objeto da Organização do Conhecimento. Assim como Radford, eles não abordam a questão do autor. Willcocks (2004) faz um panorama da obra de Foucault e uma ampla revisão bibliográfica sobre seu uso em CI. Seu recorte é focado nos regimes de verdade e nas relações de poder com ênfase nas relações de dominação. O texto apresenta os argumentos dos críticos hostis a Foucault, junto com uma síntese interessante do debate entre Foucault e Habermas (WILLCOCKS, 2004, p. 263-264). Ela demonstra espanto pelo pouco uso de Foucault em CI, e afirma que Foucault pode ser usado para superar o quadro habermasiano empregado inicialmente em CI (WILLCOCKS, 2004, p. 274). Curiosamente, ao apresentar o quadro dos procedimentos de controle do discurso, que em Foucault inclui o autor, ela deixa essa noção de fora (WILLCOCKS, 2004, p. 276).

É somente a partir do final da primeira década do século XXI, até onde pudemos perceber, que o debate sobre o autor em Foucault começa a aparecer na literatura de CI. Olsson (2010), a partir de Radford, aponta que, apesar da importância de Foucault, e das aproximações já feitas entre sua teoria e a CI, ele permanece largamente subutilizado no campo por conta de uma predominância de vozes estadunidenses e inglesas, afastadas da tradição continental pós-marxista na qual o trabalho de Foucault se desenvolve (OLSSON, 2010, p. 63). Ele critica também, nessas vozes, o foco nos processos mentais do indivíduo que busca informação ao invés de estudar os processos coletivos (OLSSON, 2010, p. 63). Seu texto apresenta uma seção sobre a morte do autor na qual ele aponta o problema do paradigma de transmissão de informação que flui do autor para o usuário (OLSSON, 2010, p. 68). Aponta também que, na visão Foucaultiana, significado não é determinado pelo autor, mas uma construção social, e que os autores não são originadores do trabalho mas, eles mesmos, construção social (OLSSON, 2010, p. 68). Apesar disso, não é claro em que medida ele diferencia, como faz Foucault, o autor do indivíduo que escreve.

Essa falta de clareza sobre a diferenciação ou não entre as duas entidades ganha uma nova dimensão no trabalho de Moulaison et al (2014). Eles tratam a função-autor como um objeto representativo da criação, confundindo-a com a noção, também foucaultiana, de fundador de discursividades: ela é “[...] recognition of intellectual debt that can be traced back to the works of the progenitor of concepts. The author-function is a demonstration and acceptance that some things are possible because of who and what has preceded them.” (MOULAISON et al 2014, p. 32). Para eles, o autor é um indivíduo (não necessariamente humano), cujo contexto, atributos e influências podem ser recuperados.

Information about attributes and about relationships goes a long way toward making personal name records reflect the person-ness of the authors they represent. They are less able, however, to indicate how those attributes and relationships were engendered if they were the result of an event in the author’s life. (MOULAISON et al 2014, p. 39)

Eles acabam por estabelecer alguma distância entre o autor como indivíduo e a função-autor ao afirmar que esses eventos podem definir aspectos da vida de um autor, trazendo mudanças nos relacionamentos e status que podem, por sua vez, afetar a função-autor (MOULAISON et al 2014, p. 40). Assim, o autor, para Moulaison et al, parece ser simultaneamente o indivíduo que escreve e um mecanismo que permite acessar esse indivíduo através de seus atributos para que possamos resgatar nosso débito intelectual para com ele.

Um grupo que se dedica de forma mais continuada a essa questão no início da década de 2010, e que é comentado por Moulaison et al (2014, p. 31, 33), é aquele formado por Smiraglia, Lee e Olson, do Information Organization Research Group da Escola de Estudos de Informação da Universidade de Wisconsin-Milwaukee. Eles apresentam uma leitura mais próxima do texto foucaultiano e articulam seus argumentos com vários de seus temas. Seu trabalho parte da análise de classificações encontradas em diferentes épocas e culturas para encontrar o bibliotecário no papel de moldar a autoridade do autor e estabelecer a estrutura e a retórica que influencia a leitura de textos específicos (SMIRAGLIA et al, 2013, p. 1-2; 2011, p. 138-139). Nesse contexto, ao analisar o papel classificatório que emerge da percepção cultural da ideia de autor, eles refutam prontamente a possibilidade de o autor representar alguma responsabilidade intelectual. Ao invés disso, ele é apresentado como a representação de um papel comunicativo essencial (SMIRAGLIA et al, 2013, p. 1; 2011, p. 137) que emerge da percepção cultural do papel de autor (SMIRAGLIA et al, 2011, p. 137), constituindo um fenômeno cultural icônico usado para limitar a livre percepção de uma obra (SMIRAGLIA, 2011, p. 137), aproximando a noção de autor do formato em que ela aparece no texto

foucaultiano como procedimento de controle do discurso. Por outro lado, eles também caracterizam o autor como o indivíduo que, no momento histórico em que a tradição ocidental se desloca da biografia para a tecnicidade, vai adotar o conceito de autoria para operar junto ao mercado editorial e, com isso, ganhar seu sustento (SMIRAGLIA et al, 2013, p. 2). Assim, em seu texto, a identificação do autor oscila entre a etiqueta construída socialmente que identifica uma classe de discurso (SMIRAGLIA et al, 2011, p. 142), e o indivíduo que escreve e vende seus textos. Faz sentido que eles indiquem como ponto focal de sua análise a tensão entre a simples atribuição e a classificação cultural.

No Brasil, um conjunto de artigos que se aprofunda na questão da autoria a partir de Foucault é aquele proveniente do grupo de pesquisadores que se organizou no PPGCI da UNESP de Marília, e que traz uma forte influência do trabalho de Smiraglia. Martínez-Ávila et al (2015) apresentam uma abordagem que mistura elementos da teoria foucaultiana com a fenomenologia husserliana. A partir dela, discutem a diferença entre o nome de autor e o indivíduo que escreve, mostrando as características da função-autor como ferramenta para seleção e organização de textos. Nessa discussão, eles fazem a crítica da noção foucaultiana de autor como se ela apontasse para a responsabilidade intelectual do indivíduo que escreve (MARTÍNEZ-ÁVILA et al, 2015, p. 1104), numa fusão entre a definição de Foucault para a função-autor (FOUCAULT, 2015, p. 46-51) e sua descrição das práticas da crítica literária herdadas da exegese cristã a partir da entrada da modernidade (FOUCAULT, 2015, p. 51-52).

Silva et al (2017) trazem uma leitura da questão da autoria orientada para o direito. Nela, atribuem à noção foucaultiana de autor uma não univocidade que é típica da variação histórica e cultural do conceito (Silva et al, 2017, p. 46), mas que não corresponde à definição apresentada por Foucault (2015, p. 46-51). A atribuição dessa não univocidade lhes permite fazer a manutenção da ambivalência entre autor como etiqueta de catalogação de obras e nome de indivíduo proprietário de uma obra e intelectualmente responsável por ela.

Também pertencentes ao mesmo grupo, Hilário et al (2018) tratam, a partir de Foucault, o autor como campo de coerência teórica ou conceitual, e discutem a emergência do regime de propriedade da obra, descrito por Foucault, em conexão com o debate sobre propriedade intelectual a partir de Stallman. Entendem, a partir da leitura de Barthes sobre a emergência do leitor como causa da morte do autor, o deslocamento do “arquiteto da escrita” para a linguagem ela mesma. Apesar disso, continuam tratando a morte do autor como um caso de apagamento do indivíduo que escreve.

3 A noção de autor no contexto da teoria foucaultiana

A função-autor se constrói no trabalho de Foucault num arco que atravessa, principalmente, três textos: “A arqueologia do saber”, de 1969, “O que é um autor?” e, também de 1969, “A ordem do discurso”. Nesse arco, ela se transforma, de um posicionamento no debate previamente existente sobre a questão da morte do autor, em um dos procedimentos de controle dentro de um sistema que descreve como nós, e as relações de poder que nos atravessam, funcionamos na rede dos discursos. Discurso, nesse contexto, é definido como um conjunto limitado de enunciados que pertencem a um mesmo sistema de formação. Aqui, enunciado e sistema de formação são dois conceitos cujas definições remetem uma à outra. Enunciado é uma formulação vista a partir de suas condições históricas de existência, enquanto formulação, por sua vez, é todo ato individual ou coletivo que faz surgir, valendo-se da materialidade, um conjunto de signos. Já a noção de sistema de formação adquire diferentes elaborações no decorrer da obra de Foucault. Inicialmente ela diz respeito à episteme, como instância exclusivamente discursiva. Porém, à medida que ele desloca sua análise do campo do saber para os campos do poder e da ética, ela passa a incorporar as visualidades através da noção de dispositivo e, posteriormente, a de prática na noção de formação discursiva. No decorrer desse texto usaremos a última para designar o sistema de formação de um discurso.

Essa definição do discurso como o ponto de contato entre enunciados que são resultado de atos de formulação e uma formação discursiva que os conecta às visualidades e às práticas, todos observados a partir de suas condições históricas de possibilidade, confere ao discurso um caráter de performance e de indeterminação. Ela situa-o em uma posição que não pertence nem à linguagem nem ao pensamento (nem à estrutura nem ao significado ou, se preferirmos, nem às palavras nem às coisas), mas sim ao caos que existe entre as duas. O receio desse caos é o que dá margem à necessidade dos procedimentos de controle descritos aqui.

Os procedimentos de controle do discurso têm, portanto, a função conter o caos e a imprevisibilidade do discurso, “conjurar seus poderes e perigos, dominar seu acontecimento aleatório, esquivar sua pesada e temível materialidade” (FOUCAULT, 2014a, p 9). Eles dão forma ao desejo por uma linguagem unívoca e consensual como espelho de um mundo objetivo sólido e simples. Esses procedimentos são divididos por Foucault em 3 categorias: exclusão, ordenação e sujeição, cada uma delas com suas próprias subcategorias.

Procedimentos de sujeição são aqueles que selecionam os sujeitos que falam. Eles rarefazem a rede de

discursos, delimitando quem pode participar dela através da determinação de suas condições de funcionamento. Estabelecem quem pode dizer o que e quando, ou pelo menos quem deve ser levado em consideração quando o faz. Alguns discursos são abertos a quase todos, enquanto outros são vetados exceto para um conjunto restrito de indivíduos. Os mais bem posicionados podem ter acesso a uma ampla gama de discursos, um peso muito grande sobre um discurso específico ou ambos. Esses procedimentos são divididos em quatro subcategorias: rituais da palavra, sociedades do discurso, grupos doutrinários e apropriações sociais.

Os instrumentos que permitem aos membros de uma sociedade o acesso ao discurso são compreendidos pelas apropriações sociais, que carregam em si as marcas das divisões dessa sociedade e seus conflitos. Nesses espaços coletivos, de acordo com a repartição dos papéis, das tarefas e dos poderes a eles associados, surgem escolas de pensamento – às vezes mais de uma no mesmo espaço, alimentando os conflitos que são constituintes de suas identidades. A aceitação de normas e verdades garante a pertença a um grupo doutrinário que vale como um sinal de pertença a uma classe social. Ao contrário das disciplinas científicas, que controlam o discurso apenas no plano formal e do conteúdo, os grupos doutrinários questionam simultaneamente o enunciado e o sujeito que fala, assim como um através do outro. Se os grupos doutrinários acolhem todos aqueles que tomam para si suas regras, as sociedades de discurso restringem os espaços de circulação dos discursos, mantendo os privilégios de seus detentores através de regras estritas de distribuição. Mesmo que sociedades de discurso explicitamente constituídas não sejam mais comuns, vários de seus mecanismos de preservação da separação e não permutabilidade de papéis entre palavra e escuta continuam em uso na forma mais difusa de “sociedade de discurso” representada pelo mercado editorial, pelo livro, e pela personagem do escritor. Essa personagem aponta constantemente para a presença em seu discurso de uma escritura e de uma criatividade que marcam a sua dissimetria com relação a qualquer outra prática discursiva. Por último, essa dissimetria é marcada também por um conjunto de práticas que carregam os signos da eficácia de um discurso e da qualificação daquele que o pronuncia. A essas práticas Foucault dá o nome de “rituais da palavra”.

Os discursos religiosos, judiciários, terapêuticos e, em parte também, políticos, não podem ser dissociados dessa prática de um ritual que determina para os sujeitos que falam, ao mesmo tempo, propriedades singulares e papéis estabelecidos (FOUCAULT, 2014a, p. 37).

Há também um conjunto de procedimentos de controle que são internos aos discursos, e funcionam no plano de sua classificação, ordenação e distribuição. Eles

aparecem como uma forma de os próprios discursos submeterem sua dimensão de acaso e de acontecimento. Esses procedimentos são relacionados a um jogo de identidade: o comentário pela repetição do mesmo; o autor na forma da individualidade e do eu; a disciplina através da reatualização permanente das regras.

Por disciplina, aqui, não se deve entender o conceito elaborado por Foucault no empreendimento genealógico em que faz a crítica das relações de poder durante os anos 1970. Ao invés disso, trata-se de um campo que reúne objetos, conceitos e técnicas, um horizonte teórico, uma certa relação entre verdadeiro e falso, num princípio que permite construir novas proposições de acordo com regras estritas. Não apenas construí-las, mas construí-las indefinidamente sem que seu sentido e sua validade dependam necessariamente de quem está na origem desse campo. Se a disciplina se constitui como um sistema anônimo de produção e controle de discursos, o autor, por outro lado, é entendido não como o indivíduo que produz o discurso, mas como uma ferramenta de agrupamento de discursos sob uma etiqueta identificada, no mais das vezes, ao nome de uma pessoa. Isso não significa que não existe o indivíduo que escreve, nem que ele não possa clamar para si a posição de autor. Significa apenas que o autor que atribuímos a um conjunto de discursos não corresponde àquele indivíduo, nem à posição de autor que ele porventura tenha reivindicado. Além disso, nem todos os discursos possuem autor. Essa distinção é dada apenas àqueles que, por alguma razão, precisam ser separados dos discursos cotidianos, como é o caso do escritor ou do compositor. A categoria do comentário diz respeito aos procedimentos que permitem que alguém pronuncie um novo discurso dissimulado na forma da repetição de um discurso antigo, enquanto se apropria da autoridade conferida por esse discurso de impacto e reputação já consolidados. Parte-se de uma assimetria dos discursos no que diz respeito à sua importância e sua riqueza em possibilidades de interpretação e projeta-se no discurso escolhido como ponto de partida os sentidos que se acredita que ele deve ter (sempre os teve ocultos) e que se pode resgatar no novo discurso.

Se os procedimentos de ordenação operam internamente, com os discursos fazendo seu próprio controle, os procedimentos de exclusão atuam externamente na zona de contato do discurso com o desejo e o poder. O papel do discurso nessa ligação não é apenas o de veículo de manifestação, mas também de objeto desse desejo e desse poder. “O discurso não é simplesmente aquilo que traduz as lutas ou os sistemas de dominação, mas aquilo porque, pelo que se luta, o poder do qual nos queremos apoderar” (FOUCAULT, 2014a, p. 10). Nessa luta, os procedimentos tomam a forma de um conjunto complexo de interdições que não cessam de se reforçar, se compenhar e se modificar, no

cenário sempre móvel da palavra proibida. Os exemplos dados por Foucault são o tabu do objeto, o ritual da circunstância e o direito privilegiado do sujeito que fala. Eles tomam também a forma de uma separação que rejeita e limita a circulação tanto das formas de discurso que não são consideradas válidas quanto dos falantes que as proferem, e correspondem, a partir do século XIX, à exclusão do arbítrio e do desejo (FOUCAULT, 2017, p. 508). Os discursos e os indivíduos marcados por essa segregação ocupam uma posição na qual não são sequer ouvidos, ou, quando o são, isso é feito através de todo um aparato de saber e uma rede de instituições que mantém seu estranhamento e seu potencial de distúrbio a uma distância segura. É assim que se dá no caso da oposição entre a razão e a loucura, e que se reflete no horror da arte que desafia a lógica do mesmo e em todas as ferramentas que usamos para mantê-la afastada, na melhor das hipóteses, nos campos neutralizadores do interessante, do estranho e do exótico. Por último, eles assumem a forma de uma oposição entre verdadeiro e falso. Essa oposição, associada por Foucault à noção de “vontade de verdade”, é na realidade mais do que um subgrupo dos procedimentos de exclusão. Ela aparece como um dos elementos centrais da crítica foucaultiana e, mesmo nos textos nos quais não é o objeto principal, constitui o pano de fundo de seu trabalho. Nessa crítica, a partir de Nietzsche, a verdade é apresentada como uma noção histórica e campo de disputa de poder, apoiada por suportes institucionais e reconduzida pelo modo como o saber é repartido, distribuído e aplicado nas sociedades.

Certamente, se nos situarmos no nível de uma proposição, no interior de um discurso, a separação entre o verdadeiro e o falso não é nem arbitrária, nem modificável, nem institucional, nem violenta. Mas se nos situarmos em outra escala, se levantarmos a questão de saber qual foi, qual é constantemente, através de nossos discursos, essa vontade de verdade que atravessou tantos séculos de nossa história, ou qual é, em sua forma muito geral, o tipo de separação que rege nossa vontade de saber, então é talvez algo como um sistema de exclusão (sistema histórico, institucionalmente constrangedor) que vemos desenhar-se. (FOUCAULT, 2014a, p. 13-14)

Podemos ver que o autor e o indivíduo que escreve são entidades claramente distintas. O indivíduo que escreve opera junto a vários desses procedimentos de controle do discurso: as sociedades do discurso às quais pertence ou é excluído, as doutrinas que segue ou refuta, os rituais da palavra que performa ou ignora, as exclusões que sofre ou impõe a outros, as disciplinas nas quais atua, os comentários que o conectam a uma cadeia de legitimações e, por último, os autores aos quais está conectado através do eventual uso de seu nome. Esse autor, por outro lado, é uma função que opera em um espaço distinto do indivíduo que escreve e ao qual ele tem pouco acesso. É uma função que é constituída coletivamente e a posteriori do trabalho do

indivíduo que escreve e que se transforma, a despeito do indivíduo que escreve, de acordo com as necessidades locais e históricas de quem a constitui. É uma chave de leitura que ajuda a fechar o campo de possibilidades de atribuição de sentido para um conjunto de textos e, no exercício desse papel, pacifica nossa relação com o caos inerente aos discursos.

Entretanto, seja pelo recurso à fenomenologia que toma como premissa o olhar do indivíduo, seja pelo imperativo ético do respeito ao direito autoral, ou mesmo pela simples contraposição às práticas da biblioteca, a literatura da CI que trata da noção de autor a partir do trabalho de Foucault é marcada pelo esforço de, a despeito do texto foucaultiano, preservar a associação entre autor e indivíduo que escreve.

Esse esforço parece estar relacionado com a herança das práticas de autenticação de textos da crítica literária desenvolvidas na entrada da modernidade a partir do modelo da exegese cristã (FOUCAULT, 2015, p. 51-52), e que tiveram forte influência nas ciências humanas. Segundo Foucault (2014, p. 49), até o Renascimento, o principal elemento de validação de um texto com valor de verdade era sua procedência. A importância e o reconhecimento de um autor eram peças fundamentais no estabelecimento da confiabilidade da informação contida em um escrito. Por outro lado, nos discursos artísticos (não apenas textuais como também visuais e sonoros) o reconhecimento da autoria, apesar de eventualmente presente, não era essencial para determinar o valor da obra. A partir do final do século XVII, com o desenvolvimento do método científico, os textos com valor de verdade passam a prescindir da autoridade de seu escritor. Uma vez que era verificada a correta aplicação do método, a confiabilidade estava estabelecida. Ao mesmo tempo, inicia-se um lento processo de construção da necessidade da identificação do autor de um discurso artístico para estimar seu valor. Pintores e compositores começam a sistematicamente assinar suas obras e toma corpo o debate que vai levar às concepções de obra e autoria que encontramos na literatura e na musicologia da modernidade. Nela, a ideia de autoria associada a obra artística que antes era vista como pertencente aos domínios da habilidade técnica e do artesanato passa a ser entendida como regime de propriedade, transgressão e imputabilidade (FOUCAULT, 2015, p. 47-48). Ao invés da associação com o clérigo ou o nobre que assumiram o patronato do trabalho, ele passa a pertencer ao artista que o concebeu, com consequências diretas para o mercado editorial que se estrutura nesse período. Para afirmar sua identidade como artista e a propriedade sobre sua obra, o escritor sentirá a necessidade de buscar a experimentação que fará com que seu trabalho dê um passo para além da execução virtuosa dos valores da tradição. Esse passo ligará, por toda a modernidade e até hoje, a ideia de

autoria à da propriedade sobre uma ideia que transgride a prática comum. Não estando mais necessariamente a serviço do clero ou da nobreza, a responsabilidade pelo potencial intelectual, artístico e político do trabalho passa para o indivíduo criador. Acreditamos que a herança cultural gerada por esse processo histórico está na base desse compromisso da CI com o indivíduo criador. Compromisso que acaba por desviar a atenção da ideia foucaultiana segundo a qual o CI estaria na posição não de organizadora e difusora dos trabalhos do autor, mas de colaboradora na constituição dessa personagem.

4 A estratégia de Foucault para os procedimentos de controle do discurso

Para formular estratégias que nos permitam tratar dessa questão, podemos observar como Foucault estrutura sua abordagem metodológica para a crítica. A partir da articulação dos procedimentos de controle do discurso, Foucault toma como exemplo a Filosofia e descreve a forma como ela responde a eles e os incorpora. Essa resposta não pode deixar de refletir nos modos de pensar os discursos sobre nossa prática assim como essa mesma prática como discurso. Isso se dá não apenas pela proximidade entre a filosofia e as disciplinas acadêmicas das humanidades, mas também pela reverberação em via de mão dupla desses saberes nas práticas e nas relações de poder subjacentes a ela. Assim, podemos encontrar, nos discursos da Ciência da Informação, traços dos mesmos temas identificados por Foucault na Filosofia: a noção de verdade ideal, o conceito de racionalidade imanente e uma ética do conhecimento, que por sua vez são reforçados pelas formas específicas de elisão da realidade do discurso: sujeito fundador, experiência originária e mediação universal.

A mediação universal assume o discurso como uma instância passiva, nada mais do que uma reverberação da verdade das coisas. O logos que daria forma à racionalidade do mundo pela elevação das singularidades ao plano do conceito, permitindo compartilhá-lo numa troca sem perdas. Nessa apresentação disfarçada de representação, o tema da experiência originária pressupõe uma cumplicidade com o sentido original do mundo que nos permitiria conhecê-lo e nos comunicarmos a seu respeito. O discurso seria uma leitura do mundo, resultado da captura de seu sentido, que nos dá acesso a ele na forma da verdade. Essa captura é realizada pelo sujeito fundador: aquele que acessa a realidade sem precisar passar pela instância mediadora do discurso. É ele que dá conteúdo a uma língua que, sem a sua intuição do real, seria apenas forma vazia. Ele funda novos horizontes de significação, como escrituras do mundo, que deverão ser continuamente estudados e decifrados e, no futuro, serão a base para novos campos

discursivos e áreas do saber. Nos três casos o discurso é um jogo, respectivamente, de troca, de leitura e de escritura, jogo que opera a elisão da realidade do discurso restringindo-o ao domínio passivo e pacífico do significante.

Tudo se passa como se interdições, supressões, fronteiras e limites tivessem sido dispostos de modo a dominar, ao menos em parte, a grande proliferação do discurso. De modo a que sua riqueza fosse aliviada de sua parte mais perigosa e que sua desordem fosse organizada segundo figuras que esquivassem o mais incontornável; tudo se passa como se tivessem querido apagar até as marcas de sua irrupção nos jogos do pensamento e da linguagem. Há, sem dúvida, em nossa sociedade e, imagino, em todas as outras mas segundo um perfil e facetas diferentes, uma profunda logofobia, uma espécie de temor surdo desses acontecimentos, dessa massa de coisas ditas, do surgir de todos esses enunciados, de tudo que possa haver aí de violento, de descontínuo, de combativo, de desordem, também, e de perigoso, desse grande zumbido incessante e desordenado do discurso (FOUCAULT, 2014, p. 17-18).

Para compreender esse temor do discurso em suas condições e efeitos, Foucault elaborará um projeto centrado em três funções correspondentes às três formas de elisão de sua realidade: questionar a vontade de verdade, restituir ao discurso o caráter de acontecimento e suspender a soberania do significante. Do ponto de vista do método, essas funções exigem uma inversão que reconheça nas instâncias tidas como fundamentais e criadoras do discurso um princípio de rarefação. Esse reconhecimento não implica que, uma vez removido o princípio de rarefação, encontraremos para além dele um discurso que corre finalmente livre de toda repressão. O desencontro e a exclusão são, tanto quanto os encontros, partes integrantes do movimento descontínuo dos discursos. Nesse movimento o discurso constitui não um jogo de significações prévias de um mundo que se apresenta para nosso entendimento, mas uma violência que impomos às coisas. Toda regularidade dos acontecimentos é construída a partir da localidade e da especificidade dessa violência. Por fim, não se deve buscar no discurso algo como um núcleo ou uma camada interior de sentido a ser desvendado. Ao invés disso, deve-se buscar na sua positividade e na regularidade dessa série aleatória de acontecimentos as condições externas de possibilidade que permitem fixar suas fronteiras. Com esses quatro princípios metodológicos: inversão, descontinuidade, especificidade e exterioridade, podemos substituir as noções tradicionais de criação, unidade, originalidade e significação, dominantes em nossas análises, pelas noções, opostas termo a termo respectivamente, de acontecimento, série, regularidade e condição de possibilidade. Isso deve nos permitir tratar o discurso não como ferramenta de representação do que estaria para além dele, mas como série distinta e regular de acontecimentos em si mesmos, como a fonte do acaso,

da descontinuidade e da materialidade que nossos procedimentos de controle tentam continuamente conjurar.

Esses princípios permitem que Foucault proponha dois conjuntos de análises; um conjunto crítico focado no princípio da inversão e um conjunto genealógico direcionado aos princípios descontinuidade, especificidade e exterioridade. O primeiro procura delimitar as formas das funções de exclusão e ordenação, como a loucura, a sexualidade e a vontade de verdade, assim como dos procedimentos de limitação dos discursos, como o autor, o comentário e a disciplina. O segundo procura descrever, nas séries da formação efetiva do discurso, os sistemas de coerção e os processos de rarefação: os rituais da palavra, as sociedades do discurso, os grupos doutrinários e as apropriações sociais. Comparando essas duas formas de análise Foucault nos diz que

[...] toda tarefa crítica, pondo em questão as instâncias de controle, deve analisar ao mesmo tempo as regularidades discursivas através das quais elas se formam; e toda descrição genealógica deve levar em conta os limites que interferem nas formações reais. Entre o empreendimento crítico e o empreendimento genealógico, a diferença não é tanto de objeto ou de domínio mas, sim, de ponto de ataque, de perspectiva e de delimitação (FOUCAULT, 2014, p. 62)

O conjunto crítico, portanto, aponta para uma ciência da informação consciente de sua posição como instância produtora de verdade, enquanto o conjunto genealógico pede a análise dos modos através dos quais ela, a partir das verdades que estabelece, funciona como base para a produção de discursos. Tanto em um quanto no outro, a percepção da função-autor como procedimento de controle do discurso, independente do indivíduo que escreve, é essencial. Isso não significa que o indivíduo que escreve não tenha sua importância, e é claro que seus direitos devem ser preservados, mas, de acordo com Foucault (2014, p. 53-54), o caráter de acontecimento dos discursos, suas condições de possibilidade, sua sequência e regularidade possuem um peso e importância que se sobrepõe à identidade do indivíduo que escreve. No processo de construção, organização e difusão do conhecimento, os atributos desse indivíduo não nos parecem ser os dados mais úteis ou interessantes para a compreensão dos discursos reunidos sob seu nome, ou para a experiência dos que tomam contato com eles.

5 Considerações finais

A insistência da literatura de CI em manter o indivíduo que escreve no foco da atenção acaba por fazer a manutenção da ideia do autor como o criador de um objeto único e original que carrega em si um significado unívoco e pré-determinado. Ao invés disso, Foucault nos sugere pensar a prática de quem escreve como um processo no qual acontecem determinadas

configurações materiais da rede dos discursos. Essas configurações podem se dar em determinada série de acontecimentos cujas condições de possibilidade permitam que ela reverbere na rede dos discursos na forma de uma regularidade. Os autores se desenham a partir da constituição dessa regularidade como procedimentos de controle que visam atenuar os efeitos caóticos e imprevisíveis daquela reverberação. Logo, quando os profissionais de CI classificam obras a partir de autores, o que eles classificam são chaves de leitura pertencentes a uma comunidade específica de leitores reunidos a partir dessa mesma chave. Chaves que, por um lado, são úteis para o mercado editorial, a indústria de conteúdo, a academia, e demais espaços onde o controle do discurso é imprescindível. Por outro lado, sua utilidade para a compreensão dos discursos reunidos sob seu nome é limitada.

Referências

- FOUCAULT, Michel. **O que é um autor?** 9. ed. Lisboa: Nova Vega, 2015.
- FOUCAULT, Michel. **A ordem do discurso:** aula inaugural no Collège de France, pronunciada em 2 de dezembro de 1970. 24. ed. São Paulo: Loyola, 2014.
- FOUCAULT, Michel. **A arqueologia do saber.** 7. ed. Rio de Janeiro: Forense universitária, 2005.
- GORDON, Ray; GRANT, David. Knowledge management or management of knowledge? Why people interested in knowledge management need to consider Foucault and the construct of power. In: **Tamara: journal of critical postmodern organization science**, v. 3(2), 2004.
- HILÁRIO, Carla Mara; MARTÍNEZ-ÁVILA, Daniel; GRÁCIO, Maria Cláudia C.; WOLFRAM, Dietmar. Authorship in science: a critical analysis from a Foucauldian perspective. In: **Research evaluation**, v. 27(2), 63-72, 2018.
- LECKIE, Gloria J.; GIVEN, Lisa M.; BUSCHMAN, John. **Critical theory for library and information science: Exploring the social from across the disciplines.** ABC-CLIO, 2010.
- MARTÍNEZ-ÁVILA, Daniel; SMIRAGLIA, Richard; LEE, Hur-Li; FOX, Melodie. What is an author now? Discourse analysis applied to the idea of an author. In: **Journal of documentation**, v. 71(5), p. 1094-1114, 2015.
- MINGERS, John; WILLCOCKS, Leslie. **Social theory and philosophy for information systems.** John Wiley and Sons, 2004.
- MOULAISON, Heather Lea; DYKAS, Felicity; BUDD, John M. Foucault, the author, and intellectual debt: capturing the author-function through attributes, relationships, and events in knowledge organization systems. In: **Knowledge Organization**, v. 41(1), 2014.
- OLSSON, Michael R. Michel Foucault: discourse, power/knowledge, and the battle for truth. In: LECKIE, Gloria J.; GIVEN, Lisa M.; BUSCHMAN, John E. (eds.). **Critical theory for Library and Information Science: exploring the social from across the disciplines**, p. 63-74. Santa Barbara-CA: Libraries Unlimited, 2010.
- RADFORD, Gary P. Positivism, Foucault, and the fantasia of the library: conceptions of knowledge and the modern library experience. **The library quarterly**, v. 62, n. 4, p. 408-424, 1992.
- SILVA, Eduardo G.; MARTÍNEZ-ÁVILA, Daniel; GRACIOSO, Luciana de S. A não univocidade do conceito de "autor" sob a perspectiva de Wittgenstein e Foucault. In: **Scire: representación y organización del conocimiento**. v. 23(2), p. 45-52, 2017.
- SMIRAGLIA, Richard P.; LEE, Hur-Li; OLSON, Hope A. The flimsy fabric of authorship. In: **Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI**. 2013.
- SMIRAGLIA, Richard P.; LEE, Hur-Li; OLSON, Hope A. Epistemic presumptions of authorship. In: **Proceedings of the 2011 iConference**. 2011. p. 137-143.
- WILLCOCKS, Leslie P. Foucault, power/knowledge and information systems: reconstructing the present. In: MINGERS, John; WILLCOCKS, Leslie. **Social theory and philosophy for information systems**, p. 238-296. Chichester: John Wiley & Sons Ltd, 2004.

Indexação e desinformação: possíveis relações

Indexing and disinformation: possible relationships

Franciele Marques Redigolo (1), Ana Roberta Pinheiro Moura (2), Renata Lira Furtado (3)

(1) Universidade Federal do Pará. R. Augusto Corrêa, 01 - Guamá, Belém - PA, 66075-110, franciele@ufpa.br, (2) anaroberta@ufpa.br, (3) renatalira@ufpa.br

Resumo

Objetivo geral: Investigar as possíveis relações entre o conceito de Indexação e a Desinformação. Objetivos específicos: Estruturar teoricamente os temas elencados na pesquisa; Processar e tratar os dados coletados; Apresentar e analisar os resultados encontrados. a pesquisa tem como indagação “Qual a relação entre a Indexação e a Desinformação?” **Metodologia:** A pesquisa está dividida em três etapas: pesquisa bibliográfica, processamento dos dados pelo Wordart apresentação em forma de nuvem de palavras. **Resultados:** Por meio de uma representação conceitual, cruza-se os termos coletados evidencia-se três possibilidades de relação entre Indexação e Desinformação: possibilidade de registro de dados falsos em sistemas de informação, possibilidade de indexar informações erradas e revocação da informação. **Conclusão:** A Desinformação se destaca como um fenômeno informacional, que associado às características do processo de indexar, podem acarretar problemas diante da revocação da informação, da presença de dados falsos e da falta de informação. Assim, percebe-se a necessidade de aprofundar estudos sobre as temáticas, na intenção de reduzir ou amenizar estas possibilidades de danos ocasionados pela Desinformação, diante da busca de informações em sistemas tecnologicamente mediados.

Palavras-chave: Indexação. Desinformação. Nuvem de palavras.

Abstract

General objective: To investigate the possible relationships between the concept of Indexing and Disinformation. Specific objectives: Theoretically structure the themes listed in the research; Process and treat the data collected; Present and analyze the results found. the research has as its question “What is the relationship between Indexing and Misinformation?” **Methodology:** The research is divided into three stages: bibliographical research, data processing by Wordart and presentation in the form of a word cloud. **Results:** By means of a conceptual representation, the collected terms are crossed and three possibilities of relationship between Indexing and Misinformation are shown: possibility of registering false data in information systems, possibility of indexing wrong information and recalling information. **Conclusion:** Disinformation stands out as an informational phenomenon, which, associated with the characteristics of the indexing process, can cause problems due to the recall of information, the presence of false data and the lack of information. Thus, there is a need to deepen studies on the themes, with the intention of reducing or mitigating these possibilities of damage caused by Misinformation, in view of the search for information in technologically mediated systems

Keywords: Indexing. Disinformation. Word cloud.

Resumen

Objetivo general: Indagar las posibles relaciones entre el concepto de Indexación y Desinformación. Objetivos específicos: Estructurar teóricamente los temas enumerados en la investigación; Procesar y tratar los datos recabados; Presentar y analizar los resultados encontrados. La investigación tiene como pregunta “¿Cuál es la relación entre Indexación y Desinformación?” **Metodología:** La investigación se divide en tres etapas: investigación bibliográfica, procesamiento de datos mediante Wordart y presentación en forma de nube de palabras. **Resultados:** Mediante una representación conceptual se cruzan los términos recogidos y se muestran tres posibilidades de relación entre Indexación y Desinformación: posibilidad de registrar datos falsos en los sistemas de información, posibilidad de indexar información errónea y recordar información. **Conclusión:** La desinformación se destaca como un fenómeno informacional que, asociado a las características del proceso de indexación, puede generar problemas por la recordación de información, la presencia de datos falsos y la falta de información. Por lo tanto, existe la necesidad de profundizar los estudios sobre los temas, con la intención de reducir o mitigar estas posibilidades de daño causado por la Desinformación, en vista de la búsqueda de información en los sistemas mediados tecnológicamente.

Palabras clave: Indexación. Desinformación. Nube de palabras.

1 Introdução

A qualidade das informações acessadas tem sido tema frequente de pesquisas associadas às necessidades do

sujeito informacional. Discussões acerca de padrões de recuperação da informação colocam em evidência as pesquisas relacionadas à organização do conhecimento, bem como o processo de Indexação atualmente

vivenciado nos ecossistemas digitais [1] organizados em sistemas de informação.

Em contrapartida na contemporaneidade, a Desinformação vem sendo abordada de forma recorrente, evidenciando o desenvolvimento de fenômenos informacionais, primordialmente associados à tecnologia da informação e comunicação, por meio de instrumentos que otimizam a produção, a disseminação e a recuperação da informação, mas que também ampliam as possibilidades de desenvolvimento desse fenômeno, comprometendo, dessa forma, a qualidade das informações.

Nesse cenário, a Indexação configurada como um dos pilares para a recuperação de informações pode se tornar uma aliada no combate à desinformação. A partir disso, surge a indagação desta pesquisa: Qual a relação entre a Indexação e a Desinformação?

A pesquisa tem como objetivo geral: Investigar as possíveis relações entre o conceito de Indexação e a Desinformação.

Realizou-se uma pesquisa bibliográfica para coletar os dados, os quais, posteriormente, foram processados pelo *software Wordart* [2], para apresentação na forma de uma nuvem de palavras, dando subsídios à construção de uma representação conceitual das possíveis relações entre Indexação e Desinformação.

A organização da pesquisa está dividida em quatro seções: I) apresentação dos procedimentos metodológicos; II) o referencial teórico; III) resultados e IV) as considerações finais.

2 Procedimentos Metodológicos

Os procedimentos metodológicos visam guiar a realização da pesquisa seguindo os objetivos estabelecidos. Esta pesquisa é considerada exploratória, e segundo Gil (2002), visa proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses. Dessa forma, o quadro 1 apresenta de forma sistematizada o percurso metodológico diante de cada objetivo específico estabelecido.

Quadro 1. *Percurso Metodológico*

<i>Objetivos específicos</i>	<i>Procedimentos Metodológicos</i>
a) Estruturar teoricamente os temas elencados na pesquisa	Pesquisa bibliográfica
b) Processar os dados coletados	<i>WordArt</i> Nuvem de palavras
c) Apresentação e análise dos dados coletados	Representação conceitual Análise de conteúdo

Fonte: Elaborado pelas autoras (2023)

Os dados coletados na execução desta pesquisa bibliográfica foram processados no *WordArt*, destacando conceitos em torno da temática Indexação e Desinformação e apresentados em forma de nuvem de palavras.

A nuvem de palavras subsidiou o desenvolvimento da representação conceitual, e por meio do cruzamento dos dados destaca-se as possíveis relações existentes analisadas por meio de Análise de Conteúdo Categorical. Bardin (2016) destaca que a Análise de Conteúdo parte da epistemologia advinda do método empírico, que pode ser reinventada de acordo com o tipo de “fala” a qual se dedica e da interpretação que se pretende fazer.

Desse modo, apresenta-se os tipos de suportes informacionais e a população alvo/sujeitos de pesquisa que podem ser analisadas, a partir da aplicação desse método. Na análise do material, busca-se classificá-lo em temas ou categorias que auxiliam a compreensão do que está por trás dos discursos (TAROUCO; MADEIRA, 2013).

Portanto, o resultado da triangulação entre Pesquisa Bibliográfica, Nuvem de Palavras e Análise de Conteúdo abarcam a metodologia utilizada nesta pesquisa, buscando investigar possíveis relações conceituais entre a Indexação e a Desinformação.

3 Referencial Teórico

Esta seção visa sistematizar de forma teórica os temas elencados na pesquisa, na busca de dados que evidenciam conceitos em torno das temáticas Indexação e Desinformação. A pesquisa bibliográfica ocorreu no Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), adotando como expressões de busca: “Indexação” AND “Conceito” e “Desinformação” AND “Conceito”.

A pesquisa bibliográfica está dividida em duas seções: a primeira, 3.1, caracteriza os conceitos em torno da temática Indexação. A segunda, 3.2, caracteriza os conceitos em torno da Desinformação e estão dispostas a seguir.

3.1 Características do conceito de Indexação

A Indexação se caracteriza como uma atividade central nas unidades de informação, “responsável por evidenciar as temáticas dos documentos pela constituição de termos representativos de seus assuntos. Esse processo gera palavras-chave, índices ou cabeçalhos de assunto” (GILLEIVA, 2012, p. 31).

Segundo o Sistema Internacional *World Information System for Science and Technology* (1981, p.84), o processo de Indexação consiste na “ação de descrever e identificar um documento de acordo com seu assunto”. No Brasil, a Associação Brasileira de Normas Técnicas

(ABNT) define a Indexação como o “ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992 p. 92).

De acordo com Fujita (2003, p. 62), “a Indexação em análise documental, sob o ponto de vista dos sistemas de informação, também é reconhecida como a parte mais importante na recuperação da informação” porque condiciona os resultados de uma estratégia de busca. O bom ou mau desempenho da Indexação reflete-se na recuperação da informação feita através de índices.

Nesse sentido, é possível caracterizar a Indexação como o “processo no qual se escolhe o termo ou os termos mais adequados para descrever o conteúdo de um documento” (BAPTISTA; ARAÚJO JÚNIOR; CARLAN, 2010, p. 70). O produto dessa Indexação são os índices e o nível da Indexação varia de acordo com as necessidades dos usuários e das unidades de informação (BAPTISTA; ARAÚJO JÚNIOR; CARLAN, 2010).

Nessa perspectiva, o responsável pela Indexação é igualmente importante dentro do sistema de informação. O indexador tem a função primordial de compreender a leitura ao realizar uma análise conceitual que represente, adequadamente, “o conteúdo de um documento para que ocorra correspondência com o assunto pesquisado pelo usuário” (RUBI; FUJITA, 2003, p. 67).

Para a operacionalização da Indexação, Robredo (2005, p. 165) estabelece quatro etapas para esse processo:

- I) Análise conceitual do conteúdo significativo do documento (Identificação do assunto);
- II) Expressão da análise (conjunto de palavras, frases ou códigos que representem o assunto);
- III) Tradução (descrição dos assuntos relevantes para a linguagem de Indexação);
- IV) Organização (Descrição padronizada dos assuntos de acordo com a sintaxe da linguagem de Indexação).

Robredo (2005) destaca as quatro etapas da Indexação e evidencia que a composição de descritores para caracterização temática dos itens é o elo que garante o acesso pelo usuário, uma vez que os termos definidos pelo indexador serão manipulados na busca no Sistema de Recuperação da Informação (SRI). Logo, diante da necessidade da organização da informação e para tornar efetivo seu uso e gerar conhecimento, se faz necessário que o saber seja representado por meio da Indexação, a fim de facilitar a sua recuperação.

A representação dos conceitos referentes aos documentos está ligada ao discurso e às atividades em um contexto. Desse modo, o indexador precisa ter uma compreensão desse discurso e desses processos. Assim, em um Sistema de Recuperação da Informação, quando “o indexador analisa um documento, deve ter em mente em qual contexto o documento foi produzido, para quem ele será destinado e como será recuperado” (HJØRLAND, 2008, p. 88).

Nesse sentido, destaca-se a necessidade de o indexador possuir conhecimento prévio sobre o assunto a ser indexado, levando em consideração os aspectos linguísticos, lógicos e cognitivos que são exigidos nas estratégias de leitura para a compreensão e interpretação do conteúdo do documento (LIMA; FUJITA; REDIGOLO, 2021).

Sousa e Fujita (2013) ressaltam, ainda, a importância do indexador manter uma interação de dois níveis: de apropriação terminológica e de apropriação conceitual. A familiaridade do indexador, via apropriação da terminologia e descritores de suas enciclopédias particulares, o permite incorporar novas informações aos sistemas, formular expressões de busca e estabelecer diálogos interdisciplinares, resultando na eficácia da recuperação da informação.

Dessa forma, o acesso à informação em sistemas de recuperação da informação está condicionado diretamente ao processo de Indexação de informações, envolvendo discussões que destacam a recuperação da informação por usuários e as políticas estabelecidas diante do processo de Indexação, que envolvam o usuário e o indexador.

3.2 Aspectos teóricos da Desinformação

A busca pelo acesso à informação é uma das tarefas mais importantes vivenciadas pela sociedade da informação e do conhecimento. No entanto, pode-se dizer que não basta apenas ter acesso a todo tipo de informação; é preciso verificar a veracidade e a qualidade da informação obtida e evitar qualquer tipo de Desinformação.

Para Moura, Furtado e Belluzzo (2019), acredita-se que a Desinformação esteja associada diretamente às novas tecnologias de informação e comunicação, o que impacta na produção, gerenciamento e disseminação da informação, comprometendo a qualidade das informações acessadas, em elementos como confiabilidade e autenticidade.

De acordo com Heller, Jacobi e Borges (2020, p.199):

O fenômeno da Desinformação, é uma ambivalência da informação e situa-se claramente no âmbito da Ciência da Informação. Seja oferecendo uma discussão robusta, seja indicando meios de enfrentamento a esse fenômeno, a Ciência da Informação, pode assumir o protagonismo de quem tem a informação como seu objeto.

A Desinformação é classificada diante de três características: Desinformação é informação; Desinformação é uma informação enganosa; Desinformação não é uma informação acidentalmente enganosa (FALLIS, 2015)

Conforme Tadjman e Mikelic (2003, p.1514) “*disinformation*”, em inglês, possui origem russa “*dezinformatsiya*” e “é usada no sentido de informação intencionalmente falsificada, particularmente lançada

por um governo para outro ou para o público". Nesse sentido, Romero-Rodríguez (2011, p.4) afirma que a Desinformação será, portanto:

[...] o resultado da manipulação voluntária, explícita e deliberada da informação, por meio de um resultado pretendido de um processo que emprega um truque específico (seja semântico, técnico ou psicológico); para enganar, desinformar, influenciar, persuadir ou controlar algo ou alguém, geralmente para benefício próprio ou de outros.

Nesse sentido, Brisola (2021) aponta que a Desinformação passa por uma máquina de informar que utiliza a verdade e partes de verdade para desinformar. Demo (2000, p.39) ao afirmar que “desinformar faz parte da informação, assim como a sombra faz parte da luz”, aponta que “a Desinformação é uma informação, apenas com sinais inverso”.

A Desinformação ainda é evidenciada por Fetzer (2004) e Piper (2002) como uma forma problemática de informação: é a informação que desinforma. Carvalho (2001) classifica a Desinformação como uma ação que pode ser observada a partir de dois aspectos: o da Desinformação que afeta o indivíduo pela ignorância e da Desinformação elaborada de modo estratégico.

O termo Desinformação é conceituado no Dicionário Michaelis (2020), como a ação de desinformar; dados falsos que induzem ao erro e ainda a privação de conhecimento sobre determinado assunto; o estado de ignorância. O Dicionário em Inglês de Oxford (SIMPSON; WEINER, 2019) define Desinformação diante de dois termos: *Misinformation* como “informação errada ou enganosa” e *Disinformation* como “informação deliberadamente falsa”.

Ferreira e Pinho Neto (2018, p. 2) sugerem que:

[...] na falta de informação, está a informação que desinforma, que confunde e que desorienta, que remete ao ruído ou à falta de informação e dá margem à alienação dos indivíduos sociais.

Posetti e Ireton (2019) explicam que há uma simbiose da má-informação com a informação incorreta, resultando, desta forma, na Desinformação. Desta maneira, é importante entender a Desinformação e suas facetas, e investigar o que pode ser realizado com a finalidade de promover a autonomia informacional do indivíduo.

4 Apresentação dos resultados

Seguindo a metodologia aplicada para esta pesquisa, esta seção busca destacar os resultados dos dados processados no *Wordart* e apresentados na forma de duas nuvens de palavras.

A figura 1 coloca em evidência os conceitos que caracterizam a Indexação, com base nas publicações de Gilleiva (2012), *Sistema Internacional World Information System for Science and Technology* (1981),

Associação Brasileira de Normas Técnicas (1992), Fujita (2003), Baptista, Araújo Junior e Carlans (2010), Rubi e Fujita (2003), Robredo (2005), Hjørland (2008), Lima, Fujita e Rediogolo (2021) e Sousa e Fujita (2013), recuperados na pesquisa bibliográfica

Figura 1. *Nuvem de palavras de indexação*



Fonte: Elaborado pelas autoras (2023)

A nuvem de palavras de Indexação coloca em evidência os conceitos recuperados, destacando características em torno do conceito de Indexação. Além disso, ressalta conceitos articulados entre recuperação da informação, descrição de conteúdo, acesso, indexador, palavras-chave, representação da informação e sistemas de informação.

Já a figura 2 evidencia os conceitos em torno da Desinformação, recuperados na pesquisa bibliográfica, com base em: Heller, Jacobi e Borges (2020), Fallis (2015), Tudjman e Mikelic (2003), Romero-Rodríguez (2011), Brisola (2021), Demo (2000), Fetzer (2004), Piper (2002), Michaelis (2020), Simpson e Weiner (2019), Ferreira e Pinho Neto (2018), Posetti e Ireton (2019) e Moura, Furtado e Belluzzo (2019).

Figura 2. *Nuvem de palavras da Desinformação*



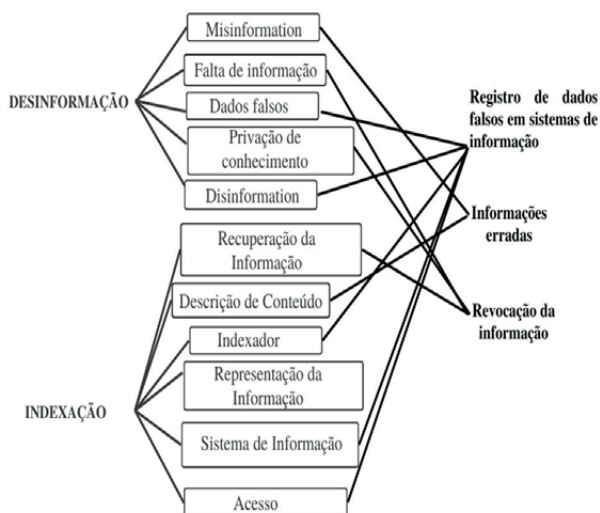
Fonte: Elaborado pelas autoras (2023)

Os termos recuperados e destacados na nuvem de palavras apresentam os conceitos relacionados à Desinformação, como informação errada, privação de conhecimento, dados falsos, *misinformation* e *disinformation*.

Os conceitos que envolvem a Indexação e a Desinformação apresentados nas nuvens de palavras

formam base para o desenvolvimento da representação conceitual, que por meio do cruzamento dos termos aponta as possibilidades de relações entre as temáticas Indexação e Desinformação, conforme disposto na Figura 3:

Figura 3. Representação conceitual entre Indexação e Desinformação



Fonte: Elaborado pelas autoras (2023)

A figura 3 apresenta, de forma conceitual, as prováveis relações entre a Indexação e a Desinformação, evidenciando três ligações: a possibilidade de registro de dados falsos em sistemas de informação e de Indexação de informações erradas e a falta de revocação da informação.

O cruzamento dos termos permitiu evidenciar possibilidades para debates entre registro de dados falsos em sistemas de informação, envolvendo o indexador e o acesso, discussões acerca da ética e o acesso da informação, estabelecendo a possibilidade do registro da Desinformação de forma proposital (*Disinformation*).

As relações expostas na Figura 3 ainda são traçadas diante da indexação de informações erradas na descrição do documento, podendo ocorrer o desenvolvimento da Desinformação (*Misinformation*), enfatizando nesse viés as questões de conhecimento prévio do indexador sobre o contexto de produção do documento e a compreensão da leitura do texto ao selecionar palavras-chaves na representação da informação.

Lopes (2002) afirma que um dos problemas comum e que geram Desinformação referem-se às buscas nos sistemas. O autor explica que, geralmente, nas bases de dados dos sistemas, os campos de título e resumo são registrados pelo indexador com termos de Indexação baseados em Linguagem Controlada. Ocorre que o processo de recuperação da informação pelo usuário acontece por meio da Linguagem Natural. Desse modo,

quando o usuário tenta buscar uma informação pela diferença entre a linguagem que foi indexada pelos agentes (Linguagem Controlada) e a que o usuário fez uso para buscar (Linguagem Natural), não consegue recuperar o dado da forma esperada.

Por fim, as possibilidades são ampliadas no viés recuperação da informação, evidenciando a provável privação de conhecimento, caso não haja disponibilidade da informação, engendrando discussões entre os aspectos: usuário da informação, indexador e a Tecnologia da Informação e comunicação. Nesse sentido, Andrade (2021) enfatiza que a falta de padronização no processo de Indexação pode gerar Desinformação, impossibilitando a recuperação de dados que gere conhecimento.

Portanto, as relações aqui traçadas podem ser alvo do desenvolvimento de novas pesquisas, ampliando discussões no escopo da Ciência da Informação, colocando em foco temáticas como o fenômeno da informação e o processo da Indexação.

5 Considerações finais

Esta pesquisa permitiu a reflexão sobre a relação entre a Indexação e a Desinformação, destacando características conceituais das temáticas. Por meio dos conceitos recuperados, construiu-se a representação conceitual entre as duas temáticas destacando possíveis relações existentes.

Diante disso, percebeu-se três possibilidades de relações: a Indexação de Informações erradas e dados falsos, de forma errônea ou proposital; e a falta de informação, o que pode gerar problemas de recuperação da Informação e a privação de conhecimento.

Nesse sentido, destaca-se a Desinformação como um fenômeno informacional, que associado às características do processo de indexar, podem acarretar problemas diante da revocação da informação, da presença de dados falsos e da falta de informação.

Assim, percebe-se a necessidade de aprofundar estudos sobre as temáticas, na intenção de reduzir ou amenizar possibilidades de danos ocasionados pela Desinformação.

Notas

[1] Boley e Chang (2007, p.8), afirmam que “um Ecossistema Digital trata-se de uma nova arquitetura de rede e um ambiente colaborativo que aborda o ponto fraco de cliente-servidor, ponto-a-ponto, rede e serviços web. Ou seja, em um Ecossistema Digital, a comunidade é aberta e não há permanente necessidade de controle centralizado ou fixo, e, dessa maneira, um agente pode ser um cliente e um servidor ao mesmo tempo”.

[2] O software está disponível no seguinte endereço: <https://wordart.com> (WORDART, 2023).

Referências

- ANDRADE, P. C. **Autorias na comunicação científica: a falta de padronização como indício de desinformação**. 2021. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Universidade Federal do Rio Grande, Rio Grande, 2021.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR 12676: informação e documentação: indexação**. Rio de Janeiro: ABNT, 1992.
- BAPTISTA, D.M.; ARAÚJO JÚNIOR, R.H.; CARLAN, E. O escopo da análise da informação. *In: ROBREDO, J.; BRÄSCHER, M.(orgs.). Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento*. Brasília: IBICT, 2010. p. 61-80.
- BARDIN, L. **Análise de conteúdo**. Lisboa: Edições 70, 2016.
- BOLEY, H.; CHANG, E. “Digital Ecosystems: Principles and Semantics”. *In: INAUGURAL IEEE INTERNATIONAL CONFERENCE ON DIGITAL ECOSYSTEMS AND TECHNOLOGIES*, 1., 2007, Cairns. *Anais [...]*. [Cairns]: IEEE, 2007. p.398-403.
- BRISOLA, A. C. C. **Competência crítica em informação como resistência à sociedade da desinformação sob um olhar freiriano: diagnósticos, epistemologia e caminhos ante as distopias informacionais contemporâneas**. Orientador: Marco André Feldman Schneider. 2021. 295 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2021.
- CARVALHO, K. Disseminação da Informação e informação de inteligência organizacional. *DataGramaZero – Revista de Ciência da Informação*, Rio de Janeiro, v. 2, n. 3, p.1-9, jun. 2001. Disponível em: <http://www.dgz.org.br>. Acesso em: 20 jan. 2016.
- DEMO, P. Ambivalências da sociedade da informação. *Ciência da Informação*, Brasília, v. 29, n. 2, p. 37-42, maio/ago. 2000. Disponível em: <http://https://www.scielo.br/j/ci/a/797VnWgmBHvsnvbJJyztKnP/?format=pdf&lang=pt>. Acesso em: 02 abr. 2019.
- FALLIS, D. What is disinformation? *Library Trends*, [S.l.], v. 63, n. 3, p. 401-426, 2015.
- FERREIRA, T. E. L. R.; PINHO NETO, J. A. S. Na contramão da informação preventiva: desinformação sobre prevenção de HIV/AIDS. *Biblionline*, João Pessoa, v. 14, n. 3, p. 3-13, 2018. Disponível em: <https://periodicos.ufpb.br/index.php/biblio/article/view/41364>. Acesso em: 02 abr. 2019.
- FETZER, J. H. **Disinformation: The use of false information**. *Minds and machines*, [S.l.], v. 14, n. 2, p. 231-240, 2004.
- FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, v. 1, n. 1, p. 60-90, jul./dez. 2003. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2089/2219>. Acesso em: 10 dez. 2022.
- GIL, A. C. Como classificar as pesquisas. *In: GIL, A. C. Como elaborar projetos de pesquisa*. São Paulo: Atlas, 2002. p. 44-45.
- GILLEIVA, I. Aspectos conceituais da indexação. *In: GILLEIVA, I.; FUJITA, M. S. L.(eds.). Política de indexação*. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 31-106.
- HELLER, B.; JACOBI, G.; BORGES, J. Por uma compreensão da desinformação sob perspectiva da Ciência da Informação. *Ciência da Informação*, Brasília, v.49, n. 2, p. 189-204. out. 2020. Disponível em: <https://revista.ibict.br/ciinf/article/view/5196>. Acesso em: 10 dez. 2022.
- HJORLAND, B. What is knowledge organization (KO)? **Knowledge Organization**. [S.l.], v.35, n.2-3, p.86-101. 2008. Disponível em: https://is.muni.cz/el/1421/jaro2016/VIKBA06/um/56249939/HJORLAND_Birger_What_is_knowledge_organization_KO_.pdf. Acesso em: 10 dez. 2022.
- LIMA, G. A.; FUJITA, M. S. L.; REDIGOLO, F. M. A importância do contexto para a indexação. **Ponto de Acesso**, Salvador, v. 15, n. 3, p.283-302, dez. 2021.
- LOPES, I. L. Uso das linguagens controlada e natural em bases de dados: revisão da literatura. **Ciência da Informação**, Brasília, v. 31, n.1, p. 41-52, jan./abr. 2002. Disponível em: <https://www.scielo.br/j/ci/a/WhYch5gHnWYPCPY3NPwbkR/?format=pdf&lang=pt>. Acesso em: 10 dez. 2022.
- MICHAELIS. **Dicionário Brasileiro da Língua Portuguesa**. São Paulo: Melhoramentos, 2020. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/>. Acesso em: 05 jul. 2022.
- MOURA, A. R. P.; FURTADO, R. L.; BELLUZZO, R. C. B. Desinformação e competência em informação: discussões e possibilidades na Arquivologia. **Ciência da Informação e Revista**, Maceió, v. 6, n. 1, p. 37-57, 2019. Disponível em: <https://www.seer.ufal.br/index.php/cir/article/view/7063>. Acesso em: 05 jul. 2022.
- PIPER, K. L. **Cartographic fictions: Maps, race and identity**. New Brunswick: Rutgers University Press, 2002.
- POSETTI, J.; IRETON, C. **Jornalismo, Fake News, & Desinformação: manual para Educação e Treinamento em Jornalismo**. [S. l.: s. n.], 2019. Disponível em: <https://blogue.rbe.mec.pt/jornalismo-fake-news-desinformacao-2352704>. Acesso em: 25 ago. 2022.
- ROBREDO, J. **Documentação de hoje e de amanhã: uma abordagem revisitada e contemporânea da Ciência da Informação e de suas aplicações biblioteconômicas, documentárias, arquivísticas e museológicas**. 4. ed. rev. ampl. Brasília: Edição do autor, 2005.
- ROMERO-RODRÍGUEZ, L. M. **La manipulación informativa y la desinformación: la anomia de los receptores y el fomento de las víctimas propiciatorias**. Almería: Universidad de Almería, 2011. Disponível em: https://www.researchgate.net/profile/Luis-Romero-Rodriguez/publication/266394270_La_manipulacion_informativa_y_la_desinformacion_La_an

omia_de_los_receptores_y_el_fomento_de_victimas_propiciatorias/links/54316e3e0cf277d58e9829b6/La-ma-nipulacion-informativa-y-la-desinformacion-La-anomia-de-los-receptores-y-el-fomento-de-victimas-propiciatorias.pdf. Acesso em: 10 dez. 2022.

RUBI, M.P.; FUJITA, M. S. L. Elementos de política de indexação em manuais de indexação de sistemas de informação especializados. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 8, n. 1, p.66-77, jan./jun. 2003.

SIMPSON, E. S. C.; WEINER, J. A. (eds.). **Oxford Dictionaries**: [Post-truth]. Oxford: Clarendon Press, 2019. Disponível em: <https://en.oxforddictionaries.com/definition/post-truth>. Acesso em: 25 ago. 2022.

SISTEMA INTERNACIONAL WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY. **Indexing principles**. Paris: UNESCO, 1981.

SOUSA, B. P.; FUJITA, M. S. L. A classificação bibliográfica no contexto do tratamento temático da

informação: um estudo com o protocolo verbal individual em bibliotecas do Instituto Federal de Educação, Ciência e Tecnologia (IF's) **Revista ACB**, Florianópolis, v. 18, n. 1, p. 796-813, 2013. Disponível em: <https://revista.acb.org.br/racb/article/view/868>. Acesso em: 10 dez. 2022.

TAROUCO, G. S.; MADEIRA, R. M. Esquerda e direita no sistema partidário brasileiro: análise de conteúdo de documentos programáticos. **Revista Debates (UFRGS)**, Porto Alegre, v.7, n.2, 2013.

TUDJMAN, M.; MIKELIC, N. Information science: Science about information, misinformation and disinformation. *In*: INFORMING SCIENCE+ INFORMATION TECHNOLOGY EDUCATION, 3., 2003, Pori. **Proceedings** [...]. Pori: [Informing Science Institute], 2003. p. 1513-1527. Disponível em: <https://proceedings.informingscience.org/IS2003Proceedings/docs/204Tudjm.pdf> Acesso em: 10 dez. 2022.

WORDART. [S.l.: s.n.], 2023. Disponível em: <https://wordart.com>

John Dewey e Birger Hjørland: Pragmatismo e lógica nos fundamentos da Organização do Conhecimento

John Dewey and Birger Hjørland: pragmatism and logic in the foundations of Knowledge Organization

José Claudio Matos

Universidade do Estado de Santa Catarina – UDESC, jose.matos@udesc.br

Resumo

Trata-se de um ensaio crítico que discute a atitude intelectual do pragmatismo na fundamentação do campo da organização do conhecimento (OC). Supõe-se, como ponto de partida, que há uma estreita relação entre a OC e a lógica - compreendida como a ciência das regras do raciocínio válido. Analisa o artigo “What is Knowledge Organization (KO)?”, de Hjørland, especialmente as duas citações que seu autor faz de John Dewey, para apresentar sua defesa do pragmatismo como um adequado fundamento para a OC. Tenta defender, como resultado, que há relação entre uma visão instrumental da lógica e uma visão instrumental da classificação, nos moldes do falibilismo epistemológico. Esta relação não foi apresentada expressamente por Hjørland, mas o ensaio argumenta que é o ponto chave na compreensão da contribuição de Dewey para o debate sobre a OC.

Palavras-chave: John Dewey, Hjørland, Organização do conhecimento, lógica, pragmatismo.

Abstract

This is a critical essay that discusses the intellectual attitude of pragmatism in the foundation of the field of knowledge organization (KO). It is assumed, as a starting point that there is a close relationship between the KO and logic - understood as the science of the rules of valid reasoning. It analyzes Hjørland's article “What is Knowledge Organization (KO)?”, especially the two quotes its author makes of John Dewey, to present his defense of pragmatism as a suitable foundation for the KO. It tries to defend, as a result, that there is a relationship between an instrumental view of logic and an instrumental view of classification, along the lines of epistemological falibilism. This relationship was not expressly presented by Hjørland, but the essay argues that it is the key point in understanding Dewey's contribution to the debate on KO.

Keywords: John Dewey, Hjørland, Knowledge Organization, Logic, Pragmatism.

Resúmen

Este es un ensayo crítico que discute la actitud intelectual del pragmatismo en la fundación del campo de la organización del conocimiento (OC). Se supone, como punto de partida, que existe una estrecha relación entre la OC y la lógica, entendida como la ciencia de las reglas del razonamiento válido. Analiza el artículo de Hjørland “What is Knowledge Organization (KO)?”, especialmente las dos citas que su autor hace de John Dewey, para presentar su defensa del pragmatismo como fundamento adecuado para la OC. Intenta defender, como resultado, que existe una relación entre una visión instrumental de la lógica y una visión instrumental de la clasificación, en la línea del falibilismo epistemológico. Esta relación no fue presentada expresamente por Hjørland, pero el ensayo argumenta que es el punto clave para comprender la contribución de Dewey al debate sobre la OC.

Palabras clave: John Dewey, Hjørland, Organización del conocimiento, lógica, pragmatismo.

1 Introdução

Este trabalho se inscreve no campo da Organização do Conhecimento (OC), em sua relação com a epistemologia de natureza mais filosófica. Trata-se de uma reflexão na forma de ensaio, em que problemas e argumentos de caráter mais especulativo são considerados em detalhe, a fim de extrair conclusões que contribuam para a compreensão dos fundamentos da OC. O assunto em discussão envolve a consideração da atitude intelectual do pragmatismo na fundamentação do campo da OC. Supõe-se, como ponto de partida, que há uma estreita relação entre a OC e a lógica -

compreendida como a ciência das regras do raciocínio válido. Certos princípios lógicos teriam relevância na construção e avaliação crítica dos Sistemas de Organização do Conhecimento (SOCs), considerados como operações formais de que tratam as teorias da lógica. No bojo desta discussão, se espera mostrar que a atitude própria da corrente pragmatista implica em uma concepção instrumental e social dos sistemas de organização de conhecimento, defendida a partir de uma concepção instrumental da própria lógica. Esta problemática é formulada a partir da análise crítica do artigo “*What is Knowledge Organization (KO)?*” (HJØRLAND, 2008). Hjørland pretende propor com

este artigo “um entendimento da OC baseado em uma explícita teoria do conhecimento” (HJØRLAND, 2008, p. 86). Hjørland propõe situar a OC na área de pesquisas da Biblioteconomia e Ciência da Informação (BCI), embora reivindique uma considerável autonomia intelectual para a OC. Prover a OC com uma reflexão epistemológica sobre seus fundamentos atende a uma demanda que a prática profissional da BCI nem sempre evidencia. Grande parte das produções neste campo se refere a pesquisas acerca da construção, melhoramento e aplicação de SOCs. Para Hjørland, tal como ele defende em seu artigo, esta demanda por fundamentação teórica é satisfeita por uma teoria falibilista-pragmatista do conhecimento.

O filósofo John Dewey, ao propor que a lógica recebe sua forma, seus métodos e seus problemas das práticas humanas a ela relacionadas, inspira Hjørland a uma importante reflexão sobre os aspectos epistemológicos com os quais a organização do conhecimento deve lidar, a fim de enriquecer sua atuação e participar do avanço da ciência. Uma pesquisa revisionária que pudesse revelar os vestígios da participação de John Dewey na BCI ainda está por ser feita. Tal pesquisa faria uma relevante contribuição ao campo e também à tradição de comentário deste pensador. Daqui para diante no presente ensaio, será usado apenas “Dewey” para se referir a John Dewey, lembrando o prezado leitor de não confundir com Melvil Dewey, autor do famoso sistema de classificação decimal.

Hjørland afirma que: “Diferentes pontos de vista necessitam diferentes sistemas de organização” (HJØRLAND, 2008, p. 95). Hjørland se refere, claramente, aos diversos tipos de SOCs, construídos e estudados pela OC. Os sistemas devem ser reconhecidos como instrumentos construídos a partir da necessidade de responder questões que foram formuladas em contextos de investigação, em vista de valores e objetivos compartilhados. O aspecto original da posição de Hjørland é ir diretamente contra a pretensão de universalidade e de valor fixo e formal dos SOCs. A posição tradicional é resumida por ele como a ideia de que “um tamanho serve em todos” (HJØRLAND, 2008, p. 94). Seria como se o valor de um sistema pudesse ser medido por critérios formais. Não se consideraria relevante o ponto de vista mantido pelas comunidades que empregam o sistema, nem os diferentes contextos de consulta da informação por eles organizada.

A visão que Dewey inspira a Hjørland é bem diferente: as características, métodos, estrutura sintática e semântica dos SOCs devem evoluir continuamente e se adaptar ao contexto, às necessidades e aos interesses das comunidades de investigação que empregam tais instrumentos. A própria lógica, para Dewey, precisa atender a este princípio: sistemas lógicos diferentes são

formulados para atender a diferentes necessidades de formalização.

Seguindo o método de análise crítica e reflexiva dirigida ao discurso de Hjørland em seu artigo, este ensaio espera esclarecer a relação entre a lógica e a OC, traçando um percurso argumentativo a partir das referências e da interpretação que este autor faz do pensamento de Dewey. O pragmatismo implica uma concepção particular acerca de diversos instrumentos da atividade intelectual sistemática, entre eles a própria lógica. A visão instrumental da classificação, defendida por John Dewey e evocada por Hjørland é, ela mesma, uma alternativa à visão tradicional da lógica, como ciência de princípios universais da correção e exatidão.

2 A OC em sentido restrito e em sentido amplo

Para responder à pergunta título de seu artigo, Hjørland faz uma distinção entre um sentido restrito e um sentido amplo da OC. Segundo ele:

No sentido restrito a Organização do Conhecimento (OC) trata de atividades como descrição de documentos, indexação e classificação efetuada em bibliotecas, bases de dados bibliográficos, arquivos, e outros tipos de ‘instituições de memória’ (HJØRLAND, 2008, p. 86).

Este é o tipo de atividade de natureza aplicada, voltada à construção e aperfeiçoamento de classificações e vocabulários controlados. Por sua vez:

No sentido mais amplo, a OC trata da divisão social do labor mental, isto é, a organização das universidades e outras instituições de pesquisa e educação superior, a estrutura das disciplinas e profissões, a organização social dos meios, a produção e disseminação de ‘conhecimento’, etc (HJØRLAND, 2008, p. 86).

No sentido comum, a OC é uma especialidade inerente a demandas que estão no coração da BCI, como campo de estudos e atividades. No sentido amplo, a OC é uma reflexão epistemológica de grande alcance, com tonalidade mais filosófica.

Hjørland afirma que “A OC tem sido principalmente uma atividade prática, sem muita teoria” (HJØRLAND, 2008, p. 87). O que o autor acentua é que a OC estreita, como atividade aplicativa e normativa, não desenvolve a si mesma. Só a OC na sua versão superior, ampla e mais reflexiva é que pode impulsionar novos desenvolvimentos no campo. Sua proposta é uma revisão da parte teórica da OC, a fim de atualizar e refinar sua fundamentação. Ele passa em revista diversas abordagens relacionadas ao assunto, a saber: A abordagem tradicional, a abordagem facetada, a recuperação da informação, a abordagem centrada no usuário, a abordagem bibliométrica e, por fim, a análise de domínio, proposta por ele mesmo. Estas abordagens

representariam etapas do desenvolvimento da OC e da própria BCI.

A crítica de Hjørland não se volta à eficiência técnica destas abordagens, mas sim à visão limitada de seus pressupostos epistemológicos, ainda assentados no que o autor chama de visão positivista do conhecimento. Para definir esta visão positivista, Hjørland cita Dewey, da passagem em que este comenta o modo de considerar a classificação como “uma transcrição vaga ou uma duplicata de algum preexistente, deliberado e já acabado arranjo da natureza” (DEWEY, 1958, p. 158). Dewey critica este modo de pensar, assim como faz Hjørland. Este identifica a abordagem fixista da classificação como sendo própria do pensamento positivista.

A persistência de uma atitude formalista e normalizadora em relação aos sistemas de organização do conhecimento é combatida por Hjørland, munido de uma perspectiva instrumental e pragmática do conhecimento, proveniente da leitura de Dewey. Dewey é um defensor do refinamento dos métodos de descoberta, prova e experimentação das teorias. O autor critica o formalismo que vê a sofisticação dedutiva como um fim em si mesmo. Reconhece que, ao longo do avanço histórico dos campos de conhecimento, há uma divisão de profissões, segundo a qual se forma uma comunidade de especialistas nas ciências formais dedutivas. Estas ciências, como a lógica e, em certo sentido, a OC, fazem a função de instrumento para as demais áreas do conhecimento. É bom que a especialização faça crescer o interesse nos estudos formais. Sobre os instrumentos de classificação ele é sutil ao explicar que:

Somente por não haver sido criado com vistas a uma aplicação determinada, por ser instrumento altamente generalizado, ele é, por igual, mais flexível, mais adaptado a usos e aplicações não previstas; esse instrumento pode sem dúvida ser empregado no trato com problemas nunca previstos (DEWEY, 1958, p. 153).

Dewey argumenta que inventores de instrumentos nem sempre são os seus principais usuários. Por isso, sua evolução depende da comunicação bem sucedida entre inventores e usuários. Os instrumentos intelectuais, como a lógica e as demais ciências dedutivas, são em geral bem mais sofisticados do que os instrumentos físicos, pois são alvo de maior manipulação teórica.

3 Pragmatismo

A fundamentação pragmatista na OC remete ao movimento intelectual do final do século XIX e início do século XX, nos Estados Unidos da América. Na sua origem se destacam os nomes de Charles Peirce, William James e John Dewey. Bertrand Russell, em sua *História do Pensamento Ocidental* (2003), menciona

Dewey como o mais original representante do pragmatismo. Diz ele:

Desde o final do século XIX, a força dominante na filosofia americana vem sendo uma forma modificada de pragmatismo. O principal representante deste movimento foi John Dewey (RUSSELL, 2003, p. 475).

Russel atribui a Dewey uma “concepção instrumental da lógica” (RUSSELL, 2003, p. 476), na qual, conforme Russell, Dewey estaria dando a entender que “sob os tradicionais e objetivos conceitos da lógica, nada genuinamente novo pode surgir no mundo. É o anseio de novidade e de expansão social que inspira essa espécie de teorização” (RUSSELL, 2003, p. 477). Embora possamos concordar com Russell sobre a originalidade de Dewey, é especialmente importante reconhecer uma espécie de linha hereditária que vai do pensamento de Peirce a James e, finalmente, a Dewey.

Este cuidado com a caracterização da atitude pragmatista revela, por exemplo, que é a Peirce que se deve atribuir uma concepção precisa e bem formulada do *falibilismo* epistemológico. Russell tenta sintetizar a opinião de Peirce, em matéria de investigação e pesquisa: “Nunca podemos estar seguros de não ter cometido um erro. A esta teoria geral da investigação Peirce chama de ‘falibilismo’” (RUSSELL, 2003, p. 445). Segundo o falibilismo, a possibilidade de erros e falhas é constante em todo processo de busca do conhecimento. Por causa disso, qualquer SOC - assim como qualquer teoria ou resultado da pesquisa - sempre deveriam estar abertos à revisão. Não existe um ponto final na investigação. Para o pragmatista a investigação é contínua, sem um ponto final determinado. O mesmo deve ser suposto dos sistemas que pretendem organizar o conhecimento. Assim, a mera eficiência formal de um sistema não é uma qualidade. Sua incompletude não é uma falha ou defeito, mas antes parte da própria natureza desses instrumentos.

Para Dewey, não é possível sustentar a noção de que categorias fixas possam ser usadas para representar a diversidade de uma realidade em constante mudança. A reconstrução contínua e a mudança qualitativa não devem ser vistas como sinais de fraqueza ou incapacidade, mas sim como o próprio modo de ser do conhecimento e de sua comunicação. Portanto, qualquer sistema de divisão, organização e classificação das ideias e realizações da cultura, precisa reconhecer seu aspecto parcial e falível, a fim de poder dar conta de sua tarefa. Mesmo que a falibilidade possa ser vista como uma fraqueza, a possibilidade de mudança e adaptação parece redimir as tentativas de organização, em relação às demandas dos usuários.

Hjørland faz duas citações de textos de Dewey no seu artigo. A primeira é da introdução ao livro de Henry Bliss, *The Organizatin of Knowledge and the System of*

Sciences (BLISS, 1929), e a segunda é do livro *Filosofia em Reconstrução* (DEWEY, 1958), inicialmente publicado em 1920. Neste livro, no capítulo VI, intitulado “O significado da reconstrução lógica”, Dewey apresenta sua visão instrumentalista da lógica e completa seu argumento com reflexões sobre a classificação. Esta conexão entre os sistemas lógicos e os sistemas de classificação e organização do conhecimento foi o que mais interessou Hjørland, pela defesa veemente ali encontrada, de que tais sistemas atendem a objetivos estabelecidos no curso da investigação. O valor de tais sistemas é avaliado por uma espécie de critério social de eficiência e validade.

Para Dewey, é a natureza da dificuldade a ser superada que determina o tipo de instrumento intelectual a ser empregado em cada tentativa de organização e classificação.

A dificuldade, isto é, o tipo de dificuldade da qual resulta ou se produz a atividade de pensar, é aquela que é incompleta e que se amplia, aquela em que o que se encontra já em existência pode ser empregado como sinal pelo que se infere o que está provavelmente para acontecer (DEWEY, 1958, p. 147).

Dewey está explicando o processo de pensamento em termos empíricos que, para ele, não se separam do aspecto formal, mas ambos se retroalimentam e dão um ao outro seu significado. Ele fala de um “senso do significado da dificuldade” (DEWEY, 1958, p. 146), pelo qual é possível antecipar resultados pelo raciocínio, a partir de evidências presentes. Esta antecipação generaliza a partir de alguma regularidade ou padrão, que resulta da classificação do problema num esquema compreensivo. Muito interessante que Dewey fale dos mecanismos de pensamento no mesmo capítulo onde fala de classificação – pelo qual se interessou Hjørland, como uma epistemologia que daria fundamento à OC.

Dewey acrescenta que “todo ato inteligente, toda função mental inteligente resulta num incremento da liberdade de ação – emancipação, independência da sorte e da fatalidade” (DEWEY, 1958, p. 149). Há no argumento de Dewey neste capítulo uma epistemologia com forte tonalidade social, cuja parte formal é a lógica concebida sobre bases experimentais, em que o processo de raciocínio é formulado instrumentalmente, isto é, a lógica como *organon*. O objetivo social para além da própria construção de conhecimento é o crescimento da liberdade, como um poder de ação inteligente sobre as coisas. Sobre sua teoria da liberdade como resultado do controle inteligente dos meios de ação, Dewey escreve o artigo “Philosophies of Freedom” (DEWEY, 1998), em 1928, um ano antes da sua contribuição ao livro de Bliss.

4 Falibilismo e classificação

Ainda é preciso perguntar: de que forma esta teoria da investigação de Dewey desemboca na discussão sobre classificação? Dewey fala da ciência institucionalizada e legitimada por padrões públicos de aceitação e crítica. Depois passa a abordar certa sociologia do conhecimento e discutir uma divisão social do trabalho, para chegar aos profissionais da informação, responsáveis pela classificação. Diz ele que: “Precisamente no grau em que o pensamento se vai desvencilhando dos fins fixados por costumes sociais, uma divisão social do trabalho vem a desenvolver-se” (DEWEY, 1958, p. 151). Discutindo a investigação em seu aspecto social, Dewey criou o cenário para vir a falar de classificação: a classificação é ocupação especializada de um grupo de profissionais que assumem esta função na divisão social do trabalho relativo ao conhecimento. Sem a classificação, não é possível ciência como um empreendimento social, institucionalizado e de grande escala. Finalmente, classificação em grande escala é uma atividade que participa do ramo empírico e reconstrutivo da lógica segundo o ponto de vista instrumental, tal como Dewey pretendia caracterizar esta ciência.

Hjørland afirma sobre a cooperação entre profissionais do conhecimento que:

As escolas de biblioteconomia têm tradicionalmente educado bibliotecários e especialistas em informação; escolas de linguagem, para propósitos especiais, têm educado tradutores; escolas de administração têm educado gestores de informação; escolas de ciência da computação tem educado engenheiros de *software* (HJØRLAND, 2008, p. 98).

Essa divisão das profissões, já comentada por Dewey, significa que a divisão do trabalho acentua a importância da classificação para o avanço do empreendimento social da busca pelo conhecimento. Hjørland faz uma defesa não só da interdisciplinaridade na formação dos especialistas em OC. Recomenda o esforço teórico em níveis fundamentais, que são comuns aos problemas que a OC tem de tratar. Entre estes problemas, está a necessidade da epistemologia em sua fundamentação.

O reconhecimento do aspecto falível do conhecimento e dos sistemas para sua organização é o ponto central de todo o argumento de Hjørland. Ele aponta para a passagem da introdução ao livro de Bliss, na qual Dewey afirma:

Uma classificação de livros, para ser efetiva no sentido prático, deve corresponder às relações de assuntos e sua correspondência pode ser assegurada unicamente pela forma como a organização intelectual ou conceitual é baseada na ordem inerente aos campos de conhecimento, que por sua vez espelha a ordem da natureza (DEWEY, apud BLISS, 1929, p. viii).

Dewey parece manifestar a exigência por uma correspondência formal entre a classificação e os estados das coisas no mundo real. A passagem, lida fora de seu contexto, transmite a impressão de que possa se chegar ao sistema final de classificação, um sistema universal completo e imutável, capaz de organizar todo o conhecimento passado e futuro. Esta promessa ficaria vinculada à noção de que o conhecimento *espelha* a ordem da natureza. No artigo “Knowledge Organization: a New Science?” (DAHLBERG, 2006), a autora descreve as duas funções principais da OC:

Aqui alguém pode distinguir entre duas aplicações de se organizar conhecimento, a saber: a) A construção de sistemas de conceitos; e, b) A correlação com, ou o mapeamento de unidades de tal sistema de conceitos com objetos da realidade (DAHLBERG, 2006, p. 12).

Perceba-se que a ideia de correlação ou *espelhamento* entre o sistema e a realidade é defendida pela autora. Uma parte da corrente tradicional da OC se compromete com essa ideia, investindo uma pesada expectativa sobre o poder e o alcance dos SOCs.

Ainda sobre a passagem de Dewey na introdução ao livro e Bliss, Hjørland comenta:

Esta citação está em acordo com a visão tradicional do conhecimento como um reflexo neutro e objetivo da realidade. Ela é, contudo, uma má representação da visão pragmática de John Dewey do conhecimento e da classificação” (HJØRLAND, 2008, p. 97).

Parece, pela interpretação de Hjørland, que haveria um choque no pensamento de Dewey, entre uma visão tradicional, formalista, absoluta do conhecimento de um lado, e uma visão construtiva, falibilista e pragmática do outro.

Toda a questão nesta parte do debate se refere a quanto à primeira citação de Dewey é uma “má representação” do seu pensamento. Hjørland chega a afirmar que: “John Dewey não aceita a metáfora do espelho” (HJØRLAND, 2008, p. 97). A interpretação defendida no presente ensaio é que Dewey diz “espelha” (*mirrors*), mas isso não deve ser entendido como uma teoria da correspondência. Nove anos antes, Dewey já havia se posicionado como um instrumentalista do conhecimento, que reconhecia o aspecto social e mutável das tentativas de explicação. Para Dewey, toda teoria envolve escolha, e toda escolha se faz na base de valores. Uma representação final e neutra da classificação vai contra os fundamentos filosóficos do pragmatismo – ou naturalismo –, eixo central da filosofia de Dewey.

Adiante em sua carreira, em 1938, Dewey publica o livro *Logic – The Theory of Enquiry*. Nesta obra ele mantém o princípio instrumentalista da lógica, afirmando que seus sistemas formais emanam das atividades de investigação efetivamente conduzidas pelas comunidades e não o contrário. Dewey afirma que

“as formas lógicas são acrescentadas aos materiais existenciais em virtude do controle exercido sobre a investigação de modo tal que esta possa alcançar seus objetivos” (DEWEY, 1974, p. 214). Esta passagem explica que a formalização lógica, assim como a formalização dos sistemas de classificação e organização, não representa uma estrutura ideal, cuja validade resida mais além dos processos efetivos de busca e utilização do conhecimento. As estruturas formais, tais como sistemas lógicos e SOCs, são tentativas de anteder demandas das comunidades que delas se beneficiam. Na medida em que haja mudanças nas teorias, no ambiente social, nos interesses das comunidades, tanto essas formas lógicas como os sistemas de classificação precisam ser revisados, modificados, às vezes até mesmo abandonados para dar lugar a outros.

Hjørland claramente abraça o falibilismo epistemológico e, em decorrência de tal falibilismo, uma proposta pragmatista para a classificação.

A implicação do falibilismo é que não podemos entender os documentos como representando conhecimento, como entendido tradicionalmente. Não devemos falar de conhecimento ou organização do conhecimento, mas sobre alegações de conhecimento e organização de alegações de conhecimento (HJØRLAND, 2008, p. 98).

Essa compreensão dos fundamentos conceituais do campo tem como consequência destacar os aspectos sociais do trabalho da OC. Quanto aos sistemas atualmente em uso e as aspirações futuras do campo da OC, a recomendação é que a atitude de constante reflexão e investigação seja levada adiante. Um princípio da continuidade da experiência, da continuidade da investigação especializada, pode ser formulado em termos de que *não há classificação suficientemente universal que não venha a dar sinais de insuficiência*, frente a novos interesses e novas áreas de investigação. Diz Dewey que “toda classificação pode ser igualmente perfeita, quando as diferenças relativas aos fins sejam levadas em conta” (DEWEY, 1958, p. 157). Assim sendo, diferentes sistemas podem conviver, e possivelmente conversar entre si. Não existe a perspectiva de um sistema ou uma linguagem formal universal, que fosse padrão de correção para os diversos sistemas particulares.

5 Conclusão

Esta visão fornece fundamento para uma abordagem da OC e da BCI em geral, segundo a qual o interesse social e as consequências sociais devem estar no horizonte das reflexões e das atividades de seus praticantes. Com isso Hjørland concordaria integralmente, a ponto de ter citado Dewey de forma tão categórica em seu artigo.

O que Hjørland, em seu artigo, não trata explicitamente, é a presença da lógica na origem de toda essa reflexão. É preciso estabelecer claramente este laço entre a lógica e a OC de caráter mais amplo e teórico, a fim de aprofundar a compreensão sobre o alcance e as limitações e as possibilidades inúmeras de aplicação dos SOCs. A origem de toda esta realização teórica, que significa aproximar o falibilismo pragmatista da discussão sobre as bases da OC, é justamente a iniciativa de Dewey de defender que a lógica tem uma função importante na sistematização e organização dos inúmeros conceitos e regras, componentes do conhecimento socialmente considerado válido.

Referências

- BLISS, Henry Evelyn. **The Organization of Knowledge and the System of Sciences**. New York: Henry Holt and Company, 1929.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, Frankfurt, v. 33, n. 1, p. 11-19, 2006.
- DEWEY, John. **Filosofia em reconstrução**. São Paulo: Companhia Editora Nacional, 1958.
- DEWEY, John. Lógica – A teoria da investigação. Capítulo IV: “O padrão da investigação”. *In: Os pensadores*. São Paulo: Abril Cultural. 1974.
- DEWEY, John. “Philosophies of Freedom”. *In: The essential John Dewey*. Indianápolis: Indiana University Press, 1998. Vol. 2: Éthics, Logic, Psychology.
- HJØRLAND, Birger. What is Knowledge Organization (KO)? **Knowledge Organization**, 35(3/2), 86-101. 2008.
- RUSSELL, Bertrand. **História do pensamento ocidental**. Rio de Janeiro. Ediouro. 2003.

A modelagem conceitual, organização do conhecimento e organização da informação: aproximações epistemológicas no universo bibliográfico

Conceptual modelling, knowledge organization and information organization: epistemological approaches in the bibliographic universe

Rhuan Henrique Alves de Oliveira (1), Fabiano Ferreira de Castro (2)

(1) Universidade Federal de São Carlos - UFSCar, Rod. Washington Luiz, s/n - Monjolinho, São Carlos - SP, rhuan.henrique.oliv@gmail.com. (2) fabianocastro@ufscar.br

Resumo

A modelagem conceitual, a partir da elaboração dos modelos conceituais do universo bibliográfico, passou a atuar e se consolidou como metodologia e objeto de pesquisa também na área da Organização do Conhecimento e da Organização da Informação. Dessa maneira, essa pesquisa procura investigar aproximações epistemológicas entre a Modelagem Conceitual, a Organização do Conhecimento e a Organização da Informação. Para tanto, utiliza-se da topologia de Capurro (2003), da Teoria do Conceito de Dahlberg e do modelo IFLA LRM, se embasando em uma revisão de literatura de caráter epistemológico, tendo como referências os trabalhos de Dahlberg (1978), Svenonius (2000), Bräscher e Café (2008), Riva, LeBoeuf e Žumer (2017), Padron (2019) e Guarino, Guizzardi e Mylopoulos (2020). Conclui-se que há muitos espaços epistemológicos em que a Modelagem Conceitual, a Organização do Conhecimento e a Organização da Informação se aproximam, como uma ocupação comum em relação aos paradigmas epistemológicos de Capurro e a unidade elementar que ambos desenvolvem. Além disso, apesar das diferentes finalidades dos produtos dessas três áreas, eles podem embasar-se mutuamente em sua construção, como no caso dos modelos conceituais para a construção de sistemas de organização do conhecimento, implementação de sistemas de informação e elaboração de regras de catalogação.

Palavras-chave: Epistemologia. Ciência da Informação. Modelos Conceituais. IFLA LRM.

Abstract

Conceptual modelling, since the elaboration of the conceptual models of the bibliographic universe, started to function and was consolidated as a methodology and research object also in the area of Knowledge Organization and Information Organization. Thus, this research seeks to investigate epistemological approaches between Conceptual Modelling, Knowledge Organization and Information Organization. For this purpose, Capurro's topology (2003), Dahlberg's Theory of Concept and the IFLA LRM model are used, based on a literature review of an epistemological nature, with references to the works of Dahlberg (1978), Svenonius (2000), Bräscher and Café (2008), Riva, LeBoeuf and Žumer (2017), Padron (2019) and Guarino, Guizzardi and Mylopoulos (2020). It is concluded that there are many epistemological spaces in which Conceptual Modelling, Knowledge Organization and Information Organization come together, like the common occupation in relation to Capurro's epistemological paradigms and the elementary unity that both develop. In addition, despite the different purposes of the products in these three areas, they can support each other in their construction, as in the case of conceptual models for building knowledge organization systems, implementing information systems and elaborating cataloging rules.

Keywords: Epistemology. Information Science. Conceptual Models. IFLA LRM.

Resumen

La modelación conceptual, basada en la elaboración de modelos conceptuales del universo bibliográfico, pasó a actuar y se consolidó como metodología y objeto de investigación también en el área de Organización del Conocimiento y Organización de la Información. Así, esta investigación busca indagar aproximaciones epistemológicas entre la Modelación Conceptual, la Organización del Conocimiento y la Organización de la Información. Para ello se utiliza la topología de Capurro (2003), la Teoría del Concepto de Dahlberg y el modelo IFLA LRM, a partir de una revisión bibliográfica de carácter epistemológico, con referencias a los trabajos de Dahlberg (1978), Svenonius (2000), Bräscher y Café (2008), Riva, LeBoeuf y Žumer (2017), Padron (2019) y Guarino, Guizzardi y Mylopoulos (2020). Se concluye que son muchos los espacios epistemológicos en los que confluyen la Modelación Conceptual, la Organización del Conocimiento y la Organización de la Información, como ocupación común en relación con los paradigmas epistemológicos de Capurro y la unidad elemental que ambas desarrollan. Además, a pesar de los diferentes propósitos de los productos en estas tres áreas, pueden apoyarse mutuamente en su construcción, como en el caso de los modelos conceptuales para construir sistemas de conocimiento, implementar sistemas de información y elaborar reglas de catalogación.

Palabras clave: Epistemología. Ciencia de la información. Modelos Conceptuales. IFLA LRM.

1 Introdução

A Modelagem Conceitual é uma metodologia advinda da Ciência da Computação, cuja propostas e estudos ocorreram, em seus primórdios, em quatro áreas dessa disciplina: Inteligência Artificial (AI), Linguagem de programação, Banco de dados e Engenharia de Software (GUARINO; GUIZZARDI; MYLOPOULOS, 2020).

Essa técnica, a partir da sua aplicação na área de bancos de dados, especificamente a partir da linguagem de modelagem Entidade-Relacionamento (E-R) elaborada por Peter Chen, foi adotada pela área da Catalogação Descritiva, que se utilizou dela para “[...] elaborar um modelo conceitual para a reestruturação sintática e semântica do domínio bibliográfico” (SOUZA, 2017, p. 27).

Esse modelo, nomeado *Functional Requirements for Bibliographic Records (FRBR)*, foi “[...] a mudança mais radical para o pensamento catalográfico de bibliotecas desde que Panizzi desenvolveu suas 91 regras para o Catálogo do Museu Britânico” (COYLE, 2016, p. 65, tradução nossa). Posteriormente foram publicados o *Functional Requirements for Authority Data (FRAD)* e o *Functional Requirements for Subject Authority Data (FRSAD)*, buscando abordar diferentes perspectivas do universo bibliográfico, formando de tal maneira a Família FR.

Diante da necessidade de consolidação desses três modelos da Família, por causa de incongruências, surge o *IFLA Library Reference Model (IFLA LRM)*, que consiste um modelo conceitual de alto nível desenvolvido a partir de uma estrutura aperfeiçoada de modelagem entidade-relacionamento (EER). O modelo aborda os dados bibliográficos em um sentido amplo e geral, explicitando princípios gerais que governam a estrutura lógica da informação bibliográfica (RIVA; LEBOEUF; ŽUMER, 2017).

Por mais que a Organização do Conhecimento e a Organização da Informação sejam processos distintos, ambos se inter-relacionam no *IFLA LRM* por meio da Modelagem Conceitual. Segundo Padron (2019, p. 80): “[...] O LRM é, portanto, um modelo de representação do conhecimento para o domínio bibliográfico que visa contribuir para uma melhor representação da informação”. Dessa maneira, a modelagem conceitual passou a atuar e se consolidou como metodologia e objeto de pesquisa também na área da Organização do Conhecimento e da Organização da Informação.

Outrossim, uma vez que modelos conceituais manipulam representações mentais conceituais (que consistem em conceitos, e, em associações que relacionam conceitos) que “[...] agentes cognitivos constroem, usam e manipulam durante a cognição de um certo domínio” (GUARINO; GUIZZARDI; MYLOPOULOS, 2020, p. 2, tradução nossa), pode-se averiguar uma relação muito próxima com questões

epistemológicas da Organização do Conhecimento, de tal modo que se pode dizer que “O conceito é a unidade elementar desenvolvida tanto pela Organização do Conhecimento como pela modelagem conceitual” (PADRON, 2019, p. 81).

Ressalta-se, conforme a conceituação de modelo teórico de Sayão (2001), que fez-se obrigatório para a Ciência da Informação, devido a sua ampla natureza interdisciplinar e buscando “mapear toda a sua realidade”, tomar como propriamente seus paradigmas e modelos de outras áreas, o que justifica as aproximações epistemológicas entre Modelagem Conceitual, Organização do Conhecimento e Organização da Informação. Tal processo é exemplificado na síntese entre as linguagens de modelagem conceitual da Ciência da Computação (Entidade-Relacionamento e Orientado-a-Objeto) e as linguagens bibliográficas da Biblioteconomia, perpetuadas nos modelos conceituais do universo bibliográfico endossados pela *International Federation of Library Associations (IFLA)*: Família FR (*FRBR*, *FRAD*, *FRSAD*), *IFLA LRM* e *FRBRoo*.

De caráter bibliográfico, essa pesquisa procura investigar aproximações epistemológicas entre a Modelagem Conceitual, a Organização do Conhecimento e a Organização da Informação por meio da topologia de Capurro (2003), da Teoria do Conceito de Dahlberg e do modelo *IFLA LRM*, se embasando em uma revisão de literatura de caráter epistemológico, tendo como referências os trabalhos de Dahlberg (1978), Svenonius (2000), Bräscher e Café (2008), Guarino, Guizzardi e Mylopoulos (2020), Riva, LeBoeuf e Žumer (2017) e Padron (2019).

2 Topologia epistemológica da Modelagem Conceitual, da Organização do Conhecimento e da Organização da Informação.

Modelos Conceituais são construções estruturais de modelos mentais que embasam-se em Linguagens de Modelagem com a finalidade de “[...] descrever dado aspecto de um domínio de interesse para propósitos de entendimento e comunicação” (GUARINO; GUIZZARDI; MYLOPOULOS, 2020, p. 4, tradução nossa). O modelo *IFLA LRM* explicita essa finalidade ao declarar que as tarefas do usuário servem como “[...] declaração do escopo funcional do modelo e confirmam sua orientação externa para as necessidades do usuário final” (RIVA; LEBOEUF; ŽUMER, 2017, p. 15).

Outrossim, esses modelos buscam um entendimento compartilhado do universo do discurso, para que uma comunidade de usuários e sistema passem a ter uma linguagem comum (PADRON, 2019).

Esses modelos tem sempre uma semântica conceitual, uma vez que as construções linguísticas das quais se utilizam denotam conceitos mediados pela

conceitualização e percepção humanas, sendo um exemplo a proposta de modelo semântico de dados, o modelo E-R, de Peter Chen (GUARINO; GUIZZARDI; MYLOPOULOS, 2020).

Por conseguinte, a inserção e a contribuição dos modelos conceituais para a Organização do Conhecimento encontram-se principalmente na construção de Sistemas de Organização do Conhecimento.

[...] modelos conceituais podem apoiar a construção de representações do conhecimento, provendo a formalização conceitual necessária a um melhor entendimento do domínio do conhecimento e comunicação entre projetistas e usuários de modelos de representação do conhecimento [...]. Percebe-se que é na construção de ontologias que a modelagem conceitual pode contribuir mais com a OC, por necessitar de uma estrutura mais elaborada e complexa que os demais SOCs (PADRON, 2019, p. 81).

Sistemas de Organização do Conhecimento consistem na implementação da Representação do Conhecimento, que por sua vez, é o produto da Organização do Conhecimento (PADRON, 2019). Em outras palavras:

A representação do conhecimento é feita por meio de diferentes tipos de sistemas de organização do conhecimento (SOC) que são sistemas conceituais que representam determinado domínio por meio da sistematização dos conceitos e das relações semânticas que se estabelecem entre eles (BRÄSCHER; CAFÉ, 2008, p. 8).

Nesses sistemas, “[...] o principal objeto de trabalho são os conceitos” (BRÄSCHER; CARLAN, 2010, p. 155). Por conseguinte, se estabelece que “O conceito é a unidade elementar desenvolvida tanto pela Organização do Conhecimento como pela modelagem conceitual” (PADRON, 2019, p. 81)

Desse modo, segundo a topologia epistemológica dos três paradigmas da Ciência da Informação de Capurro (2003), tanto a Modelagem Conceitual como a Organização do Conhecimento encontram-se nos paradigmas Cognitivo e Social, uma vez que, inseridos em tais paradigmas apresentam-se teorias de suma importância para ambas áreas de pesquisa, o que demonstra suas possíveis aproximações epistemológicas: a noção de análise de domínio (Paradigma Social) e os estudos de comportamento informacional por parte dos usuários, a partir de mapas/modelos mentais (Paradigma Cognitivo).

Outrossim, as ontologias nas quais a modelagem conceitual pode contribuir significativamente com a Organização do Conhecimento, referem-se a um tipo de Sistema de Organização do Conhecimento, que consiste em um

[...] modelo de um domínio específico do conhecimento (ex., biologia molecular, finanças, logística, doenças infecciosas) expresso em uma linguagem de representação do conhecimento (ex., RDF, OWL, F-Logic) ou modelagem conceitual (ex., UML, EER, ORM) (GUIZZARDI et al., 2011, p. 1-2).

Ressalta-se, que o uso do termo Ontologia difere do seu uso na modelagem conceitual que é utilizado de maneira bem próxima ao uso do termo na Filosofia, referindo-se a um “[...] sistema formal e filosoficamente bem-fundamentado de categorias que pode ser usado para articular conceituações e modelos em domínios específicos do conhecimento” (GUIZZARDI et al., 2011, p. 1).

Referente à Organização da Informação, se estabelece, a partir de Svenonius (2000, p. 53), que a informação, para ser organizada, precisa ser descrita, e tal descrição consiste no enunciado de propriedades de um objeto ou das relações desse objeto com outros que o identificam.

Segundo Bräscher e Café (2008), a Organização da Informação é um processo que envolve: (i) a descrição física do objeto informacional (Informação-como-coisa) e (ii) conteúdo nele incorporado (Informação-como-conhecimento). Dessa maneira, o produto desse processo é a Representação da Informação que consiste no “[...] conjunto de elementos descritivos que representam os atributos de um objeto informacional específico” (BRÄSCHER; CAFÉ, 2008, p. 5).

Além disso, conforme Svenonius (2000), um sistema cuja finalidade é a organização da informação, caso pretenda ser eficiente, deve ter alicerces em uma base intelectual. Tal base consiste em: (i) conjuntos de propósito e princípios, (ii) formalização de processos (modelagem conceitual), (iii) conhecimento científico acumulado por meio de pesquisas acadêmicas, e (iv) se dirigir às problemáticas próprias e basilares da Organização da Informação.

Portanto, a partir do terceiro item que compõe o Alicerce Intelectual de um Sistema de Informação, mapeia-se propriamente o *locus* da Modelagem Conceitual na Organização da Informação. Os modelos conceituais indicam “o que” deve ser descrito explicitando sua lógica e princípios gerais, servindo como “[...] guia ou base para formular regras de catalogação e implementar sistemas bibliográficos” (RIVA; LE BOEUF; ŽUMER, 2017, p. 10).

3 Teoria do Conceito de Dahlberg e a Modelagem Conceitual

Segundo Dahlberg (1978), os conceitos são resultados da linguagem, que por sua vez manifesta a capacidade cognitiva dos seres humanos de representarem objetos que o circundam e de comunicarem-se entre si. Dessa

maneira, pode-se dizer que a linguagem tem como finalidade principal a representação e a comunicação.

Dahlberg (1978) faz a distinção entre Objeto Individual (Conceito individual) vs. Objeto Geral (Conceito geral). O primeiro consiste em uma unidade inconfundível, única e distinto dos demais, caracterizada pelas formas do tempo e espaço, por exemplo, a Universidade Federal de São Carlos (UFSCar), a Universidade de Brasília (UnB), Sócrates etc. O segundo constitui objetos que prescindem e estão situados fora do tempo e espaço, por exemplo, universidade, humano etc.

O processo de formação de conceitos consiste na reunião e na compilação de enunciados verdadeiros a respeito de determinado objeto, fixados por um símbolo linguístico. Inversamente, pode-se dizer que “[...] é formulando enunciados sobre os atributos necessários ou possíveis dos objetos que se obtém as características dos respectivos conceitos” (DAHLBERG, 1978, p. 102).

Outrossim, a soma total dos enunciados verdadeiros sobre um objeto fornece o conceito do mesmo (DAHLBERG, 1978), conforme Quadro 1:

Quadro 1. Duas espécies de conceitos e seus enunciados

Conceito Individual	Conceito Geral
<i>IBICT (Instituto Brasileiro de Informações em Ciência e Tecnologia)</i>	<i>Instituição</i>
É uma instituição	Constituída de um grupo de pessoas
Situada no Rio de Janeiro	Que trabalham com determinada finalidade
Relacionada com a coordenação dos sistemas de informação no Brasil	Possuindo administração comum
Possui cerca de 60 funcionários, etc.	Localizada em determinado lugar
-	Durante determinado tempo, etc.

Fonte: Adaptado de Dahlberg (1978).

Dessa maneira, Dahlberg (1978, p. 102) indica que “[...] os elementos contidos nos conceitos gerais encontram-se também nos conceitos individuais, sendo, portanto, possível reduzir os conceitos individuais aos gerais e ordená-los de acordo com os conceitos gerais” (DAHLBERG, p. 102). Portanto, os conceitos gerais estruturam os conceitos individuais.

Segundo Guarino, Guizzardi e Mylopoulos (2020), os Modelos Conceituais se embasam em uma Linguagem de Modelagem (como a E-R de Peter Chen), com a

finalidade de descrever aspectos de um domínio de interesse, com propósitos de entendimento e comunicação. As unidades ou conceitos dessa conceitualização representam os universais que são cognitivamente utilizados para organizar nossa visão de mundo.

Esses universais, na conceitualização de Dahlberg (1978) são os Conceitos Gerais, que prescindem das formas tempo e espaço. Reforçando essa abordagem, Guarino, Guizzardi e Mylopoulos (2020) dizem que essas unidades conceituais refletem as regularidades que nos são cognitivamente relevantes. Da mesma maneira, no *IFLA LRM* é indicado que suas entidades representam os principais objetos de interesse dos usuários dos registros bibliográficos (RIVA; LEBOEUF; ŽUMER, 2017).

Outrossim, o relatório final do modelo *FRBR*, ao declarar a metodologia utilizada para identificação de entidades, descreve processo de conceitualização análoga:

O primeiro passo na técnica de análise de entidade é isolar os principais objetos de interesse dos usuários de um particular domínio. Esses objetos de interesse ou entidades são definidos na categoria mais alta possível. Ou seja, a análise primeiramente foca atenção não nos dados individuais mas nas “coisas” que os dados descrevem. Cada entidade definida para o modelo, dessa maneira, serve como ponto focal para um agrupamento de dados. Um diagrama de entidades para um sistema de informação administrativa, por exemplo, provavelmente identificaria “funcionário” como uma entidade que seria de interesse para usuários de tal sistema (IFLA, 2009, p. 9, tradução nossa)

Sendo assim, pode-se dizer que: (i) a partir da abstração de conceitos individuais se obtém os os conceitos gerais; e também (ii) conceitos gerais formam conjuntos de conceitos individuais.

Segundo Dahlberg (1978, p. 102) “Cada enunciado verdadeiro representa um elemento do conceito”, e que os elementos “Traduzem os atributos das coisas designadas”. Isso é expresso nos modelos conceituais E-R a partir das entidades (unidades conceituais) e seus atributos (propriedades), “[...] os atributos são o que dão sentido à descrição de uma instância de uma entidade” (RIVA; LE BOEUF; ŽUMER, 2017, p. 11).

Do mesmo modo, as relações entre conceitos conforme elaboradas por Dahlberg (1978), encontram-se na linguagem de modelagem Entidade-Relacionamento Estendido (EER), utilizada pelo IFLA LRM.

As relações hierárquicas, ou de gênero e espécie, ocorrem quando entre dois conceitos com características idênticas, há características conceituais a mais em um deles (DAHLBERG, 1978). Essa estrutura é expressa através da relação “*isA*” no modelamento formal, e implica que atributos e relacionamentos declarados em um nível superior não necessitam ser repetidos em todas as subclasses, uma vez que são herdadas (ŽUMER, 2018, p. 312).

Uma entidade pode ser declarada uma superclasse de outras entidades que, por sua vez, possuem uma relação de subclasse com essa entidade. Qualquer instância de uma entidade de subclasse é também uma instância da superclasse. Isto forma parte da estrutura do modelo entidade-relacionamento expandido e pode ser expresso como “is a” (ou isA). Por exemplo, a entidade pessoa é uma subclasse da entidade agente, e isso pode ser expresso como: pessoa isA agente (RIVA; LE BOEUF; ŽUMER, 2017, p. 19).

A entidade Agente é uma superclasse das entidades Pessoa e Agente Coletivo. As entidades Pessoa e Agente Coletivo herdam todos os atributos de sua superclasses Agente, mas também possuem atributos mais específicos que as diferenciam entre si, por isso, pode-se declarar que ‘toda Pessoa é (isA) Agente’ mas, ‘nem todo Agente é (isA) Pessoa’, porque a entidade Agente ocupa lugar superior na hierarquia, e como subclasse específica-se não somente como Pessoa mas também como Agente Coletivo.

Outro exemplo, refere-se à entidade *Res* que é

[...] a entidade superior do modelo. *Res* inclui coisas e conceitos materiais ou físicos. Tudo o que é considerado relevante para o universo bibliográfico, que é o universo do discurso neste caso, está incluído. *Res* é uma superclasse de todas as outras entidades explicitamente definidas, bem como de quaisquer outras entidades não especificamente identificadas (RIVA; LEBOEUF; ŽUMER, 2017, p. 21).

Dessa maneira, todas as outras entidades são subclasses da entidade *Res*, ou seja, agrupadas por ela, e conseqüentemente herdam todos seus atributos e relacionamentos. Todas as outras entidades do modelo são entidades desmembradas, porém possuem a mesma relação com a entidade superior *Res*, o de ser subclasse (*isA*).

As relações partitivas que segundo Dahlberg (1978) existem entre um todo e suas partes, ou, “entre um produto e os elementos que o constituem” também encontra-se no *IFLA LRM*, principalmente no que diz respeito a modelagem de Agregados e Publicações Seriadas.

No *IFLA LRM* os Agregados são definidos como “[...] uma manifestação que materializa várias expressões” (RIVA; LE BOEUF; ŽUMER, 2017, p. 95). Com isso, uma Manifestação agregada contém, além de Expressões distintas, uma Expressão agregadora que, conseqüentemente, realiza uma Obra agregadora.

Segundo o modelo, as publicações seriadas são construções complexas que combinam relações de todo/parte e relações de agregação, cuja Manifestação, que a representa em sua totalidade, é composta por outras Manifestações agregadas publicadas em série, estruturação constituída por meio do relacionamento todo/parte *LRM-R26* no nível da Manifestação. Essas Manifestações agregadas são fascículos, por suas vezes

agregados de artigos, publicados ao longo do tempo (RIVA; LE BOEUF; ŽUMER, 2017).

O relacionamento todo/parte *LRM-R26* “possui parte”, “é parte de” é um relacionamento entre duas Manifestações em que uma é componente da outra (RIVA; LE BOEUF; ŽUMER, 2017).

Dahlberg (1978) também oferece noções úteis que podem ser instrumentalizadas para a análise do nível de abstração e especificidade de entidades assim como a estruturação de sua hierarquia, como a noção de intenção do conceito e extensão do conceito. Intenção expressa a soma total das características do conceito, enquanto a extensão do conceito a soma total dos conceitos mais específicos que possui. Ou seja, a partir dessa terminologia, pode-se classificar a entidade *Res* como um conceito de intenção mínima (possuindo apenas um atributo) e extensão máxima (sendo superclasse de todas as outras unidades conceituais, ou, entidades).

4 Conclusão

A modelagem conceitual, a partir da elaboração do modelo FRBR em 1998, se consolidou como metodologia e objeto de pesquisa na área da Organização do Conhecimento e da Organização da Informação. Partindo dessa constatação, essa pesquisa, a partir de uma revisão de literatura de caráter epistemológico, teve como objetivo investigar aproximações epistemológicas entre a Modelagem Conceitual, a Organização do Conhecimento e a Organização da Informação tendo como referência os trabalhos de Dahlberg (1978), Svenonius (2000), Capurro (2003), Bräscher e Café (2008), Guarino, Guizzardi e Mylopoulos (2020), Riva, LeBoeuf e Žumer (2017) e Padron (2019).

Constatou-se que, uma vez que modelos conceituais manipulam representações mentais conceituais (conceitos e suas relações entre si) resultados da conceitualização de um domínio, há uma relação muito próxima com questões epistemológicas da Modelagem Conceitual e da Organização do Conhecimento, de tal maneira que pode concluir-se que a unidade elementar de ambas as áreas é o ‘conceito’.

Outrossim, constatou-se que a Organização do Conhecimento (no nível das entidades) e a Organização da Informação (no nível dos atributos) se inter-relacionam no modelo *IFLA LRM* por meio da Modelagem Conceitual

Referente a inserção dos modelos conceituais para a Organização do Conhecimento, constatou-se que a principal contribuição desses modelos encontra-se na construção de Sistemas de Organização do Conhecimento, pois a modelagem conceitual, por meio da semântica de suas linguagem de modelagem,

possibilitam uma formalização e estruturação conceitual que proporciona melhor entendimento e comunicação de dado domínio do conhecimento entre os sistemas e os usuários finais.

Na Organização da Informação, por sua vez, a Modelagem Conceitual tem seu lugar na formalização dos processos, o que compõe um dos itens do alicerce da base intelectual de um sistema de informação conforme Svenonius (2000). Dessa maneira, os modelos conceituais indicam ‘o que’ deve ser descrito’, explicitando a lógica e princípios gerais que governam a informação a ser representada.

Ademais, constatou-se que há diversos pontos de encontro entre o ‘conceito’ na Modelagem Conceitual e em Dahlberg (1978). As unidades ou conceitos de determinada conceitualização de domínio, que representam os universais que cognitivamente estruturam nossa visão de mundo e são a unidade elementar dos modelos conceituais são, na conceitualização de Dahlberg (1978), os Conceitos Gerais, que prescindem das formas tempo e espaço, e estruturam os Conceitos Individuais.

Também as relações entre conceitos conforme elaboradas por Dahlberg (1978), encontram-se na linguagem de modelagem Entidade-Relacionamento Estendido (EER), utilizada pelo IFLA LRM, destacando-se as relações hierárquicas, que ocorrem quando entre dois conceitos com características idênticas, há características conceituais a mais em um deles, e as relações partitivas, que existem entre um todo e suas partes.

Finalmente, constatou-se que Dahlberg (1978) oferece noções úteis que podem ser instrumentalizadas para a análise do nível de abstração e especificidade de entidades assim como a estruturação de sua hierarquia, como a noção de intenção do conceito e extensão do conceito.

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro para o desenvolvimento da pesquisa.

Referências

- BRÄSCHER, M.; CARLAN, E. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO J.; BRÄSCHER, M. O. (org.). **Passeios pelos bosques da informação: estudos sobre representação e organização da informação e do conhecimento - EROIC**. Brasília: IBICT, 2010. p. 147–176. *E-book*. Disponível em: <http://www2.senado.leg.br/bdsf/handle/id/189812>.
- BRÄSCHER, M.; CAFÉ, L. Organização da informação ou organização do conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, ENANCIB, 9. Anais... 2008. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/176535>.
- CAPURRO, R. Epistemologia e Ciência da Informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 2003, Belo Horizonte. Anais... Belo Horizonte: ENANCIB, 2003. p. 1-21. Disponível em: http://www.capurro.de/enancib_p.htm.
- COYLE, Karen. **FRBR, before and after: a look at our bibliographic models**. Chicago: American Library Association, 2016.
- DAHLBERG, I. Teoria do conceito. *Ciência da Informação*, v. 7, n. 2, 1978. Disponível em: <http://revista.ibict.br/ciinf/article/view/115>.
- GUARINO, N.; GUIZZARDI, G.; MYLOPOULOS, J. On the Philosophical Foundations of Conceptual Models. In: DAHANAYAKE, A.; HUISKONEN, J.; KIYOKI, Y.; THALHEIM, B.; JAAKKOLA, H.; YOSHIDA, N. (ed.) **Information Modelling and Knowledge Bases XXXI**. [s. l.]: IOS Press Ebooks, 2020. p.1-15. *E-book*.
- GUIZZARDI, G.; ALMEIDA, J. P.; GUIZZARDI, R. S.; BARCELLOS, M. P.; FALBO, R. Ontologias de fundamentação, modelagem conceitual e interoperabilidade semântica. In: GUIZZARDI, G.; BAIÃO, F. A.; OLIVEIRA, J. P. M. de (ed.). **Proceedings of the Iberoamerican Meeting of Ontological Research**. CEUR-WS.org, 2011. Disponível em: <http://ceur-ws.org/Vol-728/paper6.pdf>.
- IFLA. Functional Requirements For Bibliographic Records: Final Report. 2009, p. 137. Disponível em: https://cdn.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr/frbr_2008.pdf.
- PADRON, M. F. **Uma proposta de modelo conceitual para representação da música popular brasileira**. 2019. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2019.
- RIVA, P.; LE BOUEF, P.; ŽUMER, M. **IFLA Library Reference Model: um modelo conceitual para a Informação Bibliográfica**. [s. l.], p. 106, 2017. Disponível em: https://repository.ifla.org/bitstream/123456789/47/1/ifla-irm-august-2017_rev201712-por.pdf.

SAYÃO, L. F. Modelos teóricos em ciência da informação: abstração e método científico. *Ciência da Informação*, Brasília, v. 30, n. 1, p. 82–91, jan/abr 2001. Disponível em: <http://revista.ibict.br/ciinf/article/view/941>.

SOUZA, E. G. de. As Teorias Documentárias e a Organização da Informação: a Centralidade das Categorias Obra e Usuário. *In*: PINHO, F. A.; GUIMARÃES, J. A. C. (org.) **Memória, tecnologia e cultura na organização do conhecimento**. Recife: Ed. UFPE, 2017. p. 26-33. *E-book*. Disponível em: <https://isko.org.br/wp-content/uploads/2021/05/Proceedings-ISKO-Brasil-2017.pdf>.

SVENONIUS, Elaine. **The Intellectual Foundation of Information Organization**. London: MIT Press, 2000.

ŽUMER, M. IFLA Library Reference Model (IFLA LRM) - Harmonisation of the FRBR family. **Knowledge Organization**, Vol. 45, no. 4, pp.310-318, 2018. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/ko_45_2018_4_e.p

Ordo librorum hvivs operis: estrutura e fundamentos da classificação bibliográfica gesneriana

Ordo librorum hvivs operis: structure and fundamentals of gesnerian bibliographic classification

Andre Vieira de Freitas Araujo

(1) Universidade Federal do Paraná (UFPR), Av. Prefeito Lothário Meissner, 632, CEP: 80210-070 - Jardim Botânico - Curitiba (PR) - Brasil, armarius.araujo@gmail.com

Resumo

Objetivo: O objetivo do trabalho é identificar a estrutura e os fundamentos da classificação bibliográfica de Conrad Gesner (1516-1565), no intuito de situá-la nas dinâmicas da organização do conhecimento do Séc. XVI. **Metodologia:** Em perspectiva histórico-bibliográfica, o trabalho se desenvolve através dos seguintes passos: 1) leitura e tradução das passagens das *Pandectae* (1548) em que Gesner apresenta sua proposta de esquema classificatório; 2) mapeamento e demonstração de exemplos que ilustram a aplicação deste esquema; 3) breve análise e discussão do referido esquema. **Resultados:** O fólio [1v] das *Pandectae* expõe o esquema completo das classes em *ordo librorum hvivs operis* (ordenação dos livros desta obra), resultando em um verdadeiro repertório de ordem sistemática. Esse esquema busca abranger tanto as ordens naturais quanto as artificiais, tanto das coisas quanto das ciências. A estrutura documentária e a arquitetura semântica das *Pandectae* são organizadas da seguinte forma dentro de cada um dos XXI (XX) Livros/Partições: 1) cada Livro, que corresponde a uma Partição, representa uma Classe dentro das artes e ciências; 2) os Livros são subdivididos em Títulos; 3) os Títulos são divididos em Partes; 4) as Partes podem ser eventualmente subdivididas em Segmentos e Parágrafos; 5) as Partes, Segmentos e Parágrafos contêm os *Loci*. **Conclusão:** Ao contemplar os detalhes bibliográficos que passaram a compor os livros no Séc. XVI, graças à tecnologia dos tipos móveis, as *Pandectae* emergem como o primeiro esquema classificatório, desprovido de notação explícita, concebido especificamente para o universo *librario*.

Palavras-chave: Bibliografia. Classificação - História e Teoria. Classificação bibliográfica gesneriana. Organização do Conhecimento - Séc. XVI. *Pandectae* (1548).

Abstract

Objective: The aim of this work is to identify the structure and fundamentals of Conrad Gesner's (1516-1565) bibliographic classification in order to contextualize it within the dynamics of knowledge organization in the 16th century. **Methodology:** In a historical-bibliographical perspective, the work unfolds through the following steps: 1) reading and translation of passages from *Pandectae* (1548) in which Gesner presents his proposed classification scheme; 2) mapping and demonstration of examples that illustrate the application of this scheme; 3) brief analysis and discussion of the aforementioned scheme. **Results:** Folio [1v] of the *Pandectae* presents the complete scheme of classes in *ordo librorum hvivs operis* (arrangement of the books in this work), resulting in a true repertoire of systematic order. This scheme aims to encompass both natural and artificial orders, both in terms of things and sciences. The document structure and semantic architecture of the *Pandectae* are organized as follows within each of the XXI (XX) Books/Partitions: 1) each Book, corresponding to a Partition, represents a Class within the arts and sciences; 2) Books are subdivided into Titles; 3) Titles are divided into Parts; 4) Parts can be further subdivided into Segments and Paragraphs if necessary; 5) Parts, Segments, and Paragraphs contain the *Loci*. **Conclusion:** When considering the bibliographic details that came to constitute the books in the 16th century, due to the technology of movable types, the *Pandectae* emerge as the first classification scheme, devoid of explicit notation, specifically designed for the *librario* universe..

Keywords: Bibliography. Classification - History and Theory. Gesnerian bibliographic classification. Organization of Knowledge - 16th Century. *Pandectae* (1548).

Resumen

Objetivo: El objetivo del trabajo es identificar la estructura y los fundamentos de la clasificación bibliográfica de Conrad Gesner (1516-1565), con el fin de situarla en las dinámicas de la organización del conocimiento del siglo XVI. **Metodología:** En una perspectiva histórico-bibliográfica, el trabajo se desarrolla a través de los siguientes pasos: 1) lectura y traducción de los pasajes de las *Pandectae* (1548) en los que Gesner presenta su propuesta de esquema clasificatorio; 2) mapeo y demostración de ejemplos que ilustran la aplicación de este esquema; 3) breve análisis y discusión de dicho esquema. **Resultados:** El folio [1v] de las *Pandectae* expone el esquema completo de las clases en *ordo librorum hvivs operis* (ordenación de los libros de esta obra), lo que da como resultado un verdadero repertorio de orden sistemático. Es un esquema que busca contemplar la totalidad de los órdenes, natural y artificial, tanto de las cosas como de las ciencias. La estructura documental y la arquitectura semántica de las *Pandectae* se presentan de la siguiente manera dentro de cada uno de los XXI (XX) Libros/Particiones: 1) cada Libro, que corresponde a una Partición, representa una Clase dentro de las artes y las ciencias; 2) Los Libros se dividen en Títulos; 3) los Títulos se dividen en Partes; 4) las Partes pueden ser eventualmente subdivididas en Segmentos y Párrafos; 5) las Partes, Segmentos y Párrafos contienen los *Loci*. **Conclusión:** Al

considerar los detalles bibliográficos que llegaron a constituir los libros en el siglo XVI, gracias a la tecnología de los tipos móviles, las *Pandectae* emergen como el primer esquema de clasificación, desprovisto de notación explícita, concebido específicamente para el universo *librario*.

Palabras clave: Bibliografía. Clasificación - Historia y Teoría. Clasificación bibliográfica gesneriana. Organización del Conocimiento - Séc. XVI. *Pandectae* (1548).

"In a broken reality everything's shattering". (Zanias, "Simulation")

1 Introdução

No âmbito dos estudos bibliográficos no Brasil [1], Conrad Gesner (1516-1565) é mencionado como figura fundamental à constituição disciplinar da Bibliografia, o que o coloca diretamente como referência nas perspectivas histórico-epistemológicas da Organização do Conhecimento (OC).

Trabalhos anteriores, já apresentados e debatidos no contexto da ISKO Brasil e da ISKO Internacional, abordaram Gesner e a OC, explorando a dimensão histórico-epistemológica da OC e a taxonomia gesneriana (ARAUJO; CRIPPA, 2015), os aspectos semânticos das *Pandectae* (1548) (ARAUJO; CRIPPA; SABBA; 2016), a universalidade e utopia em Gesner e Paul Otlet (ARAUJO; JUVÊNCIO; CRIPPA, 2018) e a tensão bibliográfica no Séc. XVI entre a universalidade gesneriana e o anti-heretismo posseviniano (CRIPPA; ARAUJO, 2020). Aqui, vamos aprofundar aspectos da classificação bibliográfica gesneriana.

Quando se trata da classificação das espécies e dos seres, Gesner fez importantes contribuições nos campos da Botânica, com seu trabalho *Catalogus Plantarum* (GESNER, 1542) - um catálogo de plantas, como o próprio nome sugere, em quatro idiomas (latim, grego, alemão e francês) - e no campo da Zoologia, com sua obra *Historia Animalium* (GESNER, 1558), considerada o marco inicial da Zoologia Moderna.

O exercício taxonômico, comum entre eruditos do Renascimento da Europa Moderna, contribuiu diretamente para que Gesner convertesse suas reflexões sobre a classificação das espécies e dos seres à classificação dos saberes e do conhecimento em si.

Nesta perspectiva, o objetivo do trabalho é identificar a estrutura e os fundamentos da classificação bibliográfica de Conrad Gesner, no intuito de situá-la nas dinâmicas da organização do conhecimento do Séc. XVI.

Em perspectiva histórico-bibliográfica, o trabalho se desenvolve através dos seguintes passos: 1) leitura e tradução das passagens das *Pandectae* (1548) em que Gesner apresenta sua proposta de esquema classificatório; 2) mapeamento e demonstração de exemplos que ilustram a aplicação deste esquema; 3) breve análise e discussão do referido esquema.

2 Desenvolvimento

2.1 e As *Pandectae* (1548) de Conrad Gesner

Pandectarum sive partitionum universalium Conradi Gesneri Tigurini, medici & philosophiae professoris, libri XXI (*Pandectas* ou *Partições Universais* de Conrad Gesner, médico e professor de filosofia em Zurique, em 21 livros), também conhecida como *Pandectae* (1548), representa a segunda parte da obra bibliográfica seminal *Bibliotheca Universalis* (1545), de Conrad Gesner.

Enquanto a *Bibliotheca Universalis* (1545) tem uma natureza descritiva-repertorial, baseada na seleção bibliográfica e na ordenação alfabético-nominal, as *Pandectae* (1548) têm uma natureza taxonômica, indicial e semântica, fundamentadas na organização multinível do conhecimento em hierarquias universais e particulares.

Sob esta perspectiva, as fontes listadas em *Bibliotheca Universalis* (1545) estão presentes nas *Pandectae* (1548, 1549), mas, de acordo com os seus argumentos. Isto significa que as fontes estão agrupadas segundo as suas classes.

2.2 As XXI Classes das *Pandectae*

O verso da página de rosto das *Pandectae* encontra-se exposto o esquema completo das classes em *ordo librorum hvivs operis* (ordenação dos livros desta obra), resultando em um verdadeiro repertório de ordem sistemática.

O material documentário das *Pandectae* está organizado em XXI Classes, que correspondem aos Livros/Partições da obra (Quadro 1):

Quadro 1. *Classes das Pandectae (1548) e seus respectivos fólhos*

<i>Livro</i>	<i>Classe/Partição</i>	<i>Fólio</i>
Livro I	Sobre Gramática e a Filologia	1
Livro II	Sobre Dialética	43
Livro III	Sobre Retórica	49
Livro IIII	Sobre Poética	59
Livro V	Sobre Aritmética	73

Livro	Classe/Partição	Fólio
Livro VI	Sobre Geometria, Ótica e Catóptrica	77
Livro VII	Sobre Música	81
Livro VIII	Sobre Astronomia	87
Livro IX	Sobre Astrologia	95
Livro X	Sobre Adivinhação, tanto com magia lícita quanto ilícita	99
Livro XI	Sobre Geografia	107
Livro XII	Sobre Histórias	117
Livro XIII	Sobre diversas artes não eruditas, mecânicas e todas as outras úteis à vida humana	165
Livro XIV	Sobre Filosofia Natural	181
Livro XV	Sobre a Filosofia Primeira ou Metafísica e sobre a Teologia dos pagãos	237
Livro XVI	Sobre Filosofia Moral	261
Livro XVII	Sobre Filosofia Econômica	303
Livro XVIII	Sobre Política, i.e. [assuntos] civis e militares	311
Livro XIX	Sobre Jurisprudência, três índices	319
Livro XX	Sobre Medicina	---
Livro XXI	Sobre Teologia Cristã	--

Fonte: Elaborado a partir de Gesner (1548, f. [1v]).

O Livro XX, destinado a apresentar o catálogo semântico das obras sobre Medicina, não foi publicado. No entanto, o Livro XXI, que trata de Teologia, foi publicado separadamente em 1549.

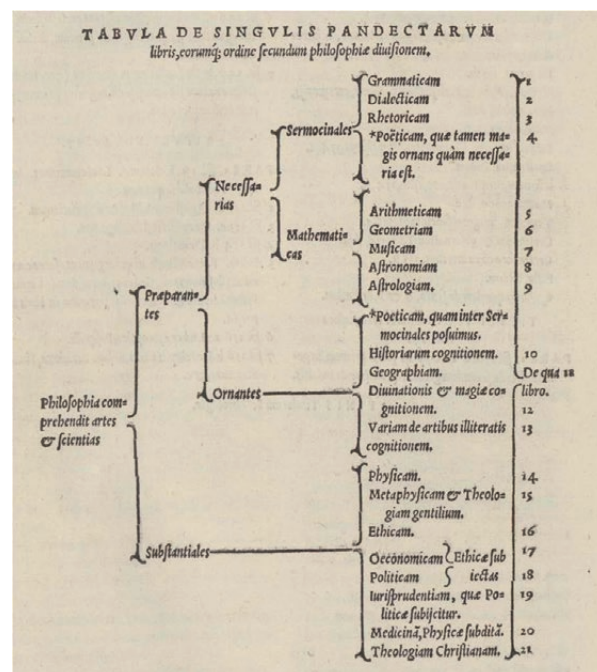
A classificação nas *Pandectae* deriva das sete artes liberais (*Trivium* - Gramática, Retórica, Dialética e *Quadrivium* - Geometria, Aritmética, Astronomia, Música) [2] e se estende até as categorias de assuntos complementares e de interesse aos estudiosos e eruditos do Renascimento.

Do ponto de vista da natureza e das escolhas relacionadas à taxonomia adotada por Gesner, é importante destacar que se trata de um esquema que busca abranger tanto as ordens naturais quanto as artificiais, tanto das coisas quanto das ciências.

2.3 Bases filosóficas e estrutura documentária do esquema classificatório

A base do esquema classificatório proposto por Gesner para distribuição dos Livros/Partições e, conseqüentemente, dos *loci* pode ser extraído da *TABVLA DE SINGVLIS PANDECTARVM libris, eorumq; ordine secundum philosophiae diuisionem* (Tabela dos livros individuais das Pandectas, e sua ordem de acordo com a divisão filosófica) (Fig. 1) que antecipa o texto das *Partitiones Theologicae* (1549).

Figura 1. *Pandectae* (1549), *Tabula de Singulis Pandectarum*



Fonte: Gesner (1549, f. [8v]).

Este esquema não só fundamenta tematicamente as classes adotadas nas *Pandectae*, mas sintetiza a sequência e a ordenação dos XXI Livros/Partições.

Com efeito, as *Pandectae* estão estruturadas de acordo com esse esquema de divisão da Filosofia, que engloba todas as artes e ciências. Em sua divisão, as disciplinas podem ser *Preparantes* (isto é, ciências preparatórias) ou *Substantiales* (isto é, ciências substanciais); as primeiras se dividem em *Necessariae* (ciências necessárias) e *Ornantes* (ciências ornamentais). As *Necessariae* por sua vez em *Sermocinales* e *Mathematicae*. As *Sermocinales* compreendem: 1) *Grammatica*, 2) *Dialectica*, 3) *Rhetorica*, 4) *Poetica*. As *Mathematicae* compreendem: 5) *Arithmetica*, 6) *Geometria*, 7) *Musica*, 8) *Astronomia*, 9) *Astrologia*. As

Ornantes: 10) *Historiarum cognitio*, 11) *Geographia*, 12) *Divinationis et magiae cognitio*, 13) *Varia cognitio de artibus illiteratis, Mechanicis, et aliss humanae vitae utilibus*. As *Substantiales* são: 14) *Physica*, 15) *Methaphysica et Theologia gentilium*, 16) *Ethica*, 17) *Oeconomica*, 18) *Politica*, 19) *Jurisprudentia*, 20) *Medicina*, 21) *Theologia Christiana* (SERRAI, 1977).

O esquema apresentado por Serrai (1977) tem semelhanças com o esquema do *Trivium* versus *Quadrivium*. Nesse sentido, o *Trivium* corresponde aproximadamente aos *Sermocinales*, enquanto o *Quadrivium* corresponde aproximadamente às quatro primeiras divisões das *Mathematicae*.

As ciências *Ornantes* parecem englobar principalmente disciplinas como História, Geografia, Magia e Artes Mecânicas - campos do conhecimento que não eram abordados nas classificações clássicas das ciências.

As ciências *Substantiales*, por outro lado, são aproximadamente referentes a:

1. A ordem das ciências no corpus de escritos aristotélicos - *Physica*, *Metaphysica*, *Ethica*, *Oeconomica* e *Politica*.
2. As três ciências correspondentes às Faculdades nas universidades medievais - *Jurisprudentia*, *Medicina* e *Theologia Christiana*.

Portanto, é evidente que, ao considerar as várias divisões das ciências, conforme estabelecido por Serrai (1977), a classificação de Gesner combina elementos relativamente tradicionais (por exemplo, nas ciências *Necessariae* sob as *Preparantes*) e elementos relativamente novos (por exemplo, nas ciências *Ornantes* sob as *Preparantes*).

Ademais, isso tem um efeito na ordem: as ciências preparatórias precedem as ciências substanciais, sendo que as ciências substanciais se enquadram no domínio da filosofia aristotélica (Física, Metafísica, Ética, Economia, Política), precedendo as três grandes ciências "profissionais" autônomas de Direito, Medicina e Teologia Cristã, sendo a Teologia considerada o ponto culminante das ciências.

Resumidamente, e de acordo com este esquema, na Filosofia,

[...] que embasa todas as artes e as ciências, se enxertam disciplinas 'preparatórias' e disciplinas 'substanciais'. As preparatórias se especificam em 'necessárias', como as matemáticas - aritmética, geometria, música, astronomia e astrologia - e em 'sermocinais' ou seja, para exposição - gramática, retórica, poesia, e dialética -, ou 'ornamentais' no sentido de úteis para aperfeiçoar todos os outros estudos, com geografia, história, magia e outras artes. As 'substanciais', por outro lado, se constituem em física, metafísica, teologia cristã e ética, na qual se encontram economia e política que, por sua vez, compreende o direito. (SABBA, 2016, p. 86).

As Partições formuladas por Gesner representariam os domínios disciplinares fundamentais na composição do conhecimento humano, baseados em doutrinas, noções e experiências.

Nesta formulação geral das disciplinas, portanto, encontramos uma ideia que, tomada diretamente de Aristóteles, forma a base para constituir todo o conhecimento do homem. A Filosofia representa o complexo das ciências e se expressa em todas aquelas formas e articulações que dão vida para as repartições singulares do saber e à toda gama de habilidades (SERRAI, 1990).

Sob esta ótica, a estrutura documentária e a arquitetura semântica das *Pandectae* são assim dispostas meticulosamente em cada um dos XXI (XX) Livros/Partições:

1. cada **Livro**, que corresponde a uma **Partição**, representa uma **Classe** dentro das artes e ciências;
2. os **Livros** são subdivididos em **Títulos**;
3. os **Títulos** se dividem em **Partes**;
4. as **Partes** podem ser eventualmente subdivididas em **Segmentos e Parágrafos**;
5. as **Partes, Segmentos e Parágrafos** contêm os **Loci**.

Para Gesner, nada é inserido em sua parte por acaso, sendo que o lugar ocupado por cada elemento de informação decorre não apenas de formas arbitrárias ou cruzadas, mas a partir da razão teórica.

Através de sua síntese, Serrai fornece uma representação mais concreta da arquitetura semântica das *Pandectae*, contribuindo para uma melhor compreensão da obra:

As *Pandectae* oferecem um quadro classificado - dividido em áreas cada vez mais estreitas e especificadas, desde os Livros aos Títulos até as Partes e aos Segmentos - nos quais nas últimas articulações são colocadas, de forma ordenada, os *loci*, ou seja, os nichos que hospedam os assuntos, isto é, os argumentos das obras semanticamente indexadas. Os assuntos que se referem às obras completas e às suas partes, que são exibidos através da estrutura enciclopédica das *Pandectae* - que se desenvolve, em 2 colunas, para mil páginas de grande formato - são aproximadamente 40.000, enquanto os cabeçalhos dos Títulos e das Partes, nos quais são articulados os 20 Livros, são pouco mais de 300. No entanto, muito maior é o número de Segmentos e de Parágrafos [...]. (SERRAI, 1990, p. 140).

No prefácio das *Pandectae*, Gesner argumenta que cada um dos Livros/Partições da obra foi dedicado aos que seriam considerados os seus patronos.

No início de cada Livro/Partição, são encontradas uma série de dedicatórias singulares, que eventualmente coincidem com os proprietários de centros editoriais e oficinas de impressão da primeira metade do Séc. XVI.

De acordo com Serrai (1990, p. 115), “[...] a ordem segundo a qual se seguem os pares Tipógrafo-Partições não é sempre casual, mas reflete, de alguma forma, um elo ideal entre a orientação da produção de um determinado tipógrafo e a especialidade semântica da classe que lhe é destinada”.

Para demonstrar exemplos da aplicação do esquema classificatório de Gesner, tomemos o *Liber I De Grammatica et Philologia* (Livro I Sobre a Gramática e Filologia), que é dedicado a Christoph Froschauer (1490-1564).

2.4 Exemplos da classificação gesneriana

Vejam os **primeiro exemplo** da classificação gesneriana, começando pelo sequenciamento do **Livro/Partição** até o **Título**.

No reto do fôlio 1 do *Liber I De Grammatica et Philologia* (Fig. 2), podemos observar as seguintes informações em sequência:

- 1) o título do Livro/Partição no topo da página: *Liber primus De Grammatica* (Livro primeiro de Gramática);
- 2) exposição geral dos Títulos que constituem o Livro/Partição I:

TITVLVS I. De philosophia, sapientia, doctrina, et pertinentibus ad eas, in genere (Título I. Sobre filosofia, sabedoria, doutrina e coisas pertencentes a elas, em geral);

II. TITVTVS II. De Grammatica in genere, de literis (Graecis et Latinis) syllabis, scriptura, etc. (Título II. Sobre a Gramática em geral, sobre as letras (gregas e latinas), sílabas, escrita, etc.);

III. Rudimenta Grammaticae (Título III. Rudimentos da Gramática);

III Grammaticae perfectiores, et Grammaticae simpliciter (Título IV. Gramáticas mais avançadas e Gramáticas simplesmente);

V. De orthographia (Título V. Sobre ortografia);

VI. De syntaxi (Título VI. Sobre sintaxe);

VII. De prosodia (Título VII. Sobre prosódia);

VIII. De etymologia: differentia vocabularum et synonymia (Título VIII. Sobre etimologia: diferença entre palavras e sinonímia);

IX. Dictionaria Latina (Título IX. Dicionários Latinos);

X. Vocabula quaedam insigniora, ordine literarum (Título X. Algumas palavras notáveis, em ordem alfabética);

XI. Proverbia (Título XI. Provérbios);

XII. Nomenclaturae rerum variarum, per ordines quosdam (Título XII. Nomenclatura de várias coisas, por certas categorias);

XIII. De varijs (Título XIII. Sobre assuntos diversos);

XIII. De puritate et elegantia linguae Latinae (Título XIII. Sobre a pureza e elegância da língua latina);

XV. De imitatione et stylo (Título XV. Sobre a imitação e estilo);

XVI. Contraria scripta Grammaticorum (Título XVI. Escritos contraditórios dos gramáticos);

XVII. Dialogi Latini (Título XVII. Diálogos Latinos);

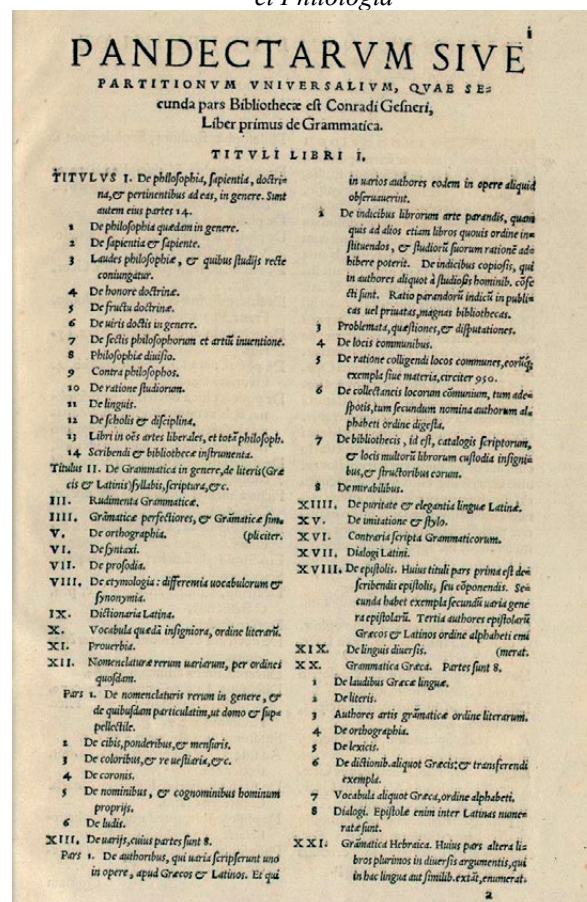
XVIII. De epistolis (Título XVIII. Sobre cartas/epístolas);

XIX. De linguis diversis (Título XIX. Sobre línguas diversas);

XX. Grammatica Graeca (Título XX. Gramática Grega);

XXI. Grammatica Hebraica (Título XXI. Gramática Hebraica). (GESNER, 1548, f. 1r).

Figura 2. *Pandectae* (1548), *Liber I De Grammatica et Philologia*

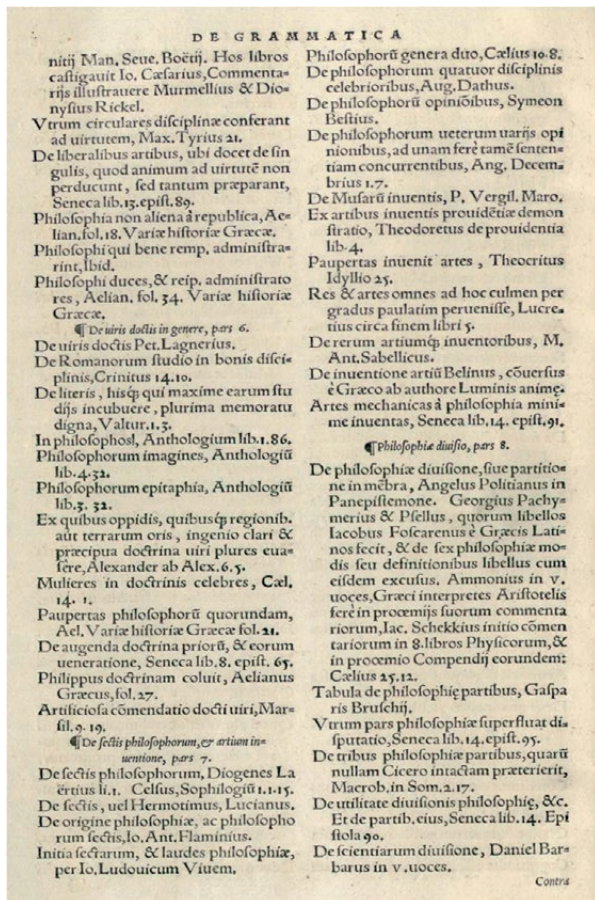


Fonte: Gesner (1548, f. 1r).

Vejam os um **segundo exemplo** do sistema de classificação de Gesner, a partir do sequenciamento do **Livro/Partição** para o **Título**, do **Título** para as **Partes** e das **Partes** para os **Loci**.

O Livro/Partição I *De Grammatica* é constituído por XXI Títulos, como vimos. O Título I *De philosophia, sapientia, doctrina, et pertinentibus ad eas, in genere* (Sobre filosofia, sabedoria, doutrina e coisas pertencentes a elas, em geral) é subdividido em 14 Partes. Dentre as Partes, temos a Parte 8 *Philosophiae diviso* (A divisão da Filosofia). Esta Parte, por sua vez, é constituída por 6 *loci* (Fig. 3):

Figura 3. *Pandectae* (1548), *Philosophiae diviso, pars 8*



Fonte: Gesner (1548, f. [3v (2 3)]).

Os 6 *Loci* da *Pars 8, Philosophiae diviso*, são:

- 1 Acerca da divisão da filosofia, ou sobre repartição dela em partes: Ângelo Poliziano em [sua obra] Panepistemon. Georg Pachymerius e Pselo, cujos livros foram traduzidos do grego para o latim por Giacomo Foscarini, e o livro, composto por eles, acerca dos seis modos ou definições da filosofia. Amônio em [sua obra] 'V voces', os intérpretes gregos de 2 Aristoteles em geral nos próemios de seus comentários, Jacob Schekkius no início de seus comentários dos oito livros da Física e no próêmio do Compêndio dos mesmos: Célio 25,12.
- 2 Tabela de partes da filosofia de Kaspar Brusch.

- 3 Seria imenso o debate acerca da existência de uma segunda partição da filosofia: Sêneca, livro 14, epístola 95.
- 4 Acerca das três partições da filosofia, as quais não teriam sido omitidas por Cícero: Macróbio, Comentário sobre o sonho de Cipião, 2, 17.
- 5 Acerca da utilidade da divisão da filosofia, etc. Também acerca das partições da mesma: Sêneca, livro 14, epístola 90.
- 6 Acerca das divisões das ciências: Daniel Barbarus em 'V voces' (GESNER, 1548, f. [3v (2 3)]).

No exemplo da *Philosophiae diviso*, as Partes contêm os *loci*, seguidos por uma breve referência aos autores e às obras que abordam os temas inseridos nesses *loci*.

Serrai e Sabba confirmam justamente este sistema presente na *Philosophiae diviso* e que se estende às *Pandectae* como um todo:

As Partições, ou classes principais são articuladas em Títulos; os Títulos estão subdivididos em Partes; e estas em Segmentos. As Partes contêm os *Loci*, seguidos de uma brevíssima referência aos autores e às obras que trataram, há pouco, dos temas que se identificam ou se colocam naqueles *Loci*. As referências são acompanhadas com indicações concernentes ao livro e ao capítulo da obra, mas que aparece muitas vezes implícita e mencionada apenas sob o nome do autor. O *Loci* são apresentados, geralmente, em sequência temática, ou seja cobrem o arco dos argumentos que competem a uma certa parte; muitas vezes, no entanto - a exemplo, nas enumerações envolvendo filósofos, animais, plantas, deuses, oráculos etc - os *Loci* aparecem em ordem alfabética. (SERRAI; SABBA, 2005, p. 56-57).

Nas diversas partes contidas nas *Pandectae*, que representam *loci*, ou assuntos, de extensão semântica bastante ampla

[...] são reportadas as citações daquelas obras, e de suas seções (muitas vezes até um único capítulo, mas às vezes até uma única página), no qual é tratado o argumento que compete àquela parte. As citações individuais, geralmente expressas através das próprias palavras do texto citado, constituem os assuntos particulares ou, portanto, seus descritores; por trás do anúncio de alguns desses assuntos, pode haver referências a vários trabalhos diferentes. (SERRAI, 1990, p. 142).

É importante destacar que estes núcleos são amplamente elencados nas *Pandectae* no formato de sentenças, e não necessariamente como termos, a exemplo de palavras-chave, como seria feito em uma denominação contemporânea.

Por outro lado, no princípio das *Pandectae*, fica evidente a proposta de refinar essas sentenças até chegar aos *loci*, que não estão mais no formato de sentenças, mas sim no formato de termos. Esses termos, inclusive, serão utilizados para configurar os índices gesnerianos.

De fato, temos aqui uma relação explicitamente estreita entre o esquema de classificação gesneriano e a indicialização gesneriana [3].

Portanto, o aspecto semântico das *Pandectae* é moldado não apenas pelo sistema de classificação adotado, mas também pela abordagem de Gesner em relação aos *loci* e índices.

Conforme observado por Sabba (2016), a classificação empregada por Gesner em suas *Pandectae* (1548) era baseada em conceitos, o que possibilitava que um mesmo objeto pudesse ser classificado de maneiras distintas, dependendo da disciplina em que estivesse inserido.

Trata-se de um ponto crucial para as classificações: enquanto nos 21 livros gesnerianos os conceitos ocupavam o lugar que lhes cabia na arquitetura sistemática, a elaboração moderna de assuntos às vezes resulta ainda inadequada, ainda excessivamente ancorada à impositação de natureza terminológica e linguística mais do que à semântica e, para sanear esses pontos críticos, foram individualizadas nos tesouros e listas de assuntos catalográficos “facetas” para qualificar e definir as relações semânticas entre os termos (mas, de fato, não entre os conceitos!). Em tempos de trocas internacionais contínuas e de abertura imediata e global dos produtos científicos é inaceitável a submissão, ainda, a esses vínculos linguísticos, mas é necessário admitir que Gesner teve a sorte de poder recorrer ao latim que, representando uma espécie de ‘língua franca’ facilitava as operações, tornando-as comuns. (SABBA, 2016, p. 87).

A complexidade do sistema de classificação de Gesner impõe problemas de três naturezas às *Pandectae* enquanto obra:

1. Apelo comercial: as *Pandectae* não tiveram o mesmo sucesso comercial que *Bibliotheca Universalis*, possivelmente por sua natureza e escopo temático. Hoje, por exemplo, as *Pandectae* são mais difíceis de serem encontradas, mesmo em bibliotecas patrimoniais. Trata-se de uma obra rara de grande valor financeiro e patrimonial;
2. Continuidade: como vimos, somente à *Bibliotheca Universalis* foi dada continuidade, por meio dos *Epitomi*;
3. Leitura e interpretação: a dificuldade de leitura e interpretação da obra pode ter sido uma das razões para o fato de que há predominância da aplicação do método bibliográfico de Gesner sob sua faceta descritiva e não semântica, como se observa em bibliografias pós-gesnerianas.

Do ponto de vista da história dos sistemas de classificação bibliográficos, o esquema proposto por Gesner destaca-se como notavelmente original, conforme assinalado por Malclès (1960), pois expande significativamente o escopo das sete artes liberais medievais.

3 Conclusão

Conrad Gesner, por meio de suas *Pandectae*, desempenhou um papel proeminente na história das classificações de natureza bibliográfica.

Sua abordagem categorial permitiu avanços significativos no campo da organização do conhecimento, uma vez que

o polímata foi capaz de reconhecer a importância de estabelecer um esquema classificatório circunscrito ao seu tempo. Sua contribuição consistiu em elaborar um esquema de classificação baseado em conceitos, o que possibilitou uma maior flexibilidade na categorização de conteúdos, levando em consideração as diferentes disciplinas em que estavam inseridos.

O esquema de classificação proposto por Gesner reflete sua concepção sobre a divisão do conhecimento, que está intimamente ligada ao pensamento clássico do homem renascentista.

O esquema de classificação gesneriano, portanto,

[...] não é simplesmente o mapa de uma série de desvios categoriais de natureza genericamente bibliográfico-literário ou bibliográfico-disciplinar; é a fachada de um edifício conceitual, enormemente complexo, cujos espaços internos foram modelados a partir de conceitos e coisas ordenadas de acordo com os modos infalíveis de uma ciência demonstrável, apoiada pela razão e sustentada pela fé. (SERRAI, 1990, p. 139).

Como vimos, há uma relação intrínseca entre o esquema de classificação gesneriano e a indicialização gesneriana. O esquema de classificação delinea a estrutura e a ordenação dos distintos tópicos e disciplinas, ao passo que a indicialização gesneriana emprega termos específicos derivados dessas classificações para erigir índices, objetivando facilitar a pesquisa e a referência às informações contidas nas obras. Por meio dessa correlação, Gesner estabelece uma abordagem sistemática para a organização e o acesso ao conhecimento nas *Pandectae*.

Ao contemplar os detalhes bibliográficos que passaram a compor os livros no Séc. XVI, graças à tecnologia dos tipos móveis, as *Pandectae* emergem como o primeiro esquema classificatório, desprovido de notação explícita, concebido especificamente para o universo *librario*.

Notas

- [1] Destacamos os estudos gesnerianos amplamente produzidos e difundidos no contexto do Fórum Internacional A Arte da Bibliografia, a partir de seus seminários internacionais e itinerantes e seus dossiês temáticos.
- [2] "Apesar de seu caráter elementar e sobretudo descritivo, [o sistema das sete artes liberais] havia atravessado toda a Alta Idade Média e conservava no Séc. XII uma popularidade que não se acabaria nem mesmo na época universitária. Sabemos que ele se baseava na distribuição das sete artes "liberais" (únicas dignas do homem livre) em duas categorias: as artes da palavra e do signo (gramática, retórica, dialética) no *trivium* e as artes do número e das coisas (geometria, aritmética, astronomia, música) no *quadrivium*. O estudo destas sete artes era propedêutico ao da disciplina suprema, Filosofia para os antigos, Santa Escritura para os doutores da Igreja". (VERGER, 2001, p. 74).

- [3] Sobre indicialização gesneriana, ver Araujo (2018), que apresenta e discute o contexto e os princípios do texto *De indicibus librorum*, de Conrad Gesner, e Araujo (2019), que ilustra e demonstra a aplicação da indicialização gesneriana.

Referências

ARAUJO, Andre Vieira de Freitas. *De indicibus librorum e a arte de indicialização em Conrad Gesner (Parte I): contexto e princípios*. **Informação & Informação**, Londrina, PR, v. 23, n. 2, p. 14-37, set. 2018. ISSN 1981-8920. DOI: <http://dx.doi.org/10.5433/1981-8920.2018v23n2p14>. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/34495>. Acesso em: 15 mar. 2020.

ARAUJO, Andre Vieira de Freitas. *De indicibus librorum e a arte de indicialização em Conrad Gesner (Parte II): ilustração e aplicação*. **Em Questão**, Porto Alegre, RS, v. 25, ed. esp., p. 137-158, 2019. ISSN 1808-5245. DOI: <https://doi.org/10.19132/1808-5245250.137-158>. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/92832>. Acesso em: 16 mar. 2020.

ARAUJO, Andre Vieira de Freitas Araujo; CRIPPA, Giulia. A dimensão histórico-epistemológica da Organização do Conhecimento: contributos da taxonomia gesneriana, Séc. XVI. In: GUIMARÃES, José Augusto Chaves; DODEBEL, Vera. (org.). **Organização do conhecimento e diversidade cultural**. Marília: ISKO-Brasil: FUNDEPE, 2015, p. 37-46. (Estudos Avançados em Organização do Conhecimento, v. 3). *E-book*. ISBN 978-85-98176-70-3. Disponível em: <http://isko-brasil.org.br/wp-content/uploads/2015/09/Organizacao-do-Conhecimento-e-Diversidade-Cultural-ISKO-BRASIL-2015.pdf>. Acesso em: 17 dez. 2017.

ARAUJO, Andre Vieira de Freitas; CRIPPA, Giulia; JUVÊNCIO, Carlos Henrique. Universality and utopia in Conrad Gesner and Paul Otlet: historical approximations. In: HADI, Widad Mustafa El (coord.). **Fondements épistémologiques et théoriques de la science de l'information-documentation**. London: ISTE Editions, 2018. p. 321-335. ISBN 978-1-78405-487-8. Actes du 11^o Colloque ISKO, Paris, França, 2017.

ARAUJO, Andre Vieira de Freitas; CRIPPA, Giulia; SABBA, Fiammetta. Semantic order in the 16th century: an introductory discussion of Conrad Gesner's *Pandectae*. In: GUIMARAES, José Augusto Chaves; MILANI, Suellen Oliveira; DODEBEL, Vera (ed.). **Knowledge organization for a sustainable world: challenges and perspectives for cultural, scientific, and technological sharing in a connected society**. Würzburg: Ergon Verlag, 2016. p. 59-66. ISBN 978-3-95650-221-7. Proceedings of the Fourteenth International ISKO Conference, Rio de Janeiro, Brasil, 2016. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/aiko_vol_15_2016_07_araujo_sabba_crippa.pdf. Acesso em: 19 mar. 2021.

CRIPPA, Giulia; ARAUJO, Andre Vieira de Freitas. Order of knowledge, selection and bibliographical tension in the 16th century: between Gesnerian universality and Possevinian anti-heretism. In: LYKKE, Marianne; SVARRE, Tanja; SKOV, Mette; MARTÍNEZ-ÁVILA, Daniel (ed.)

Knowledge Organization at the Interface. Würzburg: Ergon Verlag, 2020. p. 105-114. *E-book*. ISBN 978-3-95650-776-2. Proceedings of the Sixteenth International ISKO Conference, Aalborg, Denmark, 2020. Disponível em: <https://www.nomos-elibrary.de/10.5771/9783956507762-105/order-of-knowledge-selection-and-bibliographical-tension-in-the-16th-century-between-gesnerian-universality-and-possevinian-anti-heretism?page=1&select-row=abstract>. Acesso em: 19 mar. 2021.

GESNER, Conrad. **Catalogus plantarum latine, graeca, germanice & gallice [...]: una cum vulgaribus pharmacopolarum nominibus [...]**. Tigri: apud Christoph. Froschouerum, 1542. Disponível em: <http://www.e-rara.ch/zuz/content/titleinfo/847714>. Acesso em: 06 maio 2016.

GESNER, Conrad. **Conradi Gesneri medici Tigurini historiae animalium [...]**. Tigri: apud Christophorum Froschouerum, 1558. Disponível em: <http://www.e-rara.ch/zuz/content/titleinfo/5356265>. Acesso em: 05 maio 2016.

GESNER, Conrad. **Pandectarum sive partitionum universalium Conradi Gesneri Tigurini, medici & philosophiae professoris, libri XXI: ad lectores**. Secundus hic Bibliothecae nostrae tomus est, totius philosophiae & omnium bonarum artium atque studiorum locos communes & ordines universales simul & particulares complectens [...]. Tiguri: excudebat Christophorus Froschouerus, 1548b. Zentralbibliothek Zürich, 5.13, 2. Public Domain Mark. Disponível em: <http://www.e-rara.ch/doi/10.3931/e-rara-1936>. Acesso em: 05 maio 2018.

GESNER, Conrad. **Partitiones theologicae: pandectarum universalium Conradi Gesneri liber ultimus: Ad lectorem**. Pandectis nostris sive secundo Bibliothecae tomo, cuius libri XIX nuper editi sunt, [...]. Tiguri: Christophorus Froschouerus excudit, 1549. Zentralbibliothek Zürich, 5.13, 3. Public Domain Mark. Disponível em: <http://www.e-rara.ch/zuz/content/titleinfo/625727>. Acesso em: 05 maio 2018.

MALCLÈS, Louise Nöelle. **La bibliografía**. Buenos Aires: EUDEBA, 1960.

SABBA, Fiammetta. Natureza e origem da Bibliografia: uma perspectiva disciplinar para contemporaneidade. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 7, p. 65-98, ago. 2016. Disponível em: <http://www.revistas.usp.br/incid/article/view/118750/116234>. Acesso em: 17 fev. 2017.

SERRAI, Alfredo. **Conrad Gesner**. Editado por Maria Cochetti. Roma: Bulzoni Editore, 1990.

SERRAI, Alfredo. **Le classificazioni: idee e materiali per una teoria e per una storia**. Firenze: Leo S. Olschki Editore, 1977.

SERRAI, Alfredo; SABBA, Fiammetta. **Profilo di storia della Bibliografia**. Milano: Edizioni Sylvestre Bonnard, 2005.

VERGER, Jacques. **Cultura, ensino e sociedade no Ocidente nos séculos XII e XIII**. Bauru: EDUSC, 2001.

Organização do Conhecimento e Epistemologia Social: relações teóricas, epistemológicas e aplicadas

Knowledge Organization and Social Epistemology: theoretical, epistemological and applied relations

Ana Maria Mendes Miranda (1), Larissa Moraes Martins (2), Carlos Eduardo da Silva Carvalho (3),
Ana Cristina de Albuquerque (4)

(1) Universidade Estadual de Londrina (UEL), Londrina, anamirandamm@uel.br; (2) larissa.moraes.martins@uel.br;
(3) eduardo.carvalho98@uel.br; (4) albuati@uel.br

Resumo

Objetivo: Este trabalho tem como objetivo compreender as relações teóricas desenvolvidas entre a Epistemologia Social e a Organização do Conhecimento na última década. **Metodologia:** Trata-se de uma pesquisa descritiva, com abordagem qualitativa e delineamento bibliográfico. A coleta de dados se deu através de um levantamento bibliográfico relacionando os termos “Organização do Conhecimento” e “Epistemologia Social” em publicações de países ibero-americanos. **Resultados:** Quanto aos resultados, foram recuperados 16 textos que relacionam a Organização do Conhecimento a Epistemologia Social, essas contribuições têm como foco as reflexões epistemológicas acerca do próprio conhecimento; as sustentações teóricas e metodológicas da Organização do Conhecimento; e as contribuições aplicadas na construção de Sistemas de Organização do Conhecimento. **Conclusão:** Conclui-se que apesar de debate profícuo promovido nos trabalhos levantados, ainda é pequena a aproximação entre os campos.

Palavras-chave: Epistemologia Social. Organização do Conhecimento. Organização Bibliográfica. Margaret Egan. Jesse Shera.

Abstract

Objective: This work aims to understand the theoretical relations developed between Social Epistemology and the Organization of Knowledge in the last decade. **Methodology:** This is a descriptive research, with a qualitative approach and bibliographic design. Data collection took place through a bibliographic survey relating the terms "knowledge organization" and "social epistemology" in publications from Ibero-American countries. **Results:** As for the results, 16 texts were retrieved that relate the Knowledge Organization to Social Epistemology. These contributions focus on epistemological reflections about knowledge itself; the theoretical and methodological foundations of knowledge organization; and contributions applied in the construction of Knowledge Organization Systems. **Conclusion:** It is concluded that despite the fruitful debate promoted in the surveyed works, the approximation between the fields is still small.

Keywords: Social Epistemology. Knowledge Organization. Bibliographic Organization. Margaret Egan. Jesse Shera.

Resumen

Objetivo: Este trabajo tiene como objetivo comprender las relaciones teóricas desarrolladas entre la Epistemología Social y la Organización del Conocimiento en la última década. **Metodología:** Se trata de una investigación descriptiva, con enfoque cualitativo y diseño bibliográfico. La recolección de datos se realizó a través de un levantamiento bibliográfico relacionando los términos "organización del conocimiento" y "epistemología social" en publicaciones de países iberoamericanos. **Resultados:** En cuanto a los resultados, se recuperaron 16 textos que relacionan la Organización del Conocimiento con la Epistemología Social, estos aportes se enfocan en reflexiones epistemológicas sobre el conocimiento mismo; los fundamentos teóricos y metodológicos de la organización del conocimiento; y aportes aplicados en la construcción de Sistemas de Organización del Conocimiento. **Conclusión:** Se concluye que a pesar del fructífero debate promovido en los trabajos relevados, la aproximación entre los campos es aún pequeña.

Palabras clave: Epistemología Social. Organización del Conocimiento. Organización Bibliográfica. Margaret Egan. Jesse Shera.

1 Introdução

A ideia de Organização do Conhecimento (OC) vincula-se aos sistemas e atividades desempenhadas em, por exemplo, bibliotecas, arquivos e museus, envolvendo as atividades de classificação, indexação, descrição, catalogação, entre outras – em suma, o

tratamento de objetos por sujeitos inseridos nestes sistemas. Ainda, a Organização do Conhecimento refere-se à organização das universidades e instituições vinculadas à pesquisa e ao ensino, em sua produção e difusão de conhecimento (HJØRLAND, 2008). Apoiamo-nos, em suma, na ideia de que a Organização do Conhecimento se vincula aos estudos dos processos

de representação e recuperação de informações em diversos campos disciplinares, frisando o foco no estudo do conhecimento socializado ou registrado (BARITÉ, 2001; 2015), por meio de abstrações da realidade solidificadas em modelos de representação do mundo (BRASCHER; CAFÉ, 2008).

No que se refere a Epistemologia Social (ES), esta busca uma aproximação de elementos mais amplos da sociedade para a reflexão epistemológica, colocando uma alternativa para a discussão centrada no indivíduo (EGAN; SHERA, 1952), abrindo uma série de possibilidades para os estudos relacionados ao conhecimento humano. Conforme exposto no decorrer do presente trabalho, algumas das possibilidades da Epistemologia Social permitem explorar como coletivamente uma sociedade percebe, influencia e é influenciada pelos conhecimentos que a orbitam. Destaca-se, portanto, o olhar para a produção e para os rumos do conhecimento ao longo do tecido social (SHERA, 1972).

Ao propor uma compreensão de Conhecimento, a Epistemologia Social abre espaço para coadunar com proposições e investigações da Ciência da Informação, sobretudo no campo da Organização do Conhecimento. É interessante, portanto, observar que relações se apresentam entre a Organização do Conhecimento e a Epistemologia Social, levando-se em consideração as várias e distintas ocupações possíveis tanto em seu sentido amplo como estrito, conforme definidos por Hjørland (2008). Ainda, o estudo das relações que vêm sendo estabelecidas entre Epistemologia Social e Organização do Conhecimento pode auxiliar na compreensão de que potencialidades existem nas discussões teóricas e epistemológicas entre ambas. Tendo isto posto, este trabalho tem como objetivo refletir sobre as relações desenvolvidas entre Organização do Conhecimento e a Epistemologia Social na literatura em Ciência da Informação na última década.

É pertinente pontuar a relevância de trabalhos dessa natureza, tanto no que tange o reconhecimento dos caminhos traçados pela área, como para compreensão das relações empreendidas, e contributivas mutuamente, dos diferentes campos da Ciência da Informação. Mas, sobretudo sua pertinência para futuros estudos teóricos e epistemológicos sobre os processos de Organização do Conhecimento.

No que tange os procedimentos metodológicos, esta pesquisa pode ser caracterizada como descritiva, com abordagem qualitativa e delineamento bibliográfico. Inicialmente realizamos um levantamento bibliográfico relacionando os termos “Organização do Conhecimento” e “Epistemologia Social” em país ibero-americanos, depois filtramos a presença dos debates de autores como Jesse Shera e Margaret Egan nos textos sobre Organização do Conhecimento. Essa

busca foi realizada nas bases de dados do Google Acadêmico, La Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc), Scientific Electronic Library Online (SciELO), Portal de Periódicos da Capes, Biblioteca Digital de Teses e Dissertação do *Instituto Brasileiro de Informação em Ciência e Tecnologia (BDTD)*, Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI), nos Anais da ISKO Brasil e dos GT's 1 e 2 do Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ENANCIB) - os dois últimos foram selecionados visto a sua predominância nos temas aqui discutidos. As buscas foram realizadas considerando as publicações de 2012 a fevereiro de 2023.

Após o levantamento, foi realizada a leitura dos resumos, exclusão de trabalhos duplicados, editoriais de periódicos e textos que não tratam especificamente dos objetivos desta pesquisa. No total foram recuperados 16 textos que traçam relações entre a Organização do Conhecimento e a Epistemologia Social. No que tange a análise, esta se baseou na análise categorial de Bardin (2010), onde os textos foram lidos na íntegra, depois estabeleceu-se as categorias seguindo as tendências de debates apresentados no *corpus* da pesquisa, por fim foi identificado em cada texto quais trechos traziam relações dos dois campos e a quais categorias melhor se aderiam tanto os trechos selecionados quanto os trabalhos em sua totalidade.

As três categorias foram estabelecidas a partir das relações estabelecidas entre Organização do Conhecimento e Epistemologia Social tratando-se das seguintes: 1) epistemológica; 2) teórica; e 3) aplicada.

2 Organização do Conhecimento

Como afirmam Carlan e Medeiros (2011, p. 55), o conhecimento está em constante crescimento e expansão, e “a criação de conhecimento sobre os objetos que nos cercam constitui uma prerrogativa essencial da racionalidade humana”, sendo que o processo cognitivo de identificar as características dos objetos e agrupá-los de acordo com tais características desencadeia um processo classificatório.

Assim, se faz necessário destacar o que compreendemos por Organização do Conhecimento. Para tal, partimos da definição de Barité *et al.* (2015) presente no ‘*Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología*’, para quem a Organização do Conhecimento se dedica ao estudo das “[...] leis, princípios e procedimentos pelos quais o conhecimento especializado em qualquer disciplina, a fim de representar tematicamente e recuperar as informações contidas em documentos de qualquer tipo [...]”

(BARITÉ *et al.*, 2015, p. 120). Os autores ainda afirmam que a Organização do Conhecimento tem como foco de estudo o conhecimento socializado ou registrado.

Em concordância à reflexão de Brascher e Café (2008) quanto à finalidade da Organização do Conhecimento vincular-se à construção de modelos de representação do mundo, se constituindo em abstrações da realidade, considera-se que a representação do conhecimento “[...] se constitui numa estrutura conceitual que representa modelos de mundo, os quais, segundo Le Moigne (*apud* CAMPOS, 2004, p. 23), permitem descrever e fornecer explicações sobre os fenômenos que observamos.” (BRASCHER; CAFÉ, 2008, p. 6)

Neste sentido, Hjørland (2008) apresenta o campo como em dois âmbitos, o estrito e o amplo. No sentido estrito, o autor destaca as atividades e sistemas desenvolvidos na Organização do Conhecimento por bibliotecários, arquivistas e outros profissionais da informação como “[...] descrição, indexação e classificação de documentos realizadas em bibliotecas, bancos de dados bibliográficos, arquivos e outros tipos de “instituições de memória” [...]” (HJØRLAND, 2008, p. 86). No sentido amplo, de acordo com Hjørland (2008, p. 86) a Organização do Conhecimento trata

[...] sobre a divisão social do trabalho mental, ou seja, a organização das universidades e outras instituições de pesquisa e ensino superior, a estruturação das disciplinas e profissões, a organização social dos meios de comunicação, a produção e difusão do “conhecimento” etc.

Em abordagem similar, Barité (2001) destaca que mesmo que o processo de assimilação e construção de um novo conhecimento seja individual, ao final o resultado de tal processo será socializado de modo a tornar-se coletivo. O autor afirma então que o conhecimento “[...] também pode ser definido como ‘a acumulação social do saber humano em todas as suas expressões, isso é, o saber ou conhecimento efetivamente socializado’” (BARITÉ, 2001, p. 45, tradução livre).

Em conformidade com a compreensão apresentada por Hjørland (2008), reforça-se o entendimento da Organização do Conhecimento tanto em suas práticas materiais de organização, como também no todo que envolve a construção daquele conhecimento. Tendo em vista um campo que se direciona a pensar o conhecimento, tornam-se relevantes as compreensões epistemológicas que permeiam a compreensão de conhecimento, suas abordagens e aplicações - é nesse contexto que revisitamos a Epistemologia Social, com a finalidade de compreender na contemporaneidade suas relações com a Organização do Conhecimento.

3 O conhecimento a partir da Epistemologia Social

Em um contexto de desenvolvimento das tecnologias de informação após a Segunda Guerra Mundial e do aumento exponencial da informação científica, Egan e Shera (1952) publicaram o artigo intitulado “*Foundations of a Theory of Bibliography*”, no qual debatiam os fundamentos para uma teoria da bibliografia e identificaram uma lacuna no cenário disciplinar desse campo, a qual poderia ser preenchida com uma nova disciplina: a “Epistemologia Social” (EGAN; SHERA, 1952). Neste texto inaugural, os autores apresentam a Epistemologia Social como resultado da complexificação da produção, distribuição e utilização de produtos intelectuais. Eles evidenciam que a Epistemologia social, compreende o estudo dos

[...] processos pelos quais a sociedade como um todo busca alcançar uma relação perceptiva ou compreensiva com o meio ambiente total. mental-físico, psicológico e intelectual. A derivação do termo é facilmente perceptível. [...] A epistemologia social simplesmente eleva a disciplina da vida intelectual do indivíduo para a da sociedade, nação ou cultura. (EGAN; SHERA, 1952, p. 132)

Para os autores, essa nova disciplina seria fundamentada por quatro suposições básicas (EGAN; SHERA, 1952, p. 132-133): 1) o indivíduo pode entrar em uma relação de saber sobre o seu ambiente imediato, ou parte da totalidade de seu ambiente com o qual ele possui contato; 2) que os registros do conhecimento desenvolvidos pela humanidade permitem que o indivíduo possua aproximadamente um mesmo tipo de relação com a parte do ambiente total que está além da sua experiência imediata, mas é capaz de compreendê-la, pois os símbolos culturais e linguísticos presentes nos registros permitem que essa experiência não imediata também se torne sua experiência; 3) coordenando distintos conhecimentos de um conjunto de indivíduos, a sociedade pode conhecer mais do que o indivíduo sozinho, ou a soma de seus conhecimentos com os demais; 4) a ação social, refletindo a ação intelectual coletiva, consegue ultrapassar a ação individual, pois os seres humanos encontram-se em um ambiente amplo e complexo, que não pode ser totalmente compreendido individualmente.

A partir desses pontos é possível compreender que os seres humanos são capazes de realizar uma síntese intelectual de seu ambiente, e esse ambiente através dos registros pode incluir a experiência imediata e a experiência mediada pelos signos. Nesse contexto, Egan e Shera propunham que a Epistemologia Social se direcionasse para uma ação social inteligente, de maneira a possibilitar que sistemas bibliográficos fossem planejados e implementados em nível coletivo, de modo que componentes individuais sejam coordenados e integrados, visto que “[...] a ação coletiva, e não individual, é cada vez mais a força determinante vital no mundo para a formação da

sociedade contemporânea” (EGAN; SHERA, 1952, p. 134, tradução nossa).

Podemos dessa forma refletir, com aporte em Bozzete e Saldanha (2017), que a Epistemologia Social, para Egan e Shera, seria uma filosofia da ciência, que estuda o conhecimento e sua relação com a estrutura social e com a construção coletiva. Para Silva (2014), Shera visa compreender como o conhecimento de sujeitos individuais e grupos sociais alteram as estruturas e instituições sociais e estabelecem novos conhecimentos. Nesse sentido, a epistemologia social,

[...] promove uma fundamentação de cunho teórico ao processo de produção do conhecimento que interliga o conhecimento individual/físico e o conhecimento social/institucional que modificam estruturas e instituições sociais como forma de compreender a superação da percepção gnosiológica à fundamentação epistemológica do conhecimento [...]. (SILVA, 2014, p. 177)

É possível argumentar que a Epistemologia Social visa elevar os estudos sobre aspectos intelectuais, dos processos individuais para coletivos, relacionados a uma sociedade, nação, cultura e como elas compreendem a totalidade dos estímulos que agem sobre elas. O foco encontra-se, portanto, na “[...] produção, fluxo, integração e consumo de todas as formas de pensamento comunicado em todo o tecido social. De tal disciplina deve emergir um novo corpo de conhecimento e um novo sistema de interação entre conhecimento e atividade social.” (SHERA, 1972, p. 112).

No que tange nossos o campo da Organização do Conhecimento, a Epistemologia Social caminha na direção de compreender o conhecimento não a partir de modelos cognitivos individuais, mas de considerar tais processos a partir de suas condições culturais, coletivas e simbólicas, na qual o conhecimento desenvolve papel relevante no sentido de mediar as condições da realidade para a compreensão humana, tanto imediata quanto não imediata. Entendemos desta forma, que não só pelo seu pioneirismo na compreensão dos aspectos informacionais, mas também por sua contribuição para pensar o próprio conhecimento, muitas são as contribuições da Epistemologia Social aos estudos de Organização do Conhecimento.

4 Apresentação e análise dos resultados

Realizado o levantamento bibliográfico, conforme indicado na seção 1, os textos foram lidos e selecionados visando atender o objetivo da pesquisa. Desta forma, 16 textos compuseram o corpus da pesquisa (Tabela 1 – Anexo A). Após a leitura integral destes textos, estabeleceram-se categorias a posteriori que visam indicar quais debates têm sido realizados entre a Organização do Conhecimento e a Epistemologia Social. Como mencionado as três

categorias foram: 1) epistemológica; 2) teórica; e 3) aplicada.

Sobre tais categorias, cabe mencionar as discussões realizadas por Lamar (2007) para melhor explicar o sentido das categorias desenvolvidas neste trabalho para compreender as produções que envolvem a Organização do Conhecimento e a Epistemologia Social neste recorte. No que toca os trabalhos agrupados na categoria 1) *Epistemológica*, destacamos a preocupação com a dimensão epistemológica do conhecimento vinculado à OC. Especificamente aquelas relacionadas a compreensão do que é conhecimento, como ele pode ser socialmente compreendido no âmbito da Organização do Conhecimento, assim como no caráter do conhecimento, a forma como é “[...] produzido, difundido e consumido na sociedade [...]” (LAMAR, 2007, p. 105) nos SOC.

A despeito da dimensão teórica também presente nas discussões epistemológicas, a categoria 2) *Teórica* reúne os trabalhos que colocam suas discussões no campo da teoria, dedicados a discutir e elaborar teoria no campo da organização do conhecimento, munidos da Epistemologia Social com vias de legitimação ou validação das discussões empreendidas – de forma geral, uma discussão teórica, a qual envolve hipóteses e aparatos para observar um determinado objeto (ABBAGNANO, 2007). Conforme explica Francelin (2018), a depender da inserção teórica de um debate, a própria noção de “epistemologia” pode ganhar particularidades em seu sentido – fica desta forma evidente a diferença entre discussões diretamente direcionadas ao debate epistemológico e debates teóricos que podem vir a fazer uso dos construtos de determinada concepção para fundamentar ou legitimar seus argumentos.

Os trabalhos organizados em 3) *Aplicada* apresentam reflexões e debates sobre as potencialidades práticas das reflexões da Epistemologia Social para a Organização do Conhecimento, bem como experiências de trabalho apoiadas na Epistemologia Social diretamente em sistemas de Organização do Conhecimento, “[...] podendo contribuir para o melhoramento do bibliotecário e da Ciência da Informação” (LAMAR, 2007, p. 105), em geral realizando intercâmbios com metodologias e abordagens diversas, comumente apresentando um caráter “[...] interdisciplinar, para o qual poderão, também, contribuir, entre outras, a Linguística, a Ciência da Comunicação e a Lógica” (LAMAR, 2007, p. 105). Evidenciadas as categorias, realizou-se a seleção de trechos que serão apresentados e debatidos com a literatura nesta seção.

4.1 Epistemológica

No que concerne à categoria epistemológica, identificamos que a maioria dos textos recuperados apresentaram essa compreensão da Epistemologia Social como ponto de reflexão acerca do próprio conhecimento. Esse dado é semelhante ao encontrado por Vieira, Lucas e Araújo (2017), que mapearam as menções a Jesse Shera na Ciência da Informação. Os autores apontam que os estudos históricos e epistemológicos aparecem em destaque entre essas citações, destacando que a Epistemologia Social e a Ciência da Informação devem ser vistas como complementares, e não como campos distintos - assim como a compreensão de Shera de que a Ciência da Informação é uma evolução da Biblioteconomia.

No contexto dessa pesquisa, podemos destacar na categoria epistemológica o texto de Martínez-Ávila (2018), que inicia seu debate estabelecendo o que é o conhecimento científico, indicando que é aquele obtido por meio do método científico e sistematizado principalmente na forma de teorias. Ao levar em conta tal compreensão, os autores entendem que historicamente, um dos principais marcos científicos da Ciência da Informação foi a institucionalização, em 1928, do primeiro programa de doutorado na *Graduate Library School* (GLS) da Universidade de Chicago. O programa em questão, tinha como discentes profissionais que já possuíam outras formações, contexto que levou as pesquisas de doutorado a serem relacionadas a outras áreas do conhecimento, trazendo novas contribuições científicas para delimitar o objeto de estudo da CI, assim como delimitar sua identidade.

Para além de contribuições relacionadas a institucionalização da CI como um todo, destaca-se que a Epistemologia Social desenvolvida na *Graduate Library School*, na década de 50, pode ser considerada ainda hoje uma das primeiras e mais sérias iniciativas de oferecer uma base epistemológica sólida para o campo. Vale destacar, que essa proposta surge como uma resposta a problemas científicos de comunicação, cognição e uso de ferramentas bibliográficas utilizadas para a Organização do Conhecimento, denominada à época de organização bibliográfica (MARTÍNEZ-ÁVILA, 2018).

Para Martínez-Ávila (2018, p. 85) apoiado em Martínez-Ávila e Zandonade (2017)

Esta Epistemologia Social supõe a passagem de uma epistemologia filosófica especulativa individual (seguindo a interpretação mais anglófona do conceito de epistemologia de James Frederick Ferrier e correspondente ao que Bertrand Russell chama de teoria do conhecimento) para uma Epistemologia Social que incorpora elementos da sociologia do conhecimento. É influenciado principalmente pelo pragmatismo.

Assim, a Epistemologia Social, ao questionar o papel dos profissionais da informação e da organização

bibliográfica no processo de comunicação, propõe uma passagem da concepção filosófica do conhecimento, para uma concepção sociológica em relação aos aspectos de organização e comunicação do mesmo. Essa inversão proposta, passa especificamente por compreender a partir de uma lente coletiva o compartilhamento do conhecimento e os processos de organização caros à Biblioteconomia e à Ciência da Informação como classificações, cabeçalhos de assunto, título de disciplinas, entre outros (MARTÍNEZ-ÁVILA, 2018). É nessa compreensão de conhecimento coletivo, cujas estruturas conceituais estão em mudanças, resultados dos fluxos históricos e sociais de cada tempo, que Egan e Shera apontam para um caminho epistemológico possível ao se pensar a Organização do Conhecimento.

Em compreensão semelhante, Veronez Júnior (2021, p. 155) considera que os estudiosos da GLS foram alguns dos estudiosos que a época refletiram sobre “[...] aspectos históricos, sociais e culturais acerca do conhecimento e a forma como ele deveria ser organizado, acessado e comunicado de uma maneira integrada para a sociedade.” Nesse ínterim, Silva (2014) ainda adiciona que é por uma compreensão de forças intelectuais na produção, circulação e uso do conhecimento que a Epistemologia Social se caracteriza como um campo do conhecimento sobre o próprio conhecimento. De um lado preocupando-se com o conhecimento científico, por outro com “[...] o conjunto de conhecimentos gnosiológicos – materiais e sociais – que regem as relações sociais e promovem base epistemológica a produção do conhecimento constituindo um viés teórico-empírico.” (SILVA, 2014, p. 177)

Vale pontuar que, apesar do pioneirismo de GLS e de Egan e Shera ao pensar uma Epistemologia Social, outros autores que surgiram posteriormente também têm refletido epistemologicamente a questão do conhecimento. Goldman (2001, não paginado) por exemplo, olha para a Epistemologia Social como um “estudo das dimensões sociais do conhecimento ou da informação.” Segundo o autor, não há consenso sobre o que compreende o termo conhecimento e qual alcance a perspectiva social possui nesse contexto.

Santos, Neves e Souza (2019) se utilizam das compreensões de Shera, para argumentar acerca dos aspectos epistemológicos do conhecimento, a exemplo disso, pode-se mencionar o entendimento que apresentava Shera sobre as classificações bibliográficas, evidenciando que essas não têm apenas uma função prática organizativa, mas refletem a própria lógica de estruturação do pensamento humano. Desta forma, compreende-se que além do fator epistemológico, a Organização do Conhecimento e a Epistemologia Social têm como foco as relações de produção do conhecimento humano e a configuração das

metodologias para a criação de produtos intelectuais que possibilitem o acesso a esse conhecimento. Ou seja, Egan e Shera, propõem um estudo da formação e Organização do Conhecimento na sociedade (WILSON, 2008, p. 7).

Refletindo tais apontamentos, é possível evidenciar que a conexão entre a Organização do Conhecimento e a Epistemologia Social deva ser analisada de um ponto de vista coletivo, no ponto em que ambas se voltam a compreensão de um conhecimento produzido, registrado, organizado e comunicado às comunidades - sendo também compreendidos de forma histórica como parte das construções socioculturais daquelas comunidades. (VERONEZ JÚNIOR, 2021).

Em proposta semelhante, Araújo (2021) discute os desafios epistemológicos colocados para a Ciência da Informação na contemporaneidade. Preocupado com questões ligadas especificamente à Organização do Conhecimento, apresenta a perspectiva debatida por Furner em torno das reflexões epistemológicas para a OC. O autor enfatiza a procura por fundamentos epistemológicos realizada por Furner para a OC, destacando a aproximação da Epistemologia Social e da noção de justiça epistêmica. É notável nesse sentido, a associação realizada entre a Epistemologia Social e um sentido ético, na perspectiva de reformar práticas e instituições que historicamente se construíram discriminatórias, incluindo sistemas de Organização do Conhecimento, com vistas a construir o que Furner chama de Organização do Conhecimento crítica. Esta, por sua vez, deveria se inspirar em valores da justiça epistêmica, respeitar os direitos humanos, preferir a verdade em detrimento da relevância e ser informada pela Epistemologia Social aplicada (ARAÚJO, 2021). Ao debater os novos desafios para a epistemologia na CI, portanto, Araújo (2021) eleva a uma prateleira de grande importância as discussões realizadas por Furner com sua teorização relacionada à Epistemologia Social observando sistemas de Organização do Conhecimento.

Já o trabalho de Araújo, Guimarães e Tennis (2021), pretende reconhecer a forma como diferentes posturas epistêmicas influenciam a prática da representação do conhecimento e procura compreender as influências da pesquisa neste domínio. Para isso, analisam as publicações sobre epistemologia na Organização do Conhecimento no periódico *Knowledge Organization*. O texto aponta para a importância e influência da epistemologia na Organização do Conhecimento. Contudo, não apresenta nenhuma reflexão dedicada à Epistemologia Social, concentrando-se na verificação das tendências epistemológicas que habitam a OC.

Dentre os resultados, os autores identificaram a presença mais expressiva do pensamento pragmático nos estudos de Organização do Conhecimento, seguidos por abordagens da teoria crítica e do historicismo. A perspectiva pragmática aparece como

um “[...]estímulo para repensar o desenho de sistemas universais em OC e dar como certo o ambiente social e cultural e sua influência no processo e nas ferramentas da OC.” (ARAÚJO; GUIMARÃES; TENNIS, 2021, p. [13]). Os autores ainda ressaltam a dominância do paradigma sociocognitivo no campo da OC, uma vez que tanto o conhecimento quanto a Organização do Conhecimento se relacionam às necessidades e interesses humanos, e o conhecimento é construído em diferentes contextos

[...] e leva a diferentes reivindicações de conhecimento. Entende-se que a pesquisa e a prática da OC tendem ao modelo socialmente pluralista, o que pode estar relacionado ao pensamento pós-moderno sobre a relação entre as pessoas, o conhecimento como construção social e a influência de várias perspectivas. (ARAÚJO; GUIMARÃES; TENNIS, 2021, p. [13]).

Apresentando, portanto, uma interessante reflexão que aprofunda a discussão epistemológica no campo, e ressaltando a presença e predominância de pressupostos não vinculados diretamente à Epistemologia Social, como o sociocognitívismo, mas com um modelo pluralista e em alguma medida associado ao entorno social.

Cabe ainda destacar nessa categoria textos como de Martínez-Avila e Mello (2023) cuja proposta é inter-relacionar alguns aspectos históricos da epistemologia geral com as epistemologias feministas, destacando-se a corrente pós-estruturalista. Para tal, partem da confecção de críticas às epistemologias que apontam como herdeiras do cartesianismo e às pretensões de construção de determinada razão purificada, verdadeira e imutável, buscando os debates construídos pelos movimentos sociais para amparar sua argumentação. Nesse contexto, os autores relacionam as epistemologias feministas como a Epistemologia Social. Traçando à tona Egan e Shera (1952), os autores reiteram que as epistemologias sociais são de grande importância à Organização do Conhecimento, assim como às epistemologias de gênero, como as feministas, pois implica numa posição epistêmica social.

Vale nesse tocante destacar uma especificidade da aproximação entre a Epistemologia Social e as epistemologias sociais feministas mencionadas pelo trabalho, já que em meio ao alerta de que se concebe categorizações como sempre provisórias, destaca-se a relação com as epistemologias feministas pós-estruturalistas como principal vetor de crítica aos métodos cartesianos. Tal é a linha argumentativa desenvolvida sob a necessidade apontada de “[...] reverter os dogmatismos epistemológicos na Organização do Conhecimento” (MARTÍNEZ-AVILA; MELLO, 2023, p. 191). Em suma, os autores supracitados apresentam uma crítica às formas tradicionais de pensamento e Organização do Conhecimento, atreladas à epistemologia “tradicional” e ao cartesianismo, apoiando-se em aproximações com

a Epistemologia Social e apontando que construtos como os das “epistemologias feministas” apresentariam possibilidades de organização mais adaptadas às novas demandas da sociedade.

Já Nascimento e Guimarães (2017), dialogam com questões colocadas pela Epistemologia Social, já que partem da necessidade de abastecer o campo acadêmico com discussões e inquietações trazidas por setores da sociedade historicamente discriminados, além de contribuir para a transformação de sistemas de Organização do Conhecimento sob a perspectiva da Epistemologia Social, já que estuda termos e vocabulário de grupos que historicamente foram representados em sistemas de organização de forma profundamente discriminatória. Em meio à identificação de problemas em sistemas de Organização do Conhecimento ao lidar com as temáticas LGBTs e aos estudos do discurso interno à própria comunidade, os autores apontam a necessidade de “[...] repensar os processos cognitivos e semânticos humanos que o faz um ser essencialmente social - a linguagem.” (NASCIMENTO; GUIMARÃES, 2017, p. 352) e aproximam-se de Shera (1977) para compreender a necessidade humana de conceituar e comunicar conceituações. Ao final destacam que é na identificação de termos para

[...] nomear, organizar e celebrar as identidades, as comunidades LGBT toma para si o poder de nomear: não como intermediários de informação que se esforçam pela objetividade e neutralidade, mas como membros comprometidos de uma comunidade marginalizada, frente a construção histórica [...] (NASCIMENTO; GUIMARÃES, 2017, p. 358).

É possível ponderar que aproximações como de Nascimento e Guimarães (2017) e Martinez-Ávila e Mello (2023), fazem emergir novas compreensões da Epistemologia Social a Organização do Conhecimento. Para o Oddone (2007, p. 112) a Epistemologia Social quando

[...] referendada por um quadro teórico mais amplo e complexo, que simultaneamente a enriquece e contextualiza, a tese da Epistemologia Social transforma-se, atualiza-se, passando a constituir um corpo de conhecimentos sobre a dinâmica social da atividade intelectual dos coletivos humanos.

Quanto a essas compreensões podem ser vistas em obras de autores da Epistemologia Social como Popkewitz (1997), ainda que não tenham sido foco dos debates de Egan e Shera. Sobre isso, o autor argumenta que a Epistemologia Social se refere ao estudo da relação do conhecimento, instituições e poder. O autor propõe um uso da epistemologia como conceito histórico, social e pragmático, reforçando que

[...] não há base comum onde possamos colocar um verdadeiro consenso ou um modelo neutro permanente, segundo o qual seja possível avaliar um argumento

racional. Não há esquemas universais de raciocínio e racionalidade, mas somente epistemologias socialmente construídas que representam e incorporam relações sociais. (POPKEWITZ, 1997, p. 39)

Por fim, vale a reflexão de Veronez Júnior (2021) em sua pesquisa de dissertação quando pondera que ao estabelecer essa inter-relação entre a Organização do Conhecimento e a Epistemologia Social, como campos da CI, contribui para as investigações sobre o conhecimento, a forma como é produzido, principais comunidades que o produzem, assim como a maneira que elas estabelecem a ordem e prioridade de organização, classificação, representação e socialização para outras comunidades, além de contribuir para pensar sobre como esse conhecimento será recuperado.

4.2 Teórica

No que tange a categoria teórica, muitas menções são realizadas a Epistemologia Social com o intuito de reforçar argumentos teóricos relevantes, mas sem um aprofundamento maior sobre as contribuições da Epistemologia Social ao campo em questão. Esse dado se assemelha mais uma vez a pesquisa de Vieira, Lucas e Araújo (2017, p. 213), que durante sua análise das citações que remetem a Shera, identificaram que o autor é utilizado, na maioria das vezes, para legitimar alguma ideia, reflexão ou pensamento.

Marcondes (2021), ao discutir os fundamentos da Organização do Conhecimento, fala sobre três paradigmas epistemológicos na CI, apoiando-se em Capurro para destacá-los: o físico, o cognitivo e o social, este tendo “[...] suas origens na obra de Jesse Shera, atualmente representado pelas teorias de Bernd Frohmann, Birger Hjørland, Rafael Capurro e Søren Brier” (MARCONDES, 2021, p. 261). Mesmo sem o uso para a construção de uma perspectiva para fundamentar o trabalho, Shera aparece como um dos autores basilares no cenário da Organização do Conhecimento.

No entanto, é possível destacar o debate realizado por Macedo (2013) em sua análise sobre a Ciência da Informação e o conceito de informação como fenômeno social, onde discute as contribuições e relações entre os escritos da Epistemologia Social de Margaret Egan e Jesse Shera, da Análise de Domínio de Birger Hjørland e dos três paradigmas para a Ciência da Informação de Rafael Capurro. Neste estudo o autor reforça o papel das Bibliografias e da Documentação destacados por Egan e Shera para o desenvolvimento teórico e prático da Ciência da Informação e áreas correlatas.

Vale destacar também, a proposta de Santis (2016) que defende em seu trabalho uma Organização do Conhecimento pautada na recuperação da informação em detrimento de priorizar a representação. Para sustentar tal argumento se utiliza da compreensão de

Shera (1960) de que a recuperação deveria ser o foco de uma teoria da Organização do Conhecimento e, portanto, o objetivo que deveria centralizar os esforços dessas teorias.

É válido também mencionar a pesquisa de Tamayo (2014), na qual propõe um modelo para o ensino dos conteúdos da Organização do Conhecimento nos cursos de graduação em Biblioteconomia, Documentação e Ciência da Informação no Mercosul. O modelo é composto por três categorias principais, sendo elas: Fundamentos, Aplicação e Ética. Estas “[...] estão relacionadas entre si, pois os fundamentos impulsionam e sustentam as aplicações e estas, por sua vez, retroalimentam os fundamentos; ambas as categorias são regidas pela ética.” (TAMAYO, 2014, p. 78). A autora pontua que a Epistemologia Social deve figurar entre os conteúdos propostos para a categoria de fundamentos do modelo. Tal categoria seria dedicada a apresentar a definição, o objeto, as atividades realizadas, além da justificação, caracterização, função e diferenciação da Organização do Conhecimento, de modo a oferecer um embasamento e compreensão do campo e suas relações.

Outros autores que destacam as contribuições teóricas de Egan e Shera e da Epistemologia Social para a cientificidade do campo são Martins e Moraes (2015, p. [16]) que afirmam que

[...] a organização dos registros do conhecimento conquista espaço, não só como uma atividade pragmática, mas expande-se em um campo de estudos teóricos e aplicados e suas contribuições dão origem a uma disciplina institucionalizada tornando assim, mais acessível o uso e a apropriação do conhecimento produzido.

Veronez Júnior (2021) também discorre sobre as relações entre a Organização do Conhecimento e a Epistemologia Social, sendo esta última considerada o exame entre as relações que os seres humanos estabelecem com seu entorno, seja ele social, filosófico, cultural, tecnológico ou antropológico, visando e se manifestando no “[...] estudo do ciclo que envolve atividade de produção, tratamento, classificação, organização, circulação e o uso do conhecimento, caracterizado em sua materialidade como uma verdadeira ecologia sociotécnica do trabalho intelectual desenvolvido não só por cientistas e pesquisadores, mas também pela sociedade como um todo.” (VERONEZ JÚNIOR, 2021, p. 155)

Castanha, Veronez Junior e Dalessandro (2021) analisam a produção científica sobre Epistemologia Social no Brasil por meio de uma análise bibliométrica. Menciona-se a Organização do Conhecimento junto às outras práticas vinculadas à CI e às unidades de informação, ressaltando a relevância da Epistemologia Social, sobretudo o destaque que Hjørland possui entre os autores mais citados. Para os autores, Hjørland

relaciona a Epistemologia Social à Organização do Conhecimento e a Análise de Domínio, como uma maneira de produção, organização, classificação, representação e socialização do conhecimento sob uma perspectiva social, principalmente através dos estudos de comunidades discursivas por meio de suas práticas linguísticas e culturais.

Sobre a questão da relação entre conhecimento e linguagem, Shera (1977, p. 10) argumenta que o conhecimento e a linguagem não podem ser separados, isso porque a linguagem se apresenta como uma estruturação simbólica do conhecimento, o que permite sua comunicação. Desta maneira, a linguagem age como signo tanto sobre conhecimento em sua concepção individual, quanto coletivo.

Nesse ponto, cabe se apropriar de Oddone (2007) quando indica que os processos de mapear e descrever a inteligência coletiva, que trata-se de uma rede complexa compartilhada por diferentes agentes que promovem o pensamento, a inteligência e a informação, são sem dúvida relacionados às reflexões teóricas da Epistemologia Social. Isto porque, ela pode ser compreendida como uma ecologia sociotécnica do trabalho intelectual, ou seja, são esses processos que sustentam, organizam e disciplinam a produção, circulação e compartilhamento de conhecimentos. Esse arcabouço teórico-conceitual pode ser utilizado no contexto da Organização do Conhecimento como embasamento teórico para as pesquisas e propostas metodológicas.

4.3 Aplicada

No que se refere à categoria aplicada, Martínez-Ávila e Zandonade (2020) apresentam em seu escrito detalhes sobre o contexto intelectual em que nasceu a Epistemologia Social. Incluindo aspectos como suas influências teóricas, contextos intelectuais e características principais. Nessa incursão, evidenciam que a Epistemologia Social inicialmente tratava-se de uma abordagem direcionada a organização bibliográfica, sendo que para os autores tal abordagem tratava-se de uma filosofia tanto para os estudiosos e desenvolvedores, como para os profissionais que as aplicavam sistemas de classificação, regras de catalogação e qualquer outro mecanismo de controle bibliográfico. As contribuições de Egan e Shera se apresentam relevantes para os sistemas de Organização do Conhecimento em uma fase de inconsistências teóricas tanto para o desenvolvimento destas ferramentas, como para a própria fundamentação da Ciência da Informação.

Sobre essas contribuições, Smiraglia (2014) indica que a Escola de Chicago se tornou marca registrada do pensamento que levaria a termos hoje o campo da Organização do Conhecimento. Isto porque, Egan e Shera em um ensaio de 1950, refletem acerca da crise

de crescimento editorial do início do século XX, e a consequente dificuldade das bibliotecas de realizar o controle bibliográfico. Neste mesmo ensaio, os autores debatem a Organização Bibliográfica ponderando que existem duas maneiras de aplicá-la: interna e externa. Isso porque a organização bibliográfica pode ser percebida do ponto de vista dos bibliotecários e dos estudiosos, relacionados a sua concepção e uso. Mas pode também ser compreendida a partir do contexto de atividades intelectuais em geral (SHERA; EGAN, 1950).

Com essa análise, Shera e Egan (1950) equiparam a organização bibliográfica ao movimento da documentação, e refletem sobre o desenvolvimento de dispositivos que possam direcionar a energia intelectual na extração de informações relevantes para tarefas específicas. Essa divisão entre aspectos internos e externos, vão se apresentar como um prenúncio das noções de Wilson de domínios descritivos e exploradores.

Para Smiraglia (2014), a Escola de Chicago ainda contribui com o campo da Organização do Conhecimento ao descrever o papel social da organização bibliográfica como uma arte da comunicação, colocando a classificação como fator central para resolução da crise enfrentada pelas comunidades bibliográficas.

Também vale destacar a relevância da aproximação que Shera e Egan fazem da Sociologia do Conhecimento, para o desenvolvimento do campo de Organização do Conhecimento. Os autores apontam que tradicionalmente, as ferramentas e métodos empregados pela Biblioteconomia para a organização de suas coleções, tais como esquemas de classificação, cabeçalhos de assuntos, índices e outros dispositivos para a análise de assuntos de unidades bibliográficas, têm se desenvolvido baseados em uma suposição de que as relações entre os ramos do conhecimento são permanentes ou relativamente permanentes, tornando tais sistemas mais inflexíveis (MARTÍNEZ-ÁVILA; ZANDONADE, 2020). Esses questionamentos apontaram para novos caminhos e abordagens para pensar os sistemas de Organização do Conhecimento, a priori em bibliotecas, e repensar as próprias relações conceituais entre os domínios do conhecimento.

Em seu estudo, Brascher e Guimarães (2018), buscaram identificar a influência dos paradigmas propostos por Capurro na literatura quanto ao tratamento temático da informação (TTI) que, de acordo com os autores, se caracteriza como inserida no contexto dos estudos da Organização do Conhecimento. A relação com a Epistemologia Social aparece nos resultados referentes ao paradigma social, em especial no artigo de Fairthorne (1969) que se dedica ao papel social do bibliotecário e as atividades desenvolvidas por ele, em especial a TTI. Martínez-Ávila e San

Segundo (2020) realizam uma discussão sobre as abordagens de crowdsourcing para Organização do Conhecimento e, em especial, voltadas para o desenvolvimento de sistemas de classificação. Em meio ao debate os autores pontuam sobre aspectos da Organização do Conhecimento e a classificação em específico aos quais a Epistemologia Social pode contribuir como a questão de justificativa, autoridade, universalização, entre outros. Os autores também afirmam que a metodologia *Open Shelves Classification* (OSC) pode ser considerada como exemplo da Epistemologia Social no âmbito da Organização do Conhecimento, uma vez que “a Epistemologia Social analisa como os processos sociais levam à aquisição de conhecimento e a *Open Shelves Classification* era um sistema social para criar e adquirir conhecimento.” (MARTÍNEZ-ÁVILA; SAN SEGUNDO, 2020, p. 12).

Nesse tocante, vale mencionar mais uma vez a pesquisa desenvolvida por Vieira, Lucas e Araújo (2017, p. 212), que identifica poucas menções a contribuições aplicadas da Epistemologia Social, tais como o Catálogo Sistemático, que mesmo bastante conhecido, não é mencionado com frequência como uma contribuição de Shera a Organização do Conhecimento.

Em sentido semelhante, Veronez Júnior (2021, p. 155) reforça que ainda que Egan e Shera não tenham desenvolvido “[...] uma concepção ou projeto de ‘organização do conhecimento científico’, é salutar destacar os esforços que ambos desempenharam em termos de análises, estudos e desenvolvimento de teorias para o campo da Ciência da Informação [...]”. É possível refletir a partir desses dados e das falas dos autores, que seja de forma direta, como o Catálogo Sistemático, ou de forma indireta, como as pesquisas apresentadas, a Epistemologia Social e a Organização do Conhecimento desenvolvem processos relacionais que auxiliam na construção e modelagem de Sistemas de Organização do Conhecimentos (SOC).

5 Considerações finais

Podemos conjecturar que tanto a Epistemologia Social, quanto os estudos de Organização do Conhecimento, tiveram influências significativas no que concerne os debates teóricos, epistemológicos e metodológicos na compreensão contemporânea do conhecimento e suas formas de organizá-lo e compartilhá-lo.

No que tange aos resultados dessa pesquisa, foi possível notar as aproximações entre a Organização do Conhecimento e a Epistemologia Social, que ficam mais proeminentes nas reflexões acerca do que é conhecimento e como ele é adquirido, organizado e disseminado na sociedade. E tais debates já aparecem, mesmo que não de forma abundante, na literatura da Organização do Conhecimento e como pontuam

Castanha, Veronez Júnior e Dalessandro (2021) e Martínez-Ávila e Mello (2023), é importante que tais relações sejam cada vez mais debatidas nas produções do campo e de maneira direta refletir sobre as contribuições da Epistemologia Social em especial referente aos aspectos sociais. Vale destacar os esforços empreendidos pelo pesquisador Espanhol Daniel Martínez-Ávila, que têm movido suas pesquisas no sentido de debater o papel da Epistemologia Social ao pensar fatores fundantes da Ciência da Informação e da Organização do Conhecimento, sobretudo a partir de debates epistêmicos para esse campo.

O presente trabalho se dedicou a apresentar os estudos que relacionam os dois campos em três categorias, a epistemológica, a teórica e a prática. A categoria epistemológica apresentou debates em que os autores buscavam refletir sobre alguns conceitos chave para ambos os campos, tais como conhecimento e ética, assim como apresentar propostas alternativas às epistemologias tradicionais do campo. Dentre os resultados identificados na categoria teórica foi possível perceber que os autores se dedicavam a discutir a cientificidade e institucionalização do campo e as contribuições para a fundamentação de seus estudos e produtos. E quanto a categoria prática, esta abarcou estudos que tratavam mais especificamente sobre como tais perspectiva contribuem e impactam na mudança e desenvolvimento das atividades e processos da organização do conhecimento, lembrando que estes são os aspectos principais trabalhados nos estudos, mas as três categorias acabam por aparecer em todos os estudos uma vez que as categorias se encontram interligadas na realidade.

Ainda, vale ressaltar que estes elementos atravessam em maior ou menor medida todos os trabalhos. Destacamos a preocupação relacionada à como se produz e constitui o conhecimento, pano de fundo dos trabalhos que discutem e apresentam as questões basilares da Epistemologia Social em incursão na constituição coletiva e relacional do conhecimento, entendendo que a ação coletiva ultrapassa a ação individual (EGAN; SHERA, 1952). E, também, a preocupação e atenção dedicadas a problemas sociais e sua influência em várias dimensões da organização do conhecimento, desde os aportes epistemológicos até a concretude dos SOC, que aparecem em alguma medida ao longo dos trabalhos arrolados, mas muito destacadamente em Nascimento e Guimarães (2007), Araújo (2021) e Martínez-Ávila e Mello (2023).

Referências

- ABBAGNANO, N. **Dicionário de filosofia**. São Paulo: Martins Fontes, 2007.
- ARAÚJO, C. A. A. Novos desafios epistemológicos para a ciência da informação. **Palabra Clave (La Plata)**, Buenos Aires, v. 10, n. 2, 2021.
- ARAÚJO, P. C. de; GUIMARÃES, J. A. C.; TENNIS, J. T. A concepção de epistemologia da organização do conhecimento. **Palabra Clave (La Plata)**, Buenos Aires, v. 10, n. 2, 2021.
- BARDIN, L. **Análise de Conteúdo**. Lisboa: Ed. 70, 2010.
- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual en bibliotecología y documentación. *In*: CARRARA, K. (org.). **Educación, universidad e pesquisa**: textos completos do III simpósio em filosofia e ciência: paradigmas do conhecimento no final do milênio. Marília: Unesp; São Paulo: FAPESP, 2001. p. 35-60.
- BARITÉ, M.; COLOMBO, S.; DUARTE BLANCO, A.; SIMÓN, L.; CABRERA CASTROMÁN, G.; ODELLA, M. L.; VERGARA, M. **Diccionario de organización del conocimiento**: Clasificación, Indización, Terminología. 6. ed. Montevideo: CSIC, 2015.
- BOZZETTI, R. P.; SALDANHA, G. Jesse shera, the wars and the pieta: social epistemology as criticism of information ontology. **Brazilian Journal of Information Science: research trends**, Marília, v. 11, n. 2, p. 79-87, 2017.
- BRÄSCHER, M.; CAFÉ, L. Organização da informação ou organização do conhecimento? *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais [...]**. São Paulo: USP, 2008.
- BRÄSCHER, M.; GUIMARÃES, J. A. C. Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995. **Liinc em Revista**, Rio de Janeiro, v. 14, n. 2, p. 241-258, nov., 2018.
- CAMPOS, M. L. de A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ci. Inf.**, Brasília, v.33, p.1, p.22-32, jan./abr. 2004.
- CARLAN, E.; MEDEIROS, M. B. B. Sistemas de organização do conhecimento na visão da Ciência da Informação. **RICI**, Brasília, v. 4, n. 2, 2011.
- CASTANHA, R. G.; VERONEZ JÚNIOR, W. R.; DALESSANDRO, R. C. As relações e influências da epistemologia social no Brasil: uma análise bibliométrica. **AtoZ: novas práticas em informação e conhecimento**, Curitiba, v. 10, n. 2, 2021.
- EGAN, M. E.; SHERA, J. H. Foundations of a theory of bibliography. **The Library Quarterly**, [S. l.], v. 22, n. 2, p. 125-137, 1952.
- FAIRTHORNE, R. A. Content analysis, specification, and control. **Annual Review of Information Science (Arist)**, v.4, p. 73-109, 1969.
- FRANCELIN, M. M. Epistemologia da ciência da informação: evolução da pesquisa e suas bases referenciais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 23, n. 3, p. 89-103, 2018.
- GOLDMAN, A. I. Social epistemology. *In*: E. Z. (ed.) **Stanford encyclopedia of philosophy**. Stanford: Stanford University, 2001. Disponível em: <https://meinong.stanford.edu/entries/epistemology-social/>. Acesso em: 05 mar. 2023.

- HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, [S. l.], v. 35, n. 2/n. 3, 2008.
- LAMAR, A. R. Epistemologia social: possível origem e alguns momentos de seu percurso. **Pro-posições**, v. 18, n. 1, p. 103-113, 2007.
- MACEDO, T. B. Ciência da Informação: uma abordagem para a transformação. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTO E CIÊNCIA DA INFORMAÇÃO, 25., Florianópolis, 2013. **Anais [...]**. Florianópolis: UFSC, 2013.
- MARCONDES, C. H. Fundamentos da organização do conhecimento. **PontodeAcesso**, Salvador, v. 15, n. 3, p. 249-282, dez., 2021.
- MARTÍNEZ-ÁVILA, D. Hacia una base teórica social de la ciencia de la información. **Anuario ThinkEPI**, [S. l.], v. 12, p. 83-89, 2018.
- MARTÍNEZ-ÁVILA, D.; MELLO, M. R. G. de. Epistemologias, gênero e dogmatismo científico: Desdobramentos na Organização do Conhecimento. **LOGEION: Filosofia da Informação**, Rio de Janeiro, v. 9, n. 1, p. 182-194, set., 2022/ fev., 2023.
- MARTÍNEZ-ÁVILA, D.; SAN SEGUNDO, R. The application of crowdsourcing and the Bazaar model to the development of library classifications: an assessment of the Open Shelves Classification. **Transinformação**, Campinas, v. 32, 2020.
- MARTÍNEZ-ÁVILA, D.; ZANDONADE, T. Social epistemology in Information Studies: A consolidation. **Brazilian Journal of Information Science: Research trends**, v. 14, n. 1, p. 7-36, jan./mar., 2020
- MARTINS, G. K.; MORAES, J. B. E. de. Organização e representação do conhecimento: institucionalização como disciplina científica no âmbito da Ciência da Informação. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 16., 2015, João Pessoa. **Anais [...]**. João Pessoa: UFPB; ANCIB, 2015.
- NASCIMENTO, F. A.; GUIMARÃES, J. A. C. A contribuição da organização do conhecimento na representação da informação em contextos lgbt: interpelações acerca da linguagem. In: PINHO, F. A.; GUIMARÃES, J. A. C. (org.). **Memória, tecnologia e cultura na organização do conhecimento**. INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION (ISKO), 2017, Recife. **Anais [...]**. Recife: Ed. UFPE, 2017.
- ODDONE, N. Revisitando a "epistemologia social": esboço de uma ecologia sociotécnica do trabalho intelectual. **Ciência da Informação**, Brasília, v. 36, n. 1, p. 108-123, jan./abr. 2007.
- POPKEWITZ, Th. **Reforma Educacional: uma política sociológica – poder e conhecimento em educação**. Porto Alegre: Artes Médicas, 1997.
- SANTIS, R. de. **Sistemas de organização do conhecimento para domínios complexos: abordagem a canções populares na web semântica utilizando propriedades fuzzy**. 2016. 150 f. Tese (Doutorado em Ciência da Informação) - Universidade Federal do Rio de Janeiro (UFRJ); Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Rio de Janeiro, 2016.
- SANTOS, R. F. dos; NEVES, D. A. de B.; SOUZA, E. D. A organização do conhecimento como domínio de estudo da ciência da informação: uma reflexão a partir dos aspectos epistemológicos. In: BARROS, T. H. B.; TOGNOLI, N. B. (org.). **Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas**. Belém, 2019. **Anais [...]**. Belém: Ed. UFPA, 2019.
- SHERA, J. H. Epistemologia social, semântica geral e biblioteconomia. **Ciência da Informação**, Rio de Janeiro, v. 6, n. 1, jun. 1977.
- SHERA, J. H. **The foundations of education for librarianship**. New York: Becker-Hayes, 1972.
- SHERA, J. H. What lies ahead in classification. **Proceedings of the Allerton Park Institute**, p. 116–128, 1960.
- SHERA, Jesse. EGAN, Margaret E. Documentation in the United States. **American Documentation**, v. 1, n. 1, jan. 1950. p. 8-12
- SILVA, J. L. C. **Múltiplas interlocuções da informação no campo da Ciência da Informação no âmbito dos fundamentos técnico-pragmáticos, humanos e científicos**. 2014. 490 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal da Bahia, Salvador, 2014.
- SMIRAGLIA, R. P. **The elements of knowledge organization**. Springer, 2014.
- TAMAYO, A. M. M. Modelo para la toma de decisiones sobre los contenidos curriculares de la Organización del Conocimiento en programas académicos de Bibliotecología, Documentación o Ciencia de la Información. **Palabra Clave (La Plata)**, Buenos Aires, v. 3, n. 2, p. 71-90, abr. 2014.
- VERONEZ JÚNIOR, W. R. **Epistemologia social e organização do conhecimento: contribuições às abordagens culturais**. 2021. 189 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista (UNESP), Marília, 2021.
- VIEIRA, K. R.; LUCAS, E. R.; ARAÚJO, A. V. Jesse Shera: entre citações e bibliografia. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 22, n. 2, p. 208-226, 2017.
- WILSON, T. D. A dimensão epistemológica da ciência da informação e seu impacto sobre o ensino em arquivologia e biblioteconomia. **Brazilian Journal of Information Science**, [s.l.], v. 2, n. 1, p. 1-15, 2008.

Apêndice A – *Corpus de análise*

Tabela 1 - *Textos que compõem o corpus da análise*

<i>Ano</i>	<i>Autoria</i>	<i>Título</i>	<i>Publicação</i>
2013	Macedo	Ciência da Informação: uma abordagem para a transformação	Anais de evento
2014	Tamayo	Modelo para la toma de decisiones sobre los contenidos curriculares de la Organización del Conocimiento en programas académicos de Bibliotecología, Documentación o Ciencia de la Información	Artigo
2015	Martins e Moraes	Organização e representação do conhecimento: institucionalização como disciplina científica no âmbito da Ciência da Informação	Anais de Evento
2016	Santis	Sistemas de organização do conhecimento para domínios complexos: abordagem a canções populares na web semântica utilizando propriedades fuzzy	Tese
2017	Nascimento e Guimarães	A Contribuição da Organização do Conhecimento na Representação da Informação em Contextos LGBT: Interpelações acerca da Linguagem	Anais de Evento
2018	Martínez-Ávila	Hacia una base teórica social de la ciencia de la información	Artigo
2018	Bräsche e Guimarães	Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995	Artigo
2019	Santos, Neves e Souza	A Organização do Conhecimento como domínio de estudo da Ciência da Informação: uma reflexão a partir dos aspectos epistemológicos	Anais de Evento
2020	Martínez-Ávila e San Segundo	<i>The application of crowdsourcing and the Bazaar model to the development of library classifications: an assessment of the Open Shelves Classification</i>	Artigo
2020	Martínez-Ávila e Zandonade	<i>Social epistemology in information studies: a consolidation</i>	Artigo
2021	Veronez Júnior	Epistemologia Social e organização do conhecimento: contribuições às abordagens culturais	Dissertação
2021	Castanha, Veronez Junior e Dalessandro	As relações e influências da epistemologia social no Brasil: uma análise bibliométrica	Artigo
2021	Marcondes	Fundamentos da organização do conhecimento	Artigo
2021	Araújo	Novos desafios epistemológicos para a ciência da informação	Artigo
2021	Araújo, Guimarães e Tennis	A concepção de epistemologia da organização do conhecimento	Artigo
2023	Martínez-Ávila e Mello	Epistemologias, gênero e dogmatismo científico desdobramentos na organização do conhecimento	Artigo

Fonte: Dados da pesquisa (2023).

A Organização do Conhecimento na Biblioteconomia: uma perspectiva por meio da análise da Série Estudos Avançados em Organização do Conhecimento da ISKO

The Knowledge Organization in Undergraduate Courses in Librarianship: a Perspective through the analysis of the ISKO Advanced Studies in Knowledge Organization Series

Marta Leandro da Mata (1), Marcos Antonio de Moraes (2), Paula Regina Ventura Amorim Gonçalves (3)

(1) Universidade Federal do Espírito Santo (UFES), martaleandromata@gmail.com, (2) Universidade Estadual de Londrina (UEL), marcosmoraes@uel.br, (3) Universidade Federal do Espírito Santo (UFES), paulaventuramorim@gmail.com

Resumo

Objetivo: Analisar como as temáticas que envolvem disciplinas relacionadas à Organização e Representação do Conhecimento em Biblioteconomia são abordadas na Série Estudos Avançados em Organização do Conhecimento, a partir dos eixos temático da ISKO Brasil. **Metodologia:** Caracteriza-se como uma pesquisa de cunho exploratório e bibliográfico, com abordagem qualitativa. Para a coleta e análise de dados utilizou-se a análise de conteúdo de Bardin. **Resultados:** A exploração do material consistiu em uma análise prévia de todos os trabalhos da Série por meio dos títulos, resumos e palavras-chave de cada volume. Logo em seguida, foram selecionados aqueles que contemplavam a temática em questão, totalizando-se nove trabalhos. **Conclusão:** Foi evidenciado aspectos teóricos e práticos da organização nos cursos de Biblioteconomia, ou seja, os três eixos temáticos do evento da ISKO, demonstram questões ligadas às disciplinas e/ou temas pertinentes à formação do bibliotecário, oportunizando reflexões e contribuições para os cursos de Biblioteconomia no que se refere ao vieses teóricos e práticos na Organização do Conhecimento.

Palavras-chave: Organização do Conhecimento. Sistemas de Organização do Conhecimento. Biblioteconomia - Ensino.

Abstract

Objective: analyze how the themes involving disciplines related to the Organization and Representation of Knowledge in Librarianship are approached in the Series Advanced Studies in Knowledge Organization, based on the thematic axes of ISKO Brasil. **Methodology:** It is characterized as an exploratory and bibliographic research, with a qualitative approach. For data collection and analysis, Bardin's content analysis was used. **Results:** The exploration of the material consisted of a prior analysis of all the works in the Series through the titles, abstracts and keywords of each volume. Soon after, those that contemplated the theme in question were selected, totaling nine works. **Conclusion:** Theoretical and practical aspects of the organization in Librarianship courses were evidenced, that is, the three thematic axes of the ISKO event, demonstrate issues related to disciplines and/or themes relevant to the formation of librarians, providing opportunities for reflections and contributions to the courses of Librarianship with regard to theoretical and practical biases in the Organization of Knowledge.

Keywords: Knowledge Organization. Knowledge Organization Systems. Librarianship - Teaching.

Resumen

Objetivo: Analizar cómo los temas que involucran disciplinas relacionadas con la Organización y Representación del Conocimiento en Biblioteconomía son abordados en la Serie Estudios Avanzados en Organización del Conocimiento, a partir de los ejes temáticos de ISKO Brasil. **Metodología:** Se caracteriza por ser una investigación exploratoria y bibliográfica, con enfoque cualitativo. Para la recolección y análisis de datos se utilizó el análisis de contenido de Bardin. **Resultados:** La exploración del material consistió en un análisis previo de todas las obras de la Serie a través de los títulos, resúmenes y palabras clave de cada volumen. Poco después, fueron seleccionadas las que contemplaban el tema en cuestión, totalizando nueve obras. **Conclusión:** Se evidenciaron aspectos teóricos y prácticos de la organización en cursos de Biblioteconomía, o sea, los tres ejes temáticos del evento ISKO, evidencian cuestiones relacionadas con disciplinas y/o temas relevantes para la formación de bibliotecarios, brindando espacios de reflexión y aportes a los cursos de Biblioteconomía en lo que respecta a los sesgos teóricos y prácticos en la Organización del Conocimiento.

Palabras clave: Organización del conocimiento. Sistemas de Organización del Conocimiento. Biblioteconomía - Docencia.

1 Introdução

Os cursos de Biblioteconomia dispõem de disciplinas de formação específicas na área de Organização e Representação da Informação e do Conhecimento, bem como apresentam abordagens com foco nos Sistemas de Organização do Conhecimento, visto que, essa área é fundamental para a formação e a atuação do profissional bibliotecário.

As Diretrizes Curriculares para os cursos de Biblioteconomia (BRASIL, 2001, p. 33) sinalizam as habilidades e as competências desse profissional, no que diz respeito à: “Processar a informação registrada em diferentes tipos de suporte, mediante a aplicação de conhecimentos teóricos e práticos de coleta, processamento, armazenamento e difusão da informação”. Isto é, aquelas ligadas aos processos da Organização e Representação do Conhecimento.

Nesse sentido, acredita-se relevante compreender como tem sido abordada as questões relacionadas à formação do bibliotecário no âmbito da International Society for Knowledge Organization (ISKO), considerando sua representatividade na área de Ciência da Informação (CI).

Compreender as discussões teóricas e práticas das disciplinas de Organização e Representação do Conhecimento, bem como questões relacionadas à formação e à atuação do bibliotecário nos eixos temáticos da ISKO pode reverberar efetivamente na concepção dos projetos políticos pedagógicos dos cursos de Biblioteconomia.

Neste sentido, este trabalho teve como objetivo responder a seguinte questão: Como as temáticas que envolvem disciplinas relacionadas à Organização e Representação do Conhecimento em Biblioteconomia são abordadas na Série Estudos Avançados em Organização do Conhecimento, a partir dos eixos temático da ISKO Brasil?

2 Desenvolvimento

Organizar e relacionar, armazenar, recuperar e preservar é uma atividade inerente à existência humana, nas suas mais diversas atividades. No que se refere aos registros do conhecimento, entende-se que organizar é uma das mais importantes atividades para o desenvolvimento das sociedades ao longo da história. É por meio da organização que os saberes produzidos tornam-se mais úteis à apropriação e geração de novos conhecimentos (SOUZA, 1998). Em um sentido mais estrito, organizar o conhecimento proporciona a administração e difusão da informação em ambientes informacionais tais quais: arquivos, bibliotecas, museus, centros de informação e de documentação, entre outros (MIRANDA, 1999). No contexto de tal relevância, a

Organização do Conhecimento (OC) configura-se como um campo teórico e prático fundamental.

O campo da Organização do Conhecimento é considerado por Hjørland (2016), um campo tanto de pesquisa, quanto de ensino e prática, que se desenvolve sobretudo na área da Ciência da Informação (CI), servindo, em sua perspectiva teórico-metodológica, com afazeres dos profissionais da informação, tais como bibliotecários, entre outros profissionais da informação e correlatos.

Brascher e Carlan (2010, p. 150), entendem a OC como: “[...] um processo de modelagem que visa construir representações do conhecimento”. Dentro dessa perspectiva de processo, entende-se que a OC compreende as atividades de:

“[...] descrever, representar, arquivar e organizar documentos e representações de documentos, bem como assuntos e conceitos tanto por humanos quanto por programas de computador” (HJØRLAND, 2016, p. 1).

Quanto aos objetivos práticos da OC, deve-se levar em consideração os objetivos em comum de um determinado grupo, social o qual será atendido pelo já mencionado processo de modelagem de representações do conhecimento. Assim, é necessário desenvolver essa representação para posterior organização, de acordo com as metas e valores dos usuários, o que resultará em Sistemas de Organização do Conhecimento (SOCs), adequados ao contexto de utilização (HJØRLAND, 2008).

Os SOCs são produtos do processo de modelagem mencionados por Brascher e Carlan (2010) e Hjørland (2016), e compreendem os instrumentos utilizados principalmente pelos profissionais bibliotecários para organizar a informação e promover a gestão do conhecimento. Entre os tipos distintos de SOCs, estão os tesouros, taxonomias e ontologias. Nesse sentido, os SOCs:

[...] incluem esquemas de classificação e categorização que organizam materiais em um nível geral, cabeçalhos de assunto que fornecem acesso mais detalhado e arquivos de autoridade que controlam versões variantes de informações-chave, tais como nomes geográficos e nomes pessoais (HODGE, 2000, p. 1).

Dahlberg (2006), aponta que os SOCs são utilizados dentro do contexto de sistemas de informação e tem como funções o mapeamento de objetos informacionais e a representação temática de documentos existentes em acervos informacionais. Assim, entende-se que os SOCs são utilizados por bibliotecários nas atividades relacionadas a classificação e a indexação de documentos, processos intrinsecamente ligados a organização do Conhecimento.

Enquanto prática de ambientes informacionais, a Organização do Conhecimento é facilmente percebida na história das bibliotecas e da Biblioteconomia. (MEY,

2004). Desse modo, o ensino de processos e práticas de organização do conhecimento sempre fizeram parte importante da grade curricular de cursos de Biblioteconomia

O primeiro curso de Biblioteconomia do Brasil foi implantado no ano de 1911, na cidade do Rio de Janeiro, pela Biblioteca Nacional. De forte cunho humanístico, o primeiro curso da área foi muito influenciado pela escola francesa de Biblioteconomia, que tinha como grande expoente a École Nationale des Chartes. Somente 17 anos depois, surgia o segundo curso de Biblioteconomia do país, na cidade de São Paulo, 1929; este com maior influência da escola estadunidense, de caráter mais tecnicista em relação à escola francesa (CASTRO, 2000).

Mais recentemente, Dalessandro (2018) traça um panorama do ensino de disciplinas concernentes à OC nos cursos de Biblioteconomia no Brasil, apontando que até 2018, todos os cursos de graduação da área mantinham em suas grades curriculares disciplinas relacionadas à OC. Essas disciplinas concentram-se sobretudo nos eixos de classificação e indexação, tendo a sua nomenclatura variada, tais como: Fundamentos da Organização do Conhecimento, Representação Temática, Linguagens Documentárias, Sistemas de Organização do Conhecimento, Indexação e Resumos, Linguagens de Indexação entre outras nomenclaturas.

3 Procedimentos metodológicos

A presente investigação caracteriza-se como uma pesquisa de cunho exploratório e bibliográfico, com abordagem qualitativa em torno de questões que envolvem disciplinas e temáticas ligadas à Organização e Representação do Conhecimento em cursos de graduação em Biblioteconomia.

Para a coleta e análise de dados utilizou-se a análise de conteúdo de Bardin (2011), que é constituída por três fases, a saber: pré-análise, exploração do material e tratamento, interpretação e inferência dos resultados.

Na primeira fase, no mês de março de 2023, foi acessado o site da International Society for Knowledge Organization (ISKO) Brasil (<https://isko.org.br/>), recorrendo-se ao menu publicações, que contém a Série Estudos Avançados em Organização do Conhecimento, organizada em cinco volumes, baixando-se em seguida todos os e-books:

Quadro 1: *Série Estudos Avançados em Organização do Conhecimento*

<i>Vol.</i>	<i>Séries</i>	<i>Ano</i>
v. 5	Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas	2019
v.4	Memória, tecnologia e cultura na organização do conhecimento	2017
v. 3	Organização do conhecimento e diversidade cultural	2015
v. 2	Complexidade e organização do conhecimento, desafios de nosso século	2013
v. 1	Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade	2012

Fonte: Elaborado pelos autores com base na Série Avançados em Organização do Conhecimento,

A exploração do material consistiu em uma análise prévia de todos os trabalhos da Série por meio dos títulos, resumos e palavras-chave de cada volume. Logo em seguida, foram selecionados aqueles que contemplavam a temática em questão, ou seja, abordavam sobre disciplinas e temáticas ligadas à Organização e Representação do Conhecimento em cursos de graduação em Biblioteconomia. Nesse aspecto, foram encontrados nove trabalhos.

A segunda fase, consistiu em exploração aprofundada dos novem trabalhos selecionados. Para tanto, foi realizada a sistematização das informações contidas nos materiais e construído um quadro onde foram elencadas as informações dos trabalhos, categorizando-as a partir dos Eixos da ISKO Brasil, a saber: 1 - Dimensão Epistemológica da Organização do Conhecimento; 2 - Dimensão Aplicada da Organização do Conhecimento; e 3 - Dimensão Sociocultural da Organização do Conhecimento). Foram coletadas as seguintes informações: tema central do trabalho, título, resumo, palavras-chave e autores. Observou-se que no ano de 2012, 2013 e 2015 os trabalhos não apresentavam resumos e palavras-chaves, sendo que, nestes casos, foi realizada uma síntese e a inserção de palavras-chave pelos próprios autores da presente pesquisa para a realização da análise.

Na terceira fase, os resultados foram interpretados com maior profundidade, a partir da categorização realizada na fase anterior.

4 Resultados

No Eixo 1, Dimensão Epistemológica da Organização do Conhecimento, concentraram-se três trabalhos, em 2012, 2015 e 2017, conforme Quadro 1.

Quadro 2: *Dimensão Epistemológica da Organização do Conhecimento*

<i>Título/Data</i>	<i>Autores</i>	<i>Palavras-chave</i>
Concepções de Docentes Brasileiros sobre o Ensino da Análise de Assunto (2017)	Tartarotti, R. C. D. Dal'Evedove, P. R. Fujita, M. S. L.	Análise de assunto; Indexação; Ensino de Biblioteconomia; Brasil. Abstract
Ensino em organização e representação do conhecimento: linguagens e normas (2015)	Santos, C. A. M. Jesus, E. A. Luca, J. R.	Ensino; Representação da Informação; Linguagens documentárias; esquemas: as taxonomias e ontologias; Sistemas de organização do conhecimento
Concepções de Análise Documental: Um estudo comparativo a partir dos universos acadêmicos brasileiro e espanhol em Ciência da Informação (2012)	Guimarães, J. A. C. Salles, R.	Análise documental; Ciência da Informação; Brasil; Espanha

Fonte: Elaborada pelos autores

O trabalho “Concepções de Docentes Brasileiros sobre o Ensino da Análise de Assunto” (GUIMARÃES; SALES (2017) focou na caracterização das abordagens conceituais adotadas e trabalhadas pela comunidade docente brasileira. Como resultados, observou-se que em sua maioria, eles possuem uniformidade entre as concepções teóricas dos docentes, tanto em relação às etapas da indexação quanto ao conceito de análise de assunto.

Na publicação “Ensino em organização e representação do conhecimento: linguagens e normas” (SANTOS; JESUS, LUCAS, 2015), constatou-se que a literatura da organização e representação do conhecimento apresenta uma lacuna na pesquisa sobre metodologias de ensino-aprendizagem das linguagens e esquemas de organização da informação. Os autores acreditam que seria necessário repensar o ensino das linguagens documentárias e demais sistemas de organização do conhecimento para os profissionais da informação.

O Eixo 2, Dimensão Aplicada da Organização do Conhecimento, teve uma quantidade maior em relação aos outros dois eixos, tratando de pesquisas com viés aplicado, isto é, quatro trabalhos.

Quadro 3: *Dimensão Aplicada da Organização do Conhecimento*

<i>Título/Data</i>	<i>Autores</i>	<i>Palavras-chave</i>
TemaTres: Uma análise acerca de suas funcionalidades (2019)	Ferreira, R. E. S. Cavati Sobrinho, H.	TemaTres; Sistemas de Organização do Conhecimento; Linguagens Documentárias; Tesouro.
Atuação bibliotecária em repositórios de dados de pesquisa: da perspectiva da entrega à perspectiva do acesso (2019)	Tartarotti, R. C. D. Dal'Evedove, P. R. Fujita, M. S. L.	Dados de pesquisa; Repositório de dados; Atuação bibliotecária. Biblioteconomia de dados. Bibliotecário de dados; Organização da informação
Mapas conceituais interativos como instrumento didático no campo da Organização da Informação e do Conhecimento (2015)	Lara, M. L. G. Previdello, G. Kobashi, N. Y.	Mapas conceituais interativos; Instrumento didático; Ensino na área de Organização da Informação e do Conhecimento
Modelagem conceitual e sistemas de organização de conhecimento; uma reflexão baseada em experiências acadêmicas e profissionais no Brasil no domínio da Biblioteconomia e Ciência da Informação (2013)	Alvarenga, L.	Modelagem conceitual; Sistemas de organização do conhecimento; Biblioteconomia ; Ciência da Informação

Fonte: Elaborada pelos autores

O estudo “TemaTres: uma análise acerca de suas funcionalidades” (FERREIRA; CAVATI SOBRINHO, 2019) é centrado na ferramenta TemaTres como instrumento de criação e administração de Linguagens Documentárias (LDs) sob a perspectiva da sua utilização pelo profissional Bibliotecário. Para sua avaliação, utiliza microtesauros desenvolvidos por estudantes de graduação em Biblioteconomia, na

disciplina de Linguagens Documentárias, com “Modelo Metodológico Integrado para Construção de Tesauro” de Cervantes (2009). Constatou-se que a ferramenta fornece praticidade no desenvolvimento e na manutenção de LDs além de abranger diversas possibilidades de aplicação.

A pesquisa de Tartarotti, Dal’Evedove e Fujita (2019), “Atuação bibliotecária em repositórios de dados de pesquisa: da perspectiva da entrega à perspectiva do acesso”, destaca a Biblioteconomia de dados e o bibliotecário de dados. Salienta-se que os repositórios de dados de pesquisa trazem uma perspectiva inovadora à atuação do bibliotecário ao demandarem a observância de questões relacionadas à cobertura do sistema, metadados, acesso e qualidade dos dados, preservação digital, confiabilidade dos sistemas e interoperabilidade, os quais exigem o envolvimento de pesquisadores e profissionais na condução de pesquisas e ações dedicadas à gestão digital de dados de pesquisa e às perspectivas e desafios que se apresentam para a organização da informação.

Lara, Previdello e Kobashi (2015), com o trabalho “Mapas conceituais interativos como instrumento didático no campo da Organização da Informação e do Conhecimento”, apresentam um projeto para abordagem do ensino-aprendizagem de forma aberta, sistematizando conteúdos que dizem respeito às teorias, métodos e processos de construção da informação documentária, acessível em plataformas digitais. Neste trabalho, evidencia-se a necessidade de trazer para o aluno (e para o professor), um conjunto de referências de diferentes tipos e níveis de apresentação, de modo a complementar os estudos realizados presencialmente.

Alvarenga (2013), com a publicação “Modelagem conceitual e sistemas de organização de conhecimento: uma reflexão baseada em experiências acadêmicas e profissionais no Brasil no domínio da Biblioteconomia e Ciência da Informação”, traz reflexões sobre experiências prévias de criação e desenvolvimento de tesouros para sistemas de recuperação de informações mantidos por órgãos brasileiros, bem como no trabalho acadêmico e de pesquisa, realizado na Universidade Federal de Minas Gerais, UFMG, com alunos de graduação, mestrado e doutorado.

No Eixo 3, Dimensão Sociocultural da Organização do Conhecimento, obtiveram-se duas publicações um pouco mais recentes, em 2017 e 2019, sinalizando uma preocupação com os aspectos éticos e com os Sistemas de Organização do Conhecimento.

Quadro 4: *Dimensão Sociocultural da Organização do Conhecimento*

Título/Data	Autores	Palavras-chave
Valores éticos em Organização do Conhecimento: uma análise a partir dos estudantes de Arquivologia e Biblioteconomia da UNESP-Marília (2019)	Silva, A. P. Evangelista, V. E. Guimarães, A. C.	Ética; Organização do Conhecimento; Valores éticos; Arquivologia; Biblioteconomia
Problemas Relacionados a Biases em Sistemas de Organização do Conhecimento: Perspectivas para a Representação de Assunto (2017)	Milani, S. O. Guimarães, A. C.	Representação de assunto; Sistemas de organização do conhecimento; Biases; Dilemas éticos.

Fonte: Elaborada pelos autores.

Silva, Evangelista e Guimarães (2019), na investigação “Valores éticos em Organização do Conhecimento: uma análise a partir dos estudantes de Arquivologia e Biblioteconomia da UNESP-Marília”, questionou a postura dos estudantes de Arquivologia e Biblioteconomia dos cursos da Universidade Estadual Paulista - UNESP, enquanto futuros profissionais da área, se possuem questões relacionadas aos valores éticos referentes às atividades de Classificação, Descrição e Indexação. Constatou-se entre os participantes, que estes possuem um perfil profissional dinâmico, voltado ao acesso à informação sem barreiras, informação essa representada de maneira fidedigna por um especialista que busca sempre atualização de conhecimentos e compartilhamento de experiências.

Milani e Guimarães (2017), na pesquisa “Problemas Relacionados a Biases em Sistemas de Organização do Conhecimento: perspectivas para a representação de assunto”, por intermédio de um estudo de caráter exploratório e aplicado, apresenta exemplos para discussão sobre como sistemas de organização do conhecimento podem fomentar dilemas éticos na atuação do bibliotecário, enfocando em disciplinas relacionadas à organização do conhecimento, assim como em outros espaços abertos a reflexões éticas.

Complementando as análises apresentadas anteriormente, as palavras-chave foram agrupadas e cotejadas, gerando assim, uma nuvem das palavras-chave dos trabalhos analisados, visando demonstrar as perspectivas teóricas e práticas da ISKO no que se refere às temáticas ligadas à Organização e

Representação do Conhecimento em cursos de graduação em Biblioteconomia.

Figura 1: Nuvem a partir das palavras-chave



Fonte: Dados da pesquisa.

A partir da imagem é possível observar que a palavra-chave mais recorrente é sistema de organização do conhecimento, que apareceu quatro vezes, sendo duas vezes na dimensão aplicada, uma vez na dimensão epistemológica e uma na sociocultural, isto é, a única palavra que evidenciou-se nos três eixos da ISKO. Vale ressaltar que para Smiraglia (2002), os SOCs são, principalmente, ferramentas pragmáticas de organização, representação e recuperação do conhecimento.

Em segundo lugar, encontra-se duas vezes as seguintes palavras: linguagens documentárias, Biblioteconomia e Ciência da Informação, trazendo o foco para a área em que está essa pesquisa e suas inter-relações. Todas as outras palavras apresentadas na nuvem de palavras tiveram uma ocorrência cada uma delas, indicando os enfoques das pesquisas até o ano de 2019.

5 Conclusão

A partir dos resultados apresentados neste estudo foi possível analisar as temáticas que envolvem disciplinas relacionadas à Organização e Representação do Conhecimento publicados na Série Estudos Avançados em Organização do Conhecimento e correlacioná-las aos eixos temáticos da ISKO Brasil.

Os resultados permitiram verificar que os trabalhos publicados acerca da temática investigada nesta pesquisa estão distribuídos nos três eixos temáticos da ISKO, isto é, 1 - Dimensão Epistemológica da Organização do Conhecimento; 2 - Dimensão Aplicada da Organização do Conhecimento; e 3 - Dimensão Sociocultural da Organização do Conhecimento.

No eixo 1, Dimensão Epistemológica da Organização do Conhecimento, há trabalhos ligados aos docentes dos cursos de Biblioteconomia no que se refere ao ensino da análise de assunto, ao ensino em organização e representação do conhecimento e às concepções de Análise Documental no âmbito brasileiro e espanhol.

No eixo 2, Dimensão Aplicada da Organização do Conhecimento, em que ocorreu o maior número de trabalhos, há trabalhos que focam em ferramentas, mais especificamente, TemaTres, na atuação bibliotecária em repositórios de dados de pesquisa, recorrendo-se à formação do profissional bibliotecário, no mapa conceitual como instrumento didático na Organização da Informação e do Conhecimento e na modelagem conceitual e sistemas de organização de conhecimento; uma reflexão baseada em experiências acadêmicas.

No eixo 3, Dimensão Sociocultural da Organização do Conhecimento, o enfoque foi nos valores éticos em Organização do Conhecimento e problemas relacionados a biases em sistemas de organização do conhecimento.

De modo geral, pode-se inferir que a Série Estudos Avançados em Organização do Conhecimento abrange trabalhos nos três eixos temáticos do evento, mostrando que tanto questões ligadas às disciplinas e/ou temas ligados à formação do bibliotecário estão inclusas nos três eixos temáticos da ISKO Brasil, trazendo reflexões e contribuições para os cursos de Biblioteconomia no que se refere aos vieses teóricos e práticos na Organização do Conhecimento.

Referências

BRASCHER, Marisa; CARLAN, Eliana. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO, Jaime; BRASCHER, Marisa.(org.). **Passeios pelo bosque da informação**: estudos sobre representação e organização da informação e do conhecimento. Brasília: IBICT, 2010. p. 147-176.

BRASIL. Ministério da Educação. **Parecer CNE/CES 492/2001**: [Diretrizes Curriculares Nacionais dos cursos de Filosofia, História, Geografia, Serviço Social, Comunicação Social, Ciências Sociais, Letras, Biblioteconomia, Arquivologia e Museologia]. 2001. Disponível em: <http://portal.mec.gov.br/cne/arquivos/pdf/CES0492.pdf>. Acesso em: 12 mar. 2023.

CASTRO, C.A. **História da biblioteconomia brasileira**: perspectiva histórica. Brasília: Thesaurus, 2000.

CERVANTES, Brígida Maria Nogueira. **A construção de tesouros com a integração de procedimentos terminográficos**. 2009. 209f. Tese (Doutorado) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009.

DAHLBERG, Ingetraut. Knowledge organization: a new science? **Knowledge Organization**, [s. l.], v. 33, n. 4, p. 11-19, 2006.

DALESSANDRO, Rafael Cacciolari. **A Questão ética no ensino de Organização do Conhecimento**: uma análise de currículos de cursos de graduação em Arquivologia e Biblioteconomia no Brasil. 2018. 60f. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência da Informação – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília, 2018.

HJØRLAND, Birger. Deliberate bias in knowledge organization. **Advances in Knowledge Organization**, [s. l.], v. 11, p. 256-261, 2008.

HJØRLAND, Birger. Knowledge organization (KO). **KO Knowledge Organization**, [s. l.], v. 43, n. 6, p. 475-484, 2016.

HODGE, G. **Systems of knowledge organization for digital libraries**: beyond traditional authority files. Washington: Council on Library and Information Resources, 2000.

MEY, Eliane Serrão Alves. Bibliotheca Alexandrina. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 2, n. 1, p. 71-91, 2004.

MIRANDA, M.L.C. A organização do conhecimento e seus paradigmas científicos: algumas questões epistemológicas. **Informare: Cadernos do Programa de Pós-Graduação em Ciência da Informação**, Rio de Janeiro, v. 5, n. 2, p. 64-77, 1999.

MISCHIATI, A.C; VALENTIM, M.L.P. Reflexões sobre a ética e a atuação profissional do bibliotecário. **Transinformação**, Campinas, v. 17, n. 3, p. 209-220, 2005.

SMIRAGLIA, R. P. The progress of theory in knowledge organization. **Library Trends**, v. 50, n. 3, p. 330-349, 2002. Disponível em: <https://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=2AC74C876DB8CF8EE3BF799DEA9A3C4A?doi=10.1.1.892.1683&rep=rep1&type=pdf>. Acesso em: 05 mar. 2023.

The purpose of epistemology of Knowledge Organization

O propósito da Epistemologia da Organização do Conhecimento

Paula Carina de Araújo (1)

(1) Universidade Federal do Paraná, paulacarina@ufpr.br

Resumo

Objetiva investigar os propósitos da epistemologia da organização do conhecimento. Desenvolve um estudo descritivo uma vez que analisamos e descrevemos os dados buscando apresentar os propósitos da epistemologia da organização do conhecimento. Estudo desenvolvido a partir de uma perspectiva hermenêutica. O corpus do estudo é formado por 46 artigos publicados na revista *Knowledge Organization* no período de 1993 a 2022. O estudo discute onze propósitos da epistemologia do domínio da organização do conhecimento. Os onze propósitos apresentados completam estudos posteriores que investigaram o conceito de organização do conhecimento e sua concepção. A epistemologia orienta o processo de organização do conhecimento, pois destaca o nível conceitual, também deixa claro as posturas epistêmicas que influenciam a organização do conhecimento, é essencial para o projeto e implementação de sistemas de organização do conhecimento e ajuda a aplicar a análise de domínio, pois fornece uma visão sobre os pressupostos de teorias e é uma das onze abordagens de análise de domínio.

Palavras-chave: Epistemologia. Organização do conhecimento. Posições epistemológicas.

Abstract

This study aims to investigate the purposes of epistemology of knowledge organization. It develops a descriptive study since we analyze and describe the data seeking to present the purposes of the epistemology of KNOWLEDGE ORGANIZATION. The study is developed from a hermeneutic perspective. The first part of the research was the collection of data. The corpus of this study is formed by 46 papers published in the journal *Knowledge Organization* from 1993 to 2022. The study discusses eleven purposes of epistemology of knowledge organization domain. The eleven purposes presented complete further studies that investigated the concept of KNOWLEDGE ORGANIZATION and its conception. Epistemology guides the knowledge organization process since it highlights the conceptual level, it also makes clear the epistemic stances influencing knowledge organization, it is essential to the design and implementation of knowledge organization systems, and helps to apply domain analysis since it provides insight into the assumptions of theories and it is one of the eleven approaches of domain analysis.

Keywords: Epistemology. Knowledge organization. Epistemic stance.

Resumen

Este estudio tiene como objetivo investigar los propósitos de la epistemología de la organización del conocimiento. Desarrolla un estudio descriptivo ya que analizamos y describimos los datos buscando presentar los propósitos de la epistemología del organización del conocimiento. El estudio se desarrolla desde una perspectiva hermenéutica. El corpus de este estudio está formado por 46 artículos publicados en la revista *Knowledge Organization* desde 1993 hasta 2022. El estudio discute once propósitos de la epistemología del dominio de la organización del conocimiento. Los once propósitos presentados completan estudios posteriores que investigaron el concepto de organización del conocimiento y su concepción. La epistemología guía el proceso de organización del conocimiento ya que destaca el nivel conceptual, también aclara las posturas epistémicas que influyen en la organización del conocimiento, es esencial para el diseño e implementación de sistemas de organización del conocimiento y ayuda a aplicar el análisis de dominio ya que proporciona información sobre los supuestos, de teorías y es uno de los once enfoques del análisis de dominio.

Palabras clave: Palavras-chave traduzidas para o inglês

1 Introduction

Epistemology is the critical study of the principles, hypotheses and knowledge production of the various sciences. Furthermore, the scientific knowledge features, delimitations and methodological process in each domain are of epistemology interest. In the context of the knowledge organization domain,

"epistemology is an important part of the KNOWLEDGE ORGANIZATION armature because it reflects our assumptions about language, the primary material of Knowledge Organization Systems (KOS)". (TENNIS, 2008).

Knowledge organization seeks to offer a conceptual framework to the different practices and social activities regarded to the access to

knowledge. Knowledge organization also “intends to be an instrument for representing, managing and using information, as well as integrating the applications related to the structure, layout, access and dissemination of knowledge in society” (BARITÉ; 2001, p. 39-40, our translation).

Sales (2015a, 2015b, 2016) acknowledges that there are at least two perspectives regarded to the nature of KNOWLEDGE ORGANIZATION: KNOWLEDGE ORGANIZATION as an activity of operational nature and as a field of study that seeks, its theoretical, methodological and practical development.

The author also presents a deeper understanding of the concept regarded to IS domain. He approaches three different perspectives of the concept of KNOWLEDGE ORGANIZATION: - KNOWLEDGE ORGANIZATION as an independent scientific discipline and a subfield of a Science of Science, the same way it is stated by Dahlberg (1993, 1995, 2006); - KNOWLEDGE ORGANIZATION uses, sometimes, subjects from IS to apply to KNOWLEDGE ORGANIZATION, which is regarded to Hjørland’s thoughts; - KNOWLEDGE ORGANIZATION as a research field that deals with the development of theory, methods and, practices that connects the context of production and use of information. It considers KNOWLEDGE ORGANIZATION as a part of IS (SALES, 2015a, 2015b, 2016).

Furthermore, under a more pragmatic perspective, knowledge organization in information science is related to the processes, tools and products related to a certain knowledge that has been recorded and became socially available in order to guarantee its portability in space and its permanence in time. Such conception much more grounded in the social epistemology, which conceives knowledge as a social product under a continuous helical movement where knowledge organization is a mediating stance between the produced knowledge and the production of another knowledge (Egan & Shera, 1952; Shera, 1961; Budd, 2002; Smit & Barreto, 2002; Guimarães, 2008; Martinez-Avila & Zandonade, 2020;).

Considering the concept of epistemology and knowledge organization and its connection, this study aims to respond what are the purposes of epistemology of knowledge organization? That question and its answer is part of a broader research on the conception of epistemology of knowledge organization.

We acknowledge the conception of epistemology on the KNOWLEDGE ORGANIZATION domain as the critical study of principles, hypothesis and knowledge production in the domain. Epistemology is concerned with the scientific knowledge produced by the KNOWLEDGE ORGANIZATION domain, as well as with the application of that knowledge to design the

KOSs and to support the KNOWLEDGE ORGANIZATION process (ARAÚJO; GUIMARÃES; TENNIS, 2021).

The aim of this study is to investigate the purposes of epistemology of knowledge organization in these domain. Therefore we present the methods used to develop the study and the main results from the analysis followed by the final remarks and references.

2 Process of study

It develops a descriptive study since we analyze and describe the data seeking to present the purposes of the epistemology of KNOWLEDGE ORGANIZATION.

We study that subject from a hermeneutic perspective (ALVESSON; SKÖLDBERG, 2009; TENNIS, 2005a). “Hermeneutics as the methodology of interpretation is concerned with problems that arise when dealing with meaningful human actions and the products of such actions, most importantly texts” (MANTZAVINOS, 2016, p. 1).

The first part of the research was the collection of data. The corpus analysed for the purpose of this study is formed by 46 papers published in the journal Knowledge Organization from 1993 to 2022. The following string was used to retrieve the papers from the database Web of Science using the field topic: (epistemolog* OR “theor* of knowledge”) AND (“information organization” OR “knowledge organization”). The filter publication was applied to retrieve only papers published in the journal Knowledge Organization.

We developed the analysis in the second part of the research and we borrowed from grounded theory two systematic procedures in order to identify, analyze and describe the purposes of epistemology of knowledge organization.

3 The purpose of knowledge organization

“Science is the mode of production of knowledge that is subject to conditions of obtention ruled by epistemology, and epistemology is the nuclear discipline in the construction of scientific knowledge” (SAN SEGUNDO MANUEL; MARTÍNEZ-ÁVILA, 2014, p. 96)

Epistemological theories provide us with the most generalizable mental models (HJØRLAND, 2002a). Smiraglia (2013, p. 3) states that “epistemology is an essential tool of knowledge organization and a dimension is an expression of the extent of a space”. He adds that “in knowledge organization, epistemology represents one dimension, because it is how we can measure or express the space within our domain

ranging from the empirical to the rational, the two primary epistemological stances”.

Discussing epistemology and KNOWLEDGE ORGANIZATION as complementary domains, Smiraglia argues:

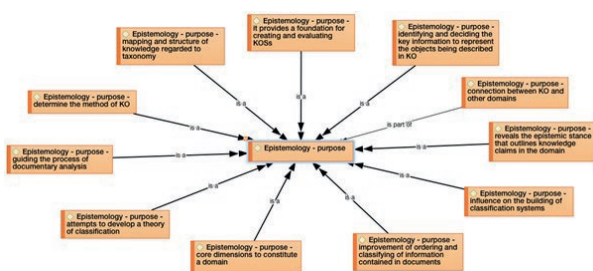
Epistemology is the science of knowledge. Knowledge organization is the science of the order of knowledge. The domain of knowledge organization is a discourse community in which inquiry takes place concerning knowledge, and its various orderings or sequences. The products of the domain are ordered segments of knowledge, and the rules either for discovering their natural orders, or the rules for imposing a useful sequence (SMIRAGLIA, 2013, p. 2).

Following the same thought, Mai (1999, p. 547) considers that “any theory of knowledge organization must further involve considerations regarding the epistemological basis of the theory and regarding the practical utilization of the theory”. Furthermore, “a knowledge organization cannot be epistemological neutral” and, because of that, “the epistemological basis should address the problems of how knowledge is generated and realized, the practical level should regard how knowledge is organized and represented”.

San Segundo Manuel and Martínez-Ávila (2014, p. 96) acknowledge that approaching KNOWLEDGE ORGANIZATION from the epistemological perspective, “one must take into consideration that this discipline deals with the foundations, criteria and validation by which scientific knowledge is justified, including the historical, political, economic, social, etc. circumstances”. Epistemology is understood as “the nuclear discipline in the construction of scientific knowledge although it is not capable to overcome its status of beliefs, interests and cultural objectives, and therefore it is not capable to require itself what it requires to other disciplines”.

Considering the arguments presented until here and the analysis of the corpus of this study, we may bring to light some purposes of epistemology of KNOWLEDGE ORGANIZATION that are represented in Figure 1 as an attribute family called “the purposes of epistemology of KNOWLEDGE ORGANIZATION”:

Figure 1. *The purposes of epistemology of Knowledge Organization*



Since Garcia-Marco and Esteban-Navarro (1993, p. 128) acknowledge that epistemology is “devoted to the study of the processes of human knowledge, its logic, origins and basis”, they state that “the study of epistemology is, therefore, essential for the design and implementation of better cognitive strategies for guiding the process of documentary analysis, particularly for indexing and abstracting scientific documents”.

Epistemology is also described as the “science centered on the study of the characteristics of scientific discourse and on the evolution of scientific paradigms” (GARCIA MARCO; ESTEBAN NAVARRO, 1993, p. 128). Therefore, it may influence the improvement of the ordering and classifying of information contained in documents, thus allowing their effective retrieval only, if it is possible to discover the conceptual framework (terms, concepts, categories, propositions, hypotheses, theories, patterns, and paradigms) of their authors from the discursive elements of texts (words, sentences and paragraphs) (GARCIA MARCO; ESTEBAN NAVARRO, 1993).

Hjørland (2018b) also recognizes that criteria for indexing are to be found in epistemological theories. The author adds that “an important function of meta-descriptions is to provide information about the nature of the object description. This means, first of all, epistemological and methodological descriptions”. That way we can state the important role of epistemology for the knowledge organization practice.

Hjørland and Hartel (2003) recognize the connection between ontological, epistemological and sociological dimension. If we take into account that various epistemological views influence the definition of a domain, its culture and practices, and its information forms, we may remember that epistemology is concerned about the evaluation of scientific paradigms as we discussed before. We know that each paradigm tends to develop, to some extent, its own terminology, theoretical view, methods and concepts. Specially, because “different epistemologies are connected to the different schools of thought that populate academe” (HJØRLAND; HARTEL, 2003, p. 242).

Furthermore, epistemology is concerned with the mapping and structure of knowledge, what is regarded to the study of the historical evolution of scientific paradigms. Garcia Marco and Esteban Navarro (1993, p. 128) consider that “it is crucial that emphasis be given to the analysis of scientific methodology and the classification strategies of nature through a branch of epistemology called taxonomy”.

One could ask why taxonomy is so relevant to IS research, the answer is because taxonomy is as aspect of documentary classification “because it shows the relationship of a scientific text to other texts and within

its scientific context”. More than that, taxonomy “helps to maintain universal classification systems, thesauri, and terminological databases according to the general evolution of science and of each particular scientific discipline” (GARCIA MARCO; ESTEBAN NAVARRO, 1993, p. 128).

Garcia Marco and Esteban Navarro (1993) add another role, or contribution from epistemology to information science and KNOWLEDGE ORGANIZATION, the development of a theory of classification. “This is because documentary classification systems are in close relationship with the two principal contemporary Western approaches to human knowledge - rationalism and logical positivism”. They believe the influence of these philosophical thoughts on the building of classification systems has not usually happened in a conscious way.

Philosophical theories of knowledge are usually a synthesis of the dominant characteristics present in a given historical period and proposals to explore new fields. As a consequence, the relationship between epistemology and information science is usually the result of the unconscious impregnation of information theorists with the principal epistemological approaches of their age and sometimes conscious efforts to adopt and adapt these ideas to the field of documentation (GARCIA MARCO; ESTEBAN NAVARRO, 1993, p. 129).

Hjørland and Hartel (2003, p. 240) explain that “different epistemologies have different views about the roles of, for example, observation, theoretical analysis, languages, traditions, sex and values in the production of knowledge”. They recognize its importance for library and IS, and acknowledge it is a central concern of domain analysis. They state that epistemology shapes information phenomena, for example, relevance. Hjørland (2002a) provides a simplified point of view of four epistemological schools regarded to “relevance criteria” and he explains what is relevant, non-relevant and low priority to empiricist, rationalist, historicist and pragmatist epistemologies.

Another role of epistemology of KNOWLEDGE ORGANIZATION is related to methods of KNOWLEDGE ORGANIZATION, which are methods of constructing systems of KNOWLEDGE ORGANIZATION such as classifications, thesauri, and processes like indexing and classifying. Every method of KNOWLEDGE ORGANIZATION (standardization, computer based KNOWLEDGE ORGANIZATION, quantitative methods, qualitative methods, text based methods, bibliometric methods, pragmatic, epistemological and critical methods, etc.) is connected to fundamental theories of epistemology (HJØRLAND, 2003). The two fundamental methods of classification (scientific classification and bibliographical classification) are analyzed through the eyes of

empiricist, rationalist, historicist and pragmatist epistemologies in Hjørland (2003).

Tennis (2008, p. 106) discusses epistemology, theory and methodology in the KNOWLEDGE ORGANIZATION domain, its connection and he states that they represent the driving force behind argument and findings in much of the conceptual work of KNOWLEDGE ORGANIZATION. His research also shows the role of epistemology regarded to methods of KNOWLEDGE ORGANIZATION. He states that epistemology shapes theory and methodology and that fact lies in the importance of labeling the kind of knowledge claims made through research narrative or research techniques”. Methodology is the combination of epistemic stance and the methods of investigation. It is also the machine used to create knowledge.

Echoing the same argument, Martínez-Ávila and Beak (2016) present methodological aspects as a fundamental part of Hope Olson’s papers since they were interested in epistemology. It is evident the connection between epistemology and methodology in the statement that epistemic assumption is linked to the method and the method is determined by the epistemology. Furthermore, Olson (1995) stated that “methodology develops from the researcher’s ontological and epistemological stance”. In this case, we refer specifically to research methods and not methods of construction KOSs.

As we move forward, we realize that one of the most important purposes of epistemology to the KNOWLEDGE ORGANIZATION domain is regarded to its nature of discovering the origin of knowledge. By observing different key subjects of KNOWLEDGE ORGANIZATION and considering the various epistemologies, we may present the similarities and differences among their methods, theories, thoughts and epistemic stances.

Thinking of that, Tennis (2008, p. 104) discusses that epistemology in the KNOWLEDGE ORGANIZATION domain “results in an epistemic stance that outlines knowledge claims. In the case of KNOWLEDGE ORGANIZATION we are concerned with assumptions about language, and how we can work with it in harmony with our conceptions of reality, how we know it, and what it means”. That is an important contribution from epistemology to the KNOWLEDGE ORGANIZATION domain, since we may understand the way that knowledge was created and applied to a specific domain. More than that, we understand the foundations and how the domain is constituted.

The same way, Hansson (2013) and Kleineberg (2013) discuss the connection between ontology and epistemology. Hansson explains that “knowledge organization systems normally presupposes some kind

of ontology, [...] they are thought to be reproducing some sort of structure, which refers to an equivalent in the world as such". Considering that he acknowledges "a very fundamental assumption; that epistemology is a sort of 'key' with which it is possible to unlock the ontological level of reality, whether natural, social, or spiritual" (2013, p. 385).

Kleineberg (2013) also presents an integrative approach based on a combination of ontology, epistemology, and methodology which he calls "constructive realism", a metatheoretical standpoint. Based on that view, knowledge is seen as both a human construction and, to some extent, a reflection of reality which is partially independent from human observers. It is proposed a parallel between a modernist (classification as ontology) and a postmodernist approach (classification as epistemology).

Ridi (2016) agrees with Kleineberg (2013) when he presents the synthesis of objectivism and subjectivism that consists in recognizing that reality is neither completely given nor completely built and that it constitutes beyond too much marked pseudo-oppositions, the paradigm today de facto dominant both in epistemology studies and in KNOWLEDGE ORGANIZATION studies. He states that his approach may share the same aspirations to the synthesis between the ontological and epistemological needs expressed by Kleineberg (2013).

In this scenario, epistemology influence a revision of the underlying concept of phenomenon, since "these epistemologically oriented theories (postmodernist approach) consider phenomena not merely as pre-given but as constructed by knowing subjects (the WHO of knowledge) [...]" (KLEINEBERG, 2013, p. 341). On the other hand, there is the claim on a primacy of ontology, which point of departure "is that reality itself underlies a certain structure which can be adopted as organizing principle for KOS's based on entities or phenomena" (KLEINEBERG, 2013, p. 342). Nevertheless, we recognize that both Hasson and Kleineberg seek the combination of ontology and epistemology "which would implicate a multidimensional knowledge concept" (KLEINEBERG, 2013, p. 352).

When Hjørland (2008) proposes to approach the fields contributing to KNOWLEDGE ORGANIZATION, he cites computer science, linguistics and natural language processing, theory of knowledge, theory of social organization etc. Hjørland (2008, p. 98) recognizes that "an understanding of the nature of knowledge, cognition, language and social organization is decisive for the understanding of KNOWLEDGE ORGANIZATION and thus for the ability to design, evaluate and use knowledge organizing processes and knowledge organizing systems.

Approaching the connection between linguistics and LIS, Hjørland (2008) acknowledges that epistemology is a deeper way to understand both fields. Furthermore, linguistics, LIS and KNOWLEDGE ORGANIZATION are influenced by changing epistemological views and interdisciplinary trends. We may state that epistemology has a role on the recognition of a satisfactory metatheory to better understand the relationship between KNOWLEDGE ORGANIZATION and other domains.

Seeking to understand the perspective of authorship from the notion of "own name," with the inflection on the philosophical discussion, Silveira and Saldanha's (2016) research also demonstrates an implicit role of epistemology in the KNOWLEDGE ORGANIZATION domain. Own name is understood as "the name of an individual objectively representing his work". Furthermore, "the "author" in a bibliographic record presupposes the meaning "own name," since it personalizes and organizes knowledge through standardized access points" (SILVEIRA; SALDANHA, 2016, p. 266).

We believe that the action of identifying authorship, establishing, and standardizing access points in a bibliographic record, is part of KNOWLEDGE ORGANIZATION and, sometimes, represents a subjective decision that is epistemologically influenced. In that manner, Silveira and Saldanha state that "the document being represented may provide data for information organization referring to authorship in a clearer manner than that of the organization of authorship knowledge, the latter being directly related to the philosophical issue of "own name" and knowledge organization" (2016, p. 270).

Domain analysis is a new approach to IS or library and information science (LIS). Hjørland and Hartel (2003, p. 239) suggest that three theories and concepts interact in the constitution of a domain: ontological, epistemological and sociological. The second dimension is "epistemological theories and concepts about knowledge and the ways to acquire knowledge, implying methodological principles about the ways objects are investigated". Hjørland (2017, p. 444) understands that "domain analysis without epistemological analyses tends to be superficial because epistemology provides insight into the assumptions of theories about user behavior".

Therefore, to Hjørland (2017) any ontology is based on epistemological assumptions. He considers that classification systems, for example, are best grounded in both ontology and epistemology. "Epistemology is of particular import to the field of KNOWLEDGE ORGANIZATION, a field which deals with the classification of existing knowledge accumulated over thousands of years of scientific inquiry" (IBEKWE-SANJUAN; BOWKER, 2017, p. 194). Epistemology

provides a foundation for creating and evaluating KOSs. This way, our observations and classifications are theory-laden.

The idea of describing things in the world in an atheoretical way is therefore naïve. All classification depends on the methodology used, which is again connected to epistemology (see further in Hjørland 2017b). Epistemology is therefore not just a requisite for discipline-based classifications, but also for phenomena-based classifications (HJØRLAND, 2017, p. 448).

Knowledge is created by humans for some specific purposes and serves some interests better than others. If we consider epistemology of KNOWLEDGE ORGANIZATION, concepts and semantic relations are not a priori or neutral, but should be examined in relation to their implications for the users they are meant to serve (HJØRLAND, 2013a).

Mazzocchi (2018, p. 66) argue that “semantics is strongly interlinked with epistemology. Our approach to KNOWLEDGE ORGANIZATION and in constructing KOSs is always more or less influenced by the epistemological stance adopted”. Furthermore, Włodarczyk (2020, p. 60) states that “[...] epistemological analyses can potentially provide the creators of KOSs with additional knowledge to broaden their view of a domain”. Hjørland (2020, p. 462) believes that “the quality of a given KOS also depends on how it has been constructed and, in the end, needs to include epistemological arguments”.

Hjørland (2018a, p. 686) seeks “to define “data” and to relate this concept to other core concepts in library and information science (LIS), including facts, information and documents”. In that matter, the author argues that “data differs among epistemological positions”. Hjørland (2018a, p. 694) discusses a specific epistemic stance, the social epistemology. The author argues “that the social epistemology originally introduced by Jesse H. Shera (1951) and since developed in the domain-analytic approach is the most fruitful approach” is the best theoretical frame for understanding data.

In summary, we may list the main purposes of epistemology of KNOWLEDGE ORGANIZATION:

- a) it is essential to the design and implementation of better cognitive strategies for guiding the process of documentary analysis (GARCIA MARCO; ESTEBAN NAVARRO, 1993);
- b) it influences the improvement of the ordering and classifying of information contained in documents through the discover of the conceptual framework of their authors from the discursive elements of texts (GARCIA MARCO; ESTEBAN NAVARRO, 1993; HJØRLAND, 2018b);
- c) it is concerned with the mapping and structure of knowledge regarded to taxonomy, an aspect of

documentary classification (GARCIA MARCO; ESTEBAN NAVARRO, 1993, p. 128);

- d) since its close relation to theory, epistemology has a role in the development of a theory of classification (GARCIA MARCO; ESTEBAN NAVARRO, 1993);
- e) every method of KNOWLEDGE ORGANIZATION is connected to fundamental theories of epistemology. Epistemic assumption is linked to the method and the method is determined by the epistemology (HJØRLAND, 2003; MARTÍNEZ-ÁVILA; BEAK, 2016; TENNIS, 2008);
- f) it influences the discovery of the origin of knowledge and reveals the epistemic stance that outlines knowledge claims in the domain (HJØRLAND, 2020);
- g) it influence a revision of the underlying concept of phenomenon, based on an integrative approach between the ontological and epistemological level (HANSSON, 2013; KLEINEBERG, 2013; HJØRLAND, 2017; RIDI, 2016; HJØRLAND, 2020);
- h) it recognizes a satisfactory metatheory to better understand the relationship between KNOWLEDGE ORGANIZATION and other domains (HJØRLAND, 2008);
- i) it has a role on identifying and deciding the key information to represent the objects being described in KNOWLEDGE ORGANIZATION, for example, when the cataloger chose the author access point of a document (own name) (SILVEIRA; SALDANHA, 2016; HJØRLAND, 2018b). It also has specific implications for standards, database semantics and in the different approaches of data in different domains (HJØRLAND, 2018a).
- j) it provides a foundation for creating and evaluating KOSs (HJØRLAND, 2017; MAZZOCCHI, 2018; HIDER, 2020; WŁODARCZYK, 2020; HJØRLAND, 2020);
- k) domain analysis without epistemological analysis tends to be superficial. So, epistemology is fundamental to domain analysis studies since it provides insight into the assumptions of theories and it is one of the eleven approaches of domain analysis. It is one of the core dimensions to constitute a domain (HJØRLAND, 2017; HJØRLAND; HARTEL, 2003).

We believe that to have a complete conception of epistemology in the KNOWLEDGE ORGANIZATION domain, we may describe the concept of epistemology of knowledge organization

(ARAÚJO; GUIMARÃES; TENNIS, 2021), the purposes of epistemology and the epistemic stances that influence the KNOWLEDGE ORGANIZATION domain. The eleven purposes presented in this section complete further studies that investigated the concept of KNOWLEDGE ORGANIZATION and its conception.

4 Conclusion

The study reviewed eleven purposes of epistemology of knowledge organization that influence the domain.

Epistemology guides the KNOWLEDGE ORGANIZATION process since it highlights the conceptual level, it also makes clear the epistemic stances influencing KNOWLEDGE ORGANIZATION, it is essential to the design and implementation of KOSs, and helps to apply domain analysis since it provides insight into the assumptions of theories and it is one of the eleven approaches of domain analysis.

Further research can focus on the same analysis of the proceedings of ISKO Conferences to create a broader understanding of the KNOWLEDGE ORGANIZATION domain.

References

- ALVESSON, M.; SKÖLDBERG, K. **Reflexive methodology**: new vistas for qualitative research. 2nd. ed. Los Angeles; London: SAGE, 2009.
- ARAÚJO, P. C. de; GUIMARÃES, J. A. C.; TENNIS, J. T. A concepção de epistemologia da organização do conhecimento. **Palavra Chave (Argentina)**, v. 10, n. 2, 2021. DOI: 10.24215/18539912e120 Acesso em: 12 abr. 2023.
- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K (Ed.). **Educación, universidad e pesquisa**. Marília: UNESP, 2001. p. 35–60.
- BUDD, John. Jesse Shera, social epistemology and praxis. **Social epistemology**, v.16, n.1, 2002. p.93-98
- DAHLBERG, I. Knowledge organization: its scope and possibilities. **Knowledge Organization**, v. 20, n. 4, p. 211–222, 1993.
- DAHLBERG, I. Current trends in knowledge organization. In: GARCIA MARCO, F. J (Ed.). **Organización del conocimiento en sistemas de información y documentación**. Zaragoza: Universidad de Zaragoza, 1995. p. 7–25.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, v. 33, n. 1, p. 11–19, 2006.
- EGAN, Margaret E.; SHERA, Jesse H. Foundations of a Theory of Bibliography. **The Library Quarterly: Information, Community, Policy**, v. 22, n. 2, p. 125–137, 1952. Disponível em: <http://www.jstor.org/stable/4304106>.
- GARCIA MARCO, F. J.; ESTEBAN NAVARRO, M. A. On some contributions of the cognitive science and epistemology to a theory of classification. **Knowledge Organization**, v. 20, n. 3, p. 126–132, 1993.
- GUIMARÃES, José Augusto Chaves. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-Americana de Ciência da Informação**, v. 1, n.1, p. 77-99, 2008.
- HANSSON, J. The materiality of knowledge organization: epistemology, metaphors and society. **Knowledge Organization**, v. 40, n. 6, p. 384–391, 2013.
- HARTEL, J. Social Epistemology as Theoretical Foundation for Information Science: Supporting a Cultural Turn, Copenhagen, August 16-17, 2017. **KNOWLEDGE ORGANIZATION**, v. 45, n. 1, p. 79–84, 2018. <https://doi.org/10.5771/0943-7444-2018-1-79>.
- HIDER, P. Towards a Sociology of KOS and More Basic KO Research. **KNOWLEDGE ORGANIZATION**, v. 47, n. 6, p. 501–510, 2020. <https://doi.org/10.5771/0943-7444-2020-6-501>.
- HJORLAND, B. Data (with Big Data and Database Semantics). **KNOWLEDGE ORGANIZATION**, v. 45, n. 8, p. 685–708, 2018a. <https://doi.org/10.5771/0943-7444-2018-8-685>.
- HJORLAND, B. Indexing: Concepts and Theory. **KNOWLEDGE ORGANIZATION**, v. 45, n. 7, p. 609–639, 2018b. <https://doi.org/10.5771/0943-7444-2018-7-609>.
- HJORLAND, B. Political Versus Apolitical Epistemologies in Knowledge Organization. **KNOWLEDGE ORGANIZATION**, v. 47, n. 6, p. 461–485, 2020. <https://doi.org/10.5771/0943-7444-2020-6-461>.
- HJØRLAND, B. Epistemology and the socio-cognitive perspective in information science. **Journal of the American Society for Information Science and Technology**, v. 53, n. 4, p. 257–270, 1 jan. 2002.
- HJØRLAND, B. Fundamentals of knowledge organization. **Knowledge Organization**, v. 30, n. 2, p. 87–111, 2003.
- HJØRLAND, B. What is knowledge organization (KO)? **Knowledge Organization**, v. 35, n. 2–3, p. 86–101, 2008.
- HJØRLAND, B. Theories of knowledge organization: theories of knowledge. **Knowledge organization**, v. 40, n. 3, p. 169–181, 2013.
- HJØRLAND, B. Domain analysis. **Knowledge Organization**, v. 44, n. 6, p. 436–464, nov. 2017.
- HJØRLAND, B.; HARTEL, J. Afterword: ontological, epistemological and sociological dimensions of domains. **Knowledge Organization**, v. 30, n. 3–4, p. 239–245, 2003.
- IBEKWE-SANJUAN, F.; BOWKER, G. C. Implications of big data for knowledge organization. **Knowledge Organization**, v. 44, n. 3, p. 187–198, 2017.
- KLEINEBERG, M. The blind men and the elephant: towards an organization of epistemic contexts. **Knowledge Organization**, v. 40, n. 5, p. 340–362, 2013.

- MAI, J.-E. A post-modern theory of knowledge organization. In: WOODS, L (Ed.). **Proceedings of the 62nd annual meeting of the American Society for Information Science**. Medford: Information Today, 1999. p. 547–556.
- MANTZAVINOS, C. Hermeneutics. In: ZALTA, E. N (ed.). **The Stanford Encyclopedia of Philosophy**. Winter 2016 ed. Stanford, CA: Metaphysics Research Lab, Stanford University, 2016.
- MARTÍNEZ-ÁVILA, D.; BEAK, J. Methods, theoretical frameworks and hope for knowledge organization. **Knowledge Organization**, v. 43, n. 5, p. 358–366, 2016.
- MARTINEZ-AVILA, Daniel; ZANDONADE, Tarcísio. Social epistemology in information studies: a consolidation. **Brazilian Journal of Information Science**, v.14, n.1, p.7-36, 2020.
- MAZZOCCHI, F. Knowledge Organization System (KOS): An Introductory Critical Account. **KNOWLEDGE ORGANIZATION**, v. 45, n. 1, p. 54–78, 2018. <https://doi.org/10.5771/0943-7444-2018-1-54>.
- OLSON, H. Quantitative versus qualitative research: the wrong question. In: CANADIAN ASSOCIATION FOR INFORMATION SCIENCE CONFERENCE, 1995, Edmonton. **Proceedings** [...], Edmonton: School of Library and Information Studies. University of Alberta, 1995.
- RIDI, R. Phenomena or noumena?: objective and subjective aspects in knowledge organization. **Knowledge Organization**, v. 43, n. 4, p. 239–253, 2016. \SALES, R. DE. A Relação entre organização do conhecimento e ciência da informação na comunidade científica brasileira: uma investigação no âmbito da ISKO Brasil. In: GUIMARÃES, J. A. C.; DODEBEI, V (eds.). **Organização do conhecimento e diversidade cultural**. Marília, SP: FUNDEP; ISKO-Brasil, 2015a. v. 1p. 73–84.
- SALES, R. de. A relação entre organização do conhecimento e ciência da informação na comunidade científica brasileira: uma investigação no âmbito da ISKO Brasil. In: GUIMARÃES, J. A. C.; DODEBEI, V (eds.). **Organização do conhecimento e diversidade cultural**. Marília, SP: FUNDEP; ISKO-Brasil, 2015a. v. 1p. 73–84.
- SALES, R. de. O diálogo entre a Organização do Conhecimento e a Ciência da Informação na comunidade científica da ISKO-Brasil. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16., 2015, João Pessoa. **Anais eletrônicos** [...]. João Pessoa, PB: UFPB, 2015b. Available in: <http://www.ufpb.br/evento/liti/ocs/index.php/enancib2015/enancib2015/paper/viewFile/2757/1003> . Access in: 10 maio 2017.
- SALES, R. de. Knowledge organization in the Brazilian scientific community and its epistemological intersection with information science. GUIMARÃES, J. A. C. ; DODEBEI, V (eds.). **Knowledge organization for a sustainable world: challenges and perspectives for cultural, scientific, and technological sharing in a connected society: proceedings of the fourteenth International ISKO Conference**. Würzburg: Ergon-Verlag, 2016.
- SAN SEGUNDO MANUEL, R.; MARTÍNEZ-ÁVILA, D. Digital as a hegemonic medium for epistemology and knowledge organization. In: INTERNATIONAL ISKO CONFERENCE, 13., Kraków, Poland. **Proceedings** [...]. Kraków, Poland: Wiesław Babik, 2014.
- SHERA, Jesse Hauk. “Social Epistemology, General Semantics and Librarianship”. **Wilson Library Bulletin**, v. 35, n. 10, 1961, pp. 767–770.
- SILVEIRA, N. C.; SALDANHA, G. S. “Own name” in knowledge organization epistemology: a philosophical-theoretical debate. **Knowledge Organization**, v. 43, n. 4, p. 265–278, 2016.
- SMIRAGLIA, R. P. The epistemological dimension of knowledge organization. **IRIS: Revista de Informação, Memória e Tecnologia**, v. 2, n. 1, p. 2–11, 2013.
- SMIT, Johanna. W.; BARRETO, Aldo de Albuquerque. Ciência da informação: base conceitual para a formação do profissional. In: VALENTIM, M.L. (Org.). **Formação do profissional da informação**. São Paulo: Polis, 2002. p. 9-23.
- TENNIS, J. T. **Conceptions of subject analysis: a metatheoretical investigation**. [s.l.] ProQuest Dissertations Publishing, 2005.
- TENNIS, J. T. Epistemology, theory, and methodology in knowledge organization: toward a classification, metatheory, and research framework. **Knowledge Organization**, v. 35, n. 2–3, p. 102–112, 2008.
- WLODARCZYK, B. KABA Subject Headings and the National Library of Poland Descriptors in Light of Wojciech Wrzosek’s Theory of Historiographical Metaphors and Different Historiographical Traditions. **KNOWLEDGE ORGANIZATION**, v. 47, n. 1, p. 56–71, 2020. <https://doi.org/10.5771/0943-7444-2020-1-56>

Referentes teóricos em indexação arquivística: estudo em artigos científicos brasileiros

Theoretical references in archival indexing: study in brazilian scientific articles

Graziela Martins de Medeiros (1), Camila Schwinden Lehmkuhl (2)

(1) Universidade Federal de Santa Catarina, Florianópolis, SC, Brasil, graziela.m@ufsc.br, (2) camila.lehmkuhl@ufsc.br

Resumo

Objetivo: A presente pesquisa objetiva levantar artigos científicos que tratam a respeito da indexação na Arquivologia visando identificar os referentes teóricos utilizados por autores brasileiros; **Metodologia:** utiliza análise exploratório-descritiva, qualitativa e quantitativa. Realiza busca nas fontes de informação BRAPCI, BENANCIB, ISKO-Brasil e ISKO-Iberico de artigos e anais de eventos publicados por autores brasileiros sobre o tema da indexação em Arquivologia; **Resultados:** Os resultados mostram que 14 publicações foram encontradas sobre o tema nas fontes pesquisadas. Dentre os autores citados pelo menos cinco vezes, Mariangela Fujita, Bibliotecária, é a autora mais citada em 7 artigos da análise e com 12 publicações diferentes. Os outros autores citados em sua maioria são dos Estados Unidos, Canadá e Europa totalizando 63,6%, enquanto os brasileiros somam 36,4% entre os mais citados; **Conclusão:** A pesquisa sobre indexação na Arquivologia no Brasil tem prevalecido a partir de 2010, as bases teóricas que estão sendo utilizadas no Brasil provêm em maioria de outros países, em especial da América do Norte e Europa, a análise dos autores mais citados mostra a interdisciplinaridade entre as áreas.

Palavras-chave: Representação da informação arquivística. Indexação arquivística. Indexação de assunto.

Abstract

Objective: The present research aims to raise scientific articles that deal with indexing in Archival Science in order to identify the theoretical references used by Brazilian authors; **Methodology:** uses exploratory-descriptive, qualitative and quantitative analysis. Searches the BRAPCI, BENANCIB, ISKO-Brasil and ISKO-Iberico information sources for articles and annals of events published by Brazilian authors on the topic of indexing in archival science; **Results:** The results show that 14 publications were found on the subject in the searched sources. Among the authors cited at least five times, Mariangela Fujita, Librarian, is the most cited author in 7 articles of the analysis and with 12 different publications. Most other cited authors are from the United States, Canada and Europe, totaling 63.6%, while Brazilians account for 36.4% among the most cited; **Conclusion:** Research on indexing in Archival Science in Brazil has prevailed since 2010, the theoretical bases that are being used in Brazil come from most other countries, especially from North America and Europe, the analysis of the most cited authors shows the interdisciplinarity between areas.

Keywords: Representation of archival information. Archival indexing. Subject indexing.

Resumen

Objetivo: La presente investigación tiene como objetivo levantar artículos científicos que traten sobre indexación en Archivística con el fin de identificar los referentes teóricos utilizados por autores brasileños; **Metodología:** utiliza análisis exploratorio-descriptivo, cualitativo y cuantitativo. Busca en las fuentes de información BRAPCI, BENANCIB, ISKO-Brasil e ISKO-Ibérico artículos y anales de eventos publicados por autores brasileños sobre el tema de la indexación en archivística; **Resultados:** Los resultados muestran que se encontraron 14 publicaciones sobre el tema en las fuentes buscadas. Entre los autores citados por lo menos cinco veces, Mariangela Fujita, Bibliotecaria, es la autora más citada en 7 artículos de análisis y con 12 publicaciones diferentes. La mayoría de los demás autores citados son de Estados Unidos, Canadá y Europa, totalizando el 63,6%, mientras que los brasileños representan el 36,4% entre los más citados; **Conclusión:** La investigación sobre indexación en Archivística en Brasil prevalece desde 2010, las bases teóricas que se están utilizando en Brasil provienen de la mayoría de los demás países, especialmente de América del Norte y Europa, el análisis de los autores más citados muestra la interdisciplinariedad entre las áreas.

Palabras clave: Representación de la información de archivo. Indexación de archivos. Indización de materias.

1 Introdução

A indexação é processo significativo para a recuperação e o acesso à informação e configura-se como um tema da agenda de pesquisa da Ciência da Informação e da Biblioteconomia. Entretanto, no campo arquivístico ainda carece de aprofundamento teórico sob a justificativa de que o documento de arquivo é organizado considerando aspectos contextuais e sua natureza orgânica, que o caracterizam enquanto objeto de estudo da Arquivologia.

Os processos de classificação e descrição arquivística são identificados na área como essenciais para representar o documento arquivístico. A classificação como aquela que refletirá a estrutura da produção documental da instituição e a descrição como aquela que descreve especificamente cada nível desta estrutura para possível recuperação da informação. Para tanto, parte-se do pressuposto da necessidade de representar o assunto e contribuir, assim, para otimizar o processo de recuperação de documentos de arquivo. Dado processo é identificado como indexação.

Diferentes autores nacionais e internacionais identificam a indexação como um processo a ser aplicado no campo dos arquivos, tais como: Lytle (1980), BCA (1992), Maurel e Champagne (1999), Hudon (1997-1998), Lévesque (2001-2002) e Ribeiro (1996, 2011).

Alguns autores indicam a utilização do processo de indexação no que chamam de “sistema de descrição e indexação”, tais como o BCA (1992) e Maurel e Champagne (1999). Ou seja, a descrição e indexação compartilham as mesmas funções, produtos e políticas. Autores como Martin (1987) e Ribeiro (2011) sinalizam a indexação como um processo a ser enfatizado no campo arquivístico e que merece ser aprofundado.

Ribeiro (2011) apresenta os fundamentos teóricos da indexação e as suas especificidades no contexto dos arquivos e afirma que “a questão da indexação nos arquivos tem sido muito pouco pensada e discutida e, por isso mesmo, a literatura sobre o assunto é escassa e com limitações evidentes.” (RIBEIRO, 2011, p. 40). Segundo a autora, existem três processos no cerne da organização arquivística: a classificação, a descrição e a indexação. Entretanto, a autora questiona a visão tradicional dos arquivos imputada no século XIX e que prevalece nos dias atuais. Relembra as modificações sofridas na área de Arquivologia com a visão pós-custodial, a qual pode influenciar o desenvolvimento desses processos. Assim, essas operações não devem ocorrer de forma meramente tecnicista:

Tradicionalmente, consideram-se três operações maiores – a classificação, a descrição e a indexação – no cerne da organização arquivística. Mas essa visão tradicional, própria do paradigma custodial, historicista e tecnicista que dominou a arquivística desde o século XIX e que ainda

prevalece sob perspectivas mais atuais, é questionada de modo diferente se nos situarmos num paradigma novo, apelidado de pós-custodial, científico e informacional, cujo enfoque se centra sobre a informação numa perspectiva dinâmica e não sobre o documento, enquanto objeto material estático e com valor patrimonial. Assim, as três operações técnicas acima referidas não se realizam de forma asséptica e meramente tecnicista, mas são integradas [...] (RIBEIRO, 2011, p. 46).

Concordamos com a autora no tocante à existência de três processos distintos, porém integrados. Nessa perspectiva, parte-se do pressuposto que a representação do documento arquivístico não ocorre apenas com base nas funções e atividades do produtor, mas também com base em seu conteúdo e assunto. São representados elementos contextuais, mas também de conteúdo e assunto dos documentos.

No Brasil, uma sinalização das interlocuções apresentadas entre a Arquivologia e a Indexação é apresentada na própria Norma Brasileira de Terminologia Arquivística (2006). Além disso, pesquisas como a de Lehmkuhl (2021) e Couture et al (2003) já abordam a função de descrição integrada à indexação.

Dessa maneira, este artigo busca levantar artigos científicos que tratam a respeito da indexação na Arquivologia visando identificar os referentes teóricos utilizados por autores brasileiros que publicaram sobre o tema.

2 A Indexação

A representação da informação pode ser de caráter descritivo ou temático, sendo o primeiro relacionado aos aspectos físicos do material e o segundo dedicado ao seu conteúdo, conforme asseveram diferentes autores da área, tais como Bradford (1961), Chaumier (1973), Foskett (1973), Cavalcanti (1978), Smit *et al.* (1987), Dias e Naves (2013), para citar alguns. De acordo com Dias e Naves (2013, p. 7):

[...] o tratamento descritivo cuida daqueles aspectos mais objetivos capazes de bem identificar, extrinsecamente, um documento: o autor, título, a editora, elementos similares. O tratamento temático, ao contrário, tem uma forte carga subjetiva, pois como o nome indica, visa caracterizar o documento do ponto de vista de seu conteúdo.

Enfatiza-se neste artigo o tratamento temático, também conhecido como Tratamento Temático da Informação (TTI). Há dois métodos indicados por Chaumier para o TTI: o resumo (representação condensada da informação) e a indexação, que visa “destacar os conceitos do texto do documento e exprimi-los com ajuda de termos da linguagem escolhida: palavras-chave, descritores ou índices de um sistema de classificação.” (CHAUMIER, 1973, p. 16).

A indexação pode ser definida como o “ato de identificar e descrever o conteúdo de um documento em termos representativos dos seus assuntos e que constituem uma linguagem de indexação (NBR 12676/1992)”. Esta definição está contemplada na Norma Brasileira de Indexação, mas representa o entendimento de diferentes autores da área. A indexação foi recentemente definida neste mesmo sentido, porém de forma mais completa:

[...] um processo formado por subprocessos ou etapas que tem como objetivo identificar o conteúdo de um documento, por meio de uma metalinguagem construída – a linguagem de indexação – com o intuito de promover a efetiva recuperação da informação. Em termos teóricos, qualquer recurso informacional é passível de ser indexado ou representado por conceitos. Esta ação viabiliza que a informação documental seja acessada e recuperada para usos distintos por sujeitos com interesses específicos. (TARTAROTTI; DAL'EVEDOVE; FUJITA, 2017, p. 2).

Percebe-se que diferentes autores que conceituam a indexação indicam o conteúdo como elemento a ser representado. Dessa maneira, vale lembrar que a indexação é um processo que pode ser realizado com base em diferentes elementos, já que o conteúdo abrange assuntos, nomes, lugares, ou seja, uma variedade de possibilidades.

Como esses aspectos são tratados por diferentes autores da literatura e podem apresentar diferentes entendimentos, apresenta-se a discussão com base em Smit *et al.* (1987), Pinto Molina (1993) e Lancaster (2004), dentre outros.

Em relação às etapas do tratamento temático, Dias e Naves (2013) indicam três: análise de assunto, seleção dos conceitos e extração. Esclarece-se que essa divisão não é um consenso, mas se referem a procedimentos fundamentais. A Norma Brasileira de indexação (NBR 12676/1992, p. 2), por exemplo, indica a existência de três etapas, que foram chamadas de “estágios”, tais sejam: “a) exame do documento e estabelecimento do assunto de seu conteúdo; b) identificação dos conceitos presentes no assunto; c) tradução desses conceitos nos termos de uma linguagem de indexação”.

Lancaster (2004) indica a existência de duas etapas principais: análise conceitual, que visa decidir sobre o que trata um documento e tradução, que “envolve a conversão da análise conceitual de um documento num determinado conjunto de termos de indexação.” (LANCASTER, 2004, p. 18).

Conforme se verifica, alguns autores indicam duas etapas, outros três. O que fica evidente é que a análise de assunto (ou análise conceitual) prevalece em todas as acepções, podendo ser considerada uma etapa fundamental do processo de indexação.

Para esclarecer nosso entendimento a respeito da indexação e seus princípios, utilizamos o texto de Mendes e Simões (2002). Respalgadas em normas de

indexação, dentre elas as normas ISO 5963:1985 e normas portuguesas, as referidas autoras afirmam que:

O processo de indexação de assuntos, dentro das técnicas documentais, não pode contar com normas que verdadeiramente apoiem a objetividade da análise do documento, determinem, com precisão a decisão correta a tomar perante as opções colocadas pela representação dos conceitos, de forma a garantir a sua coerência e conferir uniformidade aos instrumentos de pesquisa disponibilizados ao universo utilizador (MENDES; SIMÕES, 2002, p. 11).

O usuário deve utilizar a mesma linguagem de indexação utilizada no sistema (entrada) para a recuperação da informação (saída), assim a comunicação poderá acontecer de forma precisa. O que aqui chamamos de “sistema” refere-se a um sistema de informação existente em qualquer unidade de informação, representado geralmente por um modelo básico de entrada, processamento e saída.

Além do controle vocabular inerente às linguagens documentárias, há o vocabulário controlado (VC), que é “[...] a lista de termos empregados no sistema e se caracteriza pelo fato de ser, naturalmente, bem menos extenso do que o vocabulário habitual do usuário do sistema, ou do que aquele constante [nos] documentos, e também menor que o vocabulário do indexador.” (CAVALCANTI, 1978, p.29).

Um tipo de vocabulário amplamente utilizado em diferentes campos do conhecimento é o tesauro. Diferente do vocabulário controlado, o tesauro apresenta uma lista de termos e os relacionamentos entre eles. Indica, por exemplo, os termos genéricos (TGs), os termos relacionados (TRs) e os termos específicos (TEs). Logo, o diferencial está nos relacionamentos, na sua estrutura e na precisão da cadeia de conceitos apresentada.

2.1 A Indexação na Arquivologia

Quanto à Indexação inerente à Arquivologia, esta tem estado relacionada à função arquivística de Descrição. Segundo a Norma Geral Internacional de Descrição Arquivística (ISAD(G)) o objetivo da descrição arquivística é “identificar e explicar o contexto e o conteúdo de documentos de arquivo a fim de promover o acesso aos mesmos.” (2000, p. 11). A ISAD(G) foi desenvolvida no âmbito do Conselho Internacional de Arquivos (ICA) em 1994, traduzida para o português e publicada no Brasil, em sua primeira versão, nos anos 1998, e segunda edição em 2000. Seu objetivo é estabelecer diretrizes gerais para a preparação de descrições arquivísticas, podendo ser utilizada como base para o desenvolvimento de normas nacionais próprias.

Em âmbito nacional, a ISAD(G) foi então adaptada e deu origem à Norma Brasileira de Descrição Arquivística (NOBRADE). A maior adaptação

identificada é a inserção da área 8, “Área de pontos de acesso e indexação de assuntos”. Essa área é o local “onde se registra os termos selecionados para localização e recuperação da unidade de descrição”. (2006, p. 18).

Segundo a NOBRADE, o item 8.1 “Pontos de acesso e indexação de assuntos” tem como objetivo:

Registrar os procedimentos para recuperação do conteúdo de determinados elementos de descrição, por meio da geração e elaboração de índices baseados em entradas autorizadas e no controle do vocabulário adotado. (BRASIL, 2006, p. 59).

Identifica-se a partir de tal objetivo as similaridades entre o que foi apresentado anteriormente quando a base da indexação na Biblioteconomia e seus instrumentos técnicos como o vocabulário controlado.

Para a NOBRADE, “A indexação dos assuntos deve contemplar nomes de entidades, eventos, áreas geográficas, períodos e assuntos tópicos” (BRASIL, 2006, p. 59). Ou seja, representações condensadas vinculadas à recuperação da informação. Como forma de exemplificação, apresenta-se alguns exemplos de termos indexadores constantes na NOBRADE. “Agradecimentos; anticomunismo; Argentina; asilo político; censura [...] regime militar; repressão política; segurança nacional; sindicatos; Uruguai”. (BRASIL, 2006, p. 61).

Além da Norma de Descrição Arquivística Brasileira, pesquisadores da área têm se atentado à temática da indexação e a tratando como uma das funções arquivísticas em conjunto com a Descrição.

Couture et al. publicam em 2003 o livro intitulado “*Les fonctions de l’archivistique contemporaine*”. Dentre as funções apresentadas, está a *Description/Indexation*. Segundo Couture, essa função diz respeito a um:

[...] conjunto de operações para descrever as características físicas e o conteúdo do arquivo, ou, mais precisamente, registrar informações sobre a estrutura, funções e conteúdo dos documentos arquivísticos (COUTURE, 2003, p. 19, tradução nossa).

Especificamente a respeito da Indexação, Couture (2003, p. 20, tradução nossa) considera que “[...] a descrição, como uma função de arquivo, evoluiu muito nos últimos dez anos. Alguns até dizem que é: <O setor que experimentou o maior desenvolvimento [...]>”. A inserção da indexação à função de descrição para o autor, está relacionada à necessidade da Arquivologia em dialogar com os sistemas de tecnologia da informação desenvolvidos para arquivos. “A informática transforma seu questionamento e modifica significativamente sua relação com a informação” (Ibid., p. 21, tradução nossa).

Segundo Couture, na descrição será preciso utilizar técnicas de análise de conteúdo com o propósito de

selecionar, condensar, categorizar, reagrupar e organizar a informação. O autor considera que ela poderá ser de duas formas:

[...] a **condensação** que leva à anotação, o resumo indicativo e o resumo informativo; a **indexação** por assunto que fornece os termos precisos que permitem o acesso à informação (o grau de precisão é decisivo em relação à eficiência relativa dos índices), é claro que toda essa redução tem a vantagem de ser baseada em um tesouro que é o vocabulário controlado – cada termo expressando um conceito único – graças ao qual é possível direcionar escolhas uniformes na indexação. (COUTURE, 2003, p. 20, tradução nossa, grifo nosso).

Nesse sentido, é visível a aproximação da Arquivologia com o referencial teórico apresentado anteriormente, quando dos objetivos, métodos e instrumentos técnicos relacionados à indexação. Na Arquivologia, os princípios, a organicidade e o contexto dos documentos precisarão ser os guias para a aplicação do que poderíamos identificar como “indexação arquivística”.

3 Procedimentos metodológicos

Em relação ao objetivo, a presente pesquisa é de caráter exploratório-descritivo e quanto à abordagem, pode ser caracterizada como qualitativa, já que utiliza métodos qualitativos para o levantamento, descrição e análise das publicações. É também considerada quantitativa, pois a partir dos dados qualitativos, serão apresentados também dados numéricos da pesquisa.

Para o levantamento, foram pesquisados artigos publicados nas seguintes fontes: Base de dados em Ciência da Informação (BRAPCI), Base de Dados do ENANCIB (BENANCIB), Base de dados da *International Society for Knowledge Organization* no Brasil (ISKO-Brasil) e ISKO-Ibérico. Justifica-se o uso dessas bases pela importância que refletem à área e à temática desta pesquisa. A BRAPCI é a base de dados referência para a Ciência da Informação no Brasil; já a BENANCIB representa o maior evento de Ciência da Informação do país; a base da ISKO-Brasil reuni os trabalhos apresentados durante o Congresso Brasileiro em Organização e Representação do Conhecimento e que está intimamente relacionado à esta pesquisa. A seleção dos artigos pesquisados nas bases acima deu-se com base nos termos: “indexação em arquivos” e “indexação arquivística”. A busca na ISKO-Ibérico deu-se por navegação nas edições, visto que algumas delas não possuem campo de busca.

Quanto aos critérios de inclusão dos textos, foram considerados, portanto, artigos científicos e trabalhos completos publicados em eventos científicos por autores brasileiros e no idioma português. Na seleção considerou-se também a pertinência do artigo a partir da leitura do seu título, identificando aqueles que tratam especificamente da indexação em arquivos ou indexação

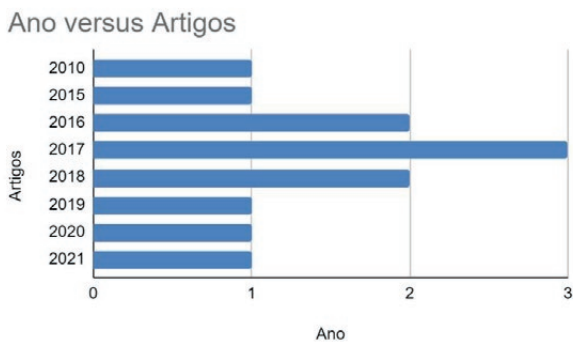
arquivística. Por fim, foram analisados 14 artigos constantes no Apêndice A.

Na análise dos artigos, procurou-se a maior incidência de referências do autor e a quantidade de obras diferentes citadas. Visando compreender o contexto dos autores, buscou-se as respectivas biografias, indicando-se sua formação entre graduação, mestrado e doutorado. A fonte para a busca de autores brasileiros foi a plataforma Lattes do CNPQ. No caso dos autores estrangeiros, procurou-se na própria universidade. Verificou-se, ainda, os países dos autores, para identificar a existência de referentes de diferentes regiões.

4 Resultados e discussão

De acordo com o resultado da busca nas fontes apresentadas, o gráfico 1 demonstra a distribuição dos 14 artigos por ano de publicação.

Gráfico 1. Artigos por ano de publicação



Fonte: Dados da pesquisa.

Verifica-se a existência de textos publicados no Brasil, nas fontes selecionadas, a partir de 2010, de forma pulverizada, com leve concentração de textos nos anos de 2016, 2017 e 2018, o que corresponde a duas ou três incidências de textos em cada ano.

Com relação às referências utilizadas por estes 14 artigos, todas foram extraídas e reunidas em uma única lista alfabética e compiladas de acordo com a quantidade de vezes que apareceram e os diferentes artigos do autor que foram citados. No quadro 1 abaixo estão identificados os autores que mais foram referenciados (coluna 1). O recorte de incidências, considerando a inviabilidade de apresentar todos os autores, mais de 40, foi feito dos que mais são mencionados até àqueles que foram apresentados pelo menos 5 vezes nesses 14 artigos. Na coluna 2 estão identificadas as diferentes obras mencionadas.

Quadro 1. Autores e menções

Autor	Incidência de referências ao autor	Diferentes Obras citadas
Fujita, Mariângela Spotti Lopes	14	12
Ribeiro, Cândida Fernanda antunes	11	6
Heredia Herrera, Antônia	9	4
Schellenberg, Theodore Roosevelt	9	2
Cruz Mundet, José Ramón	7	4
Oliveira, Isabel Cristina Borges de	6	5
Barros, Thiago Henrique Bragato	5	4
Bellotto, Heloisa Liberalli	5	1
Duchain, Michel	5	1
Duranti, Luciana	5	4
ISAD (G)	5	1

Fonte: Dados da pesquisa.

Conforme o Quadro 1, a autora com maior incidência de referências foi a Mariângela Fujita, professora da UNESP-Marília e Bibliotecária de formação. Dentre as 14 publicações citadas, são ao menos 12 diferentes, considerando aquelas em que é a primeira autora. As obras são de caráter geral sobre a indexação, políticas de indexação e linguagens documentárias, mostrando que o arcabouço teórico-metodológico no escopo da Biblioteconomia e Ciência da Informação está servindo de embasamento para publicações na área de Arquivologia, especificamente no campo da indexação. Dentre as publicações de Fujita estão: “A leitura em análise documentária”, “Manual de política de indexação para as bibliotecas universitárias da UNESP”, “A identificação de conceitos no processo de análise de assunto para indexação.”, “Política de indexação”, “A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias: um estudo de

observação do contexto sociocognitivo com protocolos verbais”, “A avaliação da indexação por meio da recuperação da informação”, “Políticas de indexação na América Latina”, “A observação da leitura documentária por meio de protocolo verbal”, “Política de indexação no contexto da política arquivística de preservação digital do Centro de Documentação e Memória da UNESP - CEDEM”, “Elementos de política de indexação em manuais de indexação de sistemas de informação especializados”, “O ensino de procedimentos de política de indexação na perspectiva do conhecimento organizacional: uma proposta de programa para a educação à distância do bibliotecário”, “Política de indexação para bibliotecas: estudo de elaboração, implantação e avaliação com pesquisa-ação integral”.

A professora Fernanda Ribeiro é a segunda com maior incidência de referências, num total de 11 em seis obras diferentes. Da Universidade do Porto-PT, a professora tem Bacharelado e Licenciatura em História e formação em Bibliotecário-Arquivista em Portugal. Com produção expressiva sobre a indexação em Arquivologia, dentre as obras citadas encontram-se publicações que tratam de classificação em arquivos, acesso à informação e fundamentos da Arquivologia. Os títulos citados são: “Indexação e controlo de autoridade em arquivos”, “O acesso à informação nos arquivos”; “Organizar e representar informação: apenas um meio para viabilizar o acesso?”, “*The use of classification in archives as a means organization, representation and retrieval information. Knowledge Organization*”, “Da arquivística técnica a arquivística científica: a mudança de paradigma”; “Indexação em arquivos: pressupostos teóricos gerais e especificidades”, “*Subject indexing in archives: theoretical and practical considerations*”, “A classificação em arquivos: processo natural ou arranjo a posteriori?”.

Heredia Herrera e Cruz Mundet, autores espanhóis, publicaram quatro obras, respectivamente, que foram citadas entre os autores que publicam sobre indexação arquivística no Brasil. Dentre as obras de Herrera está a reconhecida “*Archivística general: teoría y práctica*”, “*Archivos y archiveros entre la Guerra Civil y la España de las Autonomías*”, “*¿Qué es un archivo?*” e “*Manual de instrumentos de descripción documental*”.

Já em relação a Cruz Mundet, verifica-se que é Licenciado em História pela Universidade de Deusto (Bilbao), Doutor em História Moderna pela *Universidad Autónoma de Madrid* e Professor titular e diretor do *Master en Archivística* da Universidade Carlos III de Madrid. Dentre as obras citadas estão: “*Archivística: gestión de documentos y administración de archivos*”; “*Manual de archivística*”; “*Administración de documentos y Archivos: Textos Fundamentales*”, “*La gestión de documentos en las organizaciones*”.

Michel Duchein e Luciana Duranti foram os autores canadenses mais citados pelos autores que publicam sobre indexação em arquivos. Duchein é formado em história na Universidade de Bordéus. Trabalhou em vários arquivos, tornando-se curador-chefe em 1970 e, em seguida, inspetor-geral dos Arquivos da França em 1978. O autor foi citado por cinco autores diferentes, porém com apenas uma obra, tal seja: “*Theoretical Principles and Practical Problems of Respect des fonds in Archival Science*”. Observa-se que mesmo em artigos que tratam de indexação arquivística, a preocupação com princípios arquivísticos, o que suscita a ponderação de que a indexação em Arquivologia leva em conta os níveis de representação arquivística, conforme normas internacionais.

Luciana Duranti possui diploma de *Archivística* (equivalente a mestrado) pela *Scuola dell'Archivio di Stato di Roma* e é Ph D, *Archivista-Paleografo* pela *Scuola Speciale per Archivisti e Bibliotecari, Università di Roma* e *Dottore Lettere* (equivalente a mestrado, Artes), *Università di Roma*. A autora foi citada em cinco artigos da análise, com quatro obras diferentes, a saber: “*Origin and development of the concept of Archival Description*”; “*I Documenti Archivisti: la gestione dell'archivio da parte dell'ente produttore.*”, “*The concept of appraisal and archival theory*”, “*Diplomática: usos nuevos para una antigua ciencia*”.

Além de Fujita, outros três autores brasileiros foram citados: Isabel Cristina Borges de Oliveira, Thiago Henrique Bragato Barros e Heloísa Liberalli Bellotto. A primeira, Oliveira, possui graduação em Arquivologia pela Universidade Federal do Estado do Rio de Janeiro, graduação em História pela Universidade do Estado do Rio de Janeiro, Mestrado profissional em História, Política e Bens Culturais pela Fundação Getúlio Vargas. Dentre as obras publicadas estão: “Arquivos pessoais de valor histórico”, “Descrição arquivística: seu papel no processo de comunicação dos conteúdos informacionais dos acervos arquivísticos”, “Disseminação da informação arquivística: uma perspectiva interdisciplinar”, “Os usuários da informação arquivística”, “Descrição Arquivística: contexto arquivístico, controle de vocabulário e o usuário”.

O segundo autor brasileiro, Barros, teve incidência de cinco referências e quatro obras citadas. Importante destacar que possui Doutorado em Ciência da Informação, Graduação em Arquivologia e Mestrado em Ciência da Informação pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP). As obras do autor são: “Indexação e a Arquivística: aproximações iniciais no universo teórico da organização e representação do conhecimento”, “A representação da informação Arquivística: uma Análise do discurso teórico e institucional a partir dos contextos Espanhol, Canadense e Brasileiro”, “Os processos de representação do conhecimento arquivístico: elementos

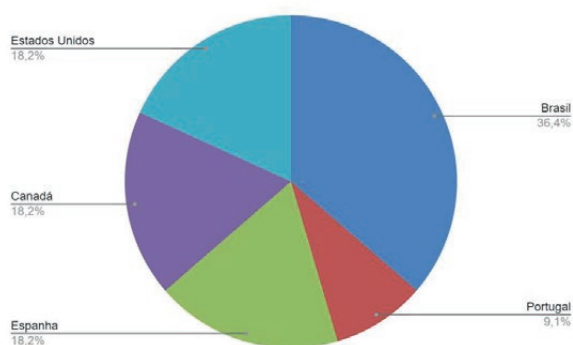
históricos e conceituais da classificação e descrição”, Arquivística, História e Ciência da Informação: diálogos e duelos”. Bellotto é Licenciada e doutora em História (USP), bacharel em Biblioteconomia (FESP) e especialista em Arquivística (*Escuela de Documentalistas*, Madri, Espanha). A autora foi citada em cinco artigos da análise, porém com uma única obra: “Arquivos permanentes: tratamento documental.”

Já entre os estadunidenses destaca-se Schellenberg, com nove incidências em duas obras citadas. Schellenberg doutorou-se em história pela Universidade da Pensilvânia. Vice-diretor dos Arquivos Nacionais dos Estados Unidos da América. As obras são de caráter mais geral da Arquivologia com os seguintes títulos: “Arquivos modernos: princípios e técnicas e ‘Documentos públicos e privados: arranjo e descrição”.

A norma ISAD(G) também se configurou como um importante referente teórico entre os autores, com incidência de cinco referências, porém com uma obra citada, que trata-se da própria Norma Internacional de Descrição Arquivística. Ao considerarmos que descrição e indexação estão relacionadas e que a indexação na Arquivologia ocorre conforme os níveis de descrição, percebe-se a relevância deste referente entre os mais citados.

Com relação ao país de origem dos autores apresentados no quadro 1, observa-se concentração em diferentes países. O gráfico 2, abaixo, demonstra a distribuição por países entre os autores citados no quadro 1.

Gráfico 2. Países dos autores mais citados



Fonte: Dados da pesquisa.

Identifica-se que o maior número de autores é do Brasil, seguido do Canadá e Espanha, e Estados Unidos e Portugal. Observa-se, também, que se considerarmos os brasileiros e não brasileiros, a maioria das referências que estão sendo utilizadas sobre a temática provém de literatura estrangeira (Brasil, 36,4% e estrangeiros 63,6%).

5 Considerações finais

Ao desenvolver este artigo foi possível verificar, na literatura analisada, que o tema da indexação em

Arquivologia ainda é pouco explorado entre autores brasileiros. Conforme se observa, embora não se tenha restringido o período de busca na metodologia, os resultados mostram apenas 14 publicações em artigos e eventos que tratam dessa temática. De acordo com a análise é possível observar que a pesquisa sobre indexação arquivística no Brasil tem prevalecido a partir de 2010, em artigos científicos e anais de eventos científicos. Outra percepção, é que as bases teóricas que estão sendo utilizadas no Brasil por autores que tratam da indexação em arquivos provêm, em maioria, de outros países e regiões, em especial da América do Norte e Europa. Embora esse recorte tenha sido feito entre os autores com no mínimo cinco incidências de referências, pode mostrar a diversidade no pensamento dos autores brasileiros.

Verifica-se na análise que a autora mais citada dentre as publicações científicas levantadas na pesquisa é Mariângela Fujita, que é Bibliotecária, ou seja, apresentando interdisciplinaridade entre as áreas. Além disso, o arcabouço teórico no campo da indexação nas áreas de Ciência da Informação e Biblioteconomia é um relevante referencial teórico para autores que publicam sobre indexação na Arquivologia. Observa-se as diferentes formações dos autores, sendo a maioria arquivistas (sete), dois historiadores e uma bibliotecária, com exceção da norma ISAD(G), não considerada entre as formações.

Ao analisar as obras citadas pelos autores, indicadas nos resultados, verifica-se em menor quantidade textos que indicam no título a especificidade da indexação em arquivos. A maior incidência é de obras que tratam, de modo geral, de fundamentos da Arquivologia, como as obras de Heredia Herrera, Cruz Mundet, Michel Duchain, Luciana Duranti e Schellenberg, as quais podem trazer aspectos da indexação arquivística em seu conteúdo ou servirem de base para fundamentação teórica dos textos. Se relacionarmos com o referencial teórico apresentado, verifica-se que os princípios arquivísticos e as bases da Arquivologia são necessários para todas as funções arquivísticas, em especial para a descrição/indexação.

Dentre os autores mais citados, aqueles que publicaram textos específicos sobre o tema, conforme indica o título são: Mariângela Fujita, com a obra: “Política de indexação no contexto da política arquivística de preservação digital do Centro de Documentação e Memória da UNESP – CEDEM”, Fernanda Ribeiro, com os títulos: “Indexação em arquivos: pressupostos teóricos gerais e especificidades” e “*Subject indexing in archives: theoretical and practical considerations*”, além de Thiago Barros, com o artigo: “Indexação e a Arquivística: aproximações iniciais no universo teórico da organização e representação do conhecimento”.

Percebe-se, portanto, carência de produções específicas sobre a indexação em Arquivologia e a necessidade do desenvolvimento de novas pesquisas nesse sentido.

Referências

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: informação e documentação: métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro: ABNT, 1992.
- BRADFORD, S. C. **Documentação**. Rio de Janeiro: Fundo de Cultura, 1961.
- BRASIL. Conselho Nacional de Arquivos (Conarq). **Nobrade**: Norma Brasileira de Descrição Arquivística. Rio de Janeiro: Arquivo Nacional, 2006. Disponível em: http://conarq.gov.br/images/publicacoes_textos/nobrade.pdf. Acesso em: 18 abr. 2023.
- BUREAU OF CANADIAN ARCHIVISTS (BCA). **Subject indexing for archives**: the report of the Subject Indexing Working Group. Bureau Canadien des Archivistes, 1992.
- CAVALCANTI, C. R. **Indexação e tesouro**: metodologia e técnicas. Brasília: ABDF, 1978
- CHAUMIER, Jacques. **As técnicas documentais**. Mem Martins: Europa-America, 1973.
- COUTURE, Carol *et al.* **Les fonctions de l'archivistique contemporaine**. Sainte-Foy (Québec) Canadá: Presses de L'Université du Québec, 2003.
- CONSELHO INTERNACIONAL DE ARQUIVOS. **ISAD (G)**: norma geral internacional de descrição arquivística: segunda edição. Rio de Janeiro: Arquivo Nacional, 2000.
- DIAS, Eduardo Wense; NAVES, Madalena Martins Lopes. **Análise de assunto**: teoria e prática. 2. ed. rev. Brasília, DF: Briquet de Lemos, 2013.
- FOSKETT, A. C. **A abordagem temática da informação**. Tradução de Antônio Agenor Briquet de Lemos. São Paulo: Polígono, 1973.
- HUDON, M. Indexation et langages documentaires dans les milieux archivistiques a l'ere des nouvelles technologies de l'information. **Archives (Quebec)**, v. 29, n. 1, p. 75-98, 1997-1998.
- LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Trad. Antonio Agenor Briquet de Lemos. 2. ed. Brasília: Briquet de Lemos, 2004.
- LEHMKUHL, Camila Schwinden. **Registros civis no Brasil frente às funções arquivísticas**. 2021. 228 f. Tese (Doutorado) - Curso de Pós-Graduação em Ciência da Informação, CIN, Universidade Federal de Santa Catarina, Florianópolis, 2021.
- LÉVESQUE, M. L'indexation: luxe ou necessite? **Archives (Quebec)**, v. 33, n. 1, p. 17-45, 2001-2002.
- LYTLE, R. H. Intellectual access to archives .1. provenance and content indexing methods of subject retrieval **American Archivist**, v. 43, n.1, p. 64-75, 1980.
- MARTIN, Russell Lewis. **Archival indexing: problems and issues**. 1987. 88f. Dissertação (Master Archives Studies) - University of British Columbia Faculty of Arts, School of Library, Archival and Information Studies, 1987. Disponível em: <https://open.library.ubc.ca/cIRcle/collections/ubctheses/831/items/1.0096790>. Acesso em: 15 mar. 2023.
- MAUREL, Dominique; CHAMPAGNE, Michel. La description et l'indexation. In: MAUREL, Dominique; CHAMPAGNE, Michel. **Les fonctions de l'archivistique contemporaine**. Sainte-Foy (Québec), Presses de l'Université du Québec, 1999, p. 255–371.
- MENDES, Maria Teresa Pinto; SIMÕES, Maria da Graça. **Indexação por assuntos**: princípios gerais e normas. Lisboa: Publicações A&B, 2002.
- PINTO MOLINA, María. **Análisis documental**: fundamentos y procedimientos. 2. ed. Madrid: EUDEMA, 1993. 270 p.
- RIBEIRO, Fernanda. Indexação em arquivos: pressupostos teóricos gerais e especificidades. In: SOUSA, Francinete Fernandes de; SANTOS, Eliete Correia dos (Org.). **A linguagem e a informação documentária**: intermediações e ressignificações possíveis. Recife: Bagaço, 2011. p. 31-59.
- RIBEIRO, Fernanda. **Indexação e controlo de autoridade em arquivos**. Porto: Câmara Municipal; Arquivo Histórico, 1996. Disponível em: <http://repositorio-aberto.up.pt/handle/10216/10721>. Acesso em: 15 mar. 2023.
- SMIT, Johanna Wilhelmina (Org.). **Análise documental**: a análise da síntese. Brasília, DF: IBICT, 1987. cap. 3, p. 38-60.
- TARTAROTTI, Roberta Cristina Dal'Evedove; DAL'EVEDOVE, Paula Regina; FUJITA, Mariangela Spotti Lopes. Avaliação da consistência da indexação em bibliotecas universitárias federais da região nordeste do Brasil. **Anales de Documentación**, Espanha, v. 20, n. 1, p. 1-19, 2017. Disponível em: <https://www.redalyc.org/articulo.oa?id=63549938004>. Acesso em: 13 mar. 2023.

Apêndice A – Lista das obras analisadas

N.	Referências
1	AZEVEDO, J. F. S.; SALES, R. A indexação e o controle de vocabulário em arquivos: uma investigação no âmbito dos arquivos públicos estaduais brasileiros. <i>Ágora</i> , v. 29, n. 58, p. 1-17, 2019. Disponível em: http://hdl.handle.net/20.500.11959/brapci/109943 . Acesso em: 07 mar. 2023.
2	BARROS, T. H. B. A indexação e a arquivística: aproximações iniciais no universo teórico da organização e representação do conhecimento. <i>Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação</i> , v. 21, n. 46, p. 33-44, 2016. DOI: 10.5007/1518-2924.2016v21n46p33. Acesso em: 07 mar. 2023.
3	CAMPOS, M. L. A. Indexação e descrição em arquivos: a questão da representação e recuperação de informações. <i>Arquivo & Administração</i> , v. 5, n. 1, 2006. Disponível em: http://hdl.handle.net/20.500.11959/brapci/51656 . Acesso em: 07 mar. 2023.
4	CANDIDO, G. G.; BARROS, T. H. B.; REDIGOLO, F. M. Indexação e descrição arquivística: relações histórico-conceituais. In: CONGRESSO BRASILEIRO DE ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO, 4, 2017, Recife. <i>Anais [...]</i> . Recife: UFPE, 2017. p. 162-169. Disponível em: https://isko.org.br/wp-content/uploads/2021/05/Proceedings-ISKO-Brasil-2017.pdf . Acesso em: 14 mar. 2023.
5	FUJITA, M. S. L.; RODRIGUEZ, S. M. T. Política de indexação em arquivos de instituições de saúde. <i>Informação em Pauta</i> , v. 3, p. 95-116, 2018. Disponível em: http://www.periodicos.ufc.br/informacaoempauta/article/view/39719/pdf . Acesso em: 07 mar. 2023.
6	MEDEIROS, G. M.; BRÄSCHER, M. A indexação de assunto em documentos arquivísticos. <i>Revista Digital de Biblioteconomia & Ciência da Informação</i> , v. 18, n. 2020, 2000. DOI: 10.20396/rdbci.v18i0.8657824 Acesso em: 07 mar. 2023.
7	NAKAGAWA, J.; FUJITA, M. S. L. O contexto da política de indexação em arquivo permanente: um estudo de observação participante da prática profissional de arquivistas. <i>Revista Brasileira de Educação em Ciência da Informação</i> , v. 8, 2021. DOI: 10.24208/rebecin.v8i.265 Acesso em: 07 mar. 2023.
8	OLIVEIRA, I. C. B. Indexação em coleções familiares quais aspectos considerar?. <i>Arquivo & Administração</i> , v. 9, n. 1, 2010. Disponível em: http://hdl.handle.net/20.500.11959/brapci/51207 . Acesso em: 07 mar. 2023.
9	OLIVEIRA, I. C. B. Padronizar, normalizar e definir pontos de acesso o desafio da indexação arquivística. <i>Arquivo & Administração</i> , v. 8, n. 1, 2009. Disponível em: http://hdl.handle.net/20.500.11959/brapci/51235 . Acesso em: 07 mar. 2023.
10	PINTO, M. C. A indexação na recuperação da informação em arquivos: uma abordagem inicial. <i>Ágora</i> , v. 26, n. 52, p. 141-150, 2016. Disponível em: http://hdl.handle.net/20.500.11959/brapci/13879 . Acesso em: 07 mar. 2023.
11	PRET, R. L.; CORDEIRO, R. I. N. Clasificación y indización de la información en Registros y Archivos de Gestión de las universidades. In: Congreso ISKO España, 12, Congreso ISKO España-Portugal, 2, 2015. <i>Anais [...]</i> . Murcia: Universidad de Murcia, 2015. p. 195-205. Disponível em: https://iskoiberico.org/wp-content/uploads/2015/12/38001-iskoep2015_actas.pdf . Acesso em: 14 mar. 2023.
12	PRET, R. L.; CORDEIRO, R. I. N. O processo de indexação nos arquivos de universidades: representação e usos no sistema de registros da universidade federal fluminense. <i>Bibliotecas. Anales de Investigación (Cuba)</i> , v. 14, n. 2, p. 242-251, 2018. Disponível em: http://hdl.handle.net/20.500.11959/brapci/59873 . Acesso em: 07 mar. 2023.
13	SOUSA, R. T. B.; ARAÚJO JÚNIOR, R. H. A indexação e criação de taxonomias para documentos de arquivo: proposta para a expansão do acesso e integração das fontes de informação. <i>Brazilian Journal of Information Science</i> , v. 11, n. 4, 2017. DOI: 10.5016/brajis.v11i4.7508. Acesso em: 07 mar. 2023.
14	TOGNOLI, Natália Bolfarini; MILANI, Suellen Oliveira; GUIMARÃES, José Augusto Chaves. Elementos de indexação para arquivos permanentes: aproximações teóricas preliminares. In: CONGRESSO ISKO ESPANHA-PORTUGAL, 3; CONGRESSO ISKO ESPANHA, 13, 2017, Coimbra. <i>Anais [...]</i> . Coimbra: Universidade de Coimbra, 2017, p. 683-692

Representação do Conhecimento Arquivístico: uma análise comparativa entre entidades de *Records in Contexts Ontology* e entidades derivadas de *Basic Formal Ontology*

Archival Knowledge representation: a comparative analysis between entities from Records in Contexts Ontology and entities derived from Basic Formal Ontology

Marieta Marks Löw (1), Rafael Port da Rocha (2) Thiago Henrique Bragato Barros (3)

(1)(3) Universidade Federal de Santa Catarina, Programa de Pós-Graduação em Ciência da Informação, R. Eng. Agrônomo Andrei Cristian Ferreira, Centro Ciências da Educação, Bloco B, Florianópolis - SC, marietalow@gmail.com, bragato.barros@ufrgs.br (2)(3) Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Ciência da Informação, Rua Ramiro Barcelos, 2705 Sala 203, Porto Alegre, rafael.rocha@ufrgs.br

Resumo

Objetivo: Discute-se as possibilidades de representação por meio de ontologias do documento arquivístico a partir de uma análise conceitual comparativa de entidades da ontologia *Records in Contexts* (RiC-O), apresentada pelo *International Council on Archives* (ICA), com entidades de ontologias derivadas da *Basic Formal Ontology* (BFO). **Metodologia:** Utiliza abordagem qualitativa, com uso de análise conceitual comparando entidades relacionadas à representação de documento arquivístico no modelo RiC-O e em entidades que representam documento em modelos que utilizam a BFO como ontologia de topo. **Resultados:** A partir de um quadro analítico das entidades observa-se um alinhamento entre as ontologias, com algumas estratégias de modelagem diferentes. Como resultado da análise é possível verificar semelhanças e divergências na representação do documento arquivístico entre os dois modelos ontológicos. O modelo do ICA apresenta entidades construídas especificamente para o contexto da descrição arquivística. Apesar de suas entidades serem mais específicas, seu modelo é pequeno e não está fundamentado em nenhuma ontologia de topo, o que pode trazer dificuldades na implementação em sistemas informacionais complexos. As entidades derivadas de BFO, por outro lado, apesar de estarem fundamentadas em uma ontologia consolidada e utilizada em diferentes domínios, não foram especializadas para abarcar as especificidades da produção documental arquivística. **Conclusão:** Com a análise conceitual realizada é possível concluir que os estudos sobre representação do documento arquivístico com ontologias precisam ser expandidos de forma a possibilitar a representação das especificidades acerca da proveniência arquivística e do contexto de produção.

Palavras-chave: Records in Contexts. Ontologia. Basic Formal Ontology. Documento arquivístico.

Abstract

Objective: We discuss the possibilities of ontological representation through ontologies of the record from a comparative conceptual analysis of entities of the Records in Contexts (RiC-O) ontology, presented by the International Council on Archives (ICA), with entities of ontologies derived from the Basic Formal Ontology (BFO). **Methodology:** A qualitative approach is used, with conceptual analysis comparing entities related to the representation of the record in the RiC-O model and in entities representing the document in models using BFO as the top ontology. **Results:** From an analytical framework of the entities, it is observed an alignment between the ontologies, with some different modeling strategies. As a result of the analysis, it is possible to verify similarities and divergences in the representation of the record between the two ontological models. The ICA model presents entities that are specifically built for the context of archival description. Although its entities are more specific, its model is small and not based on a top-level ontology, which may cause difficulties in its implementation in complex information systems. On the other hand, the entities derived from the BFO, although based on a consolidated ontology and used in different domains, were not specialized to cover the specificities of records production. **Conclusion:** With the conceptual analysis carried out, it is possible to conclude that the studies on the representation of records with ontologies need to be extended in order to allow the representation of the specificities of the archival provenance and the context of production.

Keywords: Records in Contexts. Ontology. Basic Formal Ontology. Record.

Resumen

Objetivo: Discutimos las posibilidades de representación a través de ontologías del documento de archivo a partir de un análisis conceptual comparativo de entidades de la ontología Records in Contexts (RiC-O), presentada por el International Council on Archives (ICA), con entidades de ontologías derivadas de la Basic Formal Ontology (BFO). **Metodología:** Utiliza abordaje cualitativo, con el

uso de análisis conceptual comparando entidades relacionadas a la representación de documento de archivo en el modelo RiC-O y en entidades que representan documento en modelos que utilizan BFO como ontología superior. **Resultados** A partir de un marco analítico de las entidades se observa una alineación entre las ontologías, con algunas estrategias de modelado diferentes. Como resultado del análisis es posible verificar similitudes y divergencias en la representación del documento de archivo entre los dos modelos ontológicos. El modelo ICA presenta entidades construidas específicamente para el contexto de la descripción archivística. A pesar de que sus entidades son más específicas, su modelo es pequeño y no se basa en ninguna ontología superior, lo que puede traer dificultades en la implementación en sistemas de información complejos. Las entidades derivadas de BFO, por otro lado, a pesar de estar basadas en una ontología consolidada y ser utilizadas en diferentes dominios, no fueron especializadas para abarcar las especificidades de la producción de documentos de archivo. **Conclusión:** Con el análisis conceptual realizado es posible concluir que los estudios sobre representación del documento de archivo con ontologías necesitan ser ampliados para posibilitar la representación de las especificidades sobre la procedencia archivística y el contexto de producción.

Palabras clave: Records in Contexts. Ontología. Basic Formal Ontology. Documento de archivo.

1 Introdução

O documento arquivístico trata-se de uma unidade base do arcabouço terminológico da arquivologia, produzido a partir de atividades desempenhadas pelo produtor. Os documentos acumulados por um determinado produtor formam um conjunto documental, e são organizados em subconjuntos de acordo com as diferentes funções e atividades por ele desempenhadas (HEREDIA HERRERA, 1991; MULLER; FEITH; FRUIN, 1973).

Mas para além de identificar atividades, ou processos que geram os documentos, o trabalho arquivístico envolve compreender e representar o contexto de produção destes documentos. A produção documental é consequência de ações que são realizadas pelo produtor de acordo com regras e normas locais e sociais, além de culturas e procedimentos institucionais próprios daquele produtor, o que torna cada conjunto documental único. Os documentos arquivísticos reunidos e produzidos a partir do domínio jurídico-administrativo, estão ligados uns aos outros por uma relação de organicidade (ou *archival bond*) através do contexto em que foram produzidos, e essa qualidade é o que o diferencia de outros documentos (CONSELHO NACIONAL DE ARQUIVOS, 2020)

Tradicionalmente, representar o contexto de produção dos documentos esteve associado à ordenação em que estes foram criados e mantidos na sua origem, antes de serem recolhidos a uma instituição arquivística. A explicitação desse contexto, na tradição dos séculos XIX e XX esteve associada ao desenvolvimento de atividades técnicas de classificação, arranjo e descrição da documentação, resultando em instrumentos técnicos. Com a produção documental em sistemas de informação a representação do contexto através de modelos conceituais e ontologias se torna ainda mais relevante.

Os estudos das ontologias como sistemas de organização do conhecimento voltadas para a explicitação de relacionamentos semânticos entre entidades se desenvolvem na década de 90 a partir de Gruber (1995) e Guarino (1998). O uso dos sistemas de organização do conhecimento para a documentação arquivística é um tema que vem ganhando destaque na literatura da área. Barros e Sousa (2019) discutem as aproximações

teórico-conceituais entre a organização do conhecimento e a Arquivologia. O uso das ontologias para a representação de documentos de arquivo também vem sendo discutido na literatura, como em Barros e Gomes (2018), Barros, Bastos e Santos (2022) e Löw *et al.* (2022).

O artigo discute as possibilidades de representação ontológica do documento arquivístico a partir de uma abordagem qualitativa, utilizando-se de uma análise conceitual comparativa entre entidades da ontologia *Records in Contexts* (RiC-O), apresentada pelo *International Council on Archives* (ICA), e entidades de ontologias derivadas da *Basic Formal Ontology* (BFO) [1].

A partir do levantamento das entidades, em cada modelo ontológico, que podem servir à representação do documento arquivístico e seu contexto de produção, foi realizada uma análise conceitual das entidades buscando compreender sua semântica e o relacionamento com as demais entidades do modelo, trazendo a perspectiva arquivística de proveniência e contexto de produção e buscando pontos em comum, convergências e divergências.

A análise dos dados acerca do modelo *Records in Contexts* (RiC) do ICA é apresentada na seção 2. A seção seguinte é dedicada à análise de entidades de ontologias de documentos derivadas da *Basic Formal Ontology* (BFO). Os resultados das análises são discutidos na seção 4, e a seção 5 apresenta algumas considerações e conclusão do estudo.

2 Records in Contexts: o modelo conceitual RiC-CM e a ontologia RiC-O

No âmbito da Arquivologia, está em desenvolvimento pelo Conselho Internacional de Arquivos (ICA) um modelo conceitual para o domínio da documentação arquivística. O modelo conceitual, batizado de *Records in Contexts Conceptual Model* – RiC-CM, está atualmente na versão rascunho (v. 0.2), e vem sendo desenvolvido internamente pelo EGAD, *Expert Group on Archival Description*.

A proposta do modelo apresenta os principais conceitos para representação de contexto de produção documental para a Arquivologia (Tabela 1), e, quando finalizado, terá como complemento uma ontologia com as propriedades das entidades e os seus relacionamentos apresentados em linguagem RDF. (INTERNATIONAL COUNCIL ON ARCHIVES, 2021). A ontologia está na versão 0.2, e foi lançada em fevereiro de 2021.

A tabela a seguir apresenta a hierarquia das entidades principais de RiC-CM.

Tabela 1. Hierarquia de entidade de RiC-CM

Primeiro Nível	Segundo Nível	Terceiro Nível	Quarto Nível
RiC-E01 <i>Thing</i>	RiC-E02 <i>Record Resource</i>	RiC-E03 <i>Record Set</i>	
		RiC-E04 <i>Record</i>	
		RiC-E05 <i>Record Part</i>	
	RiC-E06 <i>Instantiation</i>		
	RiC-E07 <i>Agent</i>	RiC-E08 <i>Person</i>	
		RiC-E09 <i>Group</i>	RiC-E10 <i>Family</i>
			RiC-E11 <i>Corporate Body</i>
		RiC-E12 <i>Position</i>	
		RiC-E13 <i>Mechanism</i>	
	RiC-E14 <i>Event</i>	RiC-E15 <i>Activity</i>	
	RiC-E16 <i>Rule</i>	RiC-E17 <i>Mandate</i>	
	RiC-E18 <i>Date</i>	RiC-E19 <i>Single Date</i>	
		RiC-E20 <i>Date Range</i>	
		RiC-E21 <i>Date Set</i>	
	RiC-E22 <i>Place</i>		

Fonte: Dados da pesquisa.

O grupo do ICA iniciou as discussões para a criação de um modelo conceitual para representação do domínio arquivístico em 2012. O modelo apresentou uma primeira versão rascunho em 2016, e em 2019 a segunda

versão foi divulgada de maneira prévia, para comentários, sendo republicada, ainda como rascunho (v 0.2) em julho de 2021 (INTERNATIONAL COUNCIL ON ARCHIVES, 2021).

2.1 O documento arquivístico em RiC

O conceito de documento arquivístico (*Record*) em RiC-CM é definido como:

informação inscrita ao menos uma vez, por qualquer método em qualquer suporte físico, em qualquer forma persistente e recuperável, por um agente no decurso da vida ou atividade de trabalho (INTERNATIONAL COUNCIL ON ARCHIVES, 2021, p. 24, tradução nossa).

Nesse sentido, o documento aparece como o registro de uma informação oriunda da vida ou atividade de trabalho de um agente. Logo há uma ação intencional de registrar/documentar. O documento de arquivo deve ter existido em algum momento temporal e pode ser recuperável em alguma forma persistente. Para tanto deve existir um suporte material, também nomeado em RiC-CM como “carregador físico” (*carrier*). Ao usar a expressão “informação inscrita ao menos uma vez em um suporte”, o modelo reforça a ideia de que o registro e o suporte são coisas distintas. Em RiC, essa distinção é realizada pelo uso de duas classes para representar o documento como informação (*record*), e a inscrição deste documento em um suporte, chamada de Instanciação. Dessa forma, RiC indica que o documento somente existirá se existir pelo menos uma instanciação sua, e oferece uma estratégia para a eliminação de documentos, na qual a instanciação pode ser eliminada, mas o documento, como registro, permanece.

Fica implícita a materialidade do documento e a necessidade de preservação do registro, o que implica em ações no âmbito da prática arquivística, envolvendo questões de gestão documental e custódia, autenticidade e preservação, entre outras.

A tabela a seguir apresenta as definições, na língua original, das entidades do modelo relacionadas ao conceito de documento arquivístico.

Tabela 2. Definição das entidades relacionadas à documento

Entidade	Código	Definição (no original, em inglês)
<i>Record Resource</i>	RiC-E02	<i>A Record, Record Set, or Record Part produced or acquired and retained by an Agent in the course of Activity.</i>
<i>Record Set</i>	RiC-E03	<i>One or more records that are associated by categorization and/or physical aggregation by the creator or other Agent responsible for preserving the creator's records.</i>

<i>Record</i>	RiC-E04	<i>Information inscribed at least once by any method on any physical carrier in any persistent, recoverable form by an Agent in the course of life or work Activity.</i>
<i>Record Part</i>	RiC-E05	<i>Part of a Record with discrete information content that contributes to the Record's physical or intellectual completeness.</i>
<i>Instantiated on</i>	RiC-E06	<i>The inscription of information on a physical carrier in any persistent, recoverable form by an Agent as a means of communicating information through time and space.</i>

Fonte: Dados da pesquisa.

Outros conceitos importantes, e que se relacionam ao conceito de documento arquivístico, são “recurso documental” (*Record Resource*) e “conjunto documental” (*Record Set*). Recurso documental é um conceito amplo que abrange os conceitos de documento arquivístico (*Record*), conjunto documental (*Record Set*) e parte de documento (*Document Part*), quando estes são produzidos ou adquiridos e mantidos por um agente no decurso de atividades. Aqui o termo mantido (*retained*) traz como referência a discussão teórica sobre a acumulação intencional e o processo de seleção, que está relacionado a organicidade e proveniência, na discussão sobre ordem original. Estes conceitos serão aprofundados adiante.

Já o conceito de conjunto documental (*Record Set*), que para RiC-CM significa um ou mais documentos arquivísticos associados por categorização ou agregação física pelo produtor ou outro agente responsável pela preservação dos documentos, está relacionado à organicidade e classificação documental. É um conceito importante para a teoria arquivística, pois é justamente o agrupamento dos documentos que forma um arquivo. Os documentos arquivísticos não são colecionados, ao contrário, são reunidos organicamente em função da atividade e do produtor. O agrupamento, ordenação e classificação dos documentos representam um contexto de produção único, o qual está diretamente vinculado ao contexto jurídico administrativo que regulamentou sua produção. Tradicionalmente os documentos são organizados em conjuntos e subconjuntos tais como fundo, sessão, série, dossiê ou processo. Essa entidade de RiC-CM permite expressar essas classificações.

Dois outros conceitos do modelo são complementares aos conceitos já analisados. O conceito parte de documento (*Document Part*) é definido como a parte de um documento que contenha informação distinta e que contribua para a completude física ou intelectual do

documento arquivístico. Uma parte de documento pode ser tanto uma seção de uma mensagem, por exemplo, quanto um documento avulso que é parte de outro, como uma planilha, que é parte de um relatório, ou uma sentença, que é parte de um processo judicial. Uma parte de documento, pode ela mesma, conter partes de documento, também.

A categorização dos documentos de acordo com as configurações que seus conteúdos assumem é representada, em RiC-O, pelo atributo tipo de forma documentária (*DocumentaryFormType*) que é definido como a categorização do documento com relação aos seus elementos extrínsecos e intrínsecos que juntos comunicam seu conteúdo, contexto administrativo e documental e autoridade [1]

Já o conceito de Instanciação (*Instantiation*) é definido no modelo como a inscrição de informações em um suporte físico, em alguma forma persistente e recuperável por um agente como meio de comunicar informações através do tempo e espaço. Um documento, ou parte dele, deve ter sido instanciado ao menos uma vez, ainda que tenha deixado de existir. Algumas vezes o documento poderá existir em diferentes formatos ou suportes, que correspondem a diferentes manifestações do mesmo documento, outras vezes, dependendo do contexto de produção e uso do documento, a mudança de suporte ou formato leva a que cada manifestação seja considerada como um documento diferente.

Para o modelo RiC-CM, como visto acima, um documento é uma informação inscrita ao menos uma vez, por qualquer método em qualquer suporte físico, e a cada vez que essa informação é inscrita em um suporte, uma instanciação de documento é gerada. Na ontologia RiC-O uma instanciação pode ser categorizada de duas formas. A primeira é quanto ao tipo de suporte ou carregador (*CarrierType*), como papel e disco rígido, por exemplo. É o tipo de suporte que determina as condições do ambiente de guarda e os pré-requisitos e formas de acesso e uso dos documentos. A segunda categoria está relacionada ao tipo de método de registro (*RepresentationType*), ou seja, como a mensagem é inscrita no suporte. A mensagem pode ser acessada diretamente, como num texto escrito ou gravura, ou com a mediação por algum dispositivo, como no caso de arquivos digitais, onde um computador é necessário para se ter acesso à informação, e outros tipos de mídias, como cartões perfurados, microfilme etc. A instanciação, entretanto, não deve ser confundida como mero suporte ou carregador do documento, pois são as características de uma instanciação que determinam as ações de gestão física e de preservação do suporte a serem realizadas. O tipo de suporte e o método de registros têm implicações na autenticidade do documento, assim, a forma como um documento é instanciado contribui para a contextualização de seu conteúdo.

Então é possível perceber que um documento arquivístico é apresentado, pelo modelo, como uma entidade documento (*Record*) e possui partes informacionais (*Record Part*) que são construídas a partir de regras de representação (forma documental) aceitas, o que faz com que este documento tenha uma tipologia. A instanciação será a sua materialização em um suporte, como uma folha de papel ou um disco rígido, por exemplo, representada na entidade instanciação (*Instantiation*), e será categorizada com relação ao material físico (*CarrierType*), papel ou disco rígido, e quanto ao método usado para registrar a informação nesta instanciação (*RepresentationType*), técnica de impressão ou gravação em um disco rígido.

Em RiC-CM uma instanciação pode deixar de existir, mas devem permanecer os registros de que o documento foi instanciado ao menos uma vez. Isso ocorre em razão dos modelos de gestão arquivística desenvolvidos a partir da teoria das três idades, quando documentos avaliados como não permanentes podem ser eliminados, ou seja, sua instanciação, sua materialidade física, deixa de existir, mas não os registros de sua existência. Nesses modelos o documento preservado em sua completude material e intelectual é aquele que, após uma fase de uso corrente na instituição produtora (primeira idade), ou guarda em arquivos intermediários para consulta (segunda idade), ainda preserva valor informativo e probatório e deve ser mantido em arquivo permanente, a terceira idade do documento, o que, em geral, ocorre em instituições de custódia especializadas, como instituições arquivísticas ou centros de documentação. Para RiC-CM, um documento (*Record*) ou parte de documento (*Record Part*) é também um registro de um documento ou parte dele cuja instanciação deixou de existir.

2.2 A produção do documento em RiC

O documento arquivístico se caracteriza por ser criado e acumulado no decurso de atividades de um dado produtor, e mantido em conjunto para fins informacionais e probatórios. Assim, é importante compreender as ações que levaram à criação dos documentos.

Em Ric-CM, o documento é o registro de uma atividade (*Activity*), realizada por um agente (*Agent*).

Todos os documentos arquivísticos [*Record*] que são membros de um conjunto documental [*Record Set*] podem compartilhar o atributo de terem sido acumulados pelo mesmo agente [*Agent*], ou todos compartilham a mesma forma documentária [*DocumentaryFormType*] e são criados ao longo do tempo pela mesma atividade [*Activity*]. (INTERNATIONAL COUNCIL ON ARCHIVES, 2021, p. s.p.)

Uma atividade é um fazer com propósitos humanos, e um tipo de evento (*Event*). O evento, por outro lado, é algo que acontece no tempo e espaço, com causas naturais, humanas ou uma combinação de ambas. A

produção de um documento acontece a partir de diretivas e regramentos tanto da entidade produtora, como do contexto social.

A representação dessas normativas é feita através das entidades regra e mandato. As regras (*Rule*) são entendidas no modelo como condições que regem a existência ou autoridade de um agente ou o desempenho de uma atividade, ou que contribuem para as características distintas das coisas criadas ou gerenciadas por um agente; e mandato (*Mandate*) é definido como a delegação de autoridade de um agente a outro agente para realizar uma atividade.

As entidades produtoras de documentos em RiC-CM são: pessoa (*Person*), família (*Family*) e entidades coletivas (*Corporate Body*). Essas entidades subsumem agentes (*Agent*) à medida que agem. Um agente pode ter uma ou mais identidades, e uma identidade é uma constelação de propriedades ou relações que juntas identificam o agente. Pseudônimos, heterônimos e identidades comerciais são elencados como casos de identidades alternativas de agentes. Identidade alternativa não se confunde com posição em um grupo, ou com variações de nome ou identificador de uma mesma identidade.

No caso de pessoas (*Person*), o mais comum é cada um ter apenas uma identidade social, ou a “pessoa real”, entretanto uma pessoa pode criar uma personagem alternativa para propósitos específicos, como um pseudônimo. Pessoas (*Person*) são produtoras de acervos pessoais (*Record Set*).

Um grupo (*Group*) é quando dois ou mais agentes atuam em conjunto como um único agente. O modelo identifica dois grupos, as entidades coletivas (*Corporate Body*) e as famílias (*Family*). Uma entidade coletiva (*Corporate Body*) é um grupo organizado de pessoas que atua em conjunto como um agente, e que tem posição social ou legal reconhecida, como por exemplo, uma organização produtora de documentos na qual seus membros, munidos de cargos (papéis), geram estes documentos. Já famílias (*Family*) são compostas por duas ou mais pessoas relacionadas por nascimento, ou através de casamento, adoção, união civil ou outras convenções sociais que os unam como um grupo familiar socialmente reconhecido.

Além de pessoas (*Person*), grupo (*Group*), entidades coletivas (*Corporate Body*) e famílias (*Family*), RiC-CM também introduz o conceito de mecanismo (*Mechanism*). Na atualidade, muitos documentos em órgãos públicos, como Certidões Negativas de Débito, são gerados por sistemas automatizados, que, a partir dos procedimentos especificados por pessoas em linguagem de programação, buscam as informações em bases de dados e geram documentos. RiC-CM observa a produção documental no contexto de processos ou sistemas criados por pessoas. Em RiC-CM um

mecanismo (*Mechanism*), como um sistema de software ou mecânico, age no mundo produzindo efeitos físicos e sociais, e frequentemente gera ou modifica documentos (*Record*). Um mecanismo é um agente (*Agent*), para o modelo, com isso pode adquirir a característica de ser um gerador de documentos.

Entretanto, há diferença entre gerar documento e ser o produtor de um fundo documental, embora ambos tratem de proveniência. Uma posição (*Position*) de uma pessoa (*Person*) em um grupo (*Group*) é, em RiC-CM, um papel funcional desempenhado por ela. Em uma posição (*Position*), uma pessoa pode ter a autoridade para emitir o documento e em cujo nome ou sob cuja ordem ou responsabilidade o documento é emitido. Essa autoridade de autor lhe é conferida na medida que esta é detentora de uma posição, como a função ou cargo exercido. São os normativos jurídicos que definirão as responsabilidades dos cargos e funções. RiC-CM permite a representação de que uma pessoa é autora de um documento (*AuthorshipRelation*), e da posição que esta possui em uma entidade coletiva (*PositionHoldingRelation*).

3 O documento de arquivo em *Basic Formal Ontology* e ontologias de documentos

Desenvolvida como uma ontologia realista, BFO tem como premissa filosófica que a representação ontológica ocorre a partir do conhecimento de entidades que existem no mundo real (ARP; SMITH; SPEAR, 2015). Em BFO essa realidade pode ser classificada em ocorrentes (*occurrents*), aquilo que ocorre em um momento específico no tempo, e continuantes (*continuants*), aquilo que tem continuidade no tempo. Ocorrentes são ações pontuais no tempo e no espaço, são eventos, processos, ou o ato de assinar um documento, por exemplo. Continuantes, por outro lado, são estáveis no tempo, como objetos materiais ou imateriais, e características que existem no mundo, como um documento, ou o cargo de diretor em uma empresa, por exemplo.

Desenvolvida para servir como ontologia de topo no âmbito da pesquisa científica, especialmente na área da saúde, a BFO foi especializada em diferentes ontologias de domínio e teve uma ampla aceitação, para além da área da saúde. Algumas dessas ontologias se preocuparam com a representação dos documentos e objetos informacionais. Para a CI, o documento é conceito essencial para a compreensão e representação da informação materializada em um suporte. Considerando que os documentos são registros do fazer humano em dado contexto de produção informacional, é relevante entender como as ontologias existentes representam o documento.

Analizamos duas ontologias derivadas de BFO voltadas para documentos e objetos informacionais, a fim de

compreender como o documento, enquanto conceito, está representado nessas ontologias.

A IAO é uma ontologia de entidades de informação, derivada do trabalho da *Ontology for Biomedical Investigations* (OBI) com o desenvolvimento das classes para representação de objetos digitais e entidades informacionais. Segundo Smith e Ceusters (2015), a IAO

[...] foi criada para servir como um recurso de domínio neutro para a representação de tipos de entidades de conteúdo informacional (ICE) como documentos, bases de dados e imagens digitais (SMITH; CEUSTERS, 2015, p. 1, tradução nossa).

A ontologia D-Acts, que em português poderia ser traduzida como Ontologia dos Atos Documentais, originou-se na teoria de mesmo nome, proposta por Barry Smith e na teoria dos atos sociais e entidades legais de Adolf Reinach (SMITH, 2008, 2011). Está fundamentada na concepção de que os documentos possuem estabilidade e, portanto, podem ser preservados no tempo. Os documentos são, para o autor, a materialização de ações que ocorrem no mundo real, cujo registro atesta essa ação. Ao mesmo tempo, o fato de a ação ser registrada traz consequências sociais que impactam o mundo real. O mesmo documento pode servir a diferentes atos sociais ao longo de sua existência e ter diferentes destinatários. (SMITH, 2011; SMITH; CEUSTERS, 2015)

Segundo Smith (2014), documentos são objetos, portanto, continuantes, enquanto os atos de fala são eventos, ou seja, ocorrentes, conforme a BFO. Isso significa que os documentos se mantêm idênticos através do tempo e têm a capacidade de se desprender da pessoa ou pessoas envolvidas em sua criação, adquirindo vida própria. Os documentos também podem ter vários criadores, os quais podem fazer suas contribuições para o documento em diferentes momentos. Documentos legais e administrativos podem incluir partes a serem preenchidas em momentos diferentes, por exemplo, quando sucessivas decisões foram tomadas ou uma sequência de reuniões foi realizada ou quando o documento foi visto por indivíduos sucessivos. Os documentos também podem crescer pela anexação de apêndices ou incorporação real ou virtual de outros documentos, de objetos textuais como referências cruzadas (apontando para outras partes de um único documento) ou notas de rodapé e citações (apontando para o que está fora).

Para Coelho e Almeida (2017), do ponto de vista filosófico:

A diferença de visão entre os autores da teoria dos atos de fala (Searle, 1969) e a teoria dos atos dos documentos (Smith, 2005) é que os atos da fala são declarações que existem somente no momento de sua execução, e os documentos persistem através do tempo mantendo sua identidade e, portanto, são eles que possibilitam a

permanência dos compromissos assumidos nas declarações. Desse modo, entende-se que a teoria dos atos de fala fornece uma explicação sobre como as entidades sociais começam a existir e os documentos são a base física para a existência temporal dessas entidades e de seu poder duradouro. (COELHO; ALMEIDA, 2017, p. 73–74)

A principal ontologia a tratar do documento e da produção documental a partir de BFO é a ontologia IAO. Criada como uma ontologia de entidades de informação, para esta análise consideramos os seguintes termos: entidade de conteúdo informacional (*Information Content Entity - ICE*), documento (*Document*), portador material da informação (*Material Information Bearer*) e carregador da informação (*Information Carrier*). Relacionamentos entre entidades de conteúdo informacional e processos que consomem ou produzem tais entidades também são importantes para representação do contexto.

É possível perceber que a informação é tratada nesta ontologia como algo que está contido em alguma outra entidade e não algo que possui existência *per se*. Isto porque não há uma entidade nomeada como informação. A informação está contida em uma entidade de conteúdo informacional (ICE). Esta entidade é sobre (*is about*) alguma coisa da realidade. Para IAO um documento pode ser representado como uma ICE, ou seja, um continuante genericamente dependente (*Generically Dependent Continuant*) de um portador material de informação (*Material Entity*), ou seja, uma entidade material como uma folha de papel.

A entidade de conteúdo informacional (ICE) é definida como um continuante genericamente dependente (*Generically Dependent Continuant*) que é sobre alguma coisa. A relação “é-sobre” (*is about*), pode ser definida como a porção de realidade que é denotada ou está referida na entidade de conteúdo informacional, mas que não se confunde com esta (SMITH; CEUSTERS, 2015). Um contrato, como visto, registra um acordo, ocorrido no mundo real entre as partes. O contrato não é, propriamente, um acordo, e sim, o registro desse acordo.

BFO define, também, o que é um continuante genericamente dependente: um continuante (*Continuant*) que é dependente de um ou outros continuantes independentes (*Independent Continuant*) que servem como seu portador (ARP; SMITH; SPEAR, 2015). A relação de dependência genérica estabelecida na ontologia significa que sua existência só é possível através de uma outra entidade material que lhe dê suporte, entretanto, essa entidade material pode mudar sem que a ICE deixe de existir.

ICE é um continuante genericamente dependente. Isso permite que a ICE seja migrada ou transmitida de uma entidade material para outra, por exemplo. Assim, a informação contida no contrato em papel pode ser

migrada para outro suporte, como um arquivo pdf, sem que a informação contida nele sofra perdas.

Na perspectiva arquivística, a separação entre ICE e a entidade material (*Material Information Bearer*) é conhecida e explicitada na distinção que é feita entre conteúdo e suporte do documento. Logo, há uma dependência existencial genérica entre o documento e o suporte. Deixando de existir o suporte, o documento também cessa sua existência. Entretanto, um documento pode mudar de suporte, nesse caso, continua existindo.

A entidade material que dá suporte para a ICE, como o papel do contrato, é chamada de portador material da informação (*Material Information Bearer*), definido como uma entidade material (*Material Entity*) na qual uma concretização de uma ICE pode existir. Sendo uma entidade material, é um continuante independente (*Independent Continuant*), que em BFO significa que sua identidade e existência são mantidas ainda que haja mudança em suas qualidades ou partes (ARP; SMITH; SPEAR, 2015). Isso significa que a folha de papel continua uma folha de papel, ainda que marcas de tinta tenham sido inscritas em sua superfície.

Em IAO, para além destes há a entidade que carrega a informação (*Information Carrier*), entendida como uma qualidade, ou seja, um continuante especificamente dependente, tais como as marcas de tinta em um manuscrito.

O documento (ICE) é concretizado em uma folha de papel (*Material Information Bearer*) através de um carregador da informação (*Information Carrier*), que, neste caso, seriam as marcas de tinta com letras e números. Para BFO, tais marcas de tinta são uma qualidade (*Quality*) daquela folha de papel, isso significa que as marcas possuem dependência específica de uma determinada entidade material, isto é, daquela folha de papel.

Logo, há uma dependência existencial específica entre o carregador da informação (como as marcas de tinta) e a entidade material (como o papel em que as marcas foram registradas). Deixando de existir o suporte, a primeira também cessa sua existência. Assim, um carregador de informações, comunica as informações de um documento em um portador específico, e não de qualquer documento, o qual, por sua vez, depende de um suporte material qualquer para existir (LÖW *et al.*, 2022).

Podemos deduzir que a informação contida em um documento (*Document*) como um contrato, por exemplo, é sobre (*is about*) uma ação ocorrida entre duas ou mais partes, as quais se comprometem na realização de alguma coisa no mundo real, expresso através da linguagem escrita como conteúdo do documento contrato. Esse contrato precisa ser materializado em algum suporte, como uma folha de papel (*Material Information Bearer*). As marcas de tintas com os sinais

gráficos da linguagem escrita (*Information Carrier*) e as assinaturas estarão inscritas naquela folha de papel e, ao concretizar a mensagem do contrato no suporte, poderão comunicar seu conteúdo.

Um documento é também resultado de uma ação, realizada a partir de funções e atividades do produtor. Em BFO, ações como essas são representadas com a entidade processo (*Process*), um ocorrente que existe no tempo por ocorrência ou acontecimento, tem partes temporais e depende de uma entidade material. Processos realizam entidades realizáveis, as quais são inerentes a continuantes independentes. Papéis (*Role*) são entidades realizáveis que se realizam através de processos e se concretizam em entidades materiais como organizações ou pessoas, no caso de um produtor de arquivo, por exemplo, ou de um diretor que assina um documento.

Para BFO ser capaz de gerar um documento, ou poder assiná-lo, ou, ainda, ser um produtor, no sentido arquivístico, podem ser representados, no modelo, como entidades realizáveis (*Realizable Entity*) em processos específicos do qual um portador (uma entidade material) participa.

Um produtor de documentos arquivísticos, em BFO, seria representado através de uma entidade realizável, que é inerente a uma organização (entidade material) na realização de seu negócio (processo), isto é, a organização adquire o papel de produtor. Um diretor seria uma entidade realizável que é inerente a uma pessoa (entidade material) na realização de atividades de negócio (processos) dessa organização. Esse diretor é dependente especificamente de uma pessoa. Essa dependência é específica pois o papel de diretor daquela pessoa pode deixar de existir para aquela pessoa, sem que a pessoa ao qual é inerente deixe de existir, e não pode mudar de portador, isto é, não pode existir em outra pessoa.

Em BFO, papel (*Role*) é uma entidade realizável que é possuída por seu portador devido a circunstâncias externas, físicas, sociais ou institucionais, e que pode deixar de existir sem que a composição física de seu portador seja alterada. Produtor e diretor são papéis, pois se deixarem de existir, a composição física de seus portadores (organização e pessoa) não é alterada. Ser nomeado diretor é uma circunstância externa que permite a realização do papel de diretor a uma determinada pessoa.

Um produtor, ou um diretor, não “é” alguém, mas sim, um papel (*Role*) que alguém ou alguma coisa executa em um determinado momento, por circunstâncias que são externas a si. Um diretor é alguém que em determinados momentos e com determinadas circunstâncias executa um papel de “ser diretor”. Dito de outra maneira: àquela pessoa específica, em um papel específico, como diretor, atribuído a ela, coube assinar o documento. Uma

pessoa (*Homo Sapiens*) pode ser uma entidade material que eventualmente ocupa um papel de produtor. Uma organização (*Organization*) também pode ocupar um papel de produtor.

Quanto à representação de contexto e dos conjuntos documentais, BFO não possui entidades especializadas para o agrupamento de documentos como ocorre nos arquivos. Entretanto, a ontologia possibilita que entidades sejam agrupadas por critérios determinados em agregados de entidades. A relação membro de (*member of*) estabelece uma relação meriológica entre um item e uma coleção, ou agregação de membros. Nesse sentido os documentos arquivísticos, que em BFO são entidades de conteúdo informacional, podem ser reunidos em conjuntos agregados relacionando os membros de cada conjunto como pertencentes àquele conjunto. Os documentos de uma mesma tipologia, como contrato, por exemplo poderão ser agregados como um conjunto, ou documentos relativos a uma mesma tarefa ou atividade. (LÖW *et al.*, 2022).

4 Resultados

A partir das análises feitas em cada um dos modelos ontológicos, foi estruturado um quadro analítico com as entidades e uma análise de semelhanças e diferenças considerando os seguintes critérios: entidades para representação do documento, entidades para representação do conjunto documental; entidades para representação do suporte; entidades para representação da ação geradora do documento; entidades para representação do produtor. O quadro analítico está incluído na íntegra no apêndice deste texto.

Observa-se um alinhamento entre as ontologias, com algumas estratégias de modelagem diferentes. Com relação ao documento e a materialização de sua informação em um suporte, ambos apresentam a distinção entre documento e suporte. Em RiC, Record é comparável a uma unidade de descrição do documento (INTERNATIONAL COUNCIL ON ARCHIVES, 2021), visto como uma entidade intelectual. Já instancicações (*Instantiation*) são as representações desta entidade intelectual. Em RiC um Record não existe até ser representado por pelo menos uma instanciação, mas possuir uma instanciação não é uma condição necessária para a manutenção da existência do Record. Em BFO, o documento (ICE) é genericamente dependente do suporte (*Material Information Bearer*), isto é, um documento depende de um suporte para existir. Nesse caso, o registro de um documento eliminado, em BFO, poderia ser feito pela especificação de uma outra entidade, como o perfil do documento ou seus metadados, por exemplo.

BFO especifica, ainda, o carregador da informação (como as marcas de tinta) de um documento e estabelece a dependência existencial deste carregador a um suporte

(como o papel deste documento). Já RiC permite a categorização da instanciamento do documento com relação ao material físico (*CarrierType*) e ao método usado para registrar a informação nesta instanciamento (*RepresentationType*).

Em RiC, pessoas, entidades corporativas e grupos são tipos de agentes, que são entidades que atuam ou realizam atividades no mundo e, no decorrer da execução de atividades, geram registros. Posição (*Position*), como diretor, é um tipo de agente que representa uma função em um grupo. Posição pode ser independente da pessoa que a exerce e pode ser usada para identificar os conjuntos de documentos resultantes das atividades realizadas por pessoas que ocupam ou ocuparam a posição.

Já em BFO, pessoas, entidades corporativas e grupos seriam Continuantes Independentes; portanto, subsumiriam agentes, como em RiC. Em BFO, a condição para uma pessoa, entidade corporativa ou grupos realizar uma ação ou atividade seria especificada por meio de uma Entidade Realizável. Em BFO, essa entidade realizável (como o diretor José) seria especificamente dependente do seu portador (como a pessoa José), pois esse papel de diretor pode deixar de existir em José sem que José deixe de existir e não pode existir em outra pessoa.

Como resultado da análise é possível verificar semelhanças e divergências na representação do documento arquivístico entre os dois modelos ontológicos. O modelo do ICA apresenta entidades construídas especificamente para o contexto da descrição arquivística. Apesar de suas entidades serem mais específicas, seu modelo é pequeno e não está fundamentado em nenhuma ontologia de topo, o que pode trazer dificuldades na implementação em sistemas informacionais complexos. As entidades derivadas de BFO, por outro lado, apesar de estarem fundamentadas em uma ontologia consolidada e utilizada em diferentes domínios, não foram especializadas para abarcar as especificidades da produção documental arquivísticas.

5 Considerações Finais

A partir de uma análise conceitual das entidades de RiC-CM e algumas de RiC-O foi possível compreender a semântica da representação do documento arquivístico nesse modelo ontológico. Foram analisados os conceitos acerca do documento e da sua produção, como expressados no modelo. Comparativamente foi feita uma análise de entidades de ontologias derivadas de BFO que poderiam representar o documento e sua produção, a partir de uma perspectiva arquivística.

A partir de um quadro analítico das entidades observa-se um alinhamento entre as ontologias, com algumas

estratégias de modelagem diferentes. Apesar disso, ambos os modelos parecem adequados para a representação do documento arquivístico. Entretanto as especificidades do documento arquivístico, como a representação do seu contexto de produção, precisam estar presentes em um sistema ontológico.

Com a análise conceitual realizada é possível concluir que os estudos sobre representação do conhecimento arquivístico com base em ferramentas ontológicas precisam ser expandidos de forma a possibilitar a representação das especificidades acerca da proveniência arquivística e do contexto de produção.

Notas

- [1] A pesquisa aqui apresentada foi desenvolvida, em parte, durante o mestrado em Ciência da Informação de um dos autores.

Referências

- ARP, Robert; SMITH, Barry; SPEAR, Andrew D. **Building Ontologies with Basic Formal Ontology**. Cambridge, MA: The MIT Press, 2015. *E-book*. Disponível em: <https://direct.mit.edu/books/book/4044>. Acesso em: 27 mar 2023.
- BARROS, Thiago Henrique Bragato; BASTOS, Carine Melo Cogo; DOS SANTOS, Adriana Carla Ribeiro. Sistemas de organização do conhecimento no contexto da arquivologia: aportes metodológicos para seu desenvolvimento. **Acervo**, [s. l.], v. 35, n. 2 SE-Dossiê, p. 1–20, 2022. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/1812>. Acesso em: 17 abr. 2023
- BARROS, Thiago Henrique Bragato; GOMES, Daniel Libonati. Classification and Knowledge Organization Systems: ontologies and archival classification. **Challenges and Opportunities for Knowledge Organization in the Digital Age**, [s. l.], p. 103–111, 2018.
- BARROS, Thiago Henrique Bragato; SOUSA, Renato Tarciso Barbosa de. Archival science and knowledge organization: Mapping methodological relationships. **Knowledge Organization**, [s. l.], v. 46, n. 7, p. 493–501, 2019.
- COELHO, Kátia Cardoso; ALMEIDA, Maurício Barcellos. Atos Dos Documentos E Sua Natureza Teórico-Filosófica. **Brazilian Journal of Information Science**, [s. l.], v. 11, n. 3, p. 66–78, 2017.
- CONSELHO NACIONAL DE ARQUIVOS. **Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos e-Arq Brasil (E-ARQ V2 2020)**. Rio de Janeiro: [s. n.], 2020. Disponível em: https://www.gov.br/conarq/pt-br/assuntos/noticias/conarq-abre-consulta-publica-visando-a-Atualizacao-do-e-arq-brasil/EARQ_v2_2020_final.pdf. Acesso em: 21 mar. 2023
- GRUBER, Thomas R. Toward principles for the design of ontologies used for knowledge sharing?. **International Journal of Human-Computer Studies**, [s. l.], v. 43, n. 5, p.

907–928, 1995. Disponível em:
<https://www.sciencedirect.com/science/article/pii/S1071581985710816>.

GUARINO, Nicola. **Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy.** [S. l.]: IOS press, 1998. v. 46

HEREDIA HERRERA, Antonia. **Archivística general: teoría y práctica.** 5 ed. Sevilla: Servicio de Publicaciones de la Diputación de Sevilla, 1991.

INTERNATIONAL COUNCIL ON ARCHIVES. **International Council on Archives Records in Contexts Ontology (ICA RiC-O) version 0.2.** [S. l.], 2021. Disponível em: https://www.ica.org/standards/RiC/RiC-O_v0-2.html. Acesso em: 11 abr. 2023.

INTERNATIONAL COUNCIL ON ARCHIVES. **Records in Contexts: a conceptual model for archival description (Consultation Draft v0.2).** [S. l.], 2021. Disponível em: https://www.ica.org/sites/default/files/ric-cm-02_july2021_0.pdf. Acesso em: 09 mai. 2023.

LÖW, Marieta Marks *et al.* Ontologia e documento arquivístico: análise ontológica para representação semântica do documento arquivístico em BFO. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da**

informação, [s. l.], v. 27, n. 1, p. 1–27, 2022. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/83536>. Acesso em: 23 mar. 2023.

MULLER, S.; FEITH, J. A.; FRUIN, R. **Manual de arranjo e descrição de arquivos.** 2A. ED.ed. Rio de Janeiro: Arquivo Nacional Rio de Janeiro, 1973.

SMITH, Barry. Document Acts. *In*: KONZELMANN-ZIV, A.; SCHMID, H. B. (org.). **Institutions, emotions, and group agents: contributions to social ontology.** Dordrecht: Springer, 2014. p. 19–31. Disponível em: https://doi.org/10.1007/978-94-007-6934-2_2. Acesso em: 20 out. 2020.

SMITH, Barry. Searle and de Soto: The New Ontology of the Social World. *Em*: SMITH, Barry; MARK, David; EHRLICH, Isaac (org.). **The Mystery of Capital and the Construction of Social Reality.** Chicago: Open Court, 2008. p. 35–51.

SMITH, Barry. The Ontology of Documents. *In*: OKADA, Mitsuhiro (org.). **Proceedings of the Conference on Ontology and Analytical Metaphysics**, February 24-25, 2011. [S. l.]: Keio University Press, 2011.

SMITH, Barry; CEUSTERS, Werner. Aboutness: Towards foundations for the information artifact ontology. *In*: **CEUR Workshop Proceedings.** [S. l.]: CEUR-WS, 2015.

Apêndice – Quadro comparativo de entidades de documentos em RiC e BFO

<i>Conceito</i>	<i>Entidades em RiC</i>	<i>Entidades em BFO</i>	<i>Análise comparativa</i>
Documento	<i>Record;</i> <i>Record Resource;</i> <i>Record Part;</i>	<i>Information Content Entity;</i> <i>Document;</i> <i>Information Carrier;</i>	Enquanto em RiC o documento, suas partes e até o conjunto de documentos é considerado uma fonte documental, uma evidência das atividades de um Agente. Para as ontologias de BFO o documento é um conjunto de entidades de informação que devem ser entendidos como um todo.
Conjunto documental	<i>Record Set;</i>	Relacionamento do tipo <i>member of;</i>	Em RiC há uma entidade específica para representar conjuntos documentais. Em BFO essa representação é através de um relacionamento que reúne instâncias de uma entidade.
Suporte	<i>Instantiation;</i> <i>Carrier;</i> <i>CarrierType;</i>	<i>Material Entity;</i> <i>Material Information Bearer;</i>	A representação da materialidade da informação em um suporte é feita pela instanciação do documento em um suporte, ou carregador. Em BFO essa materialidade é mais complexa, uma vez que o documento e a informação não são entidades materiais, assim, a informação é inscrita como uma qualidade na entidade material.
Ação geradora	<i>Event;</i> <i>Activity;</i> <i>Rule;</i> <i>Mandate;</i> <i>Date;</i> <i>Place;</i>	<i>Process;</i> <i>Planned Process;</i> <i>Plan;</i> <i>Directive information Entity;</i> <i>Plan Specification;</i>	E ambas as ontologias a ação geradora do documento é uma atividade desenvolvida pelo produtor, que acontece em determinado momento no tempo e no espaço. Tais atividades podem ser reguladas com normas ou procedimentos.
Produtor	<i>Agent;</i> <i>Person;</i> <i>Group Family;</i> <i>Corporate Body;</i> <i>Position;</i> <i>Mechanism;</i>	<i>Homo Sapiens;</i> <i>Organization;</i> <i>Role;</i> <i>Author Role;</i>	A forma de representação do produtor é diferente nos dois modelos ontológicos. Para RiC a ação é realizada por um agente, pessoa, família, instituição ou mecanismo, com uma dada posição ou cargo. Em BFO há uma distinção entre a pessoa que registrou o documento e o papel que ela ocupa para poder produzir o documento.

Fonte: Dados da pesquisa.

Sistemas de Organização do Conhecimento no contexto da ISKO: uma análise de domínio dos capítulos Brasileiro e Norte-Americano

Knowledge Organization Systems (KOS) in the context of ISKO: a domain analysis of the Brazilian and North American chapters

Maria Amália Cassol Lied (1), Waleska da Silva Rocha (2), Thiago Henrique Bragato Barros (3), Kamila de Andrade Moura (4)

Universidade Federal do Rio Grande do Sul (UFRGS), Rua Ramiro Barcelos, 2705 - Santana, Porto Alegre

Resumo

Objetivo: Tem como objetivo mapear, analisar temática, semântica e discursivamente os artigos publicados na área de Organização do Conhecimento, dentro dos cinco volumes publicados dos eventos da *International Society of Knowledge Organization* Brasil (ISKO-Brasil) e os oito volumes da *North American Symposium on Knowledge Organization* (NASKO). **Metodologia:** A metodologia utilizada foi a realização de um levantamento de enunciados-pivô preestabelecidos da produção científica publicada nos anais de ambos os eventos com o auxílio do *software Sketch Engine* como ferramenta e as abordagens 6 – Estudos históricos de estruturas e serviços de informação em domínios, 8 – Estudos epistemológicos e críticos de diferentes paradigmas, pressupostos e interesses em domínios, 10 – Estudos de estruturas e instituições de comunicação científica e profissional num domínio, de Hjørland (2017), e 13 – Análise do discurso da análise de domínio como aporte metodológico, de Smiraglia (2015). **Resultados:** O levantamento resultou em 131 artigos para ISKO-Brasil e 132 documentos para NASKO. Verificou-se que, mesmo que os dois *corpora* analisados estejam no escopo da ISKO, há divergências quanto ao entendimento de conceitos, bem como a relação destes com a discussão epistemológica da área; e convergências que dizem respeito aos conceitos de: 1) Análise de Domínio; 2) Sistemas de Organização; 3) Teoria do Conceito; 4) Sistemas de Classificação. **Conclusão:** A análise possibilitou vislumbrar a Organização do Conhecimento como uma área teórica e aplicada, com bases na Teoria do Conceito e na Análise de Domínio.

Palavras-chave: ISKO; Análise de domínio; Sistemas de Organização do Conhecimento; *Sketch Engine*.

Abstract

Objective: It aims to map, analyze thematically, semantically and discursively the articles published in the area of Knowledge Organization, within the five published volumes of the events of the *International Society of Knowledge Organization* Brazil (ISKO-Brazil) and the eight volumes of the *North American Symposium on Knowledge Organization* (NASKO). **Methodology:** The methodology used was to conduct a survey of pre-established pivot statements of the scientific production published in the annals of both events with the help of *Sketch Engine* software as a tool and approaches 6 – Historical studies of structures and services of information in domains, 8 – Epistemological and critical studies of different paradigms, assumptions and interests in domains, 10 – Studies of structures and institutions of scientific and professional communication in a domain, from Hjørland (2017), and 13 – Discourse analysis of domain analysis as methodological contribution, from Smiraglia (2015). **Results:** The survey resulted in 131 articles for ISKO-Brazil and 132 documents for NASKO. It was found that, even though the two *corpora* analyzed are within ISKO's scope, there are divergences regarding the understanding of concepts, as well as their relationship with the epistemological discussion of the area; and convergences concerning the concepts of: 1) Domain Analysis; 2) Organization Systems; 3) Concept Theory; 4) Classification Systems. **Conclusion:** The analysis made it possible to envision Knowledge Organization as a theoretical and applied area, based on Concept Theory and Domain Analysis.

Keywords: ISKO; Domain analysis; Knowledge Organization Systems; *Sketch Engine*.

Resumen

Objetivo: Tiene como objetivo mapear, analizar temática, semántica y discursivamente los artículos publicados en el área de Organización del Conocimiento, dentro de los cinco volúmenes publicados de los eventos de la *International Society of Knowledge Organization* Brasil (ISKO-Brasil) y los ocho volúmenes de la *North American Symposium on Knowledge Organization* (NASKO). **Metodología:** La metodología utilizada fue realizar un relevamiento de enunciados pivote preestablecidos de la producción científica publicada en los anales de ambos eventos con la ayuda del *software Sketch Engine* como herramienta y enfoques 6 – Estudios históricos de estructuras y servicios de información en dominios, 8 – Estudios epistemológicos y críticos de diferentes paradigmas, supuestos e

intereses en dominios, 10 – Estudios de estructuras e instituciones de comunicación científica y profesional en un dominio, de Hjørland (2017), y 13 – Análisis discursivo del análisis de dominio como aporte metodológico, de Smiraglia (2015). **Resultados:** La encuesta resultó en 131 artículos para ISKO-Brasil y 132 documentos para NASKO. Se constató que, si bien los dos *corpora* analizados están en el ámbito de ISKO, hay divergencias en cuanto a la comprensión de los conceptos, así como su relación con la discusión epistemológica del área; y convergencias relativas a los conceptos de: 1) Análisis de Dominio; 2) Sistemas de Organización; 3) Teoría de los conceptos; 4) Sistemas de Clasificación. **Conclusión:** El análisis permitió visualizar la Organización del Conocimiento como un área teórica y aplicada, basada en la Teoría de Conceptos y el Análisis de Dominio.

Palabras clave: ISKO; Análisis de dominio; Sistemas de Organización del Conocimiento; *Sketch Engine*.

1 Introdução

Este trabalho busca identificar e comparar as comunidades discursivas da Organização do Conhecimento (OC) dos capítulos da *International Society for Knowledge Organization* (ISKO), a partir da análise dos anais dos seguintes eventos: Congresso Brasileiro em Organização e Representação do Conhecimento (ISKO-Brasil) (2011-2019) e *North American Symposium on Knowledge Organization* (NASKO) (2007-2021). Assim, o objetivo da pesquisa é mapear, analisar temática, semântica e discursivamente os artigos publicados na área de Organização do Conhecimento, dentro dos cinco volumes publicados dos eventos da ISKO-Brasil e os oito volumes da NASKO. Para tal, a pesquisa se amparou em trabalhos dos autores Barros e Laipelt (2021) e Oliveira, Silva, Barros e Moura (2022), por se tratarem de referências norteadoras para pesquisas do gênero.

A Conferência da NASKO ocorreu pela primeira vez em 2007, com um total de oito edições desde então. O Congresso Brasileiro em Organização do Conhecimento teve seu início quatro anos mais tarde, contabilizando um total de seis eventos. Ambos permanecem regulares e ocorrem em formato bianual, aproximadamente. Observando a relevância dos eventos no cenário da ISKO, bem como a contribuição ativa da sua comunidade na produção científica da área, nesse contexto denota-se comunidades discursivas passíveis de análise para constatar como suas atuações reverberam nas conceituações e percepções da OC e dos SOC.

Ressalta-se que este trabalho constitui um estudo maior que busca analisar as tendências da literatura em Sistemas de Organização do Conhecimento (SOC) no âmbito da ISKO como um todo, financiado pelo CNPq.

2 Organização do Conhecimento

O objetivo da Organização do Conhecimento é fornecer acesso à informação em todos os campos da compreensão ou atividade humana, tanto para aqueles dentro de qualquer campo quanto aqueles além dele. O conhecimento é organizado, afinal, de modo que pode ser usado (mesmo o cientista em um laboratório colocando seus dados em alguma ordem não assim com

a intenção de voltar a essa informação). A partir desta declaração de propósito, podemos derivar várias teses sobre como o campo deve operar.

Dito isso, podemos compreender a OC como um campo associado a Ciência da Informação que se dedica a estudar processos de representação e sistematização de conceitos, obtendo como produtos os Sistemas de Organização do Conhecimento (SOC) (ontologias, taxonomias, vocabulários controlados, tesouros, entre outros) e, através destes, irá estruturar conceitos e suas relações dentro de uma área do conhecimento, construindo modelos para representação e organização da informação (BRASCHER; CAFÉ, 2010; BARROS; LAIPELT, 2021, p. 40). Dada sua importância, alguns dos enunciados-pivôs selecionados para esta análise são SOC, a saber: ontologias, tesouros e taxonomias. Sua aplicação se dá de forma distinta em diversos contextos sociais e estes possuem uma variedade de fundamentos tecnológicos que viabilizam a Organização do Conhecimento através da gestão e do acesso (BARROS, 2021, p. 64).

Conforme Barros (2021, p. 62), a análise através do discurso não delimita estritamente os termos a uma perspectiva conceitual – como costuma ser abordado tradicionalmente na OC –, pois, enquanto “aporte teórico”, viabiliza uma análise histórica, social e ideológica. Através dessa perspectiva, compreende-se que os documentos levantados para o trabalho vão além da necessidade técnica de observar o desenvolvimento terminológico dessas comunidades discursivas, bem como estão articuladas social e epistemologicamente suas tendências conceituais.

Uma vez que a delimitação e compreensão do assunto tratado influencia significativamente o andamento da pesquisa, delimitou-se os eventos e os enunciados-pivô a fim de compreender como as concepções a respeito da representação e dos SOC se desenvolvem dentro dos eventos brasileiro e norte-americano. Para tal, utilizou-se o *software Sketch Engine* como ferramenta e as abordagens 6, 8, 10 e 13 da análise de domínio como aporte metodológico.

3 Análise de Domínio enquanto aporte metodológico

A análise de domínio é uma abordagem formulada inicialmente por Hjørland e Albrechtsen (1995), que

entende a importância de estudar os domínios do conhecimento através de sua linguagem, estrutura e comunidades, uma vez que estas compartilham teorias, terminologias e paradigmas próprios. Significa analisar a atuação da CI através de perspectiva social, voltada para os contextos em que a informação circula, e os domínios especializados são terrenos férteis para entender como estas dinâmicas acontecem (LOPEZ-HUERTAS, 2015; SMIRAGLIA, 2015).

Hjørland (2002; 2017) elencou 11 abordagens para aplicação da análise de domínio, provendo ampla cobertura de análise dentro do contexto da CI. Para fins deste trabalho, as abordagens utilizadas são: 6 – *Estudos históricos de estruturas e serviços de informação em domínios*, no intuito de melhor compreender o domínio, bem como sua estrutura e organização; 8 – *Estudos epistemológicos e críticos de diferentes paradigmas, pressupostos e interesses em domínios*, no sentido de mapear como se dá a construção epistemológica da área; 10 – *Estudos de estruturas e instituições de comunicação científica e profissional num domínio*, para compreender como se configura o ciclo informacional do domínio analisado; e, com base em Smiraglia (2015), a abordagem 13 – *Análise do discurso*, para compreender o discurso institucionalizado no domínio.

4 Sketch Engine

Sketch Engine é um *software online* utilizado para análise linguística de textos através da utilização de *corpora* textuais (amostras de linguagens). Por meio de seu uso, é possível verificar o que é representativo, raro e obsoleto em um *corpus*.

Para este estudo, foram utilizados os seguintes recursos do sistema:

- a) *Word Sketch* (para análise do comportamento gramatical e colocacional da palavra);
- b) *Word Sketch Difference* (para fazer comparações entre colocações contrastantes);
- c) *Keywords* (para extração terminológica);
- d) e *Wordlist* (para análise de frequência).

Destaca-se que, para os interesses específicos desse trabalho, a busca foi realizada por *lemma*. Um *lexema* (*lemma* ou *lexeme*, em inglês), conforme Booij (2005, p. 3), é uma palavra vista através do seu “sentido abstrato”, isto é, tem-se uma noção ampla das suas variações morfológicas. Essas variações, por sua vez, são chamadas de *word forms* (“formas da palavra”, em português) que, segundo o autor (2005, p. 3), podem ser qualificadas como “sentido concreto”. Por exemplo, “As palavras concretas *andam*, *anda*, *andou*, *andando* podem ser qualificadas como formas da palavra do *lexema* [andar]” (BOOIJ, 2005, p. 3).

Realizou-se um recorte padrão de utilização dos primeiros 10 termos de cada material extraído, uma vez que não era viável apresentar todos os termos contidos nas listas e gráficos gerados pela ferramenta. Quando constatada a pouca influência do enunciado dentro de cada *corpus*, houve a redução do recorte.

Tendo em consideração que as duas comunidades discursivas analisadas – ISKO-Brasil e NASKO – possuem línguas diferentes, tal distinção apontada se constituiu em um fator determinante na busca e análise dos termos no *Sketch Engine*. Enquanto a ISKO-Brasil apresenta o português brasileiro como sua língua de maior uso (o qual é o recorte deste trabalho, uma vez que também há artigos em língua inglesa e espanhola); a NASKO tem em sua comunidade falantes da língua inglesa predominantemente norte-americana (algumas variações linguísticas advindas do inglês britânico foram constatadas). Tendo em vista que cada comunidade demonstrou ter seus próprios comportamentos gramaticais, não foi viável aplicar a mesma estratégia de busca para ambas.

A seguir, para compreensão da análise realizada, faz-se necessária a descrição da função e aplicabilidade de cada um dos recursos anteriormente elencados.

O *Word Sketch* (“esboço de palavras”, em português) permite verificar as relações de uma palavra com as outras, através de um resumo do seu comportamento colocacional e gramatical. Diversas categorias denominadas pelo *software* como relações gramaticais podem ser verificadas (SKETCH ENGINE, 2023).

Utiliza-se o *Word Sketch Difference (WSD)* (“diferenças em esboços de palavras”, em português) para estabelecer comparativamente dois *lexemas*, a fim de analisar as suas colocações e respectivas relações gramaticais (SKETCH ENGINE, 2023).

Dentre as diversas possibilidades do recurso *Keywords*, isto é, “palavras-chave” em português, destacamos a extração por *single-words* (“palavras-únicas” em português) que realiza um levantamento de termos constituídos de uma palavra e *multi-word terms* (“termos com múltiplas palavras”, em português) que lista termos que formam uma expressão. Ambas realizam extrações que são típicas de um *corpus*, documento, texto ou que definem seu conteúdo ou tópico (SKETCH ENGINE, 2023).

A *Wordlist* (“listas de palavras”, em português) é um recurso que gera listas de frequências de termos de diversos tipos (SKETCH ENGINE, 2023), voltando-se, portanto, para as métricas (quantas vezes que determinado termo aparece no *corpus*).

Por tratar-se de um programa de apoio para levantamento técnico de um *corpus*, um olhar qualitativo por parte do pesquisador se faz necessário. Desta forma, foi utilizado o recurso denominado

Concordance (“Concordância”, em português). Por meio do seu uso, é possível verificar o contexto em que as palavras estão inseridas, provendo maior fidedignidade aos termos que constituem sentido e relação ao escopo da OC.

A partir da utilização do *Sketch Engine* é possível gerar gráficos e tabelas de dados. Com o auxílio do *software*, espera-se comparar, organizar e analisar as informações extraídas de ambos os corpos textuais para então utilizar a análise de domínio.

5 Análises e discussões dos resultados

Preliminarmente, foi realizado um levantamento de enunciados-pivô preestabelecidos da produção científica da *International Society of Knowledge Organization* Brasil (ISKO-Brasil), publicados nos anais do Congresso Brasileiro em Organização e Representação do Conhecimento (2011-2019), e da *North American Symposium on Knowledge Organization* (NASKO) (2007-2021), os quais resultaram em 131 artigos e 132 documentos (constituídos por 127 artigos e cinco resumos expandidos) analisados, respectivamente. Para análise do capítulo brasileiro, é importante destacar que os artigos estão divididos em três dimensões: epistemológica, aplicada e social, cultural e política. Para fins deste trabalho, fez-se um recorte apenas das dimensões aplicada e social, cultural e política, além de utilizar somente os artigos publicados em língua portuguesa. Em relação aos enunciados-pivô, foram utilizados os seguintes para a seleção do *corpus* da ISKO-Brasil: “Indexação”, “Sistemas de Organização do Conhecimento”, “Representação Documental”, “Representação do Conhecimento”, “Representação da Informação”, “Representação e Organização do Conhecimento”, “Taxonomia”, “Tesouro” e “Ontologia”. Quanto ao capítulo norte-americano, foram empregados os respectivos termos: “Knowledge Organization Systems”, “Knowledge Organization”, “Taxonomies”, “Ontologies” e “Thesaurus”.

Após a análise inicial e seleção do *corpus* de pesquisa, o *software Sketch Engine* foi utilizado para examinar as publicações dessas duas comunidades discursivas. A ISKO-Brasil, com cinco volumes de anais publicados, conta com um *corpus* com 532,496 palavras em língua portuguesa e a NASKO, com oito volumes, constitui um *corpus* contendo 583,674 palavras em inglês. A diferença e pluralidade entre os termos no capítulo brasileiro se deve às características específicas da sua respectiva comunidade que apresenta maior dispersão terminológica que a outra.

5.1 ISKO-BRASIL

Para que fosse possível chegar na *Wordlist* do *corpus* da ISKO-Brasil, fez-se necessária a limpeza dos dados, uma vez que o *software* detecta a frequência do *corpus* como um todo. Assim, pontuações, preposições, artigos, conectores etc. foram ignorados, levando em conta somente os termos que são importantes enquanto conceito para o *corpus*, como pode ser visto no quadro 1.

Quadro 1. *Wordlist ISKO-Brasil*

	<i>Palavra</i>	<i>Frequência</i>
1º	informação	3,535
2º	conhecimento	3,339
3º	organização	2,645
4º	representação	1,687
5º	indexação	1,502
6º	análise	1,338
7º	acesso	1,199
8º	ciência	1,163
9º	pesquisa	1,113
10º	termos	1,064

Fonte: Autores.

Algumas palavras presentes na *Wordlist* são palavras que são provenientes de cabeçalhos e rodapés dos anais, isto é, fazem parte do nome dos eventos, como é o caso das quatro primeiras palavras da lista. Porém, apesar disso, elas têm importância enquanto palavras com valor para o *corpus*, uma vez que são palavras que, inclusive, fazem parte dos enunciados-pivô utilizados na formação do *corpus*. Além das quatro primeiras palavras que aparecem com maior frequência, a quinta palavra “indexação” mostra um aspecto importante no *corpus*, visto que esta não é parte de nenhum nome dos congressos, ou seja, não está nos cabeçalhos e rodapés dos artigos. O termo demonstra que há um grande interesse na pesquisa sobre a temática de indexação na ISKO-Brasil, a qual é importante para a Organização e Representação do Conhecimento e da Informação.

Em relação aos quadros desenvolvidos a partir das *Keywords*, também foi realizada a limpeza dos dados. No caso das *multi-words terms* (“termos com diversas palavras”), para que essa lista fosse possível e fiel ao *corpus*, ignorou-se as palavras que eram provenientes de cabeçalhos e rodapés dos eventos, assim como as que eram muito recorrentes por conta de ser parte de nome de revista, além de palavras em inglês (ver quadro 2).

Quadro 2. *Keywords (multi-words terms) ISKO-Brasil*

	Palavra
1º	política de indexação
2º	sistemas de organização
3º	análise de domínio
4º	análise de assunto
5º	descrição arquivística
6º	linguagem de indexação
7º	tratamento temático
8º	linguagens documentárias
9º	processo de indexação
10º	indexação de imagens

Fonte: Autores.

Dentre as 10 palavras-chave encontradas no *corpus* analisado da ISKO-Brasil, observa-se que, em sua maioria, versam sobre temas relacionados à representação temática, isto é, nas questões que giram em torno da indexação, da análise de assunto ou do tratamento temático. Em relação à arquivologia, em específico, vale destacar o termo descrição arquivística, como um processo fundamental na organização e representação do conhecimento arquivístico. Outros termos importantes de destacar nesse contexto são os referentes aos processos que organizam o conhecimento, como os sistemas de organização e as linguagens documentárias. Acerca de procedimentos metodológicos, nota-se que a “análise de domínio” aparece como uma teoria que tem sido bastante utilizada na ISKO-Brasil.

No quadro 3, o qual trata das *single-words* (“palavras-únicas”, em português), também foi feita uma limpeza dos dados, desconsiderando as palavras em inglês e nomes de autores, focando nas palavras que tenham um significado enquanto conceito importante para o *corpus*.

Quadro 3. *Keywords (single-words) ISKO-Brasil*

	Palavra
1º	indexação
2º	tesauros
3º	ontologia
4º	arquivístico
5º	tesauro
6º	isko
7º	informacional
8º	terminológico
9º	folksonomia
10º	cdd

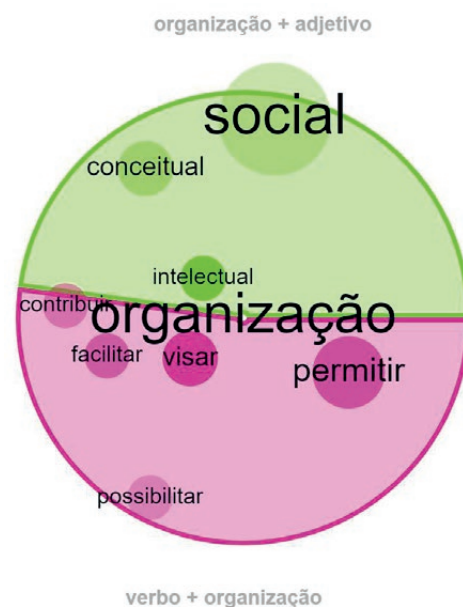
Fonte: Autores.

No quadro 3, há palavras que são parte dos enunciados-pivô do *corpus*, como é o caso de indexação, tesauros e

tesauro, ontologia. É importante destacar que parece haver um erro por parte do *software* em ter identificado a palavra “tesauro”, no singular e no plural, como palavras diferentes, uma vez que ele deveria agrupar as palavras, independentemente do seu gênero e número.

Em relação aos termos em si, observa-se que, assim como no *multi-word terms*, o que aparece em primeiro lugar é a indexação, isto é, demonstrando que há um grande interesse por parte da comunidade epistêmica da ISKO-Brasil sobre questões que envolvam indexação de materiais. Outro aspecto a se destacar dessa lista é que há termos que são sistemas de organização, como o caso de tesauro, ontologia e folksonomia, além do sistema de classificação bastante difundido em bibliotecas brasileiras, no caso, a Classificação Decimal de Dewey (CDD), demonstrando assim a importância dos SOC para a comunidade discursiva.

No que se refere a ferramenta *Word Sketch*, para fins da análise do *corpus* da ISKO-Brasil, buscou-se combinar os termos com adjetivos (termo + adjetivo) e verbos (verbo + termo). A partir disso e considerando o que já foi analisado a partir das ferramentas de *Wordlist* e *Keywords*, fez-se o recorte dos seguintes termos: organização, conhecimento, representação, informação e indexação.

Figura 1. *Word Sketch “organização”*

Fonte: Autores.

Na figura 1, a partir da relação gramatical com o verbo do termo “organização”, tem-se, de forma mais próxima, isto é, com maior relação, os verbos “permitir” e “visar”. O que se pode inferir com isso é que há uma preocupação da comunidade em possibilitar, em objetivar uma melhor organização – do conhecimento e/ou da informação. Porém, ao analisar o termo

“organização” ligado aos adjetivos, nota-se que os termos que são parte dos enunciados-pivô dessa pesquisa não são recuperados, isso se dá pelo fato de o *software* não reconhecer o termo enquanto um sintagma preposicional, uma vez que, ao utilizar esse filtro, ocorre um erro na geração dos dados. Outro aspecto importante de se destacar aqui é a ligação do termo “organização” com “social”, o que remete ao que Hjørland (2008) comenta sobre a organização social do conhecimento e como esta se organiza em diferentes âmbitos, de acordo como aquela realidade se organiza.

Figura 2. *Word Sketch* “conhecimento”



Fonte: Autores.

Em relação ao termo “conhecimento”, de acordo com o que pode ser visto na figura 2, o adjetivo que aparece mais próximo e com maior conexão é o “responsável”, porém, este é outro caso em que se trata de uma palavra que faz parte do título da 5ª edição do evento *Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas*, isto é, está em todas as páginas dos artigos deste ano do evento. Acerca dos verbos que se ligam ao termo, pode-se notar que giram em torno do que se estuda na área, como é o caso de “representar”, “organizar”, “produzir”, “compartilhar”.

Figura 3. *Word Sketch* “representação”



Fonte: Autores.

A respeito do termo “representação” (fig. 3), o adjetivo que aparece mais próximo a ele é “temático”. Através da ferramenta *Concordance*, nota-se que a “representação temática” é um tema recorrente em vários artigos do *corpus*, corroborando com o que se analisou nas ferramentas de *Wordlist* e *Keywords*, no qual há um grande enfoque da comunidade em questões que giram em torno da indexação, do tratamento temático. Outros adjetivos com os quais o termo “representação” se liga e que são importantes de se destacar, pois demonstram questões próprias, são: “documental”, “bibliográfico”, “arquivístico”, “imagético” e “gráfico”, pois tratam de materiais específicos e que demandam representações também específicas para tais. Sobre verbos e o termo “representação”, nota-se que, assim como no termo “organização”, aparecem os verbos “permitir” e “visar”, além de outros que seguem a mesma ideia, como “possibilitar”, isto é, ações que são relevantes para a representação do conhecimento e/ou da informação.

Fonte: Autores.

Figura 4. *Word Sketch* “informação”



Fonte: Autores.

Sobre o termo “informação” (fig. 4), o adjetivo que aparece mais próximo é “arquivístico”, isto é, formando o termo “informação arquivística”, além de outro termo que remete a área, que é a “informação orgânica”, demonstrando, assim, uma tendência da comunidade discursiva da ISKO-Brasil com a área da Arquivologia. Em relação aos verbos, pode-se notar que seguem a tendência da área, como no caso de “representar”, “organizar”, “recuperar”, “buscar”, entre outros.

O termo “indexação” foi escolhido, como dito anteriormente, por ser um termo que aparece com grande frequência nos artigos do *corpus*. Este é um dado que já vinha sendo observado desde a fase inicial de análise, isto é, na formação do *corpus*. A partir da ferramenta do *Word Sketch* (ver fig. 4), observou-se um enfoque na questão da “indexação automática”, o que demonstra uma evolução nos processos de indexação; a

“indexação multimodal”, a qual tem uma ligação com as ontologias; e a “indexação social” que tem uma ligação com as folksonomias e o que foi dito antes acerca do organização do conhecimento social, a partir das comunidades em que esse conhecimento é indexado, além de “linguagens de indexação próprias” que segue essa mesma linha.

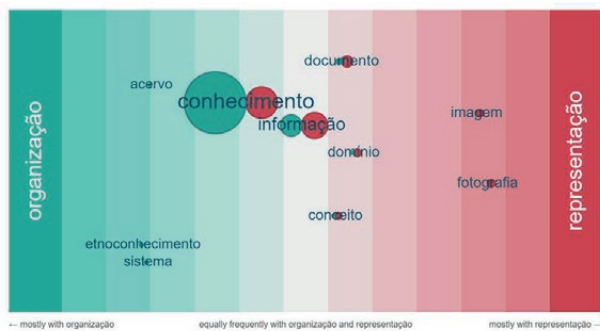
Figura 5. *Word Sketch* “indexação”



Fonte: Autores.

Com o uso da ferramenta do *Word Sketch Difference*, buscou-se suprir a falha na ferramenta de *Word Sketch* uma vez que o *software*, em língua portuguesa, não é capaz de analisar os termos enquanto sintagma preposicional, uma vez que ocorre um erro ao tentar.

Figura 6. *WSD* (organização-representação)



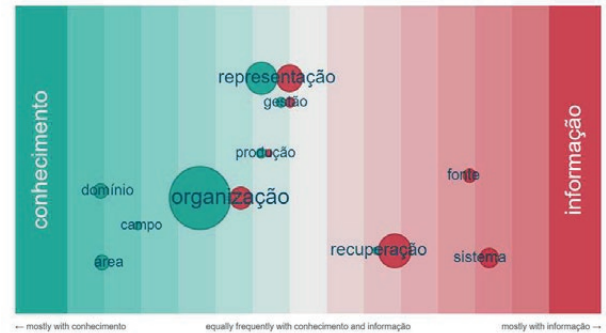
Fonte: Autores.

Na figura 6, tem-se uma análise dos termos “organização” e “representação” buscando informações a partir dessas palavras em conjunto com “de” acrescido de um substantivo. Com isso, pode-se observar que a comunidade tende a falar sobre “conhecimento” e “informação”, inclinando mais a “organização do conhecimento” e a “representação da informação”, mesmo que a palavra “informação” esteja bem no meio do gráfico, o que mostra um equilíbrio em relação ao uso dos conceitos “representação da informação” e “organização da informação”. Além disso, levando em conta o que Brascher e Café (2008) trazem sobre esses conceitos, nota-se que a organização do conhecimento

está voltada para os conceitos, enquanto a representação da informação está buscando registrar a informação.

Como forma de corroborar o que foi dito sobre “organização” e “representação”, analisou-se, com o uso do *Word Sketch Difference*, as palavras “conhecimento” e “informação” (fig. 7), a partir do filtro “... de conhecimento/informação”, isto é, formando sintagmas preposicionais.

Figura 7. *WSD* (conhecimento-informação)

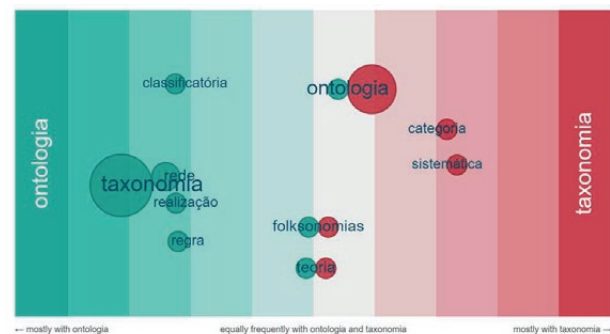


Fonte: Autores.

Assim, como no gráfico anterior, chegam-se às mesmas conclusões, uma inclinação para as questões de “organização do conhecimento”, “representação da informação” e “organização da informação”. O que pode-se destacar nesse gráfico é o fato de a palavra “recuperação” estar mais próxima do termo “informação”, assim, podendo-se inferir que a comunidade da ISKO-Brasil está, também, interessada em questões que versam a respeito da “recuperação da informação”. Além desses termos, observa-se que os dois termos também se ligam, de maneira menos representativa, com “produção” e “gestão”, demonstrando uma preocupação com o desenvolvimento e gerenciamento do conhecimento e da informação.

Ao que tange a análise dos SOC “ontologia” e “taxonomia”, a partir do *Word Sketch Difference*, utilizou-se o filtro “e ou”, assim como mostra a figura 8.

Figura 8. *WSD* (ontologia-taxonomia)



Fonte: Autores.

Sobre esses dois termos, destaca-se que, antes de o *corpus* ser analisado pelo *software*, já se observou que, por se tratarem enunciados-pivô, há uma predominância maior para trabalhos que tratem da temática de ontologias, do que de taxonomias. Outro aspecto importante de se destacar é o fato de que esses termos aparecem nos textos normalmente elencados quando se discorre a respeito de SOC, por exemplo, “ontologias, taxonomias e tesouros”. Apesar de o termo “tesouro” não aparecer nessa análise, é possível observar que “folksonomia” aparece, sendo esse outro SOC. Através da ferramenta *Concordance*, foi possível ver que se trata do mesmo caso nos quais são SOC que aparecem elencados como exemplos.

5.2 NASKO

Na *Wordlist*, foi constatado que a palavra *knowledge*, dentro do *corpus* da NASKO, é o termo de maior ocorrência (3,701 frequências). Em seguida, tal como apresentado no quadro 4, aparecem as palavras *classification*, *information*, *library*, *organization*, entre outros.

Quadro 4. *Wordlist NASKO*

	<i>Palavra</i>	<i>Tradução</i>	<i>Frequência</i>
1°	knowledge	conhecimento	3,701
2°	classification	classificação	3,398
3°	information	informação	2,587
4°	library	biblioteca	2,518
5°	organization	organização	2,472
6°	term	termo	1,627
7°	system	sistema	1,605
8°	analysis	análise	1,359
9°	category	categoria	1,090
10°	concept	conceito	1,086

Fonte: Autores.

De forma bem abrangente, é possível concluir que são palavras bastante recorrentes dentro do escopo temático tratado para esse trabalho. Estão inseridas dentro de um padrão terminológico, pois abrangem termos comumente utilizados por comunidades discursivas da OC. Para construção desse quadro, foram desconsiderados sinais, palavras irrelevantes (que somente aparecem nos cabeçalhos e não são representativos para o *corpus*, por exemplo),

preposições, autoridades etc. Essa limpeza também foi aplicada aos quadros desenvolvidos a partir da *Keywords*, pois o objetivo foi analisar tematicamente as relações e recorrências dos termos do *corpus*.

No quadro 5, vê-se, em primeiro lugar, o *lemma* relacionado a atividade de catalogação. Uma análise através do *Concordance* e do *Word Sketch*, evidencia que a temática costuma ser abordada no que diz respeito a sua prática e padronização.

Quadro 5. *Keywords (single-word) NASKO*

	<i>Palavra</i>	<i>Tradução</i>
1°	cataloge	catalogação
2°	ddc	CDD
3°	lcsch	LCSH
4°	koss	SOCs
5°	bibliographic	bibliográfico
6°	frbr	FRBR
7°	indexing	indexação
8°	classification	classificação
9°	ko	OC
10°	ontology	ontologia

Fonte: Autores.

Estão presentes, no recorte, recursos de representação desenvolvidos e recorrentemente explorados no cenário norte-americano: Classificação Decimal de Dewey (CDD) e *Library of Congress Subject Headings* (“Cabeçalhos de Assunto da Biblioteca do Congresso”, em português). Também foi identificado o modelo conceitual britânico Requisitos Funcionais para Registros Bibliográficos (FRBR). Além disso, apesar dos termos serem considerados únicos, há maior utilização de suas siglas do que seus equivalentes por extenso, com exceção da OC. Dentre os enunciados propostos, também foi localizado o termo *ontology* que indica um maior *score* dentre os SOC selecionados nos enunciados-pivô.

No que tange aos termos que compreendem mais de uma palavra, ou seja, *multi-words terms* (“termos com diversas palavras”), é observado, tal como apresentado no quadro 6, que o primeiro termo é *knowledge organization*.

Quadro 6. *Keywords (multi-words terms) NASKO*

	<i>Palavra</i>	<i>Tradução</i>
1º	knowledge organization	organização do conhecimento
2º	subject headings	cabeçalhos de assunto
3º	information science	ciência da informação
4º	controlled vocabulary	vocabulário controlado
5º	facet analysis	análise de facetas
6º	classification scheme	sistema de classificação
7º	decimal classification	classificação decimal
8º	knowledge organization system	sistema de organização do conhecimento
9º	domain analysis	análise de domínio
10º	library classification	classificação bibliográfica

Fonte: Autores.

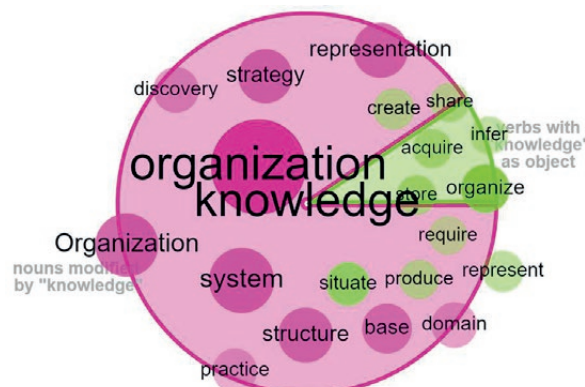
Na terceira linha, há a disciplina “Ciência da Informação” que ratifica a relação da grande área com a sua subárea Organização do Conhecimento. Identificou-se a presença do termo “Sistema de Organização do Conhecimento”, o que indicou também o interesse da NASKO pelo assunto. Dentro do recorte estabelecido, dois termos são considerados SOC (do grupo Classificação e categorização): *subject headings* e *classification schemes* (ZENG, 2008, p. 161).

Termos como *facet analysis*, *decimal classification*, *subject heading* e *library classification* evidenciam que o maior foco de estudo da comunidade discursiva é a representação. Foi identificado na nona posição a temática da Análise de Domínio, abordagem constantemente estudada e utilizada na área da OC.

De maneira complementar, o *Word Sketch* foi aplicado nos termos *knowledge*, *organization* e *system* por se tratarem de termos que contrastam com os enunciados-pivô propostos (*knowledge organization* e *knowledge organization systems*). Por isso, é relevante verificar as suas relações com os demais termos do *corpus*. Em contraste, os respectivos SOC foram analisados: taxonomia, tesouro e ontologia.

A relação mais forte de *knowledge* é com a palavra “organização”, o que demonstra que o termo está frequentemente atrelado ao nome da área “organização do conhecimento”. A palavra “sistema” também evidencia a relação com outro assunto presente nos enunciados-pivô da pesquisa, “Sistemas de Organização do Conhecimento”.

Foi analisado, a partir da figura 9, que, além da organização do conhecimento, há interesse na sua “representação”, na sua “estrutura” e nas “estratégias” empregadas para sua organização.

Figura 9. *Word Sketch “knowledge”*

Fonte: Autores.

Quanto aos termos que indicam ação sobre a palavra, isto é, os verbos, tal como se apresenta na figura 9, uma semelhança de colocações preponderantes através das palavras “organizar”, “representar” e “criar”. Outros verbos como, “compartilhar”, “produzir”, “adquirir”, “armazenar” podem ser destacados, sendo estes termos atrelados às operações básicas de tratamento de informação e conhecimento.

O termo “organização”, enquanto aquele que qualifica os demais termos, como é possível observar na figura 10, tem maior relação com “sistema”.

Figura 10. *Word Sketch “organization”*

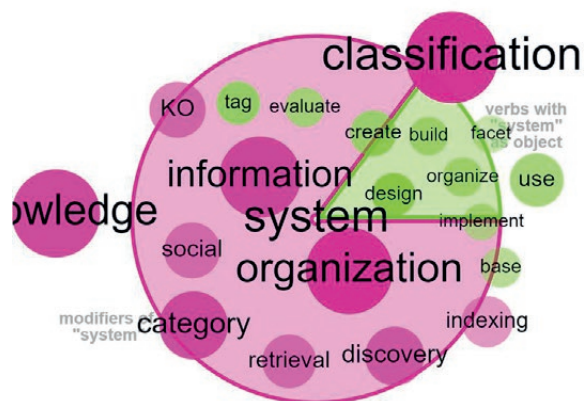
Fonte: Autores.

Outras palavras como “prática”, “literatura” e “padronização” também apresentaram uma importante aproximação do centro. É interessante destacar que a palavra *community* também está inserida no gráfico, o que já indica algum interesse de estudos acerca da NASKO como também da comunidade da organização

do conhecimento como um todo. Quanto aos verbos, estão presentes “ser”, “fazer” e “ter”.

Através do gráfico da figura 11, foram analisadas as colocações gramaticais que estão relacionadas com a palavra “sistema”. São essas, “organização”, “informação”, “classificação” e “conhecimento”.

Figura 11. *Word Sketch* “system”



Fonte: Autores.

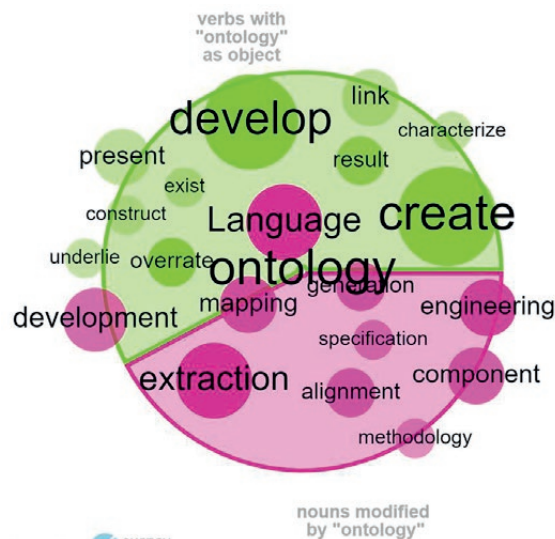
Tanto “conhecimento” como “informação”, apesar de apresentarem uma forte relação com a palavra central “sistema”, não se encontram posicionados próximos do centro. Essa circunstância denota o que uma análise mais aprofundada através da *Concordance* conseguiu elucidar, tratam-se de termos compostos de mais de uma palavra, isto é, “Sistemas de Recuperação de Informação” e “Sistemas de Organização do Conhecimento”. Dentro dessa estrutura, as palavras estão atreladas ao termo, entretanto, não se encontram posicionadas uma ao lado da outra. É possível analisar isso a partir do termo *retrieval*, o qual tem maior proximidade, mas menor relação no gráfico.

Quando investigado a respeito da palavra sistema enquanto objeto, foram elencados os seguintes verbos que se encontram diretamente relacionados aos SOC: “evoluir”, “organizar”, “projetar”, “criar”, “usar” e “basear”. Esses verbos estabelecem relação a princípio com o termo sistema, porém foi realizada uma análise mais aprofundada por meio da ferramenta *Concordance* para verificar se esses termos estavam presentes em contextos relacionados aos SOC. Dessa forma, foi possível inferir que tais verbos estão interligados com os processos de construção, aperfeiçoamento e utilização. Além disso, outro termo vinculado, o *facet*, demonstra que esses sistemas também são abordados através de uma perspectiva “facetada” por alguns autores da comunidade.

De acordo com o figura 12, a palavra “ontologia” está atribuída principalmente a termos como *Language*, *extraction*, *engineering* que mostram uma proximidade interdisciplinar da OC com tópicos das Ciências da

Computação. O que aponta um interesse da comunidade de expandir os seus estudos para outras disciplinas que são fundamentais para o desenvolvimento de um SOC, nesse caso, as ontologias. A NASKO abrange outros aspectos da ontologia dentro da sua especialização e fora dela, como: *ontology mapping*, componentes da ontologia, *ontology generation*, desenvolvimento, especificações etc. Não são estudos vastos, mas indicam um direcionamento dessa comunidade discursiva.

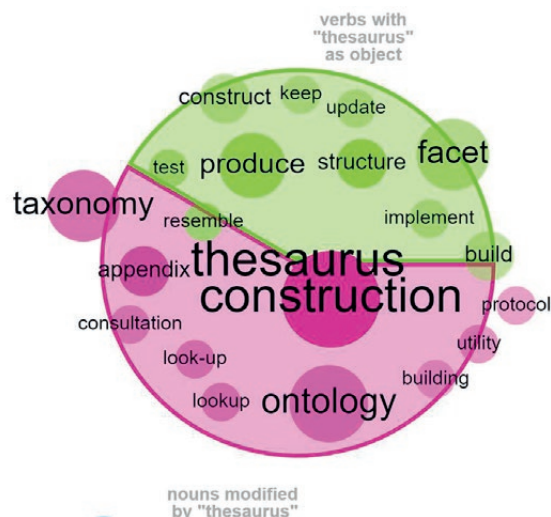
Figura 12. *Word Sketch* “ontology”



Fonte: Autores.

Quanto aos verbos que apresentam maior relação com a ontologia, foram identificadas ênfases na “criação” e “desenvolvimento” de ontologias (fig. 12).

Figura 13. *Word Sketch* “thesaurus”



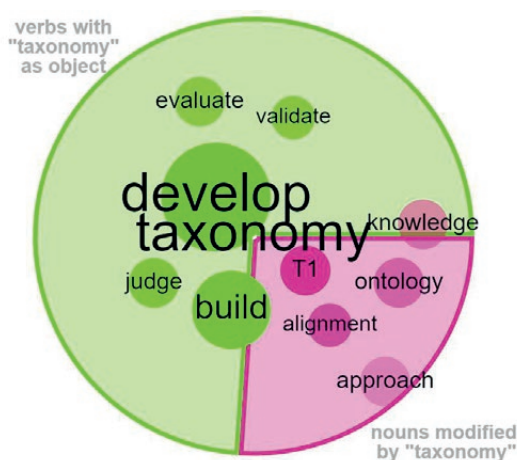
Fonte: Autores.

Além disso, conforme visualizado na figura 13, outros SOC como *taxonomy* e *ontology* aparecem relacionados em uma proporção significativa ao termo *thesaurus* (em *Word Sketch Difference* essas relações serão novamente

exploradas). Quanto aos verbos em que *thesaurus* performa enquanto objeto, foram elencados aqueles que apresentam maior recorrência: *structure* e *produce*. Novamente, termos que constituem paralelo com a produção desse tipo de esquema.

Dentre os SOC abordados, *taxonomy* é o termo que apresenta menos colocações entre as categorias de verbos e adjetivos do *Word Sketch*. Para esse caso em específico, serão exploradas duas categorias estabelecendo um recorte de cinco termos, tendo em vista que apenas os primeiros possuem uma predominância maior (fig. 14). A partir desse impasse, é possível constatar que, dentre os SOC, taxonomia é o menos explorado pelos estudos da comunidade.

Figura 14. *Word Sketch* “taxonomy”



Fonte: Autores.

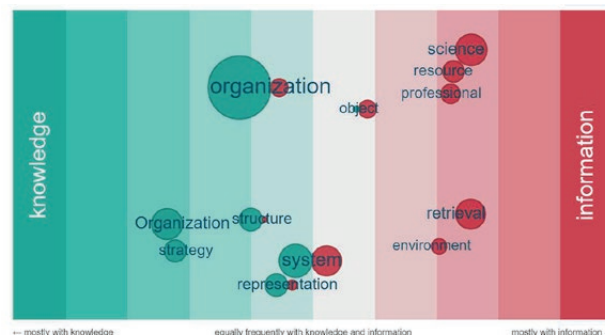
A partir disso, verificou-se a predominância dos termos T1, *ontology*, *knowledge*, *alignment* e *approach* relacionados à taxonomia. Ao empregar a análise através da ferramenta *Concordance*, foi constatado que a maior correlação é mantida pela *ontology* (ainda assim bastante baixa). No que diz respeito à relação da taxonomia como objeto dos verbos, tanto *develop* quanto *build* prevalecem como maiores agentes correlacionais.

A figura 15 apresenta os qualificadores relacionados aos termos conhecimento e informação, sendo que a classe gramatical mencionada se constitui enquanto objeto e os termos como sujeitos das orações.

A palavra *system* teve uma intercorrência equilibrada entre os dois termos, tanto no que se refere ao posicionamento no eixo central quanto na proporção identificada entre as duas esferas. Houve maior impacto relacional da palavra *organization* pelo termo conhecimento do que a informação, além disso *representation* pendeu mais para o lado de *knowledge* do que *information*. Outra palavra que se posicionou no centro foi “object”, constituindo, contudo, maior associação com a palavra informação. Dessa forma

ficou demonstrado que a materialidade possui maior relação com a informação do que o conhecimento, que muitas vezes é tratado como um termo ligado ao “abstrato”. Outros termos que contribuem para a análise são *retrieval* e *resource* que reforçam a noção de tangibilidade de *information*.

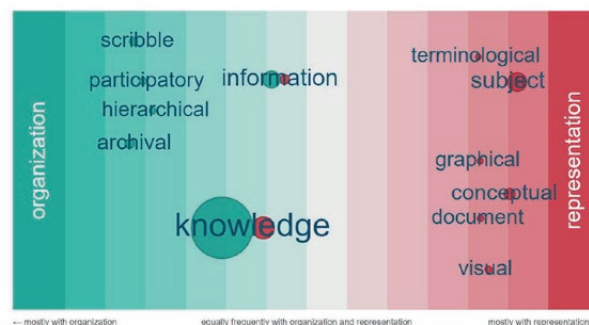
Figura 15. *WSD* (*knowledge-information*)



Fonte: Autores.

Os termos *knowledge* e *information* são qualificadores que estabelecem relações em comum entre as palavras *organization* e *representation*. Contudo, como pode ser verificado no gráfico (fig. 16) a seguir, essas relações estão mais inclinadas para a organização do que a representação.

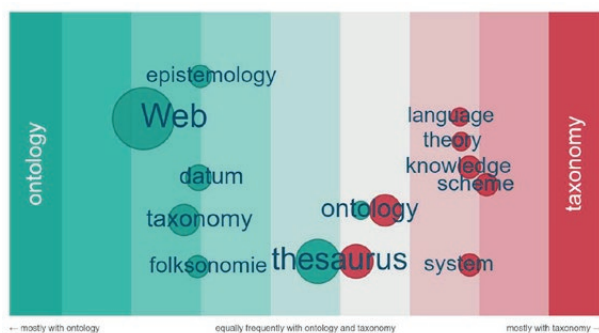
Figura 16. *WSD* (*organization-representation*)



Fonte: Autores.

“Conceitual” e “assunto” são qualificadores que estão somente associados a *representation*, o que faz alusão à representação temática (representações conceituais, representações conceituais externas e internas, representações conceituais semânticas e representação de assunto). Outros termos que qualificam a organização são hierárquico, arquivístico e participativo, ademais são termos que estabelecem presença significativa em mais de um artigo.

Para compreensão da relação entre os SOC, foco dos enunciados-pivô pré-selecionados para esse trabalho, fez-se a análise comparativa entre *ontology* e *taxonomy* (fig. 17).

Figura 17. WSD (*ontology-taxonomy*)

Fonte: Autores.

Ao longo da primeira fase do estudo, isto é, a análise manual de cada artigo, foi constatado um forte aparecimento da concatenação entre modelos de organização do conhecimento.

A análise por meio de *and/or* permitiu visualizar essas recorrências identificadas com maior clareza. No centro do gráfico, observa-se a palavra *thesaurus*, constituindo uma relação mais forte com a ontologia. Através da ferramenta *Concordance*, foi possível identificar que, nos trechos em que os termos aparecem, na maior parte dos casos, a relação estabelecida é de “igualdade” (isto também foi identificado em *taxonomy* e *thesaurus*), demonstrando que não é de interesse da comunidade evidenciar contraposições entre esses SOC.

Outro aspecto identificado é a relação entre *ontology* e *folksonomie*, que em contraposição demonstra um modelo formal de OC e outro que é bastante utilizado em contextos de informalidade. O que se pode inferir é que esses conceitos são vistos a partir do contexto em que comumente estão atrelados, isto é, a *internet*. A palavra *Web* relacionada a ontologia, e principalmente a *Web Semântica*, reafirma esse argumento.

6 Considerações finais

Com base nas análises dos dois capítulos, podemos tirar algumas conclusões a respeito das comunidades. A primeira coisa é que, apesar de ambas estarem dentro do escopo da ISKO, enquanto uma entidade internacional que funciona como uma rede de colaboração teórica, metodológica e conceitual, existem diferenças em relação à compreensão de conceitos, bem como o espaço que esses conceitos, metodologias e práticas ocupam na discussão, de encontrarmos várias vozes que se justapõem no *corpus*. Existe uma predominância de temas arquivísticos no capítulo brasileiro, aspecto que, no capítulo norte-americano, não se faz presente do mesmo modo. Isto também vale para políticas de indexação, um tema bastante presente no contexto brasileiro. Em relação às convergências, estas são muitas e as que destacamos são: 1) Análise de Domínio; 2) Sistemas de Organização; 3) Teoria do Conceito; 4) Sistemas de

Classificação. Estas indicam, portanto, que enquanto comunidade existimos a partir de um arcabouço metodológico aplicado a uma realidade empírica, ou seja, somos ao mesmo tempo teóricos e aplicados e convergimos em relação à teoria do conceito enquanto um fundamento e à análise de domínio enquanto uma abordagem metodológica.

Como se trata de uma pesquisa em desenvolvimento, continuaremos nossas análises a partir desta, apresentando outros capítulos, a fim de construirmos um panorama claro de quem realmente somos.

Com relação análise semântico-discursiva enquanto uma abordagem da análise de domínio, vemos muitas vantagens, como aquelas apresentadas ao longo do trabalho e uso de ferramenta semântica para essa atividade se prova significativo e representativo e neste momento, não percebe-se nenhum tipo de desvantagem, muito pelo contrário, comenta-se a análise efetua de outras maneiras.

Referências

- BARROS, T. H. B. A análise de discurso de Michel Pêcheux e a organização do conhecimento: possibilidades teórico-metodológicas. In: AMORIM, I. S.; SALES, R. de (org.). **Ensaio em organização do conhecimento**. Florianópolis: UDESC, 2021. p. 61-78. Disponível em: <https://dx.doi.org/10.5965/978-65-88565-20-9>. Acesso em: 04 abr. 2023.
- BARROS, T. H. B.; LAIPELT, R. do C. F. Uma análise de domínio da área de Organização e Representação do Conhecimento no contexto do periódico *Em Questão*. **Em questão**, v. 27, n. 4, p. 438-468, 2021. DOI: 10.19132/1808-5245274.438-468. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/116729>. Acesso em: 15 jan. 2023.
- BOOIJ, G. E. **The grammar of words: an introduction to linguistic morphology**. New York: Oxford University Press, 2005.
- BRASCHER, M.; CAFÉ, L. Organização da informação ou organização do conhecimento?. In: Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação, 9., 2008, São Paulo. **Anais eletrônicos [...]**. São Paulo: Escola de Comunicação e Artes da USP, 2010. Tema: Diversidade Cultural e Políticas de Informação. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/176535>. Acesso em: 14 jan. 2023.
- HJØRLAND, B. Domain analysis. **Knowledge Organization**, n. 44, v. 6, p. 436-464, 2017. Disponível em: <https://doi.org/10.5771/0943-7444-2017-6-436>. Acesso em: 18 mar. 2023.
- HJØRLAND, B. Domain analysis in information science: eleven approaches - traditional as well as innovative. **Journal of Documentation**, n. 58, v. 4, p. 422-462, ago. 2002. Disponível em: <https://doi.org/10.1108/00220410210431136>. Acesso em 15 mar. 2023.
- HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, Copenhagen, v. 35, n. 2-3, p. 86-

101, 2008. Disponível em: <https://doi.org/10.5771/0943-7444-2008-2-3-86>. Acesso em: 16 jan. 2023.

HJØRLAND, B.; ALBRECHTSEN, Hanne. Toward a new horizon in Information Science: domain-analysis. **Journal of the Association for Information Science and Technology**, v. 46, n. 6, p. 400-425, jul. 1995. Disponível em: [https://doi.org/10.1002/\(SICI\)1097-4571\(199507\)46:6<400::AID-ASI2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y). Acesso: 18 mar. 2023.

HODGE, G. **Systems of Knowledge Organization for digital libraries: beyond traditional authority files**. Washington: DLF, 2000. Disponível em: <https://www.clir.org/pubs/reports/pub91/>. Acesso em: 21 mar. 2023.

LEXICAL COMPUTING CZ S.R.O. Sketch Engine. [Brun, 202-?]. *Software online*. <https://www.sketchengine.eu/>. Acesso em: 21 mar. 2023.

LÓPEZ-HUERTAS, M. J. Domain analysis for interdisciplinary knowledge domains. **Knowledge organization**, v. 42, n. 8, 2015. Disponível em: <https://doi.org/10.5771/0943-7444-2015-8-570>. Acesso em: 18 mar. 2023.

OLIVEIRA, C. C. de; SILVA, M. C. da; BARROS, T. H. B.; MOURA, A. M. M. de. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 22., 2022, Porto Alegre. **Anais eletrônicos [...]**. Porto Alegre: UFRGS, 2022. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxiiencib/paper/view/993>. Acesso em: 4 abril. 2023.

SKETCH ENGINE. **Sketch Engine user guide**. Lexical Computing CZ s.r.o.: [Brun, 202-?]. Disponível em: <https://www.sketchengine.eu/guide/>. Acesso em: 21 mar. 2023.

SKETCH ENGINE. **What can Sketch Engine do with...?** Lexical Computing CZ s.r.o.: [Brun, 202-?]. Disponível em: <https://www.sketchengine.eu/what-can-sketch-engine-do/#top>. Acesso em: 21 mar. 2023.

SMIRAGLIA, Richard P. **Domain analysis for Knowledge Organization: tools for ontology extraction**. Waltham, MA (USA); Kidlington (UK): Elsevier, 2015.

ZENG, M. L. Knowledge Organization Systems (KOS). **Knowledge Organization**, Copenhagen, v. 35, n. 2-3, p. 160-182, 2008. Disponível em: <https://doi.org/10.5771/0943-7444-2008-2-3-160>. Acesso em: 21 mar. 2023.



**DIMENSÃO APLICADA DA
ORGANIZAÇÃO DO CONHECIMENTO**

*Applied Dimension of the
Knowledge Organization*

Análise da representação da música em bibliotecas digitais

Analysis of music representation in digital libraries

Fernanda Carolina Pegoraro Novaes (1), Jessica Beatriz Tolare (2), Walter Moreira (3)

(1) Universidade Estadual Paulista Júlio de Mesquita Filho, Avenida Hygino Muzzi Filho, 737, Marília, SP, Brasil, fernanda.pegoraro@unesp.br, (2) jessica.tolare@unesp.br, (3) walter.moreira@unesp.br

Resumo

Introdução: Os metadados podem ser encontrados dentro de um universo bibliográfico de obras literárias e textuais sobre música, formando o seu próprio universo. A música, para ser representada em domínio bibliográfico, pode utilizar ferramentas e/ou projetos auxiliares, com o objetivo de servir como diretrizes para a representação e recuperação informacional. **Objetivos:** Analisar como as bibliotecas digitais têm utilizado projetos e ferramentas auxiliares e realizar uma investigação bibliográfica de como a literatura publicada vem abordando a relação entre as bibliotecas digitais e projetos e ferramentas auxiliares. **Metodologia:** O corpus documental deste trabalho foi estruturado por meio de buscas realizadas na Association for Computing Machinery Digital Library, onde foram recuperados 16 projetos e ferramentas auxiliares. **Resultados:** Os projetos desenvolvidos no período de 1996 a 2013 abordaram: música baseada em ruído; conversão automática de partituras online; música baseada em conteúdo; trocas sobre pesquisa em Music Information Retrieval; edição de músicas e bibliotecas com recuperação baseada em áudio. No período 2014 a 2020, foram desenvolvidos projetos relacionados a conjunto de dados sobre a história do big data na música ocidental e linked open data. **Conclusão:** Foi demonstrado que a pesquisa e a construção de projetos para melhor representar e consumir a informação musical acompanhou o processo de evolução tecnológica. A sintonia tecnológica e a utilização dos mais variados recursos dessa natureza para o tratamento da informação no domínio da música fica clara com a diversidade de projetos analisados.

Palavras-chave: Bibliotecas digitais. Association for Computing Machinery Digital Library. Projetos auxiliares em música.

Abstract

Introduction: Metadata can be found within a bibliographic universe of literary and textual works on music, forming its own universe. Music, to be represented in the bibliographic domain, can use tools and/or auxiliary projects, with the objective of serving as guidelines for the representation and informational retrieval. **Objectives:** To analyze how digital libraries have used projects and auxiliary tools and carry out a bibliographical investigation of how the published literature has addressed the relationship between digital libraries and projects and auxiliary tools. **Methodology:** The documentary corpus of this work was structured through searches carried out in the Association for Computing Machinery Digital Library, where 16 projects and auxiliary tools were retrieved. **Results:** The projects developed from 1996 to 2013 addressed: music based on noise; automatic conversion of scores online; content-based music; research exchanges on Music Information Retrieval; music and library editing with audio-based recovery. From 2014 to 2020, projects related to a data set on the history of big data in western music and linked open data were developed. **Conclusion:** It was demonstrated that the research and construction of projects to better represent and consume musical information accompanied the process of technological evolution. Technological harmony and the use of the most varied resources of this nature for the treatment of information in the field of music is clear from the diversity of projects analyzed.

Keywords: Digital libraries. Association for Computing Machinery Digital Library. Auxiliary music projects.

Resumen

Introducción: Los metadatos se pueden encontrar dentro de un universo bibliográfico de obras literarias y textuales sobre música, formando un universo propio. La música, para ser representada en el dominio bibliográfico, puede utilizar herramientas y/o proyectos auxiliares, cuyo propósito es servir de guía para la representación y recuperación de información. **Objetivos:** Analizar cómo las bibliotecas digitales han utilizado proyectos y herramientas auxiliares y realizar una investigación bibliográfica de cómo la literatura publicada ha abordado la relación entre bibliotecas digitales y proyectos y herramientas auxiliares. **Metodología:** El corpus documental de este trabajo se estructuró a través de búsquedas realizadas en la Association for Computing Machinery Digital Library, donde se recuperaron 16 proyectos y herramientas auxiliares. **Resultados:** Los proyectos desarrollados desde 1996 hasta 2013 abordaron: música basada en ruido; conversión automática de puntajes en línea; música basada en contenido; intercambios de investigación sobre Music Information Retrieval; edición de música y biblioteca con recuperación basada en audio. De 2014 a 2020 se desarrollaron proyectos relacionados con un conjunto de datos sobre la historia del big data en la música occidental y datos abiertos vinculados. **Conclusión:** Se demostró que la investigación y construcción de proyectos para representar y consumir mejor la información musical acompañó el proceso de evolución tecnológica. La armonía tecnológica y la utilización de los más variados recursos de esta naturaleza para el tratamiento de la información en el campo de la música se desprende de la diversidad de proyectos analizados.

Palabras clave: Bibliotecas digitales. Association for Computing Machinery Digital Library. Proyectos musicales auxiliares.

1 Introdução

Com a evolução da tecnologia no final do século XX, a internet desenvolveu-se emparelhada à globalização, caracterizada pela integração de comunicação mundial responsável pela expansão do crescimento econômico internacional. Com a lei de oferta e procura e o desenvolvimento científico, emergiram novas ferramentas na *web* que possibilitaram tanto aos usuários comuns quanto aos especialistas em sistema de informação elaborarem o conteúdo *online*. A formação do processo de construção evidenciou a linguagem direcionada para seres humanos. O conteúdo demonstrado pelos computadores é considerado como linguagens de codificação ou marcação, como XML e HTML, pois codificam informações de renderização e *hiperlink* (BREITMAN, 2005).

A *web* sintática possui a capacidade de interpretação da informação sob responsabilidade do indivíduo. Neste cenário, a máquina é responsável pelo que se refere à viabilização do processamento dos dados. A máquina possui conteúdo informacional próprio, no qual resulta em buscas com pouca precisão (apesar de recuperar várias páginas), conflito entre linguagem natural e vocabulário controlado e ausência de organização geográfica (prevenção de páginas recuperadas individualmente de um mesmo site). Este problema estrutural deve-se à rápida, e não planejada, disseminação da tecnologia utilizada que ressalta a troca de documentos entre usuários.

Diante desse contexto e para solucionar as questões apontadas, há necessidade em aplicar um modelo semântico à internet. A *web* semântica propõe a organização informacional do meio digital, formalizando um padrão para auxiliar e simplificar o processo de acesso à informação. Para este método ser eficaz é preciso que seja descentralizado, para que possa utilizar diferentes modelos de organização (BREITMAN, 2005). Entre os modelos, os metadados são evidenciados por sua capacidade de representar domínios.

Os metadados podem ser encontrados dentro de um “universo bibliográfico” de obras literárias e textuais sobre música, formando o seu próprio universo, no qual é composto por: domínio (tipo de material que o padrão se destina), comunidade (grupo que utiliza o padrão), função (papel de desempenho), propósito (objetivo) e entre outros (RILEY, 2010).

O termo metadados antecede a *web* para descrever um conjunto de dados sobre um determinado assunto. Apenas com o surgimento de novos tipos de recursos e ambientes informacionais digitais, a partir do final do século XX, os metadados foram desenvolvidos para redefinir o tratamento, uso e recuperação das informações em meio digital.

Com isso, ao chegar no final da década de 1990, os metadados trouxeram muitas mudanças para o cenário tecnológico. No entanto, os metadados possuem um esquema de representação mais generalizado, segundo Alves e Santos (2013) e quando há um domínio muito específico a ser representado, como por exemplo a música, pode não refletir as representações que o domínio necessita. A aplicação de metadados em domínios específicos tornam-se importantes para representar adequadamente particularidades e promover a garantia do armazenamento e dos valores dos metadados em banco de dados.

Dentro do universo dos metadados pode-se observar sete principais domínios: *Cultural Objects* (obras de arte); *Datasets* (coleções de dados primários); *Geospatial data* (informações relevantes para a localização geográfica); *Moving images* (imagens em movimento ou vídeos); *Musical materials* (aborda qualquer tipo de informação musical); *Scholar texts* (pesquisas ou materiais produzidos em um processo acadêmico) e *Visual resources* (materiais de natureza artística ou documental apresentados em forma visual fixa). Eles estão inseridos em quatro principais comunidades: *Libraries*, *Archives*, *Museums* e *Information Industry* (organizações que compõem a *web* pública e a comercial (RILEY, 2010).

Partindo da consideração feita por Alves e Santos (2013) (quanto maior a especificidade de um domínio, maior deverá ser a exigência de sua representação), a música, para ser representada em domínio bibliográfico, pode utilizar ferramentas e/ou projetos auxiliares, com o intuito de servir como diretrizes para a representação e recuperação informacional.

As ferramentas e projetos auxiliares são considerados técnicas construídas e incorporadas por pesquisadores para desenvolver melhor a caracterização do material musical (Allik *et al.*, 2013).

Exemplos dessas ferramentas e projetos são:

- *Musicconn Notesuche* (Biblioteca Estadual da Baviera, Munique, Alemanha);
- *VerovioX* (Suíça);
- *Cantus Ultimusxi* (Canadá);
- *Doing Reusable Musical Data* (DOREMUS) (França) (ALBUQUERQUE; PINTO; BORBINHA, 2018).

O objetivo está em beneficiar o âmbito científico e comercial. As bibliotecas normalmente utilizam os modelos elaborados para realizar a representação da informação musical, porém algumas delas criaram e implementaram projetos auxiliares, conhecidos também como paralelos, com o intuito de melhorar a especificidade nessa representação.

Devido ao grande volume de informação na área da música, Hujran *et al.* (2020) contextualiza que o *big data* se tornou uma ferramenta necessária, seu uso consiste em coletar registros sobre os hábitos e preferências dos usuários e, dessa forma, identificar grandes sucessos, pesquisas potenciais, possibilitar o acesso ao grande conteúdo de informação em diversos formatos.

Diante desse contexto e considerando o aumento exponencial das bibliotecas digitais especializadas em música, a pesquisa apontou os seguintes questionamentos: Quais projetos e ferramentas auxiliares estão sendo utilizados por bibliotecas para realizar a representação do domínio da música? Como a literatura publicada tem abordado a relação entre projetos auxiliares e ferramentas e bibliotecas digitais?

A pesquisa possui como objetivos: analisar como as bibliotecas digitais têm utilizado projetos e ferramentas auxiliares e realizar uma investigação bibliográfica de como a literatura publicada vem abordando a relação entre as bibliotecas digitais e projetos e ferramentas auxiliares.

A necessidade em mapear os projetos e ferramentas é considerada importante para compreender como o domínio bibliográfico na área da música está sendo representado e aplicado. A importância também está na contribuição para a comunidade científica, em específico, aos que se dedicam à temática, seja por meio da terminologia, classificação de instrumentos, organização, representação e recuperação da informação na área da música.

2 Desenvolvimento e resultados

A pesquisa se caracteriza como bibliográfica, descritiva e exploratória. Para o seu desenvolvimento foi estruturado o *corpus* documental por meio de buscas realizadas na *Association for Computing Machinery Digital Library* (ACM DL).

A ACM DL é uma plataforma de pesquisa com publicações que contemplam o período desde 1936 e possui acesso restrito. Nela estão indexadas coleções de anais de conferências, artigos de jornais e revistas e entre outros tipos de materiais.

Apesar de conter um banco de dados amplo e direcionado ao domínio da computação, a plataforma pode ser considerada muito vantajosa para pesquisas que abordam as Tecnologias de Informação e Comunicação (TIC) e pesquisas no domínio da música, que aplicam e investigam a *Music Information Retrieval* (MIR).

No ano de 2020, a *Association for Computing Machinery Digital Library* (ACM DL) permitiu acesso livre de seu acervo, como incentivo aos pesquisados,

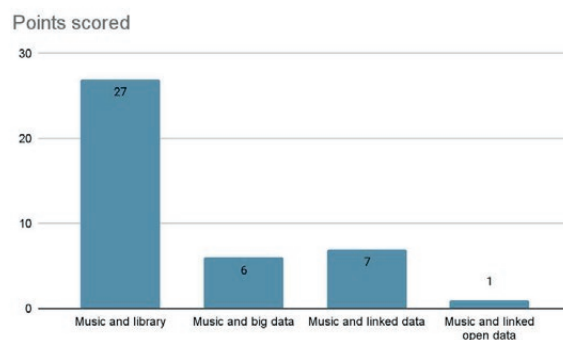
devido à situação de pandemia caracterizada pela COVID-19. A disponibilidade livre da plataforma possibilitou o acesso a inúmeros materiais importantes que proporcionaram o desenvolvimento de pesquisas e obtenção de conhecimento. Desse modo e por conta de toda a contextualização, a plataforma ACM DL foi selecionada como fonte principal de coleta de dados para o aprofundamento desta pesquisa. É importante salientar que nos anos de 2021 e 2022 a ACM DL não estava mais disponível em acesso livre, mas mesmo assim, houve tentativa de acesso na plataforma por parte dos autores. Como isso não foi possível, o escopo do *corpus* documental ficou limitado até o ano de 2020.

A seleção do *corpus* documental foi realizada a partir de buscas feitas com os seguintes termos: “*music*” and “*library*”; “*music*” and *big data*”; “*music*” and *linked data*” e “*music*” and “*linked open data*”. A plataforma é estrangeira e possui como idioma oficial o inglês, por isso, escolheu-se realizar as buscas na mesma língua. Foram utilizados os operadores booleanos como estratégia de busca e para melhorar a precisão dos resultados, por isso, as aspas nos termos.

Para localizar e selecionar os textos foram observados termos que estivessem contidos nos títulos, pois quando um assunto é inserido também como título de um trabalho, significa que há maior relevância do que quando é inserido no resumo, por exemplo. Foram eliminadas as pesquisas duplicadas e selecionados estudos que descreviam projetos e ferramentas auxiliares em bibliotecas digitais especializadas em música. Para realizar a análise dos projetos auxiliares foram desenvolvidas categorias de análise como banco de dados, biblioteca de código aberto, conjunto de dados, interface, metadados e ontologia. Cada categoria de análise engloba projetos e ferramentas auxiliares.

Ao total foram obtidos 41 trabalhos (Gráfico 1).

Gráfico 1. Resultados da busca na plataforma ACM DL



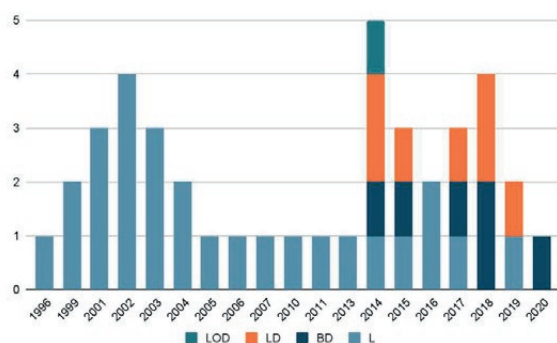
Fonte: Dados da pesquisa.

O Gráfico 1 apresenta os resultados obtidos a partir da busca feita na plataforma da ACM DL. Ao total foram encontrados 41 trabalhos, dos quais 27 foram pelos termos “*music*” and “*library*”; 6 pelos termos “*music*”

and “big data”; 7 pesquisas pelos termos “music” and “linked data” e apenas um trabalho pelos termos “music” and “linked open data”.

Os anos abordados pela pesquisa consistem entre 1996 - 2020, como apontado pelo Gráfico 2.

Gráfico 2. Trabalhos publicados divididos por anos



Legenda: L= library; BD= big data, LD= linked data e LOD= linked open data.

Fonte: Dados da pesquisa.

No Gráfico 2 é apresentado os trabalhos recuperados por anos na plataforma da ACM DL. É interessante observar as datas em que os termos aparecem. No período entre 1996 a 2013, as pesquisas referentes às bibliotecas tinham como viés discutir e solucionar questões de acesso, representação e recuperação informacional. A partir de 2014, considerando a evolução tecnológica, os projetos auxiliares começaram a serem desenvolvidos com o objetivo de estruturar grandes dados na *web*. Nesse momento, as pesquisas aprofundando o uso do *big data* tornaram-se imprescindíveis para o âmbito da música. Nesse mesmo ano, começaram as discussões sobre estabelecer comunicação e acesso entre diferentes dados da área de música, o que evidencia a presença cada vez mais constante do *big data*, *linked data* e *linked open data*.

Todas as pesquisas desenvolvidas, envolvendo o uso de *big data*, *linked open data* e *linked data* em bibliotecas digitais com a música, resultaram 16 projetos e ferramentas auxiliares. Eles podem ser divididos em categorias principais como banco de dados, bibliotecas de código aberto, conjunto de dados, interface, metadados e ontologia.

2.1 Bancos de dados

O banco de dados, segundo a definição apontada por Korth e Silberschatz (1994), é uma “coleção de dados inter relacionados, representando informações sobre um domínio específico”. Uma de suas características é agrupar informações que relacionam e abordam um mesmo assunto. Nesse estudo, a temática principal é a Música, sua organização para a disponibilização em acesso e recuperação da informação musical.

2.1.1 Big Data History of Music

O *Big Data History of Music* foi desenvolvido a fim de promover uma musicologia baseada na análise de metadados bibliográficos da música. Este projeto, a partir de sete conjuntos de dados principais de metadados, analisa e discute desafios técnicos de alinhamento dos mesmos, procurando desenvolver formas alternativas de explorar dados sobre a história da música ocidental (ROSE; TUPPEN, 2014).

Os sete principais conjuntos de dados combinados por esse projeto auxiliar são: 1. Catálogo de músicas impressas da *British Library*, com aproximadamente 1 milhão de registros bibliográficos; 2. Catálogo da *British Library* de manuscritos musicais; 3. Catálogo de músicas manuscritas de Augustus Hughes-Hughes do Museu Britânico; 4. Bibliografia do *Répertoire International des Sources* (RISM) sobre música europeia impressa entre 1500 e 1800; 5. Catálogo internacional de manuscritos musicais do *Répertoire International des Sources* no período entre 1600 a 1850; 6. Banco de dados de manuscritos de música do *Répertoire International des Sources*; 7. Banco de dados do *The Concert Programmes Project Database*, que descreve coleções históricas de programas de concertos realizados nas bibliotecas do Reino Unido.

2.1.2 Essen and Digital Tradition Folksong Dataset

Uma das buscas mais solicitadas aos bibliotecários especializados em música é baseada em notas cantaroladas ou “assoviadas”. Para essa tarefa, um acervo requer que as coleções de música estejam disponíveis *online* por meio do uso do *Optical Music Recognition* (OMR); isto é, a recuperação da música é baseada em *direct acoustic input*, entrada acústica direta. Esse tipo de recuperação é possível devido aos avanços tecnológicos que permitem a transcrição de melodias utilizando a entrada de um microfone (MCNAB *et al.*, 1996).

O *Essen and Digital Tradition Folksong Dataset* é um exemplo de projeto que executa esse tipo de recuperação. Foi criado em 1994 na Alemanha e armazena, por meio de seu banco de dados, músicas folclóricas europeias, com foco em músicas folclóricas alemãs, polonesas e chinesas. As músicas armazenadas utilizam três principais formatos, que são: TEX, PCX e *Musical Instruments Digital Interface* e os códigos utilizados para a descrição das músicas são conhecidos como *Essen Associative Code* (EsAC). Dessa forma, o *Essen and Digital Tradition Folksong Dataset* utiliza o método “prova do conceito” por possibilitar que o usuário canta utilizando um microfone e, assim, identificar e recuperar a música desejada com os seus metadados exibidos (MCNAB *et al.*, 1996).

2.1.3 Live Music Archive (LMA)

A *Live Music Archive* é considerada uma coleção crescente que contempla mais de 5.000 artistas, no qual a maior parte compositores do gênero rock; com mais de 150.000 gravações de shows ao vivo indexadas. Cada gravação acompanha metadados básicos responsáveis por descrever informações como datas, locais, listas de reprodução e a proveniência dos arquivos de áudio. A LMA descreve camadas de dados utilizando a *Computational Analysis of the Live Music Archive* (CALMA), pois permite que as camadas de dados assumam a forma de um serviço da *web* ou de uma ferramenta que faça a interação com um serviço da *web*. As conexões são obtidas por meio do *linked open data* (PAGE *et al.*, 2017).

2.1.4 MusicBrainz/LinkedBrainz/Music Ontology

O *MusicBrainz* é um banco de dados muito utilizado no domínio da música, principalmente, por descrever metadados de gravações musicais lançadas comercialmente. O *LinkedBrainz* é um *Linked data* que expõe dados do *MusicBrainz*, como triplas de RDF, capaz de mapear esquemas para coleções de ontologias localizadas no projeto *Music Ontology* (MO). Além de disseminar dados em *linked data*, no formato RDF/Turtle, as ontologias do MO também podem ser expressas nos formatos HTML, para publicar dados relacionados à música em páginas da *web* e em JSON-LD, para inserir dados relacionados à música como parte de uma API (FIELD; PHIPPEN; COHEN, 2015).

É comum e desejável que os projetos auxiliares criados no domínio da música estejam interligados, especialmente, em se tratando de uma *web* de dados vinculados. A música é considerada um domínio complexo em termos de sua representação e isso se deve, em partes, à grande quantidade de formatos com os quais se relaciona. Desse modo, a identificação e a descrição de projetos que visam a representação e recuperação da informação nesse domínio são essenciais para as requeridas ações de promoção de interoperabilidade.

2.1.5 Variations

Para promover melhor ensino aos alunos da *Indiana University School* foi desenvolvido o projeto *Variations*, responsável por desenvolver a coleção de repertórios musicais, que proporciona o acesso *online* de repertórios interpretador pelos alunos. O *Variations* indexa gravações diárias, que faz com que o seu banco de dados cresça semanalmente. Os principais gêneros musicais identificados no *Variations* são: ópera, jazz, música instrumental e rock. A recuperação de gravações específicas de peças musicais ocorre por meio do IUCAT, sistema de catálogo *online* da *Indiana University School*, baseado no *software* NOTI LMS. As gravações encontradas em formato físico na biblioteca (como o CD e outros) são representadas pelo padrão de

registro bibliográfico USMARC (DUNN; MAYER, 1999).

A atuação *VARIATIONS 2* propôs o desenvolvimento de um modelo de dados para atender às necessidades específicas do domínio da música como identificar, separar e relacionar camadas lógicas e físicas das obras musicais e suas manifestações. Esse procedimento é semelhante ao modelo desenvolvido pelo Grupo de Estudos da Federação Internacional de Associações de Bibliotecas (IFLA) sobre os Requisitos Funcionais para Registros Bibliográficos; o que os diferencia é que o *Variations 2* foi desenvolvido especificamente para informações musicais digitalizadas e o modelo da IFLA apenas para requisitos bibliográficos gerais (DUNN; ISAACSON, 2002; MINIBAYEVA; DUNN, 2002; BYRD; ISAACSON, 2003; NOTES; DUNN, 2004).

2.2 Biblioteca de código aberto

2.2.1 Early Music Online (EMO)

O projeto *Early Music Online* (EMO) disponibiliza imagens digitalizadas de 324 livros de música do século XVI da *British Library* e músicas para instrumentos de teclado e alaúde, com aproximadamente 8.500 peças de música. O padrão dos metadados é o MARCxml e são utilizadas listas de autoridades da *Library of Congress* (LC) para a padronização de entradas de assunto. O Arquivo Internacional de Autoridade Internacional (VIAF) é responsável pelo conjunto de links externos que alinha as várias autoridades da biblioteca em um único recurso; além de recuperar dados vinculados por outras bibliotecas. O *Early Music Online* utiliza outros dois recursos de dados vinculados, que são: o *DBPedia* e o *MusicBrainz* (CRAWFORD *et al.*, 2014).

2.2.2 Essentia

Essentia é um exemplo de biblioteca disponível em código fonte aberto para recuperar informações musicais baseadas em áudio. É liderada sob licença Affero GPL4 e possui algoritmos reutilizáveis para extrair recursos de áudio. Visando estabelecer uma comunicação de fácil acesso com os seus usuários, inclui ligações *Python*, que permite familiaridades com o ambiente *matlab/python*. Os algoritmos deste projeto estão representados com descritores da *Music Information Retrieval* e podem ser utilizados com o *software Sonic Visualiser*. Seu design é estruturado em promover desempenho de tempo e memória e facilidade de extensibilidade visando implementar seus algoritmos por meio de cálculos em larga escala em grandes coleções musicais (BOGDANOV *et al.*, 2013).

A *Essentia* permite como alternativa a inclusão de extratores executáveis como, por exemplo, computar descritores espectrais, rítmicos e tonais disponíveis em uma única faixa de áudio. A sua primeira versão foi lançada em código aberto para ser utilizada por pesquisadores e desenvolvedores do âmbito da música.

Como opção, a *Essentia* pode ser complementada pela Gaia, outra biblioteca lançada sob a mesma licença. A *Essentia* e a Gaia quando utilizadas juntas permitem medidas e classificaram similaridades nos resultados de análise de áudio, além de gerar nos modelos de classificação de áudio (BOGDANOV *et al.*, 2013).

2.3 Conjunto de dados

2.3.1 JazzCats: Jazz Collection of Aggregated Triples

O *JazzCats* utiliza triplas de RDF para integrar conjuntos de dados no domínio da música, especializado no gênero musical *jazz*. Ele é um projeto considerado especial por integrar juntamente aos metadados de música, informações prosopográficas e interpessoais sobre músicos, beneficiando pesquisas sobre de musicologia e desempenho musical (NURMIKKO-FULLER *et al.*, 2018).

O *JazzCats* incorpora mais três projetos de conjunto de dados diferentes, que são: *Body & Soul* (dados tabulares), *WJazzD* (dados relacionais indexados em banco de dados SQLite) e *Linked jazz* (dados vinculados disponíveis em RDF). Para facilitar as necessidades dessa coleção heterogênea que é o *jazz*, os dados de *Body & Soul* e de *WJazzD* foram convertidos também em RDF utilizando estruturas ontológicas sob medida e dois fluxos de trabalhos distintos para corresponderem às necessidades e exigências das estruturas de dados nativas (NURMIKKO-FULLER *et al.*, 2018).

2.3.2 Music to knowledge (M2K)

O *Music to Knowledge* (M2K) é um projeto desenvolvido pela *Music Digital Library* (MDL) e *Music Information Retrieval* e que integra como parte do projeto da *International Music Information Retrieval Systems Evaluation Laboratory* (IMIRSEL); considerado importante por promover padronização dos dados em música.

Os resultados obtidos são discutidos com o *Music Information Retrieval Evaluation eXchange* (MIREX), evento responsável por gerar trocas de informações sobre a recuperação da informação da música em ambiente digital. Os participantes do MIREX, utilizam o *Music to Knowledge* para prototipar e avaliar novas abordagens da *Music Digital Library*, como a classificação de gênero musical e identificação de artistas (TCHENG; EHMANN; DOWNIE, 2005).

2.4 Interface

2.4.1 Apollo

Além do uso de banco de dados, bibliotecas de código aberto, conjunto de dados e interfaces para a recuperação da informação em música, há aplicativos de *software* utilizados para sistemas comerciais, que são: *GarageBand* e *Sibelius*; e sistemas não comerciais: *OpenMusic*, *QSketcher* e *Musink*, originários da comunidade de pesquisa musical. Apesar desses cinco

software serem muito utilizados por profissionais do âmbito musical, houve a necessidade de desenvolver um novo produto que adotasse uma abordagem menos estruturada e prescritiva para gravar e desenvolver manifestações musicais.

Para atender a essa necessidade foi desenvolvida a interface *Apollo*, nome dado em homenagem ao deus grego da música. O que diferencia ele dos outros sistemas comerciais e não comerciais mencionados é o modo como permite a manipulação das manifestações musicais, possibilitando fazer por texto, imagem ou áudio, estruturando a informação em telas de forma livre e permitindo que a melodia fosse consultada baseada em texto. Em outras palavras, as manifestações estão divididas em telas e com apenas um clique do mouse há possibilidade em agrupá-las em uma tela específica ou realocá-las para uma tela diferente (BAINBRIDGE; NOVA; CUNNINGHAM, 2010).

2.4.2 Musical Instruments Digital Interface (MIDI)

Uma outra interface considerada essencial para a construção de uma biblioteca popular de música digital é a *Musical Instruments Digital Interface* (MIDI). Ela é considerada um padrão de controle e comunicação de instrumentos musicais eletrônicos, diferente do *Optical Music Recognition*, responsável pela conversão automática de partituras digitalizadas *online*. A interface identifica as entradas e saídas de notas musicais, assim como representa o padrão da escala musical ocidental para especificar a afinação. Muitos sites utilizam a MIDI para armazenamento de arquivos, como o *Midiphiles* e coleção de *jazz e folk* (BAINBRIDGE *et al.*, 1999).

2.4.3 Online Music Retrieval and Searching (OMRAS)

Online Music Retrieval and Searching (OMRAS) é também um tipo de interface utilizada para recuperar informação da música em ambiente digital. Surgiu como o propósito de analisar e resolver questões relacionadas à música polifônica baseada em conteúdo, isto é, música que consiste em uma única linha melódica encontrada normalmente na sinfonia orquestral. É financiado pela Iniciativa Internacional de Biblioteca Digital JISC/NSF (DOVEY, 2001).

2.4.4 Query By Humming (QBH)

A *Query by Humming* (QBH) é uma interface criada para que os usuários de uma biblioteca especializada em música possam realizar consultas murmurando ou cantando frases lembradas da música desejada, isto é, recupera a música “por meio de ruídos”, também considerada “prova de conceito”. Bibliotecas digitais especializadas em música são amplamente desenvolvidas por esse tipo de demonstração (consideradas como ferramentas e técnicas), por descrever especificamente um conjunto disponível de documentos musicais. Interfaces como QBH são interessantes pois registram as etapas de busca que os usuários executam, apesar de não

fornecerem informações sobre suas motivações, sucessos ou fracassos e estratégia para realizar a pesquisa. Apesar da QBH não descrever informações completas sobre o comportamento musical de seus usuários, ainda assim é considerada importante para estudos da MIR, pois fornece sugestões para futuras pesquisas em sistemas de recuperação de música (CUNNINGHAM; REEVES; BRITLAND, 2003).

2.5 Ontologia

2.5.1 Manuscript Italian Poetry in Music (MUSES)

O MUSES é um projeto que integra o modelo universal de Requisitos Funcionais para Registros Bibliográficos (FRBR) de descrição de poesia e música, para construir ontologias especializadas do domínio de poesia italiana do século XVI ao XVIII. Dentre os principais objetivos do *Manuscript Italian Poetry in Music* está a integração do conteúdo indexado com bancos de dados interoperáveis para permitir que o usuário recupere recursos de projetos de bibliotecas e arquivos existentes. Isto não apenas permitirá aos pesquisadores em música identificar a tradição dos textos poéticos, como também a anatomia morfológica de cada obra, que são as instâncias que a mesma sumiu ao longo do tempo. É possível integrar os dados da música com conteúdo da *web*, como manuscritos digitalizados, edições de músicas e arquivo de áudio e vídeo (LUZZI, 2014).

2.6 Síntese dos resultados

Considerando o exposto, é possível descrever que os projetos desenvolvidos no período de 1996 a 2013, abordaram: música baseada em ruído - prova de conceito; conversão automática de partituras *online*; música baseada em conteúdo; trocas sobre pesquisa em MIR; edição de músicas e bibliotecas com recuperação baseada em áudio.

Em se tratando do período 2014 a 2020, foram desenvolvidos projetos relacionados a conjunto de dados sobre a história do *big data* na música ocidental; *linked open data*, que é utilizado por bibliotecas digitais para promover acesso à patrimônios culturais; banco de dados musicais; bibliotecas digitais de música e conjunto de dados sobre gêneros musicais.

Sendo assim, esta pesquisa demonstrou a importância e a influência no desenvolvimento de projetos e ferramentas auxiliares em bibliotecas digitais por meio do *big data*, *link data* e *linked open data*.

3 Considerações finais

Após um aprofundamento teórico e uma análise minuciosa, observou-se que a preocupação das bibliotecas que abrigavam acervos ou bases de dados de informações sobre música até o início do século XXI, criar projetos que estabelecesse uma conexão entre o ambiente físico da biblioteca com informações de dados

online sobre o domínio da música. Pode-se tornar como exemplo disso, a procura pelo uso do *Optical Music Recognition e a Musical Instruments Digital Interface*.

Mesmo assim, os projetos auxiliares são considerados desafiadores para bibliotecas especializadas em música no enfrentamento dos desafios aos problemas presentes na aquisição de coleções (músicas disponíveis gratuitamente), arquivos convertidos em MIDI, *software* de *Optical Music Recognition* atualizado e aplicado à páginas de música digitalizadas, digitalização de áudios brutos, conhecidos como ruídos e metadados ricos para pesquisas baseadas em áudio, texto e entre outros.

A partir de meados de 2010 a discussão sobre bibliotecas analógicas especializadas em música foi encerrada, da mesma forma em que ocorreu o mesmo com os temas correlatos como em bibliotecas em ambientes digitais, conjuntos de dados, interfaces e bases de dados e entre outros. A preocupação em evidência era organizar e conectar toda a informação musical existente acessível para promover melhor a representação e a recuperação da informação da música em ambientes digitais. Nesse cenário, originou-se os estudos sobre *big data*, *linked data* e *linked open data* no domínio da música.

Por meio da análise realizada nesta pesquisa, foi demonstrado que, apesar da complexidade que a área da música possui devido aos inúmeros formatos pelos quais seus materiais podem ser expressos, a pesquisa e a construção de projetos para melhor representar e consumir a informação musical acompanhou o processo de evolução tecnológica. A sintonia tecnológica e a utilização dos mais variados recursos dessa natureza para o tratamento da informação no domínio da música fica clara com a diversidade de projetos analisados.

Desse modo, apesar do fato de que uma biblioteca, seja digital ou analógica, especializada em música apresentar questões complexas relativas à descrição, representação, organização e uso de informações musicais, há uma comunidade de pesquisadores bastante ativa e produtiva na busca de soluções para os desafios que se apresentam para investigar questões e propor soluções de representação e recuperação da informação.

Notas

[1] Gostaríamos de agradecer especialmente a ACM DL, pois devido a situação de pandemia (COVID-19), como incentivo aos pesquisadores, disponibilizou seu acervo em acesso livre até o ano de 2020. Isso permitiu acesso a inúmeros materiais importantes que, inclusive, originaram este trabalho.

[2] Não poderíamos deixar de agradecer as agências de fomento CNPq e Capes pelo financiamento de nossas bolsas de pesquisas; pois elas viabilizam a nossa permanência na pós graduação e também a realização de trabalhos como este.

Referências

- ALBUQUERQUE, M. J.; PINTO, H. S.; BORBINHA, J. A. heterogeneidade na representação da informação musical: proposta de uma ontologia para a música. *In: Actas do Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas*. 2018.
- ALLIK, A.; FAZEKAS, G.; DIXON, S.; SANDLER, M. Facilitating music information research with shared open vocabularies. *In: Extended Semantic Web Conference*. Springer, Berlin, Heidelberg, 2013. p. 178-183.
- ALVES, R. C. V.; SANTOS, P. L. V. A. C. **Metadados no domínio bibliográfico**. Niterói: Intertexto, 2013.
- BAINBRIDGE, D.; NEVILL-MANNING C. G.; WITTEN, I. H.; SMITH, L. A.; MCNAB, RODGER J. Towards a digital library of popular music. *In: Proceedings of the fourth ACM conference on digital libraries*. 1999. p. 161-169.
- BAINBRIDGE, D.; NOVAK, B. J.; CUNNINGHAM, S. J. A user-centered design of a personal digital library for music exploration. *In: Proceedings of the 10th annual joint conference on digital libraries*. 2010. p. 149-158.
- BOGDANOV, D.; WACK, N.; GÓMEZ, E.; GULATI, S.; HERRERA, P.; MAYOR, O.; ROMA, G.; SALAMON, J.; ZAPATA, J.; SERRA, X. ESSENTIA: an open-source library for sound and music analysis. *In: Proceedings of the 21st ACM international conference on Multimedia*. 2013. p. 855-858.
- BREITMAN, K. K. **Web semântica: o futuro da internet**. Rio de Janeiro: LTC-Livros Técnicos e Científicos Editora SA, 2005.
- BYRD, D.; ISAACSON, E. Music representation in a digital music library. *In: 2003 Joint Conference on Digital Libraries, 2003. Proceedings*. IEEE, 2003. p. 234-236.
- CRAWFORD, T. FIELDS, B.; LEWIS, D.; PAGE, K. Explorations in Linked Data practice for early music corpora. *In: IEEE/ACM Joint Conference on Digital Libraries*. IEEE, 2014. p. 309-312.
- CUNNINGHAM, S. J.; REEVES, N.; BRITLAND, M. An ethnographic study of music information seeking: implications for the design of a music digital library. *In: 2003 Joint Conference on Digital Libraries, 2003. Proceedings*. IEEE, 2003. p. 5-17.
- DOVEY, M. J. Adding content-based searching to a traditional music library catalogue server. *In: Proceedings of the 1st ACM/IEEE-CS joint conference on digital libraries*. 2001. p. 249-250.
- DUNN, J. W.; ISAACSON, E. J. Variations2: A digital music library system. *In: Proceedings of the 2nd ACM/IEEE-CS joint conference on digital libraries*. 2002. p. 374-374.
- DUNN, J. W.; MAYER, C. A. VARIATIONS: A digital music library system at Indiana University. *In: Proceedings of the fourth ACM conference on digital libraries*. 1999. p. 12-19.
- DUNN, J. W.; MAYER, C. A. VARIATIONS: A digital music library system at Indiana University. *In: Proceedings of the fourth ACM conference on digital libraries*. 1999. p. 12-19.
- FIELDS, B.; PHIPPEN, S.; COHEN, B. A Case Study in Pragmatism: exploring the practical failure modes of Linked Data as applied to classical music catalogues. *In: Proceedings of the 2nd International Workshop on Digital Libraries for Musicology*. 2015. p. 21-24.
- HUJRAN, O.; ALIKAJ, A.; AL-DMOUR, N.; DURRANI, U. K. Big Data and its Effect on the Music Industry. *In: Proceedings of the 3rd International Conference on Software Engineering and Information Management*. 2020. p. 5-9.
- KORTH, H. F.; SILBERSCHATZ, A. **Sistemas de banco de dados**. Makron Books, 2. ed., 1994. 904 p.
- LUZZI, C. ManUScript Italian poEtry in muSic (1500-1700) interoperable model: towards an application of FRBRoo, Linked Open Data and Semantic Web technology. *In: Proceedings of the 1st International Workshop on Digital Libraries for Musicology*. 2014. p. 1-3.
- MCNAB, R. J.; SMITH, L. A.; WITTEN, I. H.; HENDERSON, C. L.; CUNNINGHAM, S. J. Towards the digital music library: Tune retrieval from acoustic input. *In: Proceedings of the first ACM international conference on digital libraries*. 1996. p. 11-18.
- MINIBAYEVA, N.; DUNN, J. W. A digital library data model for music. *In: Proceedings of the 2nd ACM/IEEE-CS joint conference on digital libraries*. 2002. p. 154-155.
- NOTESS, M.; DUNN, J. Variations2: improving music findability in a digital library through work-centric metadata. *In: Proceedings of the 4th ACM/IEEE-CS joint conference on digital libraries*. 2004. p. 422-422.
- NURMIKKO-FULLER, T.; HAO, H.; BANGERT, D.; DOWNIE, J. S. Swinging triples: Bridging jazz performance datasets using linked data. *In: Proceedings of the 1st International Workshop on Semantic Applications for Audio and Music*. 2018. p. 42-45.
- PAGE, K. R.; WEIGL, D. M.; BECHHOFER, S.; WILMERING, T.; FAZEKAS, G. Realising a layered digital library. *In: Proceedings of the 17th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'17)*. IEEE, 2017. p. 1-10.
- RILEY, J. **Seeings Standards: a visualization of the Metadata Universe**. 2010. Disponível em: <<http://jennriley.com/metadatamap/seeingsstandards.pdf>>. Acesso em: 29 mar. 2023.
- ROSE, S.; TUPPEN, S. Prospects for a big data history of music. *In: Proceedings of the 1st International Workshop on Digital Libraries for Musicology*. 2014. p. 1-3.
- TCHENG, D.; EHMANN, A. F.; DOWNIE, J. S. Real-time genre classification for music digital libraries. *In: Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'05)*. IEEE, 2005. p. 377-377.

Análise de domínio em biodiversidade

Biodiversity domain analysis

Carla Beatriz Marques Felipe (1), Luana Farias Sales (2)

(1) Universidade Federal do Rio de Janeiro, Av. Pasteur, 250 - Urca, Rio de Janeiro - RJ, 22290-240, felipecarla12@gmail.com. (2) Instituto Brasileiro de Informação em Ciência e Tecnologia, Rua Lauro Müller, 455 - Botafogo, Rio de Janeiro - RJ, 22290-160, luanafsales@gmail.com

Resumo

A base de uma gestão de dados de pesquisa eficiente é a Organização do Conhecimento (OC). Esse fato fica cada vez mais evidente, pois, dados de pesquisa podem se configurar de forma diferente de acordo com sua área de origem, seu processo de geração e seu uso dentro do contexto da pesquisa. Assim, conhecer o domínio de geração do dado de pesquisa se torna um dos primeiros passos para uma gestão de dados eficiente. No entanto, conhecer um domínio pode não ser algo trivial. Considerando que a Organização do Conhecimento possui uma base teórica já consolidada para análise de domínio, e que o problema da presente pesquisa se coloca dentro da necessidade de gerenciamento de dados de pesquisa gerados no domínio da Biodiversidade, o objetivo da pesquisa é identificar as vertentes que constituem o domínio Biodiversidade, a partir de abordagens metodológicas de análise de domínio propostas Hjørland (2002), especialmente em três delas: abordagem de estudos epistemológicos e críticos; abordagem de estudo bibliométrico; e abordagem de estudo terminológico. Para tal, a investigação ocorreu na análise das linhas de pesquisa dos Programas de Pós-Graduação brasileiros, na busca em bases de dados para verificar a produção científica da área, e na análise de um tesouro específico do domínio. Os resultados mostram que o domínio é constituído por várias disciplinas ligadas a estudos ambientais, além de outras ligadas às humanidades. Como conclusão, tem-se que a análise de domínio pode ser uma ferramenta fundamental para a melhoria dos processos de gestão de dados de pesquisa.

Palavras-chave: Análise de domínio. Biodiversidade. Gestão de dados.

Abstract

The foundation of efficient research data management is the Organization of Knowledge (KO). This fact becomes increasingly evident, as research data can be configured differently according to their area of origin, their generation process and their use within the research context. Thus, knowing the research data generation domain becomes one of the first steps for efficient data management. However, knowing a domain may not be trivial. Considering that the Organization of Knowledge has an already consolidated theoretical basis for domain analysis, and that the problem of the present research lies within the need to manage research data generated in the field of Biodiversity, the objective of the research is identify the aspects that constitute the Biodiversity domain, based on methodological approaches of domain analysis proposed by Hjørland (2002), especially in three of them: approach of epistemological and critical studies; bibliometric study approach; and terminological study approach. For this, the investigation took place in the analysis of the lines of research of the Brazilian Postgraduate Programs, in the search in databases to verify the scientific production of the area, and in the analysis of a specific thesaurus of the domain. The results show that the domain is made up of several disciplines related to environmental studies, in addition to others related to the humanities. In conclusion, domain analysis can be a fundamental tool for improving research data management processes.

Keywords: Domain analysis. Biodiversity. Data management.

Resumen

La base de la gestión eficiente de datos de investigación es la Organización del Conocimiento (KO). Este hecho se hace cada vez más evidente, ya que los datos de investigación pueden configurarse de manera diferente según su área de origen, su proceso de generación y su uso dentro del contexto de investigación. Por lo tanto, conocer el dominio de generación de datos de investigación se convierte en uno de los primeros pasos para una gestión eficiente de los datos. Sin embargo, conocer un dominio puede no ser trivial. Considerando que la Organización del Conocimiento tiene una base teórica ya consolidada para el análisis de dominio, y que el problema de la presente investigación radica en la necesidad de gestionar los datos de investigación generados en el campo de la Biodiversidad, el objetivo de la investigación es identificar los aspectos que constituyen el dominio de la Biodiversidad, basado en enfoques metodológicos de análisis de dominio propuestos por Hjørland (2002), especialmente en tres de ellos: enfoque de estudios epistemológicos y críticos; enfoque de estudio bibliométrico; y enfoque de estudio terminológico. Para eso, la investigación pasó por el análisis de las líneas de investigación de los Programas de Posgrado brasileños, por la búsqueda en bases de datos para verificar la producción científica del área, y por el análisis de un tesoro específico del dominio. Los resultados muestran que el dominio está compuesto por varias disciplinas relacionadas con los estudios ambientales, además de otras relacionadas con las humanidades. En

conclusión, el análisis de dominio puede ser una herramienta fundamental para mejorar los procesos de gestión de datos de investigación.

Palabras clave: Análisis de dominio. Biodiversidad. Gestión de datos.

1 Introdução

Atualmente, os dados de pesquisa são produzidos em larga escala e necessitam ser gerenciados para que favoreçam o desenvolvimento da ciência por meio do seu reuso. Dado de pesquisa engloba todo o registro gerado ou empregado no campo da pesquisa, produzido nas mais variadas formas (SAYÃO; SALES, 2020). Os dados não são uniformes, haja vista que são criados dentro dos mais diversos domínios. Assim, podem se configurar de forma diferente de acordo com sua área de origem, seu processo de geração e seu uso dentro do contexto da pesquisa.

Sabe-se que existem diversas estratégias para uma boa gestão de dados, tais como a curadoria digital, a aplicação de identificadores persistentes, o uso dos metadados, além do desenvolvimento de estruturas para a disseminação dos dados de pesquisa, como os repositórios de dados e os *datajournals*, porém nada disso será eficaz se a gestão de dados não levar em consideração a natureza dos dados, bem como o domínio em que foram gerados. Para Sayão e Sales (2020, p. 32) “A heterogeneidade intrínseca aos dados de pesquisa implica também necessidade de formular políticas e estratégias de gestão de amplo espectro que englobem os vários tipos de dados coletados e analisados por uma instituição ou comunidade acadêmica”.

Os autores complementam que “isso significa dizer que estudar as propriedades do dado e como ele se manifesta em cada disciplina é condição necessária para a construção de critérios que tornarão o dado de pesquisa passível de ser selecionado, arquivado e preservado, de acordo com suas características” (SAYÃO, SALES, 2020, p.33). Isto acontece porque o dado, assim como um documento, precisa de representação, para ser compreendido, recuperado, contextualizado e reusado. É essa representação que torna o dado informação, ou sob o olhar da teoria dalberghiniana (1978) uma unidade de conhecimento -formada pela tríade referente – característica e representante. Essas características, no âmbito da gestão de dados se tornam metadados que descrevem o dado para além de sua estrutura, mas guardam também informações técnicas, administrativas e de preservação. Neste sentido, o dado é visto como um conceito que para ser gerenciado precisa estar estruturado adequadamente dentro de um Sistema de Organização do Conhecimento (SOC). Posto isto, fica claro compreender que a gestão de dados requer muito das teorias da Organização do Conhecimento.

Interessante observar que assim como a Organização de Sistemas de Organização do Conhecimento, a Gestão de Dados nem sempre é iniciada por um especialista.

Muitas vezes essa gestão é iniciada por um profissional da informação de formação genérica que precisa “entrar” em um domínio de conhecimento que não é o seu original. Por este motivo, o primeiro passo na Gestão de Dados é conhecer o domínio. Apenas conhecendo esse domínio, o profissional será capaz de identificar os ciclos da pesquisa, os tipos de dados e iniciar uma gestão coerente aos objetivos que se propõe.

No entanto, ainda assim, conhecer um domínio pode não ser algo trivial, desse modo, é preciso pensar boas estratégias de investigação do domínio, que possam ser usadas, especialmente quando o domínio em questão é relativamente novo e complexo, como o domínio da Biodiversidade.

Muito se tem falado acerca das mudanças climáticas ocorridas em nosso planeta. Os estudos sobre Biodiversidade contemplam cada vez mais as questões climáticas e ambientais. Nesse contexto, as informações sobre Biodiversidade, especialmente a brasileira, se revelam como fundamentais para a preservação e o desenvolvimento de toda a sociedade. Como exemplo, podemos citar a iniciativa Saberes do Cerrado, um projeto desenvolvido em parceria entre o Jardim Botânico de Brasília, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Universidade Federal de São Carlos (UFSCAR), cujo objetivo é fornecer informações acerca do Bioma do cerrado, valorizando socialmente aspectos ligados à Biodiversidade.

Mas afinal, o que vem a ser a Biodiversidade? Segundo a Convenção da Biodiversidade, trata-se da diversidade biológica do planeta (CONVENTION ON BIOLOGICAL DIVERSITY, 2006). Conforme Alho (2012, p. 151), a Biodiversidade “refere-se à diversidade biológica para designar a variedade de formas de vida em todos os níveis, desde micro-organismos até flora e fauna silvestres, além da espécie humana”. Em vista disso, pode-se afirmar que engloba toda a diversidade biológica presente na terra. Para gerenciar e tratar dados e informações em qualquer domínio é necessário conhecer profundamente tal área, desvendando questões como: quais são os focos das pesquisas, que disciplinas formam esse domínio, entre outras.

A área de Organização do Conhecimento pode oferecer subsídios teóricos consistentes para o conhecimento estrutural e semântico de qualquer campo do saber., especialmente a partir das propostas de análise de domínio de Hjørland (2002). Assim, a questão que se coloca é: *Como se valer da abordagem da análise de domínio oriunda da Organização do Conhecimento para mapear a área de Biodiversidade, adquirindo um*

conhecimento substancial para dar suporte à gestão de dados e informações nessa área?

A análise de domínio consiste na investigação acerca de como os domínios são formados, sobretudo em sua vertente epistemológica. Na Ciência da Informação, auxilia a Organização do conhecimento, entre outras coisas, na elaboração de Sistemas de Organização do Conhecimento (SOC).

A presente pesquisa é parte da pesquisa doutoral sobre compatibilização semântica para Sistemas de Organização do Conhecimento voltados para dados em Biodiversidade. Assim, a análise de domínio auxilia na proposta final de compatibilização, visto que para propor uma compatibilização semântica, primeiro é necessário compreender como o domínio foi constituído;

Este trabalho visa relatar a análise do domínio Biodiversidade realizada durante a pesquisa de tese, cuja sustentação metodológica foi construída com base na teoria de análise de domínio proposta por Hjørland (2002), com o foco em identificar o que vem a ser o domínio Biodiversidade, conforme entendido pelos pesquisadores de Pós-graduação no Brasil, a partir da produção científica da área e de sua representação temática.

Antes de apresentar os resultados da análise de domínio realizada, é preciso primeiro conhecer seu conceito e abordagens voltadas para o seu desenvolvimento.

2 Análise de domínio

Conhecer como são constituídos os domínios para promover melhores estratégias de organização da informação, é uma das atividades desenvolvidas por pesquisadores da Ciência da Informação. Nessa circunstância, a análise de domínio tem auxiliado os pesquisadores a desenvolverem essa tarefa de forma efetiva. Segundo Freitas e Albuquerque (2017, p. 2),

a abordagem teórico-metodológica da Análise de Domínio, especificamente nos estudos relacionados à Ciência da Informação, tem como um de seus objetivos, auxiliar no processo de pesquisa e construção de instrumentos para a organização do conhecimento.

No âmbito da Ciência da Informação, vem sendo desenvolvida internacionalmente desde a década de 90 do século passado (GUIMARÃES, 2014). A Análise de Domínio permite conhecer as teorias, instituições e necessidades informacionais de dada comunidade.

Trata, segundo Smiraglia (2011), sobre o estudo das áreas do conhecimento, geralmente por meio de sua literatura e comunidade científica, objetivando compreender como sucedem os novos conhecimentos de dada comunidade. Para Danuello (2007, p. 50),

A análise de domínio é um processo pelo qual se pode verificar o que é importante ou significativo sobre algum campo do esforço humano, identificando elementos que permitem

analisar um contexto científico, tais como tendências, padrões, objetos (pessoas), processos e relacionamentos existentes.

Desse modo, a análise de domínio permite conhecer e identificar os direcionamentos e princípios que constituem o domínio, bem como reconhecer quem são os atores que atuam.

Em relação à sua aplicação e elaboração, Hjørland (2002) disserta que a análise de domínio pode ser desenvolvida através das seguintes abordagens:

1. Produção de guias de literatura e de entradas de assunto [índices];
2. Produção de classificações especiais;
3. Pesquisa em indexação e recuperação [em áreas de] especialidades;
4. Estudos de usuários empíricos;
5. Estudos bibliométricos;
6. Estudo históricos;
7. Estudos de gênero e de documentos;
8. Estudos críticos e epistemológicos;
9. Estudos terminológicos de linguagem para propósitos especiais, estudos do discurso;
10. Estudos em estruturas e instituições em comunicação científica;
11. Análise de domínio em cognição profissional e inteligência artificial (HJØRLAND, 2002, p. 422).

Todas essas abordagens possibilitam conhecer algum aspecto de um domínio, quando aplicadas, e podem ser usadas conjuntamente, a depender do objetivo da investigação. Cabe ao pesquisador discernir, de acordo com os seus objetivos, quais dessas abordagens é pertinente para o seu trabalho.

Acreditando que a análise de domínio pode auxiliar nos aspectos ligados à Organização do Conhecimento, sobretudo em compatibilização semântica, o objetivo do presente trabalho foi realizar uma análise de domínio sobre a Biodiversidade, pela aplicação de três abordagens sugeridas por Hjørland (2002): abordagem de estudos epistemológicos e críticos; abordagem de estudo bibliométrico; e abordagem de estudo terminológico. A seguir serão descritos os procedimentos metodológicos.

3 Procedimentos metodológicos

A presente pesquisa pode ser caracterizada como uma pesquisa exploratória e descritiva, de acordo com seus objetivos. Também é classificada como uma pesquisa documental, uma vez que utiliza fontes primárias como insumos para a coleta de dados.

A área da Biodiversidade foi escolhida por se tratar de uma área de pesquisa que impacta a vida de todo ser vivo do planeta, sendo responsável por fornecer alimentos, remédios, energia, cosméticos, dentre outros. Também engloba estudos sobre as mudanças climáticas, algo de bastante relevância na atualidade. Porém, para que os objetivos estabelecidos sejam alcançados, é fundamental compreender o que é a área de Biodiversidade. Para esta compreensão, foi utilizada a abordagem metodológica de Análise de Domínio de Hjørland (2002) – abordagem muito usada no âmbito das pesquisas em Ciência em Informação.

Desse modo, foram selecionadas três abordagens da análise de domínio: abordagem de estudos epistemológicos e críticos; abordagem de estudo bibliométrico; e abordagem de estudo terminológico. Do ponto de vista de Hjørland (2002), essas abordagens devem ser realizadas em conjunto para garantir uma análise de domínio bem elaborada.

Conforme Hjørland (2002), os estudos epistemológicos e críticos permitem conhecer as teorias, as metodologias e as aplicações de um domínio. Esta abordagem é mais básica, mas necessária. Consoante o autor, quando esta abordagem não é empregada, a análise de domínio tende a ser mais rasa.

Para o desenvolvimento da abordagem de estudos epistemológicos e críticos aplicou-se a análise dos escopos dos Programas de Pós-Graduação em Biodiversidade brasileiros. Este procedimento se justifica por não haver cursos de graduação específicos em Biodiversidade no país e a Pós-graduação concentrar grande número de pesquisadores de uma área. Na maioria dos casos, são responsáveis pela elaboração e pela manutenção de eventos científicos e periódicos científicos.

Posto isto, a busca foi efetivada na Plataforma Sucupira¹. Foram selecionados os programas que oferecem cursos de mestrado e doutorado acadêmicos, uma vez que ao conter um curso de doutorado, entende-se que o Programa já está mais consolidado. No total, foram encontrados 98 programas, distribuídos nas áreas de Botânica, Ecologia, Oceanografia e Zoologia. Notou-se serem estes os domínios que formam a Biodiversidade no país. Posterior à pesquisa, os programas foram priorizados, e aqueles que não apresentaram ementas das áreas de concentração, desconsiderados. Sendo assim, foram analisados 62 programas.

O estudo bibliométrico buscou demonstrar como acontece a produção científica de dado domínio e/ou assunto. Assim, neste momento, a pesquisa teve uma abordagem quantitativa, optando-se pela base de dados Scopus, que apresenta significativa relevância no meio acadêmico, sendo considerada uma das maiores e mais importantes bases de dados existentes na contemporaneidade. Conforme Sweileh (2020, p. 2, tradução nossa), a “Scopus é a maior base de dados científica disponível. Tem mais de 23.000 revistas indexadas em todas as disciplinas”. Dessa maneira, a investigação permitiu compreender a Biodiversidade numa perspectiva global, visto que são indexados periódicos de todo o mundo. A busca ocorreu com o termo Biodiversity, no campo Título, porque tem a função de apresentar ao leitor o tema central da pesquisa. Não se aplicou recorte temporal nesta busca.

A busca retornou 27.304 itens. Como o objetivo do estudo é compreender como a área se desenvolve, efetuou-se o recorte de artigos de periódicos, excetuando outros tipos de fontes de informação, tais

como livros e preprints. Assim, desenvolveu-se a análise a partir de 18.547 artigos, tendo sido considerados: as/os áreas de conhecimento dos artigos, anos de publicação, periódicos mais destacados e palavras-chave usadas nos artigos. Cabe frisar que a própria Scopus gera os gráficos para o diagnóstico dos dados recuperados, sendo a exploração sucedida por meio dessas ilustrações.

Por fim, no que concerne ao estudo terminológico, optou-se pelo ThesBio², um tesouro específico da área da Biodiversidade desenvolvido pela rede BHL Scielo, que objetiva facilitar a representação da informação e a recuperação da informação na Biodiversidade. “A Rede BHL SciELO tem como objetivo contribuir para a indexação, qualificação, publicação, acesso e interoperabilidade de informação científica em Biodiversidade. Integra as Redes BHL e SciELO”. Além de disponibilizar informações sobre a Biodiversidade, a rede desenvolveu também o tesouro, tencionando auxiliar a organização da informação sobre Biodiversidade. Nesta parte da pesquisa, analisou-se as classes temáticas, as disciplinas e os termos que formam o tesouro. Ainda, procurou-se identificar se os termos presentes no tesouro têm ligação com as temáticas identificadas, tanto nos programas de pós-graduação, quanto na análise na Scopus.

4 Resultados

Ao observarmos os Programas de Pós-Graduação no Brasil, a produção indexada na *Scopus* e o ThesBio, podemos compreender, de forma específica, quais as particularidades presentes nas pesquisas sobre Biodiversidade e, assim, compreender o que vem a ser Biodiversidade. Esta seção será dividida de acordo com as análises, e por fim, será mostrado o mapa conceitual a ser constituído pela junção dos resultados das análises.

4.1 O que é Biodiversidade de acordo com os Programas de Pós-Graduação

Como citado acima, a busca na plataforma Sucupira retornou os Programas de Pós-Graduação em Biodiversidade, divididos em grandes áreas do conhecimento: *Botânica, Ecologia, Oceanografia e Zoologia*.

Em *Oceanografia* estão as pesquisas voltadas para Oceanos e mudanças climáticas, Sistemas costeiros e oceânicos, Ciência antártica, Química da água, Geologia marinha, Ecossistemas Aquáticos; Química da Água; Sistema Carbonato e Acidificação dos Oceanos; Interações físico-biológicas e biogeoquímica, Geoquímica Ambiental e Marinha, Evolução e Morfodinâmica dos Ambientes Depositionais Costeiros e Paleontologia, Paleoclimatologia e Paleooceanografia.

Em *Botânica*, as pesquisas têm seu foco em Sistemática Vegetal, Biologia Bioquímica vegetal, Ecologia Vegetal, Morfologia vegetal e Biodiversidade e Sociedade.

Os estudos em *Zoologia* concentram-se em Zoologia Aplicada; Etimologia; Sistemática e Biogeografia; Morfologia e Fisiologia Animal; Ecologia Animal; e Conservação. Sobre a Zoologia aplicada, é correto afirmar que estas pesquisas são desenvolvidas com o foco no cultivo, manejo e controle populacional das espécies e pesquisas.

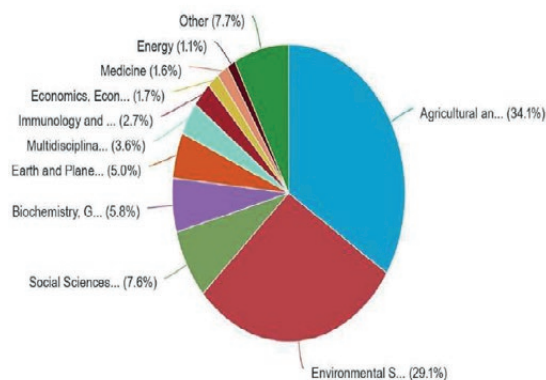
Ecologia mostrou-se a área mais ampla, englobando estudos sobre Ecologia Aplicada; Conservação; Ecologia de Ecossistemas e comunidades; Ecologia de Populações; Agroecologia; Limnologia; Ecologia e Biodiversidade Terrestre; Ecologia Aquática; Ecossistemas Marinhos; Ecologia Humana; Ecologia e Manejo de Recursos Pesqueiros; Biogeoquímica; Ecologia de Paisagem; Contaminação Ambiental; Sistemática e Evolução; Biogeografia; Ecologia e Genética de Populações; Ecotoxicologia e Qualidade Ambiental; Biologia e Ecologia de Organismos Costeiros; Processos naturais e antrópicos em Sistemas Costeiros; Oceânicos e Gestão e Conservação da Zona Costeira e Oceânica, Recursos Naturais; Ecologia Vegetal e Animal; e Gestão e Educação Ambiental.

Pode-se perceber que as áreas aqui demonstradas apresentam interseccionalidade umas com as outras, sendo a Ecologia que apresenta maior abrangência em seu objeto de estudo. A única que não apresenta algo especialmente voltado para a Biodiversidade ou Sociedade em geral é a Zoologia, porém percebe-se que as investigações sobre Biogeografia e Biologia animal podem contribuir para o que se entende, até aqui, sobre Biodiversidade.

4.2 O que é Biodiversidade de acordo com a literatura

O estudo bibliométrico realizado demonstra uma ligação com as temáticas dos programas de pós-graduação, porém com algumas informações novas, como a presença das Ciências Sociais e da Medicina ligadas à Biodiversidade. A seguir, as áreas do conhecimento presentes nos artigos recuperados que estudam a Biodiversidade (Gráfico 1).

Gráfico 1 - Documentos por área temática



Fonte: Elaborado pelas Autoras.

Conforme o gráfico 1, a área do conhecimento que mais produziu sobre Biodiversidade foi Ciência Agrícola e Biológica, com 34,1% dos artigos recuperados, seguida de Ciência Ambiental, com 29,1% dos artigos, depois Ciências Sociais, com 7,6%. A Biologia aparece aqui como a maior incidência, presume-se que seja porque, assim como nos dados dos Programas de Pós-Graduação, também está presente na maioria das pesquisas, sendo que aqui aparece em conjunto com a Ciência Agrícola. Acredita-se que a Ciência Agrícola se faz presente na literatura acerca da Biodiversidade porque versa sobre plantas e alimentação e pode impactar os aspectos ligados à mudança climática. Por sua vez, o que o banco Scopus denomina Ciência Ambiental, tem total ligação com as temáticas encontradas nos Programas de Pós-Graduação.

Também são encontradas áreas como Bioquímica, junto à Genética e Biologia Molecular (5,8 %), Ciências da Terra e Planetárias (5,0 %), áreas Multidisciplinares (3,6 %), Imunologia e Microbiologia (2,7 %), Economia, Econometria e Finança (1,7 %), Medicina (1,6), e Energia (1,1 %). Com exceção de Medicina, Economia e Energia, todas as outras áreas estão presentes dentro dos estudos dos Programas de Pós-Graduação. Não foi possível descobrir quais são as áreas multidisciplinares que estão presentes no gráfico.

Com os dados da pesquisa, ainda é possível constatar a presença de outras áreas, apresentadas como Outras, que juntas formam cerca de 7,7% da produção, essas áreas são: Artes e Humanidades, Ciências da Decisão, Negócios, Gestão e Contabilidade, Ciência da Computação, Química, Matemática, Engenharia Química, Farmacologia, Toxicologia e Farmacêutica, Física e Astronomia, Neurociência, Veterinária, Ciência de Matérias, Psicologia, Enfermagem e áreas ligadas à Saúde. Acredita-se que estas áreas, que não possuem ligação direta com as temáticas dos Programas de Pós-Graduação, estão presentes como áreas secundárias, uma vez que não têm o foco nos estudos ambientais, porém fazem com que os estudos ambientais possam funcionar.

Os periódicos que apresentam maior incidência de produção são: *Biodiversity And Conservation*; *Plos One*; *Biological Conservation*; *Conservation Biology e Ecological Indicators*, todos com mais de 200 artigos recuperados. O único periódico multidisciplinar é o *Plos One*, os outros periódicos estão voltados para os estudos ambientais. Tem-se ainda os periódicos: *Journal Of Applied Ecology*, *Forest Ecology And Management*; *Biodiversity*; *Proceedings Of The National Academy Of Sciences Of The United States Of America* e o *Agriculture Ecosystems And Environment*, com mais de 150 trabalhos recuperados. Percebe-se a presença de um periódico intitulado Biodiversidade, que segundo a Scopus versa sobre Ciências Ambientais e trata dos temas Conservação, Ecologia e Mudanças globais e

planetárias. O gráfico 2 (Apêndice A) mostra a porcentagem de produção dos periódicos mais citados.

Por fim, foi realizada a identificação dos termos mais utilizados como palavras-chave nos artigos. A análise dessas palavras-chave é importante para observar quais os termos mais empreendidos pelos pesquisadores na Biodiversidade. Os termos com o maior número de citações são termos ligados às ciências ambientais, tais como: riqueza de espécies, animais, conservação, diversidade de espécies, não humanos, proteção ambiental, ecossistema, gestão de conservação, alterações climáticas, conservação de recursos naturais, taxonomia, aves, conservação da biodiversidade, ecologia, humano, floresta, classificação, planejamento de conservação, filogenia, serviço de ecossistema, genética e outros. As palavras recuperadas nos artigos têm fusão com as 4 grandes áreas encontradas nos programas de Pós-Graduação e suas pesquisas. Também são encontrados termos ligados à Agroecologia e Química, isso deve ocorrer porque as pesquisas em Ecologia envolvem essas temáticas e os estudos ligados à Química ocorrem na Oceanografia, Ecologia e Botânica.

Após conhecer a produção de literatura da Biodiversidade, se faz importante entender como ocorre a representação da informação na área. Sabe-se que a Biodiversidade faz uso dos mais variados SOC, mas quais termos e aspectos do domínio são contemplados nesses SOC? A análise da representação foi elaborada com base em um tesouro.

4.3 O que é Biodiversidade de acordo com a representação do domínio

Através da análise do ThesBio, pode-se perceber que a temática da Biodiversidade está sendo desenvolvida por meio de disciplinas ligadas ao Meio Ambiente, assim como às Ciências Humanas, Médica e Tecnológica.

Na capa do tesouro, encontra-se o que podemos identificar como classes temáticas que formam o tesouro, são elas: Biossegurança e Biotecnologia; Ciências Agrárias; Ciências Ambientais; Denominações Geográficas; Disciplinas; Fenômenos e Processos; Organismos e Política e Gestão. Essas classes temáticas não possuem notas de escopo para a sua definição, alguns dos termos, quando apresentados isolados, como por exemplo Biossegurança, possuem uma nota de escopo e a fonte bibliográfica de onde foi retirada a informação.

Quando se analisou a classe temática Disciplina, foram recuperados os seguintes termos: Agronomia, Biologia, Ciências Médicas, Direito, Educação, Engenharia Química, Engenharia Genética, Oceanografia, Pesquisa, Tecnologia da Informação e Zootecnia. Foram encontradas disciplinas que ainda não tinham sido citadas nem pelos programas de Pós-Graduação e nem pelos artigos do Scopus, como Direito, Educação,

Engenharia Química, Engenharia Genética, Pesquisa, Tecnologia da Informação e Zootecnia.

Quanto à observação e comparação dos termos presentes no ThesBio e nas áreas dos Programas de Pós-Graduação e nos artigos no Scopus, aparentemente percebe-se um cruzamento entre as temáticas sobre Ecologia, Oceanografia, Botânica, Zoologia, Ciências Médicas e Saúde, Agronomia e Ciências Sociais, aqui representadas pelo Direito e os aspectos da Tecnologia. No entanto, seria preciso um estudo conceitual mais aprofundado para identificar esses cruzamentos com mais clareza, o que ficou inviabilizado em um primeiro momento pela ausência de definição na maior parte dos termos.

Por se tratar de um tesouro, os termos são apresentados de maneira mais específica do que nas ementas dos Programas. Como exemplo, o termo “Conservação”, que está presente como objeto de estudo nos Programas de Pós-Graduação e como objetivo de produção científica presente na Scopus, no ThesBio, este termo pode ser encontrado como: Biologia da Conservação, Conservação da Diversidade Biológica, Conservação da Natureza, Conservação de Recursos, entre outros. Isso mostra que a diversidade não se encontra apenas no tema, mas também na quantidade de nomenclaturas (signos linguísticos) existentes para nomear um mesmo referente, ou seja, há uma diversidade terminológica que atrapalha a comunicação dentro deste domínio e isso é mais uma evidência da necessidade de padronização semântica.

Como se pode perceber, dependendo do contexto investigado, a Biodiversidade se apresenta de várias formas, mas sempre centralizada nas questões ambientais. Para demonstrar o entendimento do que vem a ser Biodiversidade, de acordo com a junção dos dados coletados, um mapa conceitual é apresentado no (Apêndice B). É preciso enfatizar que um mapa conceitual nunca está 100% pronto e que pode sofrer alterações.

5 Conclusão

No contexto da presente pesquisa, ficou evidente que as abordagens propostas por Hjørland (2002) para a Análise de Domínio podem funcionar como uma excelente ferramenta para a compreensão dos assuntos cobertos por uma área, sendo algo basilar para a melhoria no processo de gestão de dados de pesquisa.

Compreender o contexto em que esses dados são gerados é fundamental para contemplar as reais necessidades informacionais dos usuários. Ao analisar esse domínio, ficou claro que os temas cobertos pela área vão além dos estudos ambientais, sendo identificados estudos interdisciplinares envolvendo áreas como Educação, Direito, Tecnologia da Informação e Ciências Médicas.

Acredita-se que esses resultados só foram possíveis devido à seleção das três abordagens escolhidas, sugeridas por Hjørland (2002): abordagem de estudos epistemológicos e críticos; abordagem de estudo bibliométrico; e abordagem de estudo terminológico. Uma abordagem sozinha poderia ter omitido alguns conceitos importantes para a compreensão da área de Biodiversidade. Uma análise ainda mais profunda e detalhada do domínio poderia ainda ser realizada a partir da aplicação das outras abordagens sugeridas por Hjørland, em conjunto, pois permitiria identificar também aspectos históricos, caracterizar os usuários, além de outros direcionamentos sobre formação das instituições relacionadas ao domínio.

Por fim, registra-se aqui que o presente trabalho é parte de uma pesquisa doutoral sobre compatibilização semântica entre vocabulários na área de Biodiversidade, onde a análise de domínio foi a primeira fase desenvolvida na pesquisa, como primeiro passo para a concepção de uma compatibilização semântica no contexto da gestão de dados de pesquisa.

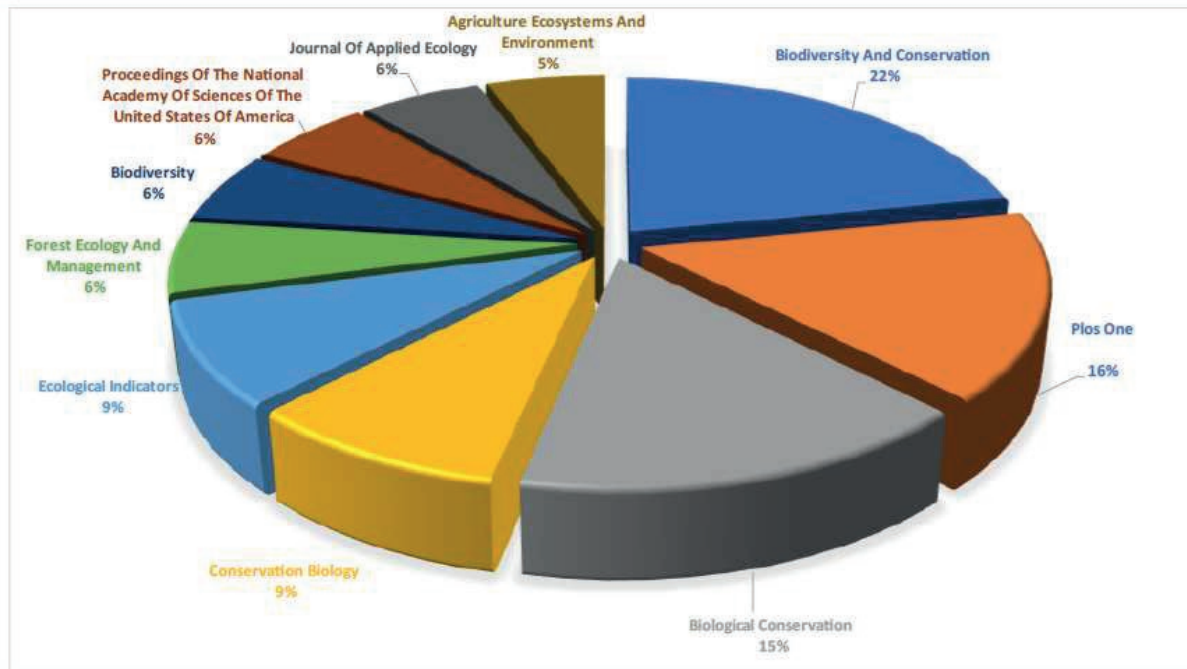
Notas

- [1] Plataforma Sucupira. Disponível em: <https://sucupira.capes.gov.br/sucupira/>. Acesso em: 15 jan. 2023.
- [2] ThesBio: Thesuros em Biodiversidade. Disponível em: <http://thesaurus.bhlscielo.org/vocab/index.php>. Acesso em: 15 fev. 2023.

Referências

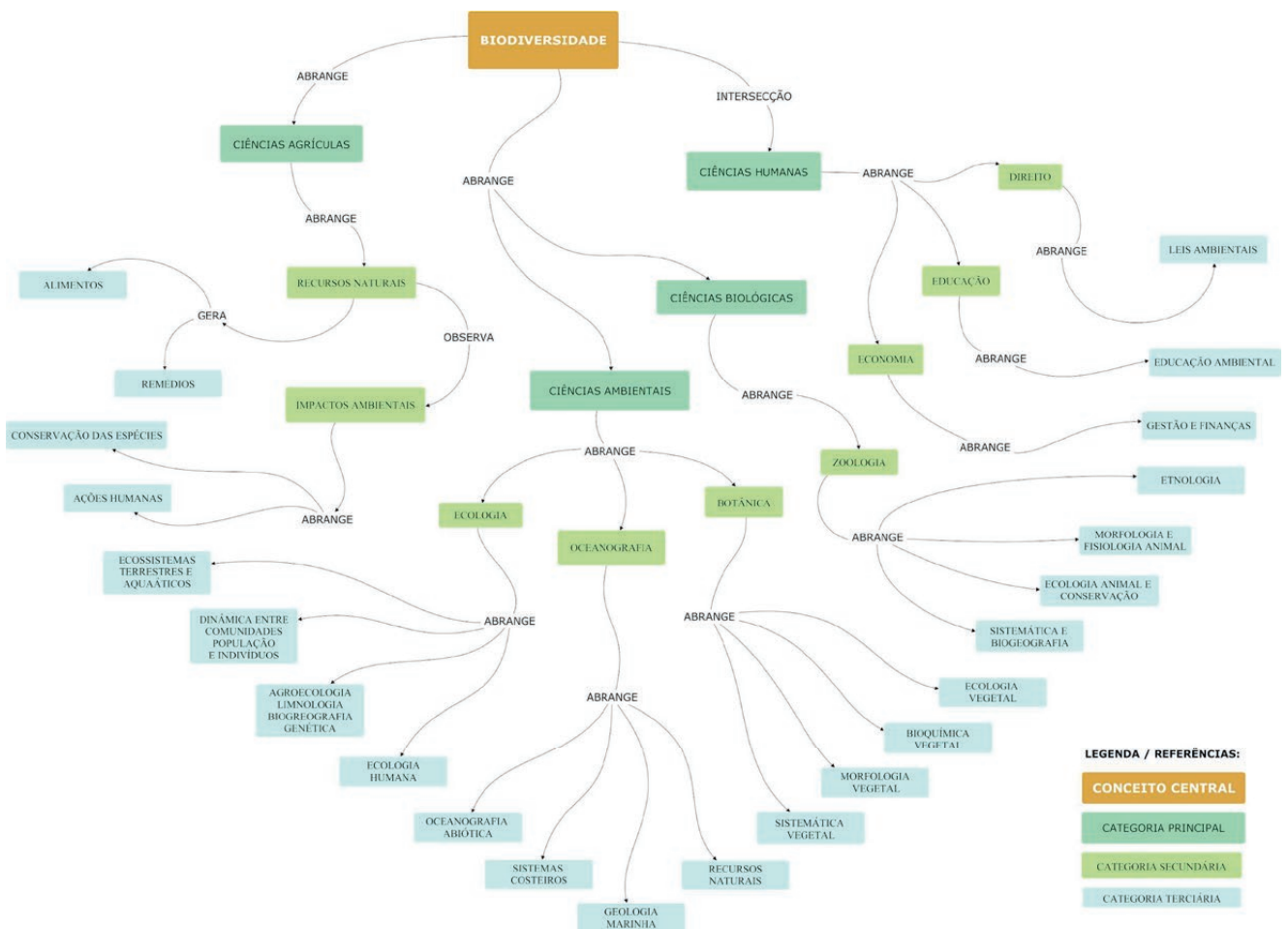
- ALHO, Cleber José Rodrigues. Importância da biodiversidade para a saúde humana: uma perspectiva ecológica. **Estudos avançados**, São Paulo, v. 26, n.74, p.151-165, 2012. Disponível em: https://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142012000100011. Acesso em: 30 jan. 2022.
- CONVENTION ON BIOLOGICAL DIVERSITY. **Article 2. Use of Terms**. 2006. Disponível em: <https://www.cbd.int/convention/articles/?a=cbd-02>. Acesso em: 05 fev. 2023.
- DAHLBERG, Ingetraut. Teoria do conceito. *Ciência da informação*, v. 7, n. 2, 1978.
- DANUELLO, Jane Coelho. **Produção científica docente em tratamento temático da informação no Brasil**: uma abordagem métrica como subsídio para a análise do domínio. Orientador: José Augusto Chaves Guimarães. 2007. 122f. Dissertação (Mestrado) – UNESP, Marília, 2007.
- FREITAS, Lidiane Marques; ALBUQUERQUE, Ana Cristina de. As abordagens da análise de domínio como aporte metodológico para a classificação arquivística. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v.10, n. 2, p. 1-19, ago./dez. 2017.
- GUIMARÃES, José Augusto Chaves. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, v. 41, n. 1, p.13-21, jan./abr., 2014.
- HJØRLAND, Binger. Domain analysis in information science: eleven approaches - traditional as well as innovative. **Journal of Documentation**, v. 58, n. 4, p. 422-462, 2002.
- SAYÃO, Luís Fernando; SALES, Luana Farias. Afinal, o que é dado de pesquisa? **Biblos**, v. 34, n. 02, p.32-51, jul./dez. 2020.
- SMIRAGLIA, Richard P. Domain coherence within Knowledge Organization: People, Interacting Theoretically, Across Geopolitical and Cultural Boundaries. *In*: MCKENZIE, Pam; JOHNSON, Catherine A.; STEVENSON, Sarah (ed.). ANNUAL CAIS/ACSI CONFERENCE, 39., 2011, Fredericton, Canada. **Exploring interactions of people, places, and information**. Fredericton: University of New Brunswick, 2011. 6 p.
- SOBRE a Rede BHL Scielo. BHL Network. Disponível em: <https://www.bhlscielo.org/> Acesso em 02.fev. 2023.
- SWEILEH, Waleed M. Research trends and scientific analysis of publications on burnout and compassion fatigue among healthcare providers. **Journal of Occupational Medicine and Toxicology**, v. 15, n. 23, p. 1-10, 2020. Disponível em: <https://occup-med.biomedcentral.com/articles/10.1186/s12995-020-00274-z>. Acesso em: 10 jan. 2023.

Apêndice A – Produção dos periódicos mais citados



Fonte: Elaborado pelas Autoras.

Apêndice B – Mapa conceitual: Biodiversidade



Fonte: Elaborado pelas Autoras.

Aspectos constitutivos e atributos da fotografia documental

Constitutive aspects and attributes of documentary photography

Adriana Aparecida Lemos Torres (1), Benildes Coura Moreira dos Santos Maculan (2)

(1) Programa de Pós-graduação em Gestão e Organização do Conhecimento (PPGGOC); Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, adrianalemos.ufmg@gmail.com (2) benildes@gmail.com

Resumo

A fotografia constitui importante documentação para a transmissão de informações e para a memória. Contudo, trata-se de documento complexo, com características e linguagem próprias. O objetivo deste artigo foi pesquisar os atributos da fotografia documental que devem ser considerados na representação desse documento iconográfico. A metodologia utilizada foi a Análise de Conteúdo de Bardin (1977). Como resultado da revisão bibliográfica e das análises e categorizações realizadas, apresentam-se breve histórico da fotografia; as principais características da fotografia e os seus atributos. Conclui-se que a fotografia documental constitui documentação relevante e para a sua representação com vistas à recuperação das informações imagéticas devem-se considerar os atributos intrínsecos e extrínsecos mapeados, incluindo os aspectos técnicos, e realizar a representação descritiva e a temática.

Palavras-chave: Fotografia. Atributos. Representação de fotografias. Indexação.

Abstract

Photography constitutes important documentation for the transmission of information and for memory. However, it is a complex document, with its own characteristics and language. The objective of this article was to research the attributes of documentary photography that should be considered in the representation of this iconographic document. The methodology used was Bardin's Content Analysis (1977). As a result of the bibliographic review and the analyses and categorizations carried out, a brief history of photography is presented; the main characteristics of the photograph and its attributes. It is concluded that documentary photography constitutes relevant documentation and for its representation, in order to recover image information, the intrinsic and extrinsic attributes mapped, including the technical aspects, must be considered, and the descriptive and thematic representation must be carried out.

Keywords: Photography. Attributes. Representation of photographs. Indexing.

Resumen

La fotografía constituye importante documentación para la transmisión de informaciones y para la memoria. Sin embargo, tratase de documento complejo, con características y lenguaje propios. El objetivo de ese artículo fue buscar los atributos de la fotografía documental que deben ser considerados en la representación de ese documento de infografía. La metodología utilizada fue el Análisis de Contenido de Bardin (1977). Como resultado de la revisión bibliográfica y de los análisis y categorizaciones hechas, presentase breve histórico de la fotografía; las principales características de la fotografía y sus atributos. Concluirse que la fotografía documental constituye documentación relevante y para su representación mirando a la recuperación de las informaciones de imágenes debe-se considerar los atributos intrínsecos y extrínsecos mapeados, incluyendo los aspectos técnicos, y realizar la representación descriptiva y temática.

Palabras clave: Fotografía. Atributos. Representación de fotografías. Indexación.

1 Introdução

A fotografia constitui um tipo documental e, segundo Boccato e Fujita (2006), transmite informações e registra o passado. Manini (2011), baseada nos princípios da teoria peirciana, propõe três eixos conceituais para a fotografia documental, também denominado tripé fotográfico: 1) a fotografia como imagem, 2) a fotografia como memória e, 3) a fotografia como informação.

Com linguagem própria, os documentos fotográficos se prestam a informar, registrar, documentar, ilustrar,

auxiliar na memória, na gestão, na ciência, na comunicação, entre outras áreas.

Para a representação da informação imagética em acervos fotográficos físicos ou virtuais, tendo em vista a recuperação da informação, é imprescindível a consideração dos atributos específicos desse tipo documental, que se diferem dos documentos textuais, conforme destacado na literatura específica sobre representação de imagens. Sobre essa questão, Santos e Azevedo (2021) destacam ser preciso que o profissional da informação tenha um conjunto de saberes especializados que o tornem capaz de analisar imagens

e identificar nelas elementos descritivos e temáticos. Desse modo, segundo os autores, esse profissional deve ter “conhecimentos sobre a natureza, aspectos técnicos da fotografia, tais como equipamento utilizado, e contextualização das obras” (SANTOS; AZEVEDO, 2021, p. 148).

Este estudo é parte de pesquisa de mestrado já concluída e apresenta a análise dos resultados da revisão bibliográfica. O objetivo é apresentar os aspectos constituintes da fotografia documental, suas características relevantes e os seus principais atributos.

Para tanto, depois desta breve introdução, este artigo se organiza da seguinte maneira: a seção 2 apresenta a metodologia, em que está explicada a revisão bibliográfica e a metodologia de Análise de Conteúdo; a seção 3 trata dos Resultados obtidos, com as inferências e interpretações realizadas, trazendo um breve histórico da fotografia, assim como as principais características e atributos da fotografia; por fim, apresentam-se as Considerações finais.

2 Metodologia

Trata-se de pesquisa de natureza aplicada e descritiva, com abordagem qualitativa do fenômeno estudado, com foco em interpretações e inferências. O método de análise utilizado foi a Análise de Conteúdo, proposto por Bardin (1977), que é um conjunto de técnicas para análise de conteúdos de documentos para a inferência de conhecimentos. O método foi aplicado a partir de três etapas básicas, a saber: 1) pré-análise (exploração da literatura), 2) exploração do material, e 3) tratamento dos resultados, inferência e interpretação, sem uso de indicadores quantitativos.

3 Resultados

Com os documentos recuperados na revisão bibliográfica, passou-se à primeira etapa do método da Análise de Conteúdo, de pré-análise, que consistiu em selecionar os documentos para a análise e delimitar o propósito da análise a partir da seguinte questão norteadora: Quais os atributos devem ser considerados na representação de documentos iconográficos do tipo fotografias, para representar os aspectos intrínsecos e extrínsecos desse tipo de documento?

Na segunda etapa, de exploração do material, os documentos selecionados para análise são profunda e detidamente examinados, buscando quadros de referência e verificando o alinhamento dos conteúdos ao propósito pretendido. Também foi averiguado se esse conteúdo cobria o escopo investigado. Ressalta-se que os procedimentos das duas primeiras etapas não são lineares, mas são interdependentes.

Por fim, na terceira etapa, de tratamento dos resultados, inferência e interpretação, foi feita a codificação de cada documento e uma leitura técnica para identificação das unidades de registro e de contexto, de modo que a análise se restringisse ao escopo, tendo significado dentro do contexto, para delimitar a interpretação. Também foi feito o agrupamento de conteúdos, a partir de três categorias estipuladas *a priori*: Delineamento conceitual da fotografia; Características da fotografia; Atributos intrínsecos e extrínsecos da fotografia. Nessa etapa também se faz a inferência, quando há uma operação intelectual que busca extrair dos dados informações e correlações, que são descritos e graficamente apresentados (tabelas, quadros, figuras etc.). Esse conteúdo dá respaldo às interpretações, quando há a construção do sentido a partir do uso de aportes teórico-metodológicos oriundos da literatura de especialidade.

A seguir apresentam-se os conteúdos resultados das ações de inferência e interpretação das categorias analisadas: a) Delineamento conceitual da fotografia (3.1); b) Características da fotografia (3.2); c) Atributos intrínsecos e extrínsecos da fotografia (3.3).

3.1 Delineamento conceitual da fotografia

Na literatura sobre a fotografia encontram-se várias denominações para retratá-la, tais como “iconografia saída da câmera escura” (BORGES, 2005, p. 11) e “iconografia fotográfica” (KOSSOY, 2001, p. 28); e, “ao ser tratada como ícone (espelho do real), a fotografia é considerada mimética em relação ao seu referente, espelhando assim a realidade fotografada” (SANTOS; AZEVEDO, 2021, p. 144). A fotografia também é uma “alternativa a mais para a leitura da realidade” (CANABARRO, 2005, p. 26), “representação visual” (CANABARRO, 2005, p. 33), “texto não verbal” (BENTES PINTO, 2008, p. 319). Nos seus primórdios, foi considerada como “reprodução fiel do real” (BORGES, 2005, p. 24), e ainda se veem os vestígios dessa crença. É considerada como “traço do ocorrido” (MERLO; BRANDÃO, 2013, p. 113), “cópia de um referente” (RODRIGUES, 2007, p. 67) e “disseminadora de mensagens” (LIMA; SILVA, 2002, p. 2).

A fotografia é a “arte de escrever com a luz” (RODRIGUES, 2007, p. 69), conforme a origem grega da palavra, *phos* ou *photo* quer dizer luz, e *graphein* quer dizer marcar, escrever, registrar (DICIONÁRIO ETIMOLÓGICO, 2008-2018). O termo fotografia significa “a arte de fixar a luz de objetos mediante a ação de certas substâncias” (BELLONE, 1997 *apud* BORGES, 2005, p. 37-38). Lima e Silva (2002) sintetizam o conceito de fotografia, ressaltando a sua natureza relacionada à luz e à tecnologia.

No campo de estudos da Ciência da Informação, Santos e Azevedo (2021, p. 144) destacam “a relação entre

fotografia e linguagem”, que, segundo os autores, “pressupõe a presença de um código que medeia a comunicação entre um emissor (fotógrafo) e um receptor (diferentes formas ou demandas de leituras da fotografia)”.

Para entender conceitualmente a fotografia, é necessário explorar e conhecer sua origem. A fotografia surgiu no século XIX, no contexto da Revolução Industrial, período de muitas transformações sociais, culturais e científicas. Sua origem relaciona-se ao desenvolvimento de novas tecnologias e ao apoio de homens da ciência, da política e do comércio.

O primeiro registro fotográfico é atribuído ao francês Joseph Nicéphore Niépce (1765-1833), em 1826. E a primeira espécie de câmera fotográfica é atribuída ao francês Louis-Jacques-Mandé Daguerre (1789-1851), que desenvolveu o daguerreótipo em 1837/38. Este fato contribuiu para a comercialização da fotografia.

Alguns estudiosos (BORGES, 2005; MAUAD, 1996; TACCA, 2005 e outros) defendem que a fotografia não pode ser considerada invenção de apenas uma pessoa, pois é resultado de diversas pesquisas, emprego de distintas tecnologias e esforços de vários pesquisadores. Nesse sentido, Borges (2005) ressalta os esforços individuais e/ou coletivos de pessoas de diferentes lugares da Europa e das Américas, que se dedicaram a pesquisas de diversos processos físico-químicos com o objetivo de captar e fixar imagens na câmera escura. Dentre eles está incluído Antoine Hercule Romuald Florence (1804-1879), francês radicado no Brasil.

Florence desenvolveu pesquisas sobre reprodução de imagens mediante processos químicos e os denominou *photographie*, termo que se tornou usual somente após 1839 (BORGES, 2005, p. 38). A autora destaca que a atribuição da invenção da fotografia aos nomes de Niépce e Daguerre foi decorrente da divulgação da experiência desses dois franceses em 1839, por François Arago, membro do Parlamento Francês, na Academia de Ciências e de Belas Artes da França.

Destaca-se que a fotografia, desde a sua origem, foi uma invenção fascinante, mas não acessível a todos. Tanto a criação quanto a popularização da fotografia foram marcadas por transformações sociais, culturais e científicas. No aspecto social e cultural, a fotografia possibilitou a democratização da memória social (de identidade) e coletiva (evolução da sociedade). No aspecto científico, é um mecanismo de transmissão de conhecimento. Os interesses comerciais, as mudanças de comportamentos, o desenvolvimento tecnológico e as mídias proporcionaram a popularização e a explosão de imagens fotográficas.

Assim, a partir de meados do século XIX e início do século XX, a intensificação de imagens nos jornais e revistas e as inovações tecnológicas que favoreceram a

popularização das câmeras fotográficas contribuíram para a disseminação e o uso das fotografias.

Desde a década de 1990, com o desenvolvimento do processo digital, a fotografia passa por um processo de massificação em grandes proporções. No início da explosão digital, o papel ainda permaneceu como suporte fotográfico. Aos poucos, ele foi substituído pelo armazenamento de fotografias no computador ou nas câmeras digitais.

No início do século XXI, a expansão das redes sociais e as tecnologias digitais inauguram “a nuvem” como novo suporte fotográfico (PAGANOTTI, 2016, p. 25). Castro (2022, p. 5) afirma que “antes estáticas, fotografias ganham mobilidade e fluidez por meio da circulação digital e online e da velocidade imposta a ela em redes sociais e aplicativos”. A inserção de câmeras fotográficas nos aparelhos celulares também contribuiu para a grande produção e disseminação de imagens fotográficas.

Conceitualmente refletindo sobre a fotografia, sabe-se que ela transformou a maneira de as pessoas verem o mundo. E também alterou a forma de transmissão da informação, conforme defendem Fagá e Costa (2014) e Kossoy (2001). A fotografia se tornou um novo recurso para a transmissão da informação e do conhecimento, antes “transmitidos unicamente pela tradição escrita, verbal e pictórica” (KOSSOY, 2001, p. 26). Ela é também instrumento de comunicação, sendo empregada como reforço de identidade, gritos ideológicos, prova documental e de valor histórico, assim como pode inventar um novo e fantástico mundo. Por exemplo, “dentre os documentos produzidos, que atuam como prova de atividades realizadas durante e após períodos de graves violações dos direitos humanos, há uma numerosa ocorrência de fotografias” (DI PIETRO, 2020, p. 2).

3.2 Características da fotografia

Pelo fato de a fotografia ser um artefato cultural e social, para a sua análise, deve-se considerar o contexto de produção, o propósito, o fotógrafo, as tecnologias utilizadas no processo fotográfico e o contexto de recepção.

No conjunto de imagens que traduzem o mundo, a fotografia pode servir como uma alternativa a mais de leitura da realidade. Enquanto produto cultural, é uma construção feita por um sujeito mediador, o fotógrafo, que seleciona pessoas e elementos e os enquadra na bidimensionalidade de um espaço a ser recortado. Entre esse sujeito e o retratado está a tecnologia, que permite a fixação da cena escolhida. Visto a fotografia ser um produto cultural, a sua construção faz parte de um determinado contexto histórico, que influencia na construção do olhar do fotógrafo, nas representações sociais impressas e no equipamento tecnológico empregado para a tomada da imagem (CANABARRO, 2005, p. 26).

A fotografia não é uma imagem isenta, pois é resultante de muitas variáveis contextuais. Rodrigues (2007) afirma que o fotógrafo, ao escolher um referente, irá registrá-lo de acordo com a sua visão de composição, e acrescenta que as fotografias são produzidas segundo distintos propósitos (cultural, religioso, político, ideológico, educacional, científico, comercial, publicitário, artístico). Isso tem relação, entre outros, com as características sociais, culturais, religiosas e econômicas de cada sociedade ou grupo.

O caráter social das fotografias é corroborado por Canabarro (2005, p. 24), que afirma que as imagens “são todas plausíveis à ação dos campos, das influências, das filiações, das referências, das determinações sociais e dos códigos de leitura, não estando atreladas ao determinismo tecnológico, pois, antes de tudo, são produtos sociais”. Desse modo, como referência social, a “história social procura ver a fotografia como um documento histórico [...] como um fragmento da história da sociedade”, imagem apropriada para “acompanhar a dinâmica das famílias [...], assim como a evolução ou contrastes de espaços físicos ao longo do tempo.” (FERREIRA, 2021, p. 25).

A fotografia atua, sobretudo, como representação. Segundo Santaella (1998, p. 159), a representação é uma “noção comumente aceita por quase todos os teóricos da imagem – pressupõe a preexistência de um objeto representado que seja da ordem da realidade visível” ideia também estabelecida por Santos e Azevedo (2021).

Shatford Layne (1986, p. 39) trata os documentos iconográficos como “representações” que contêm “informações na forma de imagens ou imagens representacionais”. Borges (2005, p. 80) também atesta o caráter representacional da fotografia: “sabe-se que uma dada imagem é uma representação do mundo que varia de acordo com os códigos culturais de quem a produz”.

Ao mesmo tempo em que a fotografia apresenta semelhanças com outras iconografias (constitui um registro visual), ela guarda especificidades em vários aspectos. Dentre os diversos aspectos possíveis, destacam-se as seguintes características: natureza, processo de produção, gênese, movimento, temporalidade e espaço, suportes, elementos constitutivos, linguagem e elementos componentes, que estão discutidas a seguir.

Desde a invenção da fotografia, existem discussões quanto à sua *natureza*, ora próxima às artes, ora mais próxima às tecnologias. Enquanto uns concebiam-na como uma técnica precisa e exata que permitiria ao homem moderno realizar seu sonho de conquista e domesticação da natureza, outros a encaravam como uma estética inteiramente nova que viria revolucionar o mundo das artes (BORGES, 2005). A partir dos anos 20 e 40 do século XX, a fotografia foi conceituada como

uma “imagem híbrida” (BORGES, 2005, p. 39). Segundo a autora, a concepção de imagem híbrida é “uma amálgama de natureza, técnica e cultura, cuja análise não se reduz a um único centro” (BORGES, 2005, p. 40).

Ao considerar o *processo de produção*, com base em Santaella (1998) e Rodrigues (2007), a fotografia é uma imagem técnica, por originar-se de procedimento automático de captação da imagem por meio da utilização de equipamentos. Nesse aspecto, a fotografia assemelha-se às “imagens matemáticas” (como holografias, infogramas), que, por sua vez, são produzidas por processos matemáticos (captadas por meio de softwares e outros procedimentos), também considerados automáticos. No entanto, diferencia-se das “imagens artesanais” ou “imagens tradicionais” (como pinturas, gravuras, desenhos, esculturas), em que se utilizam instrumentos manuais tais como pincéis, cinzéis e lápis, para a sua produção (SANTAELLA, 1998, p. 159-160; RODRIGUES, 2007, p. 69).

Souza e Silva (2022) afirmam que a fotografia inaugurou a era das imagens técnicas e foi, desde sempre, uma imagem numérica. Essa dimensão do cálculo é manifestada na óptica, mecânica, química ou eletrônica presentes na estrutura da técnica fotográfica.

Quanto à *gênese* fotográfica, trata-se de físico-química, segundo Borges (2005), pois para a invenção da fotografia foi necessário o desenvolvimento de dois processos: o físico (ótica) e o químico (ou mais tarde eletrônico) (PAGANOTTI, 2016). O processo físico seria responsável pela captação da imagem, e, o processo químico, pela fixação da imagem captada.

Quanto ao *movimento*, a fotografia – ao lado de outros tipos de imagens tais como pinturas, gravuras, desenhos, esculturas, litografias e charges – constitui o tipo imagem fixa, em regra¹, diferenciando-se das imagens animadas ou em movimento (vídeos e animações), conforme classificação de Costa (2008).

Quanto à *temporalidade*, de acordo com Kossoy (2007), a fotografia tem, simultaneamente, dois tempos: o tempo da criação e o da representação.

Uma única fotografia é dois tempos: o tempo da criação, o da primeira realidade, instante único da tomada do registro no passado, num determinado lugar e época, quando ocorre a gênese da fotografia; e o tempo da representação, o da segunda realidade, onde o elo imagético, codificado formal e culturalmente, persiste em sua trajetória na longa duração. O efêmero e o perpétuo, portanto (KOSSOY, 2007, p. 133).

Entler (2007), com base em Barthes (1984), apresenta dois tempos da fotografia: o tempo do referente e o tempo do olhar. Essa concepção coincide, respectivamente, com o tempo da criação da fotografia e o tempo da representação apresentados por Kossoy (2007).

O tempo (cronológico, época, data, momento em que se deu o registro) e o espaço (geográfico, local onde se deu o registro) constituem as coordenadas de situação, conforme Kossoy (2001). São também elementos importantes para a contextualização da imagem fotográfica.

O **espaço**, diferentemente do tempo, “é efetivamente visível: a imagem fotográfica é ela própria um espaço, uma superfície que oferece a representação de um outro espaço, aquele que faz parte do que chamamos de realidade”, explica Entler (2007, p. 29-30). Contudo, a representação de espaços na fotografia passa por seleções (fração de espaço), reduções de escala (do real para o suporte) e de planos (do tridimensional para o bidimensional).

Quanto ao plano do registro fotográfico, há uma transposição da realidade (componentes tridimensionais) para o plano bidimensional, como afirma Kossoy (2001, p. 36). Isso é corroborado por Entler (2007, p. 32) ao afirmar que, com o ato fotográfico, “o mundo tridimensional e dinâmico será traduzido para a imagem bidimensional e estática”.

Conforme Kossoy (2001), o ato do registro dá origem à materialização da fotografia. Essa materialização se dá por meio de diferentes **suportes** de armazenamento das imagens. Portanto, neste quesito, a materialização é múltipla.

O conceito de suporte é apresentado por Oliveira (2014) como matéria de produção e apresentação da imagem fotográfica:

o suporte pode ser entendido como a matéria com a qual é produzida e apresentada a imagem fotográfica: o filme ou o sensor digital inserido nas câmeras para a produção das imagens, e, depois, dentre várias possibilidades de apresentação, os suportes mais recorrentes são o papel ou a tela de computador (OLIVEIRA, 2014, p. 135).

Ao tratar das características da fotografia, é importante ainda considerar os aspectos constituintes e os elementos componentes da fotografia.

Os *elementos constitutivos* da fotografia são apresentados por teóricos como Feijó ([19–]), Kossoy (2001), Canabarro (2005), Borges (2005), Santos e Azevedo (2021) e Guerra e Marques Junior (2022).

Kossoy (2001) apresenta três elementos, a saber: o assunto, o fotógrafo e a tecnologia. Borges (2005) e Canabarro (2005) também apresentam os mesmos elementos: o retratado, o fotógrafo e a tecnologia.

Feijó ([19–]), por sua vez, aponta a existência dos aspectos objetivos (domínio de técnica e dos equipamentos), dos aspectos subjetivos (vivência, sensibilidade, criatividade do fotógrafo) e a inter-relação entre os dois aspectos, que dá origem à *linguagem* fotográfica.

Quadro 1. Elementos constitutivos da fotografia

<i>Autor (es)</i>	<i>Elementos constitutivos da fotografia</i>		
Kossoy (2001)	Assunto	Fotógrafo	Tecnologia
Canabarro (2005)	Retratado	Fotógrafo	Tecnologia
Borges (2005)	Retratado	Fotógrafo	Tecnologia
Feijó ([19–])	-	Aspectos subjetivos	Aspectos objetivos

Fonte: Os autores (2023).

Segundo Feijó ([19–]), a base da linguagem fotográfica é o processo técnico, que determina os elementos da linguagem, a saber: planos, foco, movimento, forma, ângulo, cor, textura, iluminação, aberrações, perspectivas, equilíbrio e composição. Esses elementos podem ser manipulados pelo estudo e pela pesquisa ou pela própria intuição do fotógrafo.

Para Ribeiro (2022), a linguagem fotográfica é composta por um conjunto de estratégias visuais que embasam hierarquicamente as escolhas do fotógrafo.

Smit (1996) apresenta o termo “expressão fotográfica” como aspecto relevante a ser representado pela fotografia, ou seja, a consideração dos atributos técnicos. Manini (2002) faz a mesma defesa, modificando apenas o termo “expressão fotográfica” por “dimensão expressiva”.

A historiadora Borges (2005, p. 16) admite que, muitas vezes, no emprego da fotografia como fonte histórica ou recurso pedagógico, “os cuidados necessários para a compreensão das particularidades da linguagem fotográfica são, frequentemente, desconsiderados”. Se incompreendida a linguagem fotográfica, a fotografia é concebida como um dado natural, como uma duplicação do real, da mesma forma que se pensava no século XIX.

Quanto aos **elementos componentes** da fotografia, teóricos apontam para três aspectos, a saber: suporte (objeto físico), conteúdo informacional (imagem) e expressão fotográfica, conforme disposto no Quadro 2.

Quadro 2. Elementos componentes da fotografia

<i>Autor</i>	<i>Elementos componentes da fotografia</i>		
Kossoy (2001)	Objeto físico (Matéria)	Objeto-imagem (Expressão)	-
Lacerda (1993)	Objeto (Forma)	Imagem (Conteúdo)	Expressão (como a imagem é mostrada)
Smit (1996)	Suporte	Conteúdo informacional	Expressão fotográfica
Manini (2011)	Suporte	Conteúdo informacional	Dimensão expressiva

Fonte: Os autores (2023).

Como mencionado anteriormente, o contexto da fotografia sempre deve ser considerado, mas não se restringindo apenas ao de sua produção, mas também o contexto em que a fotografia é vista e/ou “consumida”.

Guerra e Marques Junior (2022) também destacam a importância de se considerar o “para quem” a informação fotográfica se destina, ou seja, o usuário da informação. Defendem ainda que as determinações contextuais, as práticas embutidas no próprio ato fotográfico e na publicação configuram aspectos que podem influir no valor informacional da fotografia.

Dubois (1998) indica a recepção da fotografia como um dos processos do ato fotográfico (produção, recepção e/ou contemplação), questão que também é discutida por Santos e Azevedo (2021). Esse processo faz parte do ciclo informacional e, além de relacionar-se com os demais processos do ciclo, em grande medida depende de alguns, como da organização, da representação e do compartilhamento da informação.

Ribeiro (2022) considera como fundamental a posição do espectador em relação à fotografia jornalística. Segundo o autor, “o espectador tem a possibilidade de construir, adaptar e acoplar novas dimensões semânticas à fotografia” (RIBEIRO, 2022, p. 4).

O olhar do espectador e/ou usuário da informação é relevante, porque a fotografia pode apresentar múltiplos significados, considerando tanto a bagagem sociocultural desses sujeitos quanto o caráter polissêmico da fotografia.

As principais características da fotografia estão dispostas no Apêndice A.

3.3 Atributos intrínsecos e extrínsecos da fotografia

Para discorrer sobre os atributos da fotografia é necessário entender como os diferentes equipamentos culturais (arquivos, bibliotecas e museus) reconhecem a fotografia. No arquivo, “a documentação fotográfica é considerada comprobatória” e “deve ser organizada a partir da análise documentária, descrição e arranjo”; na biblioteca, a fotografia é considerada um documento especial e, em geral, o documento fotográfica fica “desvinculado do restante do acervo bibliográfico”, porém, recebendo tratamento conforme as regras e normas biblioteconômicas; já no museu, o documento fotográfico “fica disposto em exposição e guardado para pesquisa” (SANTOS; AZEVEDO, 2021, p. 151).

A fotografia constitui um documento complexo, conforme as suas características expostas. Para ser representada, com objetivo de recuperação da informação imagética, é preciso considerar os seus atributos intrínsecos e extrínsecos, incluindo os aspectos técnicos.

Bocato e Fujita (2006) defendem que os atributos constituem propriedades que não se limitam apenas aos

aspectos visuais, mas incluem aspectos cognitivos, afetivos ou interpretativos relacionados à imagem.

Segundo Albuquerque (2015), os atributos intrínsecos são deduzidos do próprio documento, pois estão relacionados com as suas características físicas. Os extrínsecos são atributos adquiridos de outras fontes.

Bräscher e Café (2008) consideram necessários os aspectos descritivos e temáticos para a representação da informação. Para as autoras, os dois tipos de descrição – dos atributos físicos e temáticos – são complementares e se aplicam tanto a documentos textuais verbais quanto a documentos iconográficos.

Com base em Bentes Pinto (2008), Lacerda (1993), Lima (2014), Manini (2002) e Shatford Layne (1994), apresentam-se os principais atributos do documento do tipo iconográfico: biográficos, físicos, histórico-artísticos, temáticos, visuais e relacionais, conforme dispostos no Quadro 3.

Quadro 3. Atributos da fotografia

Atributos	Categorias
Biográficos	Autor (fotógrafo); Autor (entidade); Título; Data; Local.
Físicos	Suporte; Tamanho; Resolução; Formato; Objetos/elementos (adicionais).
Histórico-artísticos	Tipo (ex: documental, artística, publicitária etc); Acervo.
Temáticos	Conteúdo temático/assunto.
Visuais	Cor; Expressão fotográfica (aspectos técnicos).
Relacionais	Código/Nº de registro; Documentos anexos; Referências Bibliográficas.

Fonte: Os autores (2023).

Esse conjunto de atributos é considerado para a representação descritiva e temática das fotografias.

4 Considerações finais

A imensa produção e contínua utilização da fotografia enquanto documento que transmite informações e contribui para a memória aponta o desafio de organização e representação desse tipo iconográfico, para fins de recuperação. A fotografia constitui documentação rica e relevante, mas também complexa e desafiante.

Após análise de literatura específica sobre a fotografia, foi apresentado um conjunto de características fundamentais e os principais atributos desse tipo documental. O Apêndice A mostra resumidamente o conjunto de características.

O mapeamento dos atributos da fotografia e a consideração dos aspectos técnicos são essenciais para trabalhar com a representação de imagens fotográficas.

Considera-se que os atributos intrínsecos e extrínsecos da fotografia são determinantes para a sua representação descritiva e temática. É relevante a consideração desses dois elementos de modo a se realizar uma representação mais abrangente e rica.

Além dos aspectos descritivos e temáticos, é importante a consideração da linguagem fotográfica, expressa a partir da técnica fotográfica. Entender a linguagem da fotografia contribui para a análise e representação do seu conteúdo informacional.

Este estudo contribui para a pesquisa sobre a fotografia documental na Ciência da Informação, trazendo insumos para apoiar uma reflexão mais aprofundada sobre o tema. Estudos futuros em acervos fotográficos são necessários para empiricamente testar a representação descritiva e temática simultânea, bem como estudos sobre a utilização da representação dos aspectos técnicos da fotografia (conforme já apontado por Smit).

Notas

[1] Com as tecnologias atuais em câmeras fotográficas e em *smartphones*, já existem fotografias com movimento.

Referências

- ALBUQUERQUE, Ana Cristina de. Tratamento temático da informação e a documentação museológica: aspectos e reflexões referentes à classificação. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB)*, 16., 2015, João Pessoa. **Anais [...]**. João Pessoa: ANCIB, 2015. p. 1-22.
- AVANCINI, Atílio. A imagem fotográfica do cotidiano: significação e informação no jornalismo. **Brazilian Journalism Research**, Marília, v. 7, p. 50-68, 2011. Disponível em: <https://bjr.sbpjor.org.br/bjr/article/view/285>. Acesso em: 19 maio 2023.
- BARDIN, Laurence. **Análise de conteúdo**. Lisboa: Edições 70, 1977.
- BARTHES, Roland. **A câmara clara: notas sobre a fotografia**. Rio de Janeiro: Nova Fronteira, 1984.
- BENTES PINTO, Virginia. Indexação morfossemântica de imagens no contexto da saúde visando à recuperação de informações. **História, Ciências, Saúde-Manguinhos**, Rio de Janeiro, v. 15, n. 2, p. 313-330, jun. 2008. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-59702008000200005&lng=en&nrm=iso. Acesso em: 19 maio 2023.
- BOCCATO, Vera Regina Casari; FUJITA, Mariângela Spotti Lopes. Discutindo a análise documental de fotografias: uma síntese bibliográfica. **Cadernos de Biblioteconomia, Arquivística e Documentação**, Lisboa, v. 2, n.1, p. 84-100, 2006.
- BORGES, Maria Eliza Linhares. **História & Fotografia**. 2. ed. Belo Horizonte: Autêntica, 2005.
- BRÄSCHER, Marisa; CAFÉ, Lígia. Organização da Informação ou Organização do Conhecimento? *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB)*, 9., 2008, São Paulo, **Anais [...]**. São Paulo: ANCIB, 2008.
- CANABARRO, Ivo dos Santos. Fotografia, história e cultura fotográfica: aproximações. **Estudos Ibero-Americanos**, Porto Alegre, v. 31, p. 23-39, 2005.
- CASTRO, Rodrigo Galvão. Novos territórios da fotografia: mobilidade, impermanência e ubiquidade da imagem. *In: V Grão Fino: Semana de Fotografia, 2022, Campina Grande-PB*. **Anais [...]**: Campina Grande-PB: Ano IV, Número 1, 2022, p. 1-10.
- COSTA, Luzia Sigoli Fernandes. **Uma contribuição da teoria literária para a análise de conteúdo de imagens publicitárias do fim do século XIX e primeira metade do século XX, contemplando aspectos da natureza brasileira**. 2008. 261f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2008.
- DI PIETRO, Laila Figueiredo. Ocorrências típicas de fotografias relacionadas à violação dos direitos humanos das ditaduras militares Latino-Americanas. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 25, p. 1-22, 2020. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2020.e67924>. Acesso em: 19 maio 2023.
- DICIONÁRIO ETIMOLÓGICO. **Fotografia**. 2008-2018. Disponível em: <https://www.dicionarioetimologico.com.br/fotografia/>. Acesso em: 19 maio 2023.
- DUBOIS, Philippe. **O ato fotográfico e outros ensaios**. Campinas: Papirus, 1998.
- ENTLER, Ronaldo. A fotografia e as representações do tempo. **Galáxia**, São Paulo, v. 14, p. 1, 2007.
- FAGÁ, Isadora Trombeta; COSTA, Luzia Sigoli Fernandes. Análise de conteúdo de imagens fotográficas: fatos históricos da UFSCAR ao longo dos seus 40 anos. *In: SOUZA, Lucília Maria Abrahão e; FUJITA, Mariângela Spotti Lopes; GRACIOSO, Luciana de Souza. A imagem em Ciência da Informação: reflexões teóricas e experiências práticas*. Marília: Cultura Acadêmica, 2014. p. 179-206.
- FEIJÓ, Cláudio Augusto Margarido de Mattos. **Linguagem fotográfica**. Londrina: UEL, [19-]. Disponível em: <http://www.uel.br/pos/fotografia/wp-content/uploads/downloads-2013-05/linguagem-fotografica.pdf>. Acesso em: 19 maio 2023.
- FERREIRA, Susana Patrícia da Silva. **A fotografia como lugar de memória**. 2021. 69f. Dissertação (Mestrado em Patrimônio Cultural) - Instituto de Ciências Sociais, Universidade de Minho. Braga, Portugal, 2021.
- FILIPPI, Patrícia de; LIMA, Solange Ferraz de; CARVALHO, Vânia Carneiro de. **Como tratar coleções de fotografias**. São Paulo: Arquivo do Estado: Imprensa Oficial do Estado, 2002.
- GUERRA, Claudia Bucceroni; MARQUES JUNIOR, Jairo André. É possível uma distinção entre os conceitos de ilustrar

e informar em fotografias?. **Logeion: Filosofia da Informação**, Rio de Janeiro, v. 9, p. 416-426, 2022.

INSTITUTO DO PATRIMÔNIO HISTÓRICO E ARTÍSTICO NACIONAL (IPHAN). **Dicionário IPHAN de Patrimônio Cultural**: Cadernos de pesquisa e documentação do IPHAN. 3. ed. Rio de Janeiro: IPHAN, COPEDOC, 2008.

KOSSOY, Boris. **Fotografia & História**. 2. ed. São Paulo: Ateliê Editorial, 2001.

KOSSOY, Boris. **Os tempos da fotografia**: o efêmero e o perpétuo. Cotia: Ateliê Editorial, 2007.

LACERDA, Aline Lopes. Os sentidos da imagem: fotografias em arquivos pessoais. **Acervo**, Rio de Janeiro, v. 2, n. 1/2, p. 41-54, jan./dez. 1993.

LE GOFF, Jacques. **História e memória**. Campinas: UNICAMP, 1996.

LIMA, Cláudia Albuquerque de; SILVA, Nerivanha Maria Bezerra da. Imagens Equivalentes. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO (INTERCOM), 25., 2002, Salvador. **Anais [...]**. Rio de Janeiro: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2002. p.1-17.

LIMA, Vânia. **Documentação iconográfica**. São Paulo: USP, 2014. Material instrucional. Disponível em: <https://edisciplinas.usp.br/mod/resource/view.php?id=141246>. Acesso em: 19 maio 2023.

MANINI, Miriam Paula. **Análise documentária de fotografias**: um referencial de leitura de imagens fotográficas para fins documentários. 2002. 231f. Tese (Doutorado em Ciências da Comunicação) - Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2002.

MANINI, Miriam Paula. Imagem, memória e informação: um tripé para o documento fotográfico. **Domínios da Imagem**, Londrina, v. 4, n. 8, p. 77-88, maio 2011.

MAUAD, Ana Maria. Através da imagem: fotografia e história interfaces. **Revista Tempo**, Niterói, v. 1, n. 2, p. 73-98, 1996. Disponível em: http://www.historia.uff.br/tempo/artigos_dossie/artg2-4.pdf. Acesso em: 19 maio 2023.

MERLO, Márcia; BRANDÃO, Romário. Fotografia: traços da História, da Memória e da Moda. **Modapalavra E-periódico**, [S.l.], v. 12, n. 1, 2013.

OLIVEIRA, Tatiana Pontes. Analógico e digital: a fotografia entre meios. **Iara**: Revista de Moda, Cultura e Arte, [S.l.], v. 7, 2014.

PAGANOTTI, Caio. **Evolução e revolução do suporte fotográfico**. 2016. 103f. Dissertação (Mestrado em Meios e Processos Audiovisuais) – Universidade de São Paulo, São Paulo, 2016.

RIBEIRO, Matheus Tagé Veríssimo. Gamificação Fotográfica: memes e a lógica das imagens-fluxo. **Triade: Comunicação, Cultura e Mídia**, [S.l.], v. 10, p. e022015, 2022.

RODRIGUES, Ricardo Crisafulli. Análise e tematização da imagem fotográfica. **Ciência da Informação**, Brasília, v. 36, n. 3, p. 67-76, set./dez. 2007.

SANTAELLA, Lúcia. Três paradigmas da imagem: gradações e misturas. In: OLIVEIRA, Ana Cláudia Meialves de; BRITO, Yvana Carla Fechine de (org.). **Imagens técnicas**. São Paulo: Hacker Editores, 1998. p. 167-178.

SANTOS, Alessandra de Souza; AZEVEDO, Dúnya Pinto. A fotografia-documento e a importância de saberes especializados dos profissionais da informação para a memória. **Informação em Pauta**, Fortaleza, v. 6, n. Especial, p.141-158, 2021. DOI: 10.36517/2525-3468.ip.v6iespecial.2021.62659.141-158. Acesso em: 19 maio 2023.

SHATFORD LAYNE, Sara. Analyzing the subject of a picture: a theoretical approach. **Cataloging & Classification Quarterly**, New York, v. 6, n. 3, p. 39-62, 1986.

SHATFORD LAYNE, Sara. Some issues in the indexing of images. **Journal of the American Society for Information Science**, [S.l.], v. 45, n. 8, p. 583-588, 1994.

SMIT, Johanna. A representação da imagem. **Informare**: Caderno do Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, v. 2, n. 2, p. 28-36, jul./dez. 1996.

SOUZA E SILVA, Wagner. Dimensões Numéricas da Fotografia. In: 45º Congresso Brasileiro de Ciências da Comunicação - INTERCOM, 2022, João Pessoa. **Anais [...]**. João Pessoa, 2022.

TACCA, Fernando de. Imagem fotográfica: aparelho, representação e significação. **Psicologia & Sociedade**, [S.l.], v. 17, n. 3, p. 9-17, set./dez. 2005.

Apêndice A – Características e concepções da fotografia

<i>FOTOGRAFIA</i>		
<i>Aspecto Avaliado</i>	<i>Característica</i>	<i>Embasamento Teórico</i>
Outras nomenclaturas	“Iconografia saída da câmera escura”.	BORGES, 2005.
	“Iconografia fotográfica”.	KOSSOY, 2001.
	“Instrumento de preservação”, “Linguagem”; “Valor de documento”; “Informação”, “Informatização”; “Suporte de memória”; “Imagem” e “Representação”.	IPHAN, 2008.
Quanto à produção	Imagem técnica.	RODRIGUES, 2007; SANTAELLA, 1998.
Quanto à gênese	Físico-química.	BORGES, 2005; DUBOIS, 1998; PAGANOTTI, 2016.
Quanto à natureza	Imagem híbrida (técnica e estética).	BORGES, 2005; KOSSOY, 2001.
	Polissêmica.	BORGES, 2005; LIMA; SILVA, 2002; RODRIGUES, 2007.
Quanto ao movimento	Imagem fixa.	COSTA, 2008.
Quanto à temporalidade	Dois tempos: o da criação e o da representação.	BARTHES, 1984; ENTLER, 2007; KOSSOY, 2007.
Quanto ao plano do registro	Bidimensional; Escala reduzida; Recorte de espaço.	CANABARRO, 2005; ENTLER, 2007; KOSSOY, 2001.
Quanto à constituição (elementos constitutivos)	O assunto (o retratado), o fotógrafo e a tecnologia.	CANABARRO, 2005; KOSSOY, 2001.
Quanto às dimensões do registro fotográfico	Matéria (objeto físico) e Expressão (objeto-imagem).	KOSSOY, 2001.
	Objeto (forma), imagem (conteúdo) e expressão.	LACERDA, 1993.
Quanto ao suporte (materialização)	Múltipla: placas de cobre, chapas de metal, chapas de vidro, filmes negativos e positivos (slides), papel (diversos tipos), eletrônico (armazenado em artifícios tecnológicos).	FILIPPI; LIMA; CARVALHO, 2002; OLIVEIRA, 2014; PAGANOTTI, 2016.
Quanto à linguagem	Texto não verbal. Linguagem visual.	BENTES PINTO, 2008.
	Linguagem específica (linguagem fotográfica).	BORGES, 2005; FEIJÓ, [19-]; LACERDA, 1993; LIMA; SILVA, 2002; MANINI, 2002; RODRIGUES, 2007; SMIT, 1996.
Quanto aos atributos	Biográficos, histórico-artísticos, físicos, temáticos, visuais e relacionais.	BENTES PINTO, 2008; LACERDA, 1993; LIMA, 2014; MANINI, 2002; SHATFORD LAYNE, 1994.
Quanto às concepções	Artefato cultural e social.	CANABARRO, 2005; LIMA; SILVA, 2002; RODRIGUES, 2007.
	Representação.	BORGES, 2005; LACERDA, 1993; MANINI, 2011; SANTAELLA, 1998.
	Imagem: “Espelho do real” “Transformação do real” “Traço do real”.	DUBOIS, 1998; MANINI, 2011.
	Memória.	BORGES, 2005; CANABARRO, 2005; KOSSOY, 2001; LACERDA, 1993; LE GOFF, 1996; MANINI, 2011.
	Informação.	AVANCINI, 2011; BOCCATO; FUJITA, 2006; FAGÁ; COSTA, 2014; LACERDA, 1993; MANINI, 2002; 2011.

Fonte: Os autores (2023).

Base de registros de tesauros: um mapeamento de software gestor de tesauros e tecnologias de acesso

Thesaurus Database: a mapping of thesaurus software management and access technologies

Miguel Ivan Magarzo Arias (1), José Carlos Francisco dos Santos (2)

(1) Universidade Estadual de Londrina - UEL, Londrina PR, miguemag@gmail.com. (2) Faculdades Londrina, Londrina - PR, jose.cf.santos@faculdadeslondrina.com.br

Resumo

Introdução: O tesouro é um tipo de vocabulário controlado e usado como instrumento de mediação entre sistema, indexadores e usuários. Está voltado para atuação em áreas especializadas e correlatas, promovendo a construção epistemológica do conhecimento. Torna-se relevante a importância de bases de registros de Sistemas de Organização do Conhecimento (SOC). O uso dessas bases não é somente para acesso humano, mas também para a troca de dados entre os sistemas de informação. **Objetivo:** Mapear os registros dos tesauros disponibilizados na *Bartoc*, com a abrangência nas tecnologias aplicadas e nos softwares. **Metodologia:** 1) Teórica, e 2) Aplicada. A etapa teórica caracteriza-se por uma pesquisa bibliográfica e documental de cunho exploratório e descritivo. Já na etapa aplicada, qualificou-se pela realização de três ações: 1) Planejamento do instrumento de coleta, 2) Coleta manual, e 3) Coleta automatizada. **Resultados:** Dos 788 registros de tesauros na *Bartoc*, observa-se que, na grande maioria dos tesauros, não foi possível encontrar o ano de criação e a inacessibilidade é evidente. Os dados não disponíveis na base da *Bartoc* são raramente encontrados nos acessos aos tesauros. **Conclusão:** Considera-se de grande relevância o mapeamento dos registros da *Bartoc*, sendo possível observar que existe um grande número de tesauros que apresentam um formato Web de acesso, porém não fica claro se fazem uso interno de algum software gestor.

Palavras-chave: Tesouro. Bartoc. Software gestor de tesauros.

Abstract

Introduction: The thesaurus is a type of controlled vocabulary used as an instrument of mediation between the system, indexers and users, it is aimed at acting in specialized and related areas, promoting the epistemological construction of knowledge. The importance of Knowledge Organization Systems (KOS) databases is relevant, since the use of these databases is not only for human access, but also for the exchange of data between information systems. **Objective:** Mapping the records of thesauri available at Bartoc covering the applied technologies and software. **Methodology:** 1) Theoretical and 2) Applied. The theoretical stage is characterized by an exploratory and descriptive bibliographic and documentary research. Already in the applied stage three actions are carried out: 1) Planning the collection instrument, 2) Manual collection, and 3) Automated collection. **Results:** Of the 788 thesauri records in the Bartoc, it is noted that the vast majority of thesauri could not find the creation year and the inaccessibility is evident. Data not available at Bartoc are rarely found by accessing thesauri. **Conclusion:** It is considered very relevant the mapping of Bartoc records, being possible to observe that there are a large number of thesauri that present a Web format access, but it is not clear if they make internal use of some software manager.

Keywords: Thesaurus. Bartoc. Thesaurus software management.

Resumen

Introducción: El tesouro es un tipo de vocabulario controlado usado como un instrumento de mediación entre el sistema, los indexadores y los usuarios, está dirigido para la actuación en áreas especializadas y relacionadas, promoviendo la construcción epistemológica del conocimiento. Resulta relevante la importancia de bases de registros de Sistemas de Organización del Conocimiento (SOC), ya que el uso de esas bases no es solamente para el acceso humano, sino también para el intercambio de datos entre los sistemas de información. **Objetivo:** Mapear los registros de los tesauros disponibles en la *Bartoc* cubriendo las tecnologías aplicadas y los softwares. Metodología: 1) Teórica y 2) Aplicada. La etapa teórica se caracteriza por una investigación bibliográfica y documental de tipo exploratorio y descriptivo. Ya en la etapa aplicada se realizan tres acciones: 1) Planificación del instrumento de recopilación, 2) Recopilación manual, y 3) Recopilación automatizada. **Resultados:** De los 788 registros de tesauros en la *Bartoc*, se observa que la gran mayoría de los tesauros no fue posible encontrar el año de creación y la inaccesibilidad es evidente. Los datos no disponibles en la *Bartoc* son raramente encontrados mediante los accesos a los tesauros. **Conclusión:** Se considera muy relevante el mapeo de los registros de la *Bartoc*, siendo posible observar que existe un gran número de tesauros que presentan un formato de acceso Web, pero no queda claro si hacen uso interno de algún software gestor.

Palabras clave: Tesouro. Bartoc. Software de gestión de tesauros.

1 Introdução

A informação indica um conteúdo documentário que se encontra disponível nos mais variados meios e suportes e, quando incorporada aos sistemas de informação, essa se acumula e se agrega em uma estrutura ou repositório. Para a recuperação da informação, é importante considerar que o sistema de informação que a disponibiliza, esteja amparado por um vocabulário capaz de operar os dois modos de comunicação, ou seja, a linguagem do sistema de informação e a linguagem de busca do usuário de áreas especializadas.

Torna-se, portanto, indispensável que a ciência, seja ela básica ou aplicada, mantenha sempre atualizado o seu discurso metalinguístico com vistas às necessidades que se renovam. Ao produzir novos conhecimentos, os especialistas, em uma determinada matéria, criam novos conceitos que necessitam de denominação e de identificação. As unidades resultantes do processo de denominação são os termos científicos e técnicos que, uma vez identificados e sistematizados em instrumentos terminológicos, veiculam os saberes e os fazeres consolidados por uma ciência ou tecnologia. O interesse pelo estudo da construção de linguagens para propósitos especiais tem crescido muito nas últimas décadas, motivado principalmente pelos avanços científicos e tecnológicos.

Nesse contexto, o tesouro é, ao lado de vocabulários técnico-científicos e outros recursos terminológicos, um importante instrumento de apoio às pesquisas científicas nas áreas de conhecimento. Em unidades de informação, o tesouro é considerado uma modalidade de linguagem documentária e tem a função de subsidiar, ao mesmo tempo, os processos de representação do conteúdo do documento e de recuperação temática da informação.

Fujita (1992, p. 24) evidencia que “a função de controle de vocabulário do tesouro, tanto na entrada (representação do conteúdo do documento), quanto na saída (representação da questão de busca), servirá igualmente ao bibliotecário-indexador e ao usuário do sistema de informação”. Desse modo, o tesouro pode ser visto como um instrumento de mediação capaz de permitir que bibliotecários indexadores, bibliotecários de serviço de referência e informação e usuários de um sistema de informação especializado compartilhem de um mesmo vocabulário.

Nessa perspectiva, no âmbito da Organização do Conhecimento, o tesouro como instrumento de mediação entre sistema, indexadores e usuários, está voltado para atuação em áreas especializadas e correlatas, promovendo a construção epistemológica do conhecimento. Torna-se relevante a importância de bases de registros de Sistemas de Organização do Conhecimento (SOC), o qual tem origem nos Registros de Terminologias (RT) nos anos de 2000. Destaca-se o uso dessas bases não somente para acesso humano, mas

também para a troca de dados entre os sistemas de informação.

Diante desse cenário, o presente artigo tem como objetivo geral, mapear os registros dos tesouros disponibilizados na *Bartoc* com a abrangência nas tecnologias aplicadas e nos softwares. Já os objetivos específicos são:

- Identificar os registros dos tesouros cadastrados na base de dados *Basel Register of Thesauri, Ontologies & Classifications (Bartoc)*;
- Mapear os softwares gestores de tesouros e tecnologias de acesso dos registros da *Bartoc*;
- Analisar os dados coletados relacionados ao software gestor, formato de apresentação e área temática dos tesouros.

Para o cumprimento dos objetivos específicos traçou-se os procedimentos metodológicos sedimentado em duas etapas: 1) Teórica, e 2) Aplicada. A etapa teórica caracteriza-se por uma pesquisa bibliográfica e documental de cunho exploratório e descritivo. Já na etapa aplicada, qualificou-se pela realização de três ações: 1) Planejamento do instrumento de coleta, 2) Coleta manual, e 3) Coleta automatizada.

O estudo está sistematizado em duas seções sendo a primeira para o desenvolvimento do estudo e, a segunda, para a conclusão. A primeira está organizada com as subseções:

- Tesouro principais abordagens;
- *Basel Register of Thesauri, Ontologies & Classifications (Bartoc)*;
- Procedimentos metodológicos; e
- Análises e Resultados.

2 Mapeamento de tecnologia aplicada aos tesouros

A construção de um tesouro é uma atividade intelectual exaustiva de grande complexidade. Nesse processo, são tomadas as decisões sobre a disseminação do tesouro. Recorre-se à base de registros de Sistemas de Organização do Conhecimento (SOC) para mapear as tecnologias de software e formatos de acesso aos tesouros nela registrados. Serão abordados, nas duas subseções iniciais, os principais aspectos teóricos dos tesouros e da *Bartoc*. Na terceira subseção, serão apresentados os procedimentos metodológicos e, na quarta subseção, as análises e resultados do mapeamento.

2.1 Tesouro principais abordagens

O tesouro é um instrumento de controle terminológico usado para transferir a linguagem natural dos

documentos para um sistema linguístico, estruturado por linguagem documentária de termos relacionados que garante um domínio específico de conhecimento (UNESCO, 1973). Por esse motivo, o tesouro é uma variedade de vocabulário controlado que é aplicada normalmente para padronizar cabeçalhos de assuntos ou descritores. Em qualquer circunstância que seja requerida uma padronização terminológica, adota-se o tesouro (LANCASTER, 1987). Também é definido como uma “linguagem documentária controlada e dinâmica que contém termos relacionados genérica e semanticamente, cobrindo um domínio específico do conhecimento.” (GOMES, 1990, p. 16; GOMES, 2004, n.p).

Para Novellino (1996, p. 40), o tesouro foi pensado “[...] como instrumento facilitador da comunicação dentro do sistema, padronizando as linguagens de indexação e de recuperação, a partir da terminologia da área representada”. Barité (2011) argumenta que o Tesouro é um tipo de Sistema de Organização do Conhecimento que se apresenta como estrutura e representa os termos e suas relações. Dentre as vantagens da aplicação do tesouro, está a “[...] a possibilidade de expressar o conjunto completo de relações associativas entre conceitos e não apenas relações genéricas.” (NOVELLINO, 1996, p. 41).

Os tesouros viabilizam diversos usos. Os usos vão desde a indexação, até servir de alicerce para a recuperação da informação. Qualificam-se como Sistema de Classificação Bibliográfica e se revelam diante da quantidade de registros informacionais especializados. Os tesouros são ferramentas utilizadas para a atividade de indexar, porque sua linguagem caracteriza-se por termos que o compõem mediante suas relações que também podem ser hierárquicas e cruzadas viabilizando novas relações entre os termos (MOREIRA; MOURA, 2006).

Segundo Lancaster (2004), o tesouro apresentou duas influências basilares: da América do Norte e da Europa (Reino Unido). A América do Norte inspirou a indexação alfabética de assuntos em que o arranjo dos termos é alfabético. No Reino Unido, mediante a Classificação Bibliográfica, o arranjo alfabético é hierárquico, no formato de índice.

Nesse sentido, atua ora como instrumento de controle terminológico, ora como instrumento de recuperação da informação. O tesouro é instrumento de recuperação da informação pela necessidade de linguagens hierárquicas de conceitos, indicando relações genéricas, bem como outras relações. Como um instrumento de controle terminológico usado em sistemas informacionais, pois representa a linguagem natural de registros informacionais, tanto de indexadores, quanto de usuários em um sistema de linguagem documentária, utilizado na indexação e recuperação.

2.2 Basel Register of Thesauri, Ontologies & Classifications (Bartoc)

A quantidade de SOC disponível na *Web* e os potenciais de aplicações a partir de padrões e protocolos para a representação digital evidencia a necessidade de Registros de Terminologia (RT). O RT identifica, em estrutura padronizada e de forma sistematicamente, os dados de SOC. Além da identificação, apresenta-se informações de acesso, *Web services*, *API* de conjunto de vocabulários para ser utilizados por sistemas e serviços de informação. O RT originou-se em 2000, e alguns, raramente, podem conter os conceitos, termos, relações semânticas. Os registros podem servir como instrumento terminológico para acesso automático por meio de aplicativos. Vale destacar ainda que o RT almeja não somente o acesso humano quanto a interação por máquinas (ZENG; CHAN, 2010).

A principal finalidade da *Bartoc* é servir de um local para disponibilizar e dar visibilidade aos Sistemas de Organização do Conhecimento (SOC). A *Bartoc* foi idealizada por *Andreas Ledl* e nominada de *Basel Register of Thesauri, Ontologies & Classifications (Bartoc)* na *Basel University Library, Switzerland*. A base de registros está estruturada para incluir registros de diversos tipos de SOC, e ainda com abrangência de qualquer área do conhecimento (ABOUT..., [entre 2013 e 2023]). Na Figura 1, apresenta-se um recorte da codificação em *JSON* extraído da *Bartoc*.

Figura 1. Recorte da codificação em *JSON*

```
[
  {
    "@context": "https://gbv.github.io/jskos/context.json",
    "ACCESS": [
      {
        "uri": "http://bartoc.org/en/Access/Free"
      }
    ],
    "ADDRESS": {
      "code": "20540",
      "country": "United States",
      "ext": "101 Independence Ave SE",
      "locality": "Washington",
      "region": "DC",
      "street": "Library of Congress"
    },
    "API": [],
    "FORMAT": [
      {
        "uri": "http://bartoc.org/en/Format/Online"
      },
      {
        "uri": "http://bartoc.org/en/Format/Printed"
      }
    ],
    "concepts": [],
    "contributor": [
      {
        "uri": "https://coli-conc.gbv.de/login/users/2062a0cd-5f0",
        "prefLabel": {
          "en": "Sabrina Gaab"
        }
      }
    ],
    "created": "2013-09-03T13:16:00Z",
    "creator": [

```

Fonte: Os autores (2023).

Na tag *<FORMAT>*, apresentado na Figura 1, é um exemplo de dado descritivo desse registro o qual

remonta ao acesso ao tesouro com os valores `<http://bartoc.org/en/Format/Online>` e `<http://bartoc.org/en/Format/Printed>` para `<uri>`.

Entre as finalidades da *Bartoc*, a primeira é descrever os SOC, conforme o autor supracitado, por conseguinte, a outra finalidade é a de dar o acesso. Como forma de obter a cobertura temática dos registros de vocabulários controlados da sua base, utiliza-se de outros vocabulários controlados como o *EuroVoc* (Tesouro multilíngue da União Européia), a Classificação Decimal de Dewey (CDD). No início do projeto da *Bartoc* (2013), a escolha pela CDD se deu pela grande abrangência de utilização de forma internacional nas bibliotecas, como também pela facilidade de compreensão pelos diversos tipos de usuários. A clareza no uso foi por meio do uso dos três primeiros níveis para descrição dos vocabulários controlados (LEDL; GNOLI, 2017).

2.3 Procedimentos metodológicos

O estudo está sedimentado em duas etapas: 1) Teórica, e 2) Aplicada. A etapa teórica caracteriza-se por uma pesquisa bibliográfica e documental de cunho exploratório e descritivo. Já na etapa aplicada, qualificou-se pela realização de três ações: 1) Planejamento do instrumento de coleta, 2) Coleta manual, e 3) Coleta automatizada. No planejamento, iniciou os trabalhos de conhecimento da *Bartoc*, a partir da identificação dos registros dos tesouros, instanciada realizada no mês de Fevereiro de 2023, identificou-se um total de 788 registros. A estratégia de busca adotada para compor o corpus de análise na referida base foi determinada pela seleção dos instrumentos caracterizados como “*Thesaurus*” no campo “*KOS Type*”.

Foi desenvolvida uma planilha eletrônica com os campos de pesquisa: a) N°; b) *Name*; c) *Links*; d) *Links-API*; e) *Software*; f) *Acesso*; g) *OBS*; h) *Created*; i) idiomas (*ar, fa, bg, ca, hr, cs, da, nl, en, et, fi, fr, de, el, hu, it, lv, lt, mk, mt, pl, pt, ro, sr, sk, sl, es, sv, ko, ru, no, mi, eu, gl, hi, ka, km, la, lo, ms, my, nb, nn, sw, te, th, uk, vi, zh, ja, tr, id, zl, bo, dz, pi, ta, jv, sa, si, he, is, ga, af, sq, hy, az, be, bn, bs, ny, eo, gu, ht, ha, ig, kn, kk, mg, ml, mr, mn, ne, pa, so, st, su, tg, ta, ur, uz, cy, yi, yo, zu*). O idioma foi padronizado de acordo com a norma ISO 639-1, a inserção no cabeçalho da planilha de um novo idioma foi realizada à medida que procedeu-se à coleta de dados.

Na ação de coleta manual, o preenchimento da planilha foi realizado manualmente, utilizando-se de um procedimento para a limpeza nos registros para separar o *link* de acesso ao referido tesouro. Posteriormente à primeira etapa do desenvolvimento da planilha, iniciou-se o acesso à cada registro na *Bartoc* para coleta dos dados dos registros. Em seguida, acessou-se o *link* do tesouro, a partir do registro, para complementar a coleta de dados não existentes nos registros da *Bartoc*.

Na ação de coleta automatizada, procurou-se pelos formatos disponíveis para troca de dados na *Bartoc*, nesse caso o formato disponibilizado foi o *JavaScript Object Notation (JSON)*. Utilizando a linguagem *PHP*, desenvolveu-se um algoritmo para acessar o arquivo *TXT* com os *links* dos registros da *Bartoc* e processá-los em vetores para início da extração de informações complementares da coleta manual.

O primeiro passo foi acessar cada um dos *links* o qual retorna com os registros em formato *JSON*. O acesso foi por meio de requisição ao *Web service* da *Bartoc* usando o prefixo `<http://bartoc.org/api/data?uri=>` + a *uri* do registro. Com os dados em vetores, foi realizada a extração a partir das *tag* `<prefLabel>`, `<publisher>`, `<format>`, `<subject>`, `<startDate>`, `<languages>`. Como forma de representação desse processo apresentase na Figura 2 um recorte da codificação.

Na linha 69 da Figura 2, é realizado o processo de abertura do arquivo com as *uri* dos registros dos tesouros e transformados em um vetor nas linhas 71 a 73.

O processo de acesso ao *Web service* da *Bartoc* é realizado na linha 81 e 83 da Figura 2. Já na linha 84, realiza-se a transformação em um vetor para extração dos dados.

Figura 2. Recorte da codificação em PHP do algoritmo

```

69 $arquivo = fopen ('arq.txt', 'r');
70 $result = array();
71 while(!feof($arquivo)){
72     $result[] = explode("|",fgets($arquivo));
73 }
74 fclose($arquivo);
75
76 for ($i=0; $i< count ($result); $i++){
77     $html .= '<tr>';
78     $iplus= $i+1;
79     $html .= '<td>'. $iplus. '</td>';
80     //busca o arquivo json
81     $json_url = "https://bartoc.org/api/data?uri=";
82     //echo $json_url;
83     $json = file_get_contents(trim ($json_url));
84     $data = json_decode($json, TRUE);
85
86     //campo 1
87     if (isset ($data[0]['prefLabel']['en'])){
88         $html .= '<td>'. $data[0]['prefLabel']['en']. '<
89     }
90     else{
91         $html .= '<td>'. " ". '</td>';
92     }

```

Fonte: Os autores (2023).

Após processo de extração dos dados, exportou-se esses dados para uma planilha, a qual foi integrada como complementar a planilha de dados da coleta manual. Destaca-se que essa integração se deu por meio da criação de um identificador de cada registro.

2.4 Análises e Resultados

A partir de uma análise das áreas temáticas relacionadas aos tesouros registrados na *Bartoc*, seguindo a classificação temática CDD, foram identificados um total de 141 tesouros, conforme o detalhamento do Quadro 1.

Quadro 1. Total de áreas temáticas que desenvolveram um tesouro

Grande Área (CDD)	Total
Social sciences	49
Computer science, information & general works	43
History & geography	19
Arts & recreation	15
Science	13
Literature	1
Philosophy & psychology	1
Total	141

Fonte: Os autores (2023).

Pode-se observar que a área com maior número de registros de tesouros, no total 49, são as Ciências Sociais; seguido da Ciência da Computação com 43 registros; História e Geografia com 19 tesouros; Artes e recreação com 15 registros e Ciências com 13 registros de tesouros. Optou-se, para efeitos deste estudo, o primeiro nível da CDD, primeiro sistema utilizado pela *Bartoc* para o tratamento temático dos SOC, por esse motivo destaca-se a quantidade relativamente baixa de SOC classificado com essa notação.

No que se refere ao total de tesouros identificados que usam um software gestor, do total de 788 tesouros registrados na *Bartoc*, foram identificados um total de 102 tesouros, conforme detalhado no Quadro 2.

Quadro 2. Total de tesouros identificados que usam um software gestor

Software	Total Tesouros
TemaTres	71
PoolParty	19
Skosmos	5
MultiTes	4
ISGAT	1
OpenTheso	1
VocBench	1
Total	102

Fonte: Os autores (2023).

O software mais usado é o *TemaTres*, num total de 71 tesouros; seguido do *PoolParty* com 19 tesouros; o *Skosmos* usado em 5 tesouros e finalmente o *MultiTes* usado em 4 tesouros.

É importante mencionar que atualmente na *Bartoc* não existe um campo que especifique o nome do software usado e que o processo de identificação foi realizado de forma manual acessando a cada tesouro.

Um grande número de tesouros com acesso *online* apresenta uma interface *Web*, porém não especifica se faz uso de algum software específico.

A análise também permitiu realizar um mapeamento temporal da intensidade de uso, conforme mostrado no Quadro 3.

Quadro 3. Intensidade do uso de softwares gestores de tesouros

Ano	MultiTes	PoolParty	Skosmos	TemaTres	Total
2001				3	3
2003				1	1
2004				1	1
2006		1		1	2
2007	1	1		3	5
2008				1	1
2009				3	3
2010		2		5	7
2011				2	2
2012		1		1	2
2013		7	1	2	10
2014				2	2
2015			1	4	5
2017				2	2
Total	1	12	2	31	46

Fonte: Os autores (2023).

Foi possível identificar que a intensidade de uso de softwares gestores de tesouros foi intensificada a partir do ano de 2001, sendo o *TemaTres* o software mais usado.

Salienta-se que o tesouro mais antigo identificado que atualmente usa um software é o *Defense Technical Information Center Thesaurus*, que data do ano de 1966, muito provavelmente a sua versão eletrônica foi publicada vários anos mais tarde com a popularização do *TemaTres*.

Em relação ao ano de criação dos tesouros, destaca-se os dois tesouros datados de 1962 *Thesaurus of Old English* e *Thesaurus of ASTIA Descriptors*. O *Thesaurus of Old English* encontra-se disponível de forma *online*, decorrente de atualizações realizadas nas suas versões e disponível para navegação na *Web*. Já o *Thesaurus of ASTIA Descriptors* em contrapartida está disponível somente em PDF na sua versão original. O contraste entre o PDF e a disponibilização online, destaca-se pela navegabilidade e acessibilidade aos termos e relações semânticas deles com a possibilidade de interagir com os termos e suas relações.

Finalmente, no que concerne ao tipo de formato disponível dos tesouros, a maioria indica que tem formato *online*, conforme visualizado no Quadro 4, onde são visualizados os resultados acima de 100 registros.

Quadro 4. *Top 7 dos formatos disponíveis dos tesouros*

No.	Formato	Total
1	Online	533
2	Printed	220
3	PDF	209
4	SKOS	194
5	Spreadsheet (Excel, ODS...)	149
6	XML	129
7	CSV	110

Fonte: Os autores (2023).

Nesse sentido, é importante salientar que cada tesouro pode disponibilizar vários tipos de formato, sendo em razão disso que os valores apresentados no Quadro 4 resultam maiores ao número total dos 788 registros de tesouros atualmente existentes na *Bartoc*. Apesar de que a grande parte dos tesouros indicam disponibilizar o formato *online*, o que ao nosso entendimento pressupõe um acesso via plataforma *Web*, muitos deles disponibilizam somente uma versão PDF *online*.

3 Conclusão

Considera-se que os resultados desse estudo cumpriram com os objetivos, o qual tinha inicialmente como proposta, mapear os registros dos tesouros disponibilizados na *Bartoc* com a abrangência nas tecnologias aplicadas e nos softwares. Diante dos resultados observados, destaca-se a grande quantidade de tesouros disponíveis nos formatos digitais primeiramente o *online* e posteriormente o PDF. Isso demonstra a relevância de evidenciar o processo de construção de tesouros por meio da organização dos termos em software gestor. A funcionalidade de um software gestor é de minimizar os esforços humanos em atividades operacionais, onde a tecnologia gerencia de forma eficiente. Logo, o esforço humano pode ser destinado para as atividades intelectuais que têm maior complexidade cognitiva.

Outra questão de impacto é que após uma análise inicial dos 788 registros de tesouros na *Bartoc*, pode-se concluir que na maioria dos tesouros cadastrados na base não foi possível encontrar o ano de criação do tesouro nem no cadastro nem na documentação disponibilizada. Outro assunto que emerge nos registros foi a inacessibilidade a alguns tesouros por meio dos *links* cadastrados, seja porque o *link* está quebrado, o site está fora do ar ou tem acesso restrito. Também foi possível observar que existe um grande número de tesouros que apresentam um formato *Web* de acesso, porém não fica claro se fazem uso interno de algum software gestor.

A forte presença do *TemaTres* como o software mais usado na gestão dos tesouros, demonstra a sua consolidação como um importante gestor, talvez esse fato seja consequência tanto da sua divulgação como do sucesso na implementação dos projetos de tesouros como também indica que, no momento da escolha do software, a preferência é pelo software livre em contrapartida do uso de um software proprietário como o *PoolParty*.

Nesse sentido, reitera-se que a construção de tesouros, por ser uma atividade exaustiva, o uso de ferramentas tecnológicas de informação, auxilia no processo de controle dos termos e suas relações semânticas. Bem como a disponibilização em um modelo com maior navegabilidade e abstraindo-se das funcionalidades da *Web* semântica. Por outro lado, a disponibilização em diversos formatos de troca de dados, possibilita a interação entre sistemas de informação.

O estudo apresentado está em refinamento os dados de pesquisa, portanto, vale destacar que, em trabalhos futuros, serão desenvolvidas análises detalhadas em relação aos registros e às reais formas de acesso ao tesouro, principalmente em relação ao acesso por meio de um software gestor de tesouro. A análise da frequência relacionada aos idiomas dos tesouros é outra importante contribuição na perspectiva de fomentar a construção de SOC multilíngue. Por outro lado, existem possibilidades de explorar os tesouros que estão disponíveis na própria *Bartoc*. O mapeamento da abrangência temática dos tesouros é relevante para a compreensão da construção epistemológica das áreas do conhecimento e o fomento do desenvolvimento dessas áreas especializadas.

Referências

- ABOUT. [S. l.]: Basel University Library, Switzerland, [entre 2013 e 2023]. Disponível em: <https://bartoc.org/about>. Acesso em: 20 abr. 2023.
- BARITÉ, M. Sistemas de organização do conhecimento: uma tipologia atualizada. **Informação & Informação**, Londrina, v. 16, n. 2, p. 122-139, dez. 2011. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/9952>. Acesso em: 25 fev. 2023.
- FUJITA, M. S. L. **Linguagem documentária em Odontologia**: uma aplicação do sistema de indexação PRECIS. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 1992. 3v.
- GOLUB, K.; TUDHOPE, D.; ZENG, M. L.; ŽUMER, M. Terminology registries for knowledge organization systems: Functionality, use, and attributes. **Journal of the Association for Information Science and Technology**, [S. l.], v. 65, n. 9, p. 1901-1916, 2014.

GOMES, H. E. **Manual de elaboração de tesouros monolíngües**. Brasília: Programa Nacional de Bibliotecas das Instituições de Ensino Superior, 1990. 78 p.

GOMES, H. E. (Coord.). Elaboração de tesouro documentário. *In: Biblioteconomia, Informação & Tecnologia da Informação (BITI)*. Rio de Janeiro, [2004]. Disponível em: <http://www.conexaorio.com/bit/index.htm>. Acesso em: 20 fev. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964**: information and documentation - thesauri and interoperability with other vocabularies - part 1: thesauri for information retrieval. Genebra, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 639-1**: Codes for the representation of names of languages - Part 1: Alpha-2 code. 1nd ed. Genève: International Standard Organization, 2002.

LANCASTER, F. W. **Construção e uso de tesouros**: curso condensado. Tradução de César Almeida de Menezes Silva. Brasília: IBICT, 1987. 114 p.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. rev. atual. Brasília: Briquet de Lemos, 2004.

LEDL, Andreas; GNOLI, Claudio. Indexing KOSs in BARTOC by a disciplinary and a phenomenon-based classification: preliminary considerations. *In: Faceted classification today: theory, technology and end users: proceedings of the International UDC Seminar*, p. 14-15, September 2017. London: United Kingdom, 2017. p. 109-117.

MOREIRA, M. P.; MOURA, M. A. Construindo tesouros a partir de tesouros existentes: a experiência do tci - tesouro em ciência da informação. **DataGramZero**, [S.l.], v. 7, n. 4, 2006. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/6670>. Acesso em: 15 mar. 2023.

NOVELLINO, M. S. F. Instrumentos e metodologias de representação da informação. **Informação & Informação**, Londrina, v. 1, n. 2, p. 37-45, dez. 1996. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/1603>. Acesso em: 26 fev. 2023.

UNESCO. **Guidelines for the establishment and development of monolingual thesauri**. Paris, 1973. 37p.

ZENG, Marcia Lei; CHAN, Lois Mai. Semantic interoperability. *In: Encyclopedia of library and information sciences*. CRC Press, 2010. p. 4645-4662.

Big Data e dados pessoais: questões legais emergentes na perspectiva da Ciência da Informação

Big data and personal data: emerging legal issues from an Information Science perspective

Gislaine Cristina da Conceição (1), Januário Albino Nhacuongue (2)

(1) Universidade Federal de São Carlos, PPGCI - campus São Carlos, Rodovia Washington Luís s/n, km 235
(1) gislaine.conceicao88@gmail.com, (2) januario@ufscar.br

Resumo

Esta pesquisa de mestrado em andamento tem por objetivo analisar as discussões acerca de aspectos legais no uso de dados na Ciência da Informação, no Brasil, nos últimos 10 anos. Desse modo, procura responder à seguinte questão norteadora: qual vem sendo a atenção dedicada aos aspectos legais e de privacidade no uso de dados pessoais nas pesquisas científicas dos últimos dez anos da Ciência da Informação, no Brasil? O quadro teórico de referência foi estabelecido a partir dos conceitos: *big data*; ciclo de vida de dados; dados e privacidade. Quanto à metodologia, foi feita uma análise de domínio dos materiais bibliográficos referentes a dados e privacidade na Ciência da Informação, no Brasil, nos últimos 10 anos. O foco da análise de domínio foi concentrado na identificação dos autores e suas filiações institucionais; de assuntos tratados e respectivas teorias de sustentação; de áreas de aplicação dos estudos. Após a análise de 173 artigos escritos por 243 autores, a pesquisa mostra que 75% dos estudos tratam sobre a Fase da Coleta de Dados, abordando aspectos sobre privacidade, integração, qualidade, disseminação e os direitos autorais no uso de dados pessoais. Constatou-se que as instituições UNESP, UFRJ, UNB, UFPE, FGV e UFMG com 39,5% lideram o *ranking* em publicações e afiliações de autores que abordam a Categoria Coleta de Dados. Conclui-se que há a necessidade de explorar mais os aspectos legais em todas as fases do ciclo de vida de um dado (CVD), principalmente, na era da informação.

Palavras-chave: Big Data. Dados Pessoais. Privacidade. Ciclo de Vida de Dados. Lei Geral de Proteção de Dados (LGPD).

Abstract

This master's research in progress aims to analyze the discussions about legal aspects in the use of data in Information Science, in Brazil, in the last 10 years. In this way, it seeks to answer the following guiding question: what attention has been devoted to legal and privacy aspects in the use of personal data in scientific research in the last ten years of Information Science in Brazil? The theoretical frame of reference was established from the concepts: *big data*; data lifecycle; data and privacy. As for the methodology, a domain analysis of bibliographic materials referring to data and privacy in Information Science, in Brazil, in the last 10 years was carried out. The focus of the domain analysis was concentrated on identifying the authors and their institutional affiliations; of treated subjects and respective supporting theories; areas of study application; After analyzing 173 articles written by 243 authors, the research shows that 75% of the studies deal with the Data Collection Phase, addressing aspects of privacy, integration, quality, dissemination and copyright in the use of personal data. It was found that the institutions UNESP, UFRJ, UNB, UFPE, FGV and UFMG with 39.5% lead the ranking in publications and affiliations of authors that deal with the Data Collection Category. It is concluded that there is a need to further explore the legal aspects in all stages of the data life cycle (DLC) mainly in the information age.

Keywords: Big data. Personal data. Privacy. Data Lifecycle. General Data Protection Law (LGPD).

Resumen

Esta investigación de maestría en curso tiene como objetivo analizar las discusiones sobre aspectos legales en el uso de datos en Ciencias de la Información, en Brasil, en los últimos 10 años. De esta manera, busca responder a la siguiente pregunta orientadora: ¿qué atención se ha prestado a los aspectos legales y de privacidad en el uso de datos personales en la investigación científica en los últimos diez años de Ciencias de la Información en Brasil? El marco teórico se estableció a partir de los conceptos: *big data*; ciclo de vida de los datos; datos y privacidad. En cuanto a la metodología, se realizó un análisis de dominio de materiales bibliográficos referentes a datos y privacidad en Ciencias de la Información, en Brasil, en los últimos 10 años. El enfoque del análisis de dominio estuvo en la identificación de los autores y sus afiliaciones institucionales; los temas abordados y las respectivas teorías de apoyo; estudiar áreas de aplicación; Después de analizar 173 artículos escritos por 243 autores, la investigación muestra que el 75% de los estudios se ocupan de la Fase de Recolección de Datos, abordando aspectos de privacidad, integración, calidad, divulgación y derechos de autor en el uso de datos personales. Se constató que las instituciones UNESP, UFRJ, UNB, UFPE, FGV y UFMG con el 39,5% lideran el *ranking* en publicaciones y afiliaciones de autores que se ocupan de la Categoría de Recopilación de Datos. Se concluye que existe la necesidad de explorar más a fondo los aspectos legales en todas las etapas del ciclo de vida de los datos (CVD), especialmente en la era de la información.

Palabras clave: Grandes datos. Datos personales. Privacidad. Ciclo de vida de los datos. Ley General de Protección de Datos (LGPD).

1 Introdução

Na Sociedade da Informação, as tecnologias modernas permitem a coleta e o tratamento rápido de um conjunto cada vez mais vasto de dados pessoais. Embora esse processo traga vantagens em diferentes contextos de utilização de dados, como tomada de decisões estratégicas em empresas, em algumas vezes, ocorre em concomitância com violações de privacidade. Pode-se dar como exemplo a forma como os dados são reunidos, pois, são coletados de perfis em redes sociais, transações comerciais e monitoramento de navegação na Internet, etc., muitas vezes sem o conhecimento do usuário, e utilizados para as mais diversas finalidades (LIMA, 2019).

À medida que as redes de computadores e o uso da Internet foram crescendo e se intensificando, novas formas de comunicações e registros de informações tornaram-se possíveis, facilitando o acesso e a recuperação das mesmas. Neste novo cenário, onde computadores e dispositivos tecnológicos são cada vez mais utilizados para ampliar as habilidades de comunicação humana, surge a preocupação com a privacidade, na medida em que há uso indevido de informações e dados pessoais em vários contextos (GRISOTO; SANTANA; SANTARÉM SEGUNDO, 2015).

O desenvolvimento tecnológico ocorrido nas últimas décadas, por um lado, levou não apenas a um aumento no fluxo de dados e informações, mas também proporcionou melhorias na tecnologia da informação. Por outro lado, impulsionou o risco de exposição e uso indevido ou abusivo desses dados (LIMA, 2019).

O acesso a dados vem transformando todas as esferas da atividade humana, principalmente nos últimos anos, pois as alternativas para coletar, armazenar e recuperar dados cresceram exponencialmente, superando até mesmo nossa capacidade de lidar com esses novos volumes de dados, perspectivas, diversidade e velocidade de acesso. Este novo cenário está dando espaço para o Big Data (SANTANA, 2016).

De acordo com Dumbill (2012, p.3):

Big Data são dados que excedem a capacidade de processamento de sistemas de bancos de dados convencionais. Os dados são muito volumosos, se movem muito rápido, ou não se encaixam nas estruturas das arquiteturas de banco de dados.

A informatização, conforme entende Doneda (2006), operou mudanças de ordem quantitativa e qualitativa no tratamento de dados: uma maior quantidade de dados passou a ser processada em menos tempo e resultados mais valiosos passaram a ser obtidos a partir do uso de novos métodos, algoritmos e técnicas de processamento. Com o crescimento desordenado e em larga escala da produção de dados e informações, surge a preocupação com a privacidade acerca do seu uso.

Nessa perspectiva, o conceito de privacidade se transforma para dar origem à disciplina da proteção de dados pessoais, tendo em vista o crescimento da importância da informação, o aumento das situações violadoras da privacidade e a proporção da população afetada por tais situações. Nas palavras de Rodotà (2008), o direito à privacidade é entendido como "o direito de manter o controle sobre as próprias informações", que é construído no eixo "pessoas-informação-circulação-controle" ao invés de "pessoa-informação-sigilo", em torno do qual foi construído o conceito clássico de privacidade.

A discussão proposta neste trabalho está inserida na problemática da privacidade e procura responder à seguinte questão norteadora: qual vem sendo a atenção dedicada aos aspectos legais e de privacidade no uso de dados pessoais nas pesquisas científicas dos últimos dez anos da Ciência da Informação, no Brasil? O objetivo é analisar as discussões acerca de aspectos legais e de privacidade no uso de dados na Ciência da Informação, no Brasil, nos últimos 10 anos.

A justificativa principal para esta abordagem encontra-se tanto no próprio objeto da Ciência da Informação (informação), como no aumento de dados e dos mecanismos da sua utilização pelo setor público e privado atualmente.

A demanda por acesso e análise de dados privados e públicos tem aumentado, ganhando a atenção de diversos segmentos da sociedade. A situação atual mostra que empresas lucram, governos influenciam as decisões relacionadas aos seus cidadãos e pessoas são induzidas a aceitar conceitos, agravados pelo fato de que os dados que financiam essas ações nem sempre estão envolvidos de princípios de ética e justiça (KEMPER; KOLKMAN, 2018).

O Observatório Febraban de junho de 2021 buscou investigar a visão da sociedade sobre a segurança de dados no Brasil e os crimes envolvendo violação de informações pessoais. O levantamento mostrou que os brasileiros estão atentos a essa questão, sobretudo quanto ao uso que as empresas privadas fazem dos seus dados pessoais. E em alguns casos consideraram a utilização dos dados aceitável, como por exemplo para a prevenção de crimes, de golpes em compras e fraudes com dados bancários. Outra constatação do estudo é que as pessoas passaram a adotar ou intensificar hábitos de proteção em relação a senhas, fornecimento de dados, aceitação de cookies, uso de cartão de crédito, entre outros.

No entanto, apesar da consciência sobre a importância de proteção de dados pessoais por parte de alguns entrevistados, a realidade social e a complexidade tecnológica mostram que a maior parte da população não possui mecanismos de controle e impedimento do uso indevido dos seus dados. Ademais, as violações da

privacidade podem acontecer tanto em usos criminosos (por exemplo, em ataques cibernéticos) como em usos legais (por exemplo, por governos ou instituições detentoras) de dados pessoais. Isso reforça a necessidade de mais estudos sobre o assunto.

2 Referencial Teórico

2.1 Big Data

O avanço tecnológico possibilitou que as organizações capturassem e gerissem trilhões de bytes de informações de diferentes fontes, como dados de consumidores e fornecedores. Este panorama indica novas possibilidades de uso de dados.

Este novo cenário vem sendo definido como Big Data e de acordo com Davenport (2014):

Big data se refere a dados que são grandes demais para um único servidor, muito diversos para se adequar a uma base de dados estruturada em linhas e colunas, ou cujo fluxo seja tão intenso que não permita adequação a um Data Warehouse estático (DAVENPORT, 2014).

Com o aumento exponencial do acesso aos dados, todas as áreas de atuação humana estão passando por um crescimento nos últimos anos, em função da coleta, do armazenamento e da recuperação de dados, superando a capacidade humana de lidar com estas novas perspectivas de volume, variedade e velocidade de acesso aos dados. (SANTANA, 2016).

Para Davis e Patterson (2012), embora a sociedade, de modo geral, esteja entusiasmada com os benefícios potenciais de desenvolvimento de novos produtos e serviços baseados em tecnologias de big data, o tamanho, a variedade e a velocidade das informações disponíveis levantam novas questões sobre transparência, correlação e agregação no uso de dados.

Shah, Peristeras e Magnisalis (2020) vão na mesma linha, ao considerar que, atualmente, setores governamentais implementam políticas centradas em dados para, entre vários objetivos, promover boa governança, transparência, serviços inovadores, etc. Contudo, na prática, o gerenciamento do big data ao longo do ciclo de vida dos dados ainda é um tremendo desafio.

Na nossa análise, as questões inerentes à big data e seu uso potencial, especialmente no âmbito digital, são recentes e ainda não nos permitem ter uma visão completa de todos os aspectos envolvidos. Por isso, concordamos com Santana (2016), sobre a necessidade de uma estrutura que permita a percepção de momentos e características, que, no caso dos dados, podem ser considerados como cíclicos e recorrentes, proporcionando a reflexão sobre um ciclo de vida dos dados.

2.2. Ciclo de vida de dados

A Ciência da Informação tem como dever contribuir para o novo cenário de acesso e uso intensivo de dados, explorando fatores e características que permitam a ampliação do equilíbrio entre os usuários do processo e as entendidas envolvidas no gerenciamento. Esses fatores e características podem ser compreendidos a partir do ciclo de vida de dados.

Rahul e Banyal (2020) consideram que o ciclo de vida dos dados apresenta todo o processo de dados no sistema, através de uma estrutura de alto nível para planejar, organizar e gerenciar todos os aspectos durante suas fases (coleta, armazenamento, usabilidade, compartilhamento e descarte dos dados). Logo, para uma implementação bem-sucedida do big data, é necessário manter o ciclo de vida dos dados sob um sistema ou modelo seguro, completo e eficaz de gerenciamento de dados.

O problema é que existem vários modelos de ciclo de vida de dados, com diferentes características e aplicações. Por exemplo, Sant'Ana (2013) identificou 18 iniciativas de modelos de ciclo de vida de dados entre 1997 e 2013 (SAYÃO; SALES, 2015).

Shah, Peristeras e Magnisalis (2020), por meio de uma Revisão Sistemática da Literatura, identificaram e analisaram 76 modelos de ciclos de vida de dados para propor uma estrutura de ciclo de vida de dados para governos. Os modelos analisados eram referentes a diversos domínios de aplicação, como pesquisa científica, Web Semântica, bancos de dados, qualidade de dados, processamento de dados, etc. Além disso, envolviam diferentes etapas do ciclo de vida, algumas confusas e difíceis de associá-las às práticas de gestão em organizações.

O modelo de ciclo de vida de dados para governos orientados a dados, proposto por Shah, Peristeras e Magnisalis (2020), envolve 14 fases: planejamento, coleta, preparação, análise, visualização, acesso, compartilhamento, uso e reuso, arquivamento, descarte, governança, qualidade, armazenamento e segurança ou proteção de dados. Para cada fase existem funções principais a serem consideradas. Por exemplo, na fase de coleta deve-se apresentar um programa de conscientização sobre proteção de dados durante a coleta de dados; coletar metadados com base em padrões de metadados, para garantir a interoperabilidade, etc.

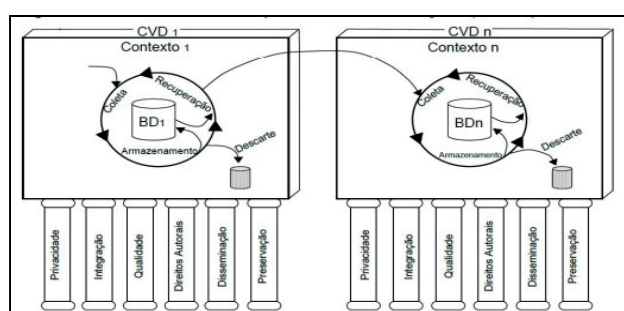
Embora o modelo dos autores possa ser utilizado em várias áreas do setor público, como saúde, agricultura, educação; seja iterativo e envolva a completude do em relação aos desafios do big data (volume, variedade, velocidade e veracidade), na nossa análise, seria de difícil compreensão e aplicação no contexto desta pesquisa.

O modelo de Ciclo de Vida dos Dados proposto por Sant'Ana (2016) foi desenvolvido levando em consideração as características e especificidades da Ciência da Informação.

Conforme afirma Santana no Ciclo de Vida dos Dados (2016, p.8):

[...] se faz necessário estruturar esta análise, e para tanto, propõe-se a utilização de uma delimitação de fases (momentos em que distintas necessidades e competências são necessárias) envolvidas no acesso e uso dos dados, mantendo-se como ponto central os próprios dados e para tanto se propõe o uso do Ciclo de Vida dos Dados - CVD como forma de evidenciar os diferentes momentos e fatores envolvidos neste processo.

Figura 1 - *Ciclo de Vida dos Dados para Ciência da Informação – (CVD-CI)*



Fonte: Sant'ana (2016).

2.2.1 Fases do Ciclo de Vida dos Dados e fatores paralelos

A figura 1 ilustra as fases do ciclo de vida dos dados que devem ser consideradas na perspectiva epistemológica e pragmática da Ciência da Informação: coleta, armazenamento, recuperação e descarte. Para cada fase do ciclo de vida devemos considerar os seguintes fatores norteadores: privacidade, integração, qualidade, direitos autorais, disseminação e, preservação.

Coleta: a princípio, é necessário obter dados que possam ser utilizados para atender a uma necessidade específica ou demanda esperada relacionada a um contexto específico. Neste primeiro momento são necessárias medidas para implementar um plano de ação, análise de viabilidade e realização da coleta de dados. Outras questões importantes nesta fase incluem: como os dados podem ser coletados? Em que formato estão? Quais são os dados necessários? Nesta etapa, o fator privacidade recai principalmente sobre as características e fontes de dados e o processo envolvido na sua coleta. Quanto à integração, torna-se necessário considerar os principais atributos para identificação dos dados, incluindo aspectos da sua granularidade. Nesta fase também se torna imprescindível considerar os aspectos inerentes à propriedade intelectual e à ética no tratamento de dados, por exemplo, respeitando os limites estabelecidos na LGPD. A disseminação nesta fase diz respeito a identificação de elementos contextuais

dos dados que possibilitem a localização e interpretação na fase de recuperação. A preservação, por sua vez, diz respeito a estruturas de armazenamento e a linguagens de representação (SANTANA, 2016).

Armazenamento: Posteriormente, depois da Coleta dos Dados, pode ser necessário e útil manter os mesmos disponíveis de alguma forma para referência futura. Neste momento passa-se, então, a uma segunda fase que é aquela em que os esforços são no sentido de manter estes dados em uma determinada base de dados. Esta fase recebe o nome de armazenamento. Outras questões importantes nesta fase incluem: Qual estrutura (física e lógica) será utilizada para seu armazenamento? Estes dados podem representar um risco a privacidade dos indivíduos ou instituições neles referenciados de alguma forma? (SANTANA, 2016).

Recuperação: Esta fase tem como objetivo a viabilização da recuperação dos dados. Dentre outras questões fundamentais desta fase pode-se destacar: quais dos dados armazenados serão disponibilizados? Existe algum público-alvo específico? Durante o processo de recuperação quais são os riscos à privacidade dos indivíduos ou entidades referenciados pelos conteúdos recuperados? O acesso será feito diretamente a base em que se encontra armazenado ou será necessário retornar a fase de armazenamento para definição de novas estruturas de armazenamento específicas para recuperação? (SANTANA, 2016).

Descarte: Nesta fase são analisados quais dados são importantes para continuar sendo mantidos em uma base e dados e dos demais são descartados, ou seja, excluídos. Quais dados já não são mais necessários? Estes dados estão replicados em outras bases? Como garantir e explicitar que estes dados foram realmente excluídos e não simplesmente ocultados? Para o descarte foi considerada a necessidade de preservação em seus diversos aspectos? (SANTANA, 2016).

De acordo com Santana (2016, p. 134-137) se estamos considerando que um dado não está disponível na fase de recuperação, o principal fator envolvido é o "disseminação", ou seja, este dado não está disponível para que outros colem, no entanto, o motivo que leva a esta falha na disseminação pode estar relacionado a diversos motivos, exemplificando:

- [1] o dado não foi disseminado por uma questão de erro na gestão dos dados, portanto seria também uma forma de considerar o fator "qualidade";
- [2] o dado não foi disseminado por ser considerado um dado sensível, portanto seria também uma forma de considerar o fator "privacidade";
- [3] o dado não foi disseminado por não ter dados complementares que permitissem sua vinculação a pesquisa que foi realizada pelo usuário em outro conjunto de dados, portanto seria também uma forma de considerar o fator "integração";

[4] o dado não foi disseminado por não ter formalmente autorização de detentores para sua publicação, portanto seria também uma forma de considerar o fator "legislação", e;

[5] o dado não foi disseminado por não ter sido mantido íntegro, seja em sua forma física ou lógica, pelo tempo necessário, portanto seria também uma forma de considerar o fator "preservação".

2.3 Dados e Privacidade

O aumento do volume de dados tem modificado a maneira como pesquisas, governança, socialização e negócios estão sendo realizados. Podem ser produzidos por diferentes classes sociais, governos, indústrias, institutos de pesquisa e universidades. Eles são coletados, armazenados e divulgados de diferentes maneiras e com diferentes organismos (ANDRADE; GONÇALEZ; BERTI JUNIOR; BAPTISTA; CONEGLIAN, 2020).

De acordo com Milagre, Santarém Segundo (2015, p.2): Dado é qualquer elemento identificado em sua forma bruta que, por si só, não conduz a uma compreensão de determinado fato ou situação.

Para Souza e Almeida (2021, p. 4):

[...] os dados não apresentam significados próprios, porém, representam a matéria prima como uma série de observações, medidas ou fatos que podem ser representados como números, palavras, sons ou imagens ao qual é produzida a informação.

Desta maneira, é possível dizer que o dado é o estado bruto de uma futura informação, pois o mesmo após ser coletado, processado e analisado será transformado em uma informação que será utilizado por uma organização.

Na vida cotidiana, é comum que as pessoas usem e forneçam seus dados pessoais para diversos fins, principalmente aqueles que resultam em compras pela internet e em estabelecimentos comerciais físicos (VIGNOLI; VERONEZ JUNIOR; SILVA; MONTEIRO, 2020).

De acordo com a Lei nº 12.527/2011 - Lei de Acesso à Informação (LAI) – e a Lei nº 13.709/2018 - Lei Geral de Proteção de Dados Pessoais (LGPD) - os dados pessoais são "[...] informações relacionadas a uma pessoa identificada ou identificável". (Brasil, 2011, 2018). De acordo com a legislação brasileira, os dados pessoais são originários de qualquer indivíduo que possa ser identificado de forma única por meio de dados e informações. Exemplos de dados pessoais incluem números de documentos como o Cadastro Geral (RG), o Cadastro de Pessoas Físicas (CPF), o próprio nome, nomes dos pais e outras informações que possam identificar uma pessoa física (VIGNOLI; VERONEZ JUNIOR; SILVA; MONTEIRO, 2020).

De acordo com Silveira, Avelino e Souza (2016, p.3):

Dado pessoal pode ser definido como qualquer informação relativa a uma pessoa singular identificada ou identificável. O Relatório do Fórum Econômico Mundial qualificou o dado pessoal como as informações e metainformações criadas por e sobre as pessoas, abrangendo: dados oferecidos voluntariamente (exemplo: perfil na rede social), dados observados (como: dados de localização ao usar os celulares) e dados inferidos (exemplo: análise de informações oferecidas ou observadas com a finalidade de construir uma pontuação de crédito) (SILVEIRA; AVELINO; SOUZA, 2016, p.3).

A privacidade é alterada à medida que ocorrem as transições sociais, políticas, culturais e econômicas, partindo do pressuposto de que o seu conceito é a inviolável, pois busca garantir o direito do indivíduo à vida privada e disponibilização de seus dados e a utilização de suas informações por terceiros (GRISOTO; SANT'ANA; SEGUNDO, 2015).

O direito à privacidade é um direito humano de primeira dimensão que se configura no direito do indivíduo de estar só, considerando-se a possibilidade que deve ter toda pessoa de excluir do conhecimento de terceiros aquilo que a ela só se refere, e que diz respeito ao seu modo de ser no âmbito da vida privada. O direito à privacidade é universal, inalienável e indivisível e está intimamente ligado à dignidade humana. (GALVÃO; RICARTE, 2012, p.98)

A criação do Marco Civil da Internet dispõe:

[...] sobre princípios, garantias, direitos e deveres para o uso da internet no Brasil e determina as diretrizes para atuação da União, dos Estados, do Distrito Federal e dos Municípios, busca garantir dentre outras coisas, a segurança no acesso e na disponibilização dos dados. (BRASIL, 2014).

Medidas estão sendo tomadas nos termos da Lei de Acesso à Informação e da Lei Marco de Cidadania visando garantir a integridade e a dignidade humana na proteção de suas informações e no direito de acesso à informação pública (GRISOTO; SANT'ANA; SEGUNDO, 2015).

Mendes (2008, p. 17-18) discorre que a:

A evolução do direito à privacidade prosseguiu para se adaptar às novas transformações sociais ocasionadas pela revolução da tecnologia da informação que possibilitou a coleta e o processamento dos dados pessoais dos cidadãos de forma pioneira. Além de adquirir um caráter positivo e de ser reconhecido no âmbito internacional, o direito à privacidade transformou-se para ensejar o nascimento da disciplina de proteção de dados pessoais, à medida que surgiram novos desafios ao ordenamento jurídico a partir do tratamento informatizado dos dados.

No entanto, o conceito de privacidade é bastante complexo e de difícil definição, envolvendo questões subjetivas, ou seja, valores relacionados à dignidade indivi-

dual e humana, e, portanto, requer reflexão de diversos campos do conhecimento. (VASALOU; JOINSON; HOUGHTON, 2014).

3 Procedimentos metodológicos

O campo da Ciência da Informação, segundo Bawden e Robinson (2012), pode ser entendido a partir da sua preocupação com o estudo da informação humana registrada, com foco nos componentes da cadeia de comunicação, investigados sob a perspectiva da análise de domínio. Nesse sentido, este trabalho utiliza a análise de domínio para investigar as discussões acerca de aspectos legais e de privacidade no uso de dados na Ciência da Informação, no Brasil.

Segundo Guimarães (2014), as abordagens teóricas e aplicadas da Ciência da Informação utilizam a análise de domínio para caracterizar e avaliar a ciência, identificando as condições da sua construção e socialização.

Smiraglia (2015) apresenta a análise de domínio como paradigma metodológico para a organização do conhecimento e como conjunto de técnicas e ferramentas para identificação de bases de conhecimento específicas para os diversos setores científicos.

Na Ciência da Informação, a análise de domínios de conhecimento como comunidades discursivas ganhou espaço a partir dos anos 1995, com a proposta de Hjørland e Albrechtsen. Mais tarde, o delineamento das onze abordagens para análise de domínio, de Hjørland (2002), clarificou o quadro para sua aplicação em estudos de organização do conhecimento do campo. Essas abordagens são: 1) Produção de guias de literatura e entradas de assunto; 2) Construção de classificações especiais e tesaurus; 3) Indexação e recuperação de especialidades; 4) Estudos empíricos de usuários; 5) Estudos bibliométricos; 6) Estudos históricos; 7) Estudos de Documentos e Gênero; 8) Estudos Epistemológicos e Críticos; 9) Estudos Terminológicos, linguagem para propósitos especiais (na língua inglesa, Special Purpose Languages - SPL), bases de dados semânticas e estudos de discurso; 10) Estruturas e Instituições na Comunicação Científica, e 11) Cognição Científica, Especialistas do Conhecimento e Inteligência Artificial: (TENNIS, 2003; GUIMARÃES, 2014; Hjørland, 2017).

Smiraglia (2015) fez um levantamento de estudos analíticos conduzidos e relatados na literatura do domínio e observou que, apesar do crescimento do paradigma de análise de domínio na organização do conhecimento ao longo do tempo, com embasamento nas 11 abordagens de Hjørland, alguns contextos continuavam fora da análise. Por exemplo, guias de literatura; especialidade de indexação; e estudo de gênero e comunicação científica. Consequentemente, apresentou a seguinte taxonomia revisada de abordagens: 1) Guias de assuntos ou guias literaria; 2) Classificações especiais e

tesaurus; 3) Estudos empíricos de usuários; 4) Estudos informativos; 5) Estudos históricos; 6) Documentação e estudos de gênero; 7) Estudos epistemológicos e críticos; 8) Estudos terminológicos; 9) Semântica do banco de dados; 10) Análises de discurso e 11) Cognição, conhecimento especializado e Inteligência Artificial.

Neste trabalho, para análise de domínio, primeiramente foi feita a coleta de materiais bibliográficos referente a dados e privacidade na CI, nas seguintes fontes de pesquisa: Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI); Scientific Electronic Library Online (SciELO); Biblioteca Digital Brasileira de Teses e Dissertações (BDTD); Anais do Congresso Brasileiro de Organização e Representação do Conhecimento organizado pela ISKO-Brasil; Anais do Encontro Nacional de Pesquisa em Ciência da Informação - ENANCIB.

Os critérios de inclusão dos estudos são: publicação dos últimos dez anos; publicação em português, inglês e espanhol; acesso ao texto completo; o estudo aborda o tema na área de modulação e no foco do domínio.

A pesquisa foi feita com base nos seguintes termos, com auxílio de operadores booleanos e de proximidade (AND, OR, NOT, WITH e NEAR): privacidade; dados pessoais; segurança cibernética; vigilância, lei de acesso à informação e lei geral de proteção de dados pessoais.

A análise do domínio se concentrou na abordagem Estudos históricos. Tennis (2003) nos mostra que o principal problema no uso da análise de domínio não está no processo investigativo do objeto, mas sim na especificação da clareza do objeto a ser investigado – o domínio. Como solução, o autor propõe dispositivos analíticos – eixos, de modo a delinear o que o analista de domínio está estudando. O eixo Áreas de Modulação fornece os parâmetros sobre o nome e a extensão do domínio, respondendo à pergunta: como se chama o domínio e o que abrange? O eixo Graus de Especialização qualifica e define a intenção do domínio, diminuindo a sua extensão.

Assim, quanto ao parâmetro nomeação, o domínio da pesquisa é produção científica da Ciência da Informação. A extensão do domínio é formada por estudos sobre aspectos legais no uso de dados pessoais. Considerando a diversidade de abordagens possíveis neste domínio¹, a intenção ou o foco do domínio é: autores e suas filiações institucionais; assuntos tratados e respectivas teorias de sustentação; área de aplicação do estudo; relações de citação e ou colaboração entre autores.

As áreas de aplicação do estudo foram especificadas a partir das etapas do ciclo de vida de dados: coleta, armazenamento, recuperação e descarte (SANTANA, 2016).

4 Resultados e Discussões

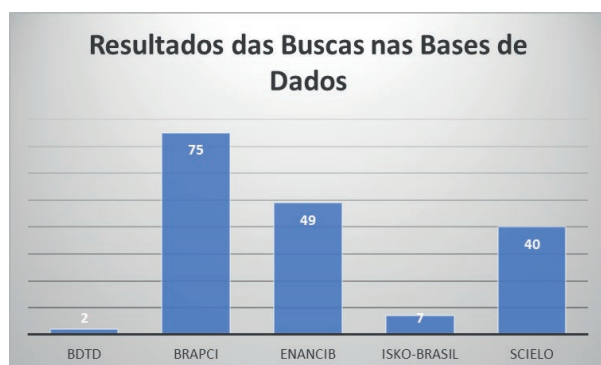
O objetivo de identificar como os trabalhos apresentados nas Bases de Dados Referenciais de Artigos de Periódicos em Ciência da Informação tratam a temática sobre aspectos legais no uso de dados pessoais foi alcançado utilizando como método o estudo qualitativo e quantitativo, tendo a literatura disponível como subsídios para a definição do processo de análise de domínio, além da coleta de dados, por meio da qual os trabalhos apresentados nas Bases de Dados foram categorizados, com a devida identificação de informações sobre os autores, suas respectivas instituições, assuntos tratados, respectivas teorias de sustentação e área de aplicação do estudo.

A partir da análise dos trabalhos apresentados nas Bases de Dados BRAPCI, SciELO, BDTD, ISKO-Brasil e ENANCIB na forma de Comunicações Orais e de Pôsteres, ao longo da trajetória dos últimos dez anos (2012-2022). Do total de publicações recuperadas (350), buscou-se identificar aqueles que tiveram os termos privacidade; dados pessoais; segurança cibernética; vigilância, lei de acesso à informação e lei geral de proteção de dados pessoais, no título, resumo e/ou palavras-chave. Algumas das publicações não atenderam a esses critérios ou não havia uma sobreposição entre eles (177).

Na base de dados BRAPCI foram recuperados 75 artigos, seguidos da ENANCIB com 49 artigos, a Scielo com 40, ISKO-Brasil com 7 e a BDTD com 2 artigos recuperados.

A partir da análise dos trabalhos apresentados nas Bases de Dados BRAPCI, SciELO, BDTD, ISKO-Brasil e ENANCIB na forma de Comunicações Orais e de Pôsteres, ao longo da trajetória dos últimos dez anos (2012-2022), resultou em 173 artigos. O Gráfico 1 apresenta a distribuição dos artigos nas categorias em estudo.

Gráfico 1 - Resultados das Buscas nas Bases de Dados



Fonte: Elaborado pelos autores.

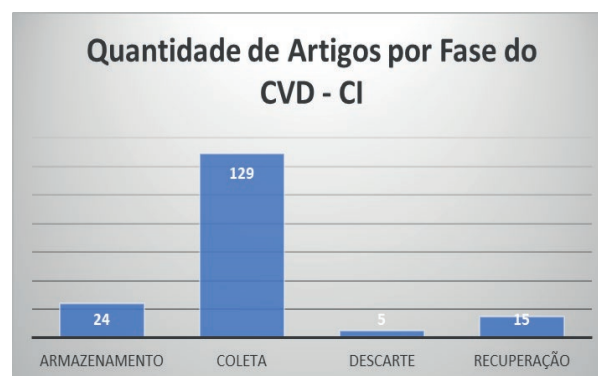
Assim, para este estudo, considerando a inclusão definida e critérios de exclusão na metodologia, foram ana-

lisadas 173 publicações. O Gráfico 2 apresenta a distribuição dos artigos nas categorias em estudo.

A partir da compilação dos dados coletados, foi possível identificar que mais 75% dos trabalhos dizem respeito a Fase Coleta de Dados.

Mediante análise dos artigos percebe-se que 39% dizem respeito aos Fatores Privacidade, seguido de Integração com 24,3%, já Qualidade ocupa a terceira posição com 19%, Disseminação vem em quarto lugar com 12,2% enquanto Direitos Autorais tem 4,2% e Armazenamento 1,3% das publicações analisadas.

Gráfico 2 - Quantidade de Artigos por Fase do CVD - CI



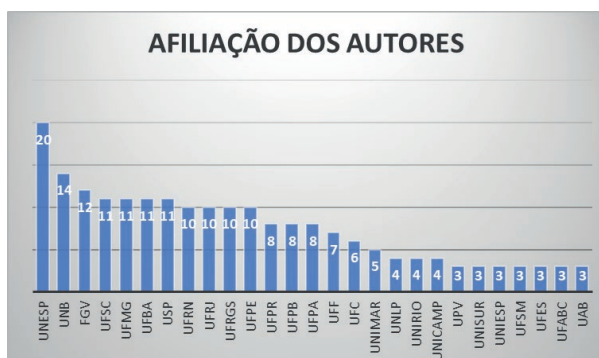
Fonte: Elaborado pelos autores.

Desta forma, a análise de domínio será focada na Fase Coleta de dados, onde será realizada a geração de gráficos com informações de distribuição temporal, de autores e suas filiações institucionais; assuntos tratados e respectivas teorias de sustentação; área de aplicação do estudo; relações de citação e ou colaboração entre autores. Em seguida, foi feita uma análise histórica de tendências das discussões nos últimos dez anos.

A Categoria Coleta de Dados conta com duzentos e quarenta e três autores distribuídos por cinquenta e cinco instituições.

O Gráfico 3 demonstra as Instituições que mais possuem autores afiliados que abordam sobre o tema são UNESP (Universidade Estadual Paulista) com vinte autores afiliados, seguida da UNB (Universidade de Brasília) com quatorze autores, a FGV (Fundação Getúlio Vargas) com doze autores afiliados. Já a UFSC, UFMG, UFBA possuem onze autores afiliados por instituições, seguidas da USP, UFRN, UFRJ, UFRGS, UFPE com dez autores afiliados por instituições e as demais seguem com oito ou menos autores por instituições.

Gráfico 3. Afiliação dos Autores

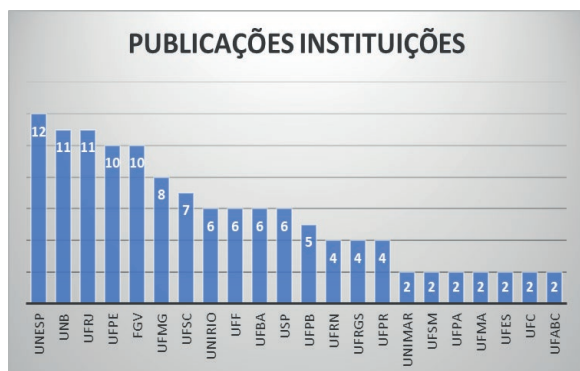


Fonte: Elaborado pelos autores.

Com as informações coletadas nas bases de dados, cinquenta e cinco instituições tiveram publicações referentes a Coleta de Dados. No Gráfico 4 está representado apenas as instituições que tiveram duas ou mais publicações sobre o tema.

Nota-se que que a Unesp teve doze publicações sobre a Fase da Coleta de Dados, seguidas das instituições UNB (Universidade de Brasília) e UFRJ (Universidade Federal do Rio de Janeiro) com onze publicações cada uma, já a UFPE (Universidade Federal de Pernambuco) e a FGV (Fundação Getúlio Vargas) contribuíram com dez publicações cada uma. As demais instituições tiveram uma variação de oito a dois artigos.

Gráfico 4. Publicações por Instituições



Fonte: Elaborado pelos autores.

A partir da compilação de dados, foi possível identificar que mais 90% dos trabalhos dizem respeito a Fase Coleta de Dados. Desta forma, a análise de domínio será focada na Fase Coleta de dados, onde será realizada a geração de gráficos com informações de distribuição temporal, de autores e suas filiações institucionais; assuntos tratados e respectivas teorias de sustentação; área de aplicação do estudo; relações de citação e ou colaboração entre autores. Em seguida, foi feita uma análise histórica de tendências das discussões nos últimos dez anos.

Segundo Smiraglia (2015), em domínios não científicos, a análise bibliométrica e outras formas de docu-

mentação devem ser usadas para analisar a comunicação dentro do domínio. Em comunidades acadêmicas, como as de ciências sociais e humanas, a análise de citações fornece vestígios de evidências de relações temáticas e teóricas. Nesse contexto, os dados das publicações foram organizados de forma que fosse possível apresentar padrões de reconhecimento explícito de cada publicação. Assim, quantitativamente, o Gráfico 5 mostra a frequência de publicações segundo o ano.

Gráfico 5. Frequência anual de publicações



Fonte: Elaborado pelos autores.

A categoria Coleta, conforme Gráfico 6 foi concebida abrangendo estudos de análise, estudo comparado, relato de experiência, estudo de caso. Esta categoria conta com 129 artigos escritos por 238 autores.

A apresentação de citações de trabalhos dessa categoria demonstra como se deu a identificação na categoria Coleta;

[...] Em um primeiro momento, depara-se com a fase de obtenção dos dados que pode ser identificada como aquela em que ocorrem: a definição das necessidades informacionais que irão nortear as escolhas e definições sobre quais dados são necessários; são estabelecidas estratégias sobre como localizar e avaliar estes dados; são escolhidos os mecanismos que serão utilizados para sua obtenção, e; são elaboradas as metodologias e ferramentas necessárias para consecução destes dados (SANTANA, 2016, p.9).

[...] A coleta é a obtenção, recepção ou produção de dados pessoais independente do meio utilizados (documento papel, documento eletrônico, sistema de informação, etc.) (SEF, 2021, p.2).

Gráfico 6. Publicações por Autores Categoria Coleta



Fonte: Elaborado pelos autores.

O Gráfico 6 traz o ranking dos vinte e três autores que mais publicaram sobre a Categoria Coleta. O primeiro lugar é liderado pelo autor Ricardo Santana, com 9 publicações, o segundo lugar é liderado pelas autoras SIEBRA, S. e VENTURA, K. com cinco publicações sobre o tema Coleta de Dados. Já o terceiro lugar é ocupado pelos autores BARROS, D., BEZERRA, A., JARDIM, J.M., MALIN, A.M., PINHEIRO, M. e RODRIGUES, F.A., com três publicações por autores

6 Conclusão

O trabalho trata-se de um estudo em andamento, até o presente momento tem-se como resultado o mapeamento parcial dos trabalhos das Bases de Dados Referenciais de Artigos de Periódicos em Ciência da Informação sobre a temática dos aspectos legais no uso de dados pessoais durante o período de 2012 a 2022.

Com a categorização temática estruturada a partir das etapas do ciclo de vida de dados segundo Santana (2016), no âmbito da Ciência da Informação permitiu-se identificar principais focos de interesse das Bases de Dados, no período em estudo, bem como verificar tendências, distribuições de modalidades e rankings de autoria e de instituições.

Através desta análise pode-se identificar que os trabalhos das Bases de Dados estão focados 75% no processo envolvendo a Coleta de dados. Nota-se a carência de estudos nas temáticas das fases do Ciclo de Vida dos Dados como Armazenamento, Recuperação e Descarte de Dados Pessoais. Essas temáticas são fundamentais para compreender os aspectos legais no uso de dados pessoais e traçar possíveis direções de ações no âmbito da Ciência da Informação.

Este estudo está em desenvolvimento e é um ponto de partida importante para o encaminhamento de futuras pesquisas, tendo em vista a necessidade de um olhar crítico e reflexivo sobre aspectos que envolvem o uso de dados e informações para subsidiar ações estratégicas de vários segmentos da sociedade na atualidade.

Notas

[1] A pesquisa em análise de domínio pode ser considerada metateórica, na medida em que incorpora muitos polos teóricos ao mesmo tempo, e de forma deliberada, a fim de fornecer a triangulação metodológica como um meio de validade de construto (SMIRAGLIA, 2015, P. 7).

Referências

ANDRADE, M. C.; GONÇALEZ, P. R.; BERTI JUNIOR, D. W.; BAPTISTA, A. A.; CONEGLIAN, C. S. Ciência responsável dos dados: imparcialidade, precisão, confidencialidade, e transparência dos dados. **Informação & Informação**, v.

25, n. 2, p. 26-48, 2020. DOI: 10.5433/1981-8920.2020v25n2p26 Acesso em: 08 set. 2022.

AKERLOF, G., A. The market for “lemons”: quality uncertainty and the market mechanism. **The Quarterly Journal of Economics**, Cambridge, v. 84, n. 3, p. 488-500, Aug. 1970.

BORGMAN, C. L. **Big data, little data, no data: scholarship in the networked world**. Massachusetts London, England: The MIT Press Cambridge, 2015.

BRASIL. Lei n.º 12.527, de 18 de novembro de 2011. **Lei de Acesso à Informação (LAI)**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília: Presidência da República, 2011.

BRASIL. Lei n.º 12.965, de 23 de abril de 2014. **Marco Civil da Internet**. Brasília, DF, abr. 2014. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm. Acesso em: 08 set. 2022

BRASIL. Lei n.º 13.709 de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD)**. Brasília, DF, ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2018/2018/lei/113709.htm. Acesso em: 08 set. 2022.

DATAONE. **What is DataONE? Online**. Disponível em: <https://www.dataone.org/what-dataone>. Acesso em: 20 jun. 2018.

DAVENPORT, T. H. **Big data at work: dispelling the myths, uncovering the opportunities**. Harvard: Harvard Business School Publishing, 2014.

DAVIS, K.; PATTERSON, D. **Ethics of Big Data**. Sebastopol, CA: O'Reilly Media, Inc., 2012.

DONEDA, D. **Da privacidade à proteção de dados pessoais**. Rio de Janeiro: Renovar, 2006.

GALVÃO, M. C. B.; RICARTE, I. L. M. **Prontuário do Paciente**. 1. ed. Rio de Janeiro: Guanabara Koogan, 2012. v. 1, 344p.

GRISOTO, A. P.; SANT'ANA, R. C. G.; SEGUNDO, J. E. S. A questão da privacidade no contexto da ciência da informação: uma análise das teses e dissertações do programa de pós graduação em ciência da informação da Unesp Campus de Marília. **Revista Ibero-Americana de Ciência da Informação**, v. 8 No 2, n. 2, p. 165-181, 2015. DOI: 10.26512/rici.v8.n2.2015.2066 Acesso em: 09 abr. 2022.

GUIMARÃES, J. A. C. “Análise de domínio como perspectiva metodológica em organização da informação”. **CI Inf.**, vol. 41 no. 1, jan./apr. 2014, p.13-21, <http://revista.ibict.br/ciinf/article/view/1415/1593>. Acesso em: 13 ago. 2022

HJØRLAND, BIGGER E ALBRECHTSEN, HANNE. “Rumo a um novo horizonte na Ciência da Informação: domínio Glushko, Robert J. (ed.). A disciplina da organização. The MIT Press, 2013, 540 p. análise”, **Journal of the American Society for Information Science**, vol. 46, n.º. 6, 1995, pág. 400- 425,

[https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/\(SICI\)1097-4571\(199507\)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y](https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1097-4571(199507)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y). Acesso em: 13 ago. 2022.

HJØRLAND, BIRGER. “Análise de Domínio em Ciência da Informação: onze abordagens – tradicionais e Harvey, Ross. Curadoria digital: um manual de como fazer. Neal-Schuman Publishers, Inc., 2010. Inovativa”. **Jornal de Documentação**, vol. 58, n.º. 4, 2002, p. 422-462 <https://doi.org/10.1108/00220410210431136>. Acesso em: 13 ago. 2022

LIMA, C. F. de. **O Profiling e a Proteção de Dados Pessoais**. 2019. Disponível em: <https://www.lume.ufrgs.br/bitstream/handle/10183/199951/001100797.pdf?sequence=1&isAllowed=y>. Acesso em: 15 jul. 2022.

MENDES, L. S. **Transparência e privacidade: violação e proteção da informação pessoal na sociedade de consumo**. 2008. 156 f. Dissertação (Mestrado em Direito)–Faculdade de Direito, Universidade de Brasília, DF, 2008. Disponível em: <https://repositorio.unb.br/bitstream/10482/4782/1/DISSERTACAO%20LAURA.pdf>. Acesso em: 9 set. 2022.

MILAGRE, J.; SANTARÉM SEGUNDO, J. E. A propriedade dos dados e a privacidade na perspectiva da Ciência da Informação. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 20, n. 43, p. 47-76, 2015. DOI: 10.5007/1518-2924.2015v20n43p47. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2015v20n43p47>. Acesso em: 9 set. 2022.

RAHUL, K.; BANYAL, R. Data Life Cycle Management in Big Data Analytics. **Procedia Computer Science**, v. 173, 2020. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1877050920315465#cebibsec1>. Acesso em 28 abr. 2023.

SALES, L. F.; CAVALCANTI, M. T. Seleção e avaliação de coleções de dados digitais de pesquisa: uma possível abordagem metodológica. **Informação & Tecnologia (ITEC)**, v. 2, n. 2, p. 88-105, jul./dez. 2015. Disponível em: <http://periodicos.ufpb.br/index.php/itec/article/view/34134/17537>. Acesso em: 14 mai. 2023.

SANT’ANA, R. C. G. Ciclo de vida dos dados e o papel da ciência da informação. In: **Encontro Nacional De Pesquisa Em Ciência da Informação**, 14., 2013, Florianópolis. Anais [...]. Florianópolis, 2013. Disponível em: <http://www.brapci.inf.br/index.php/article/view/0000013772/fac7d2fecdc2c83f1e9208204b2fef10>. Acesso em: 14 mai. 2023.

SANTANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Infor-**

mação, [S.l.], v. 21, n. 2, p. 116–142, dez. 2016. ISSN 1981-8920. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/27940>. Acesso em: 25 mar. 2022. DOI: <http://dx.doi.org/10.5433/1981-8920.2016v21n2p116>.

SAYÃO, L. F.; SALES, L. F. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. Disponível em: <http://carpedien.ien.gov.br/bitstream/ien/1624/1/GUIADEDA DOSDEPESQUISA.pdf>. Acesso em: 14 mai. 2023.

SEF - Secretaria de Estado da Fazenda. **O Ciclo de Vida do Tratamento dos Dados Pessoais**. 2021. Disponível em: <http://www.fazenda.mg.gov.br/transparencia/lgpd/LGPD-SEF-Ciclo-de-Vida-fases.pdf>. Acesso em: 19 nov. 2022.

SHAH, S. I. H.; PERISTERAS, V.; MAGNISALIS, L. Da-LiF: a data lifecycle framework for data-driven governments. **Journal of Big Data**, n. 8, v. 89, 2021. <https://doi.org/10.1186/s40537-021-00481-3>. Disponível em: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-021-00481-3>. Acesso em 3 mai. 2023.

SIEBRA, S. A.; BORBA, V. R. Análise, tendências e perspectivas da produção científica em curadoria digital. In: OLIVEIRA, H. P. C.; VIDOTTI, S. A. B. G. (org.). **Informação e tecnologias: desenhando fronteiras científicas**. João Pessoa: Editora Ufpb, 2018. p. 97-112.

SILVEIRA, S. A.; AVELINO, R.; SOUZA, J. A privacidade e o mercado de dados pessoais. **Liinc em Revista**, v. 12, n. 2, p. 217-230, 30 nov. 2016.

SMIRAGLIA, R. P. **Domain Analysis for Knowledge Organization: Tools for Ontology Extraction**. Oxford: Chandos Publishing, 2015.

SOUZA, M.; ALMEIDA, F. G. O comportamento do termo dado na ciência da informação. **Ciência da Informação em Revista**, v. 8, n. 2, p. 39-54, 2021. DOI: 10.28998/cirev.2021v8n2c. Acesso em: 08 set. 2022. TENNIS, J. T. Two axes of domains for domain analysis. *Knowledge Organization*, v.30, n.3/4, p.191-195, 2003.

VASALOU, A.; JOINSON, A.; HOUGHTON, D. Privacy as a Fuzzy Concept: A New Conceptualization of Privacy for Practitioners. **Journal of the Association for Information Science and Technology**, v. 66, n. 5, May 2014.

VIGNOLI, R. G.; VERONEZ JUNIOR, W. R.; SILVA, A. R.; MONTEIRO, S. D. Mediação da informação e a proteção da privacidade e de dados sensíveis por bibliotecários. **Informação & Informação**, v. 25, n. 4, p. 280-305, 2020.

A catalogação de pinturas no cenário *Linked Data*

Cataloging of paintings from the Linked Data perspective

Luiza Estefano (1), Fabiano Ferreira de Castro (2)

(1) Universidade Federal de São Carlos (UFSCar), Rod. Washington Luiz, s/n – Monjolinho, São Carlos, SP, luizae@estudante.ufscar.br (2) fabianocastro@ufscar.br

Resumo

Ao questionar-se sobre como a descrição de pinturas, considerando o cenário *Linked Data* está sendo discutida na literatura científica da comunidade da Ciência da Informação, busca-se identificar estudos que abordam a catalogação descritiva de obras de arte, com vistas a: a) apresentar os conceitos e as especificidades da pintura; b) analisar e apresentar o estado da arte sobre a catalogação de obras de arte, com destaque para a pintura. Utilizou-se a Revisão Sistemática de Literatura (RSL), para mapear a produção científica sobre o tema. Como resultados, são apresentados os projetos e as iniciativas da descrição de pinturas de acordo com o cenário *Linked Data*. Conclui-se, que existem poucos trabalhos sobre o tema e que estes procuram criar modelos para catalogação de pinturas. Entende-se não ser necessária, pela existência de um padrão de descrição já adaptado a toda tipologia de dados, o *Resource Description and Access* (RDA).

Palavras-chave: Pinturas. Catalogação. Metadados. *Linked Data*.

Abstract

When questioning how the description of paintings, considering the *Linked Data* scenario, is being discussed in the scientific literature of the Information Science community, we seek to identify studies that address the descriptive cataloging of works of art with a view to: a) present the concepts and specificities of painting; b) analyze and present the state-of-the art on the cataloging of works of art, with an emphasis on painting. The Systematic Literature Review (SLR) was used to map scientific production on the subject. As a result, projects and initiatives from the description of paintings are presented according to the *Linked Data* scenario. It is concluded that there are few works on the subject and that they seek to create models for cataloging paintings. It is understood that it is unnecessary due to the existence of a description standard already adapted to all types of data, the *Resource Description and Access* (RDA).

Keywords: Paintings. Cataloging. Metadata. *Linked Data*.

Resumen

Al cuestionar cómo la descripción de pinturas, considerando el escenario de *Linked Data*, está siendo discutida en la literatura científica de la comunidad de las Ciencias de la Información, buscamos identificar estudios que aborden la catalogación descriptiva de obras de arte, con miras a: a) presentar los conceptos y especificidades de la pintura; b) analizar y presentar el estado del arte en la catalogación de obras de arte, con énfasis en la pintura. Se utilizó la Revisión Sistemática de Literatura (SLR) para mapear la producción científica sobre el tema. Como resultado, se presentan proyectos e iniciativas a partir de la descripción de pinturas según el escenario *Linked Data*. Se concluye que existen pocos trabajos sobre el tema y que buscan crear modelos para la catalogación de pinturas. Se entiende que no es necesario, por la existencia de una descripción estándar ya adaptada a todo tipo de datos, el *Resource Description and Access* (RDA).

Palabras clave: Pinturas. Catalogación. Metadatos. Datos vinculados.

1 Introdução

Como uma forma de evitar a não recuperação de obras de arte, os estudos da imagem e sua representação começaram a despertar o interesse de profissionais, preocupados com o seu processamento, tratamento e formas de acesso, seja nos ambientes informacionais convencionais ou digitais.

A representação das pinturas, por meio do processo de catalogação é considerada uma tarefa complexa e que exige habilidades específicas do sujeito construtor da representação.

É por meio do processo de catalogação descritiva que o recurso informacional pode assim ser mais bem descrito, com atributos que o torna único e individualizado, garantindo sua recuperação eficiente e precisa, nos ambientes informacionais disponíveis e na confecção de metadados metodologicamente padronizados (CASTRO, 2020).

A utilização de metadados para a catalogação de pinturas, sozinha, ainda não garante uma boa recuperação das informações de uma forma mais adequada e completa, isso porque necessita-se de uma

maior integração com fontes de dados e ambientes informacionais heterogêneos.

Uma forma de auxílio em sua descrição, é a proposta de um conjunto de práticas denominado *Linked Data*, que consiste na “interligação de dados representados com base em um formato padrão, partindo da utilização de links que permitem relacionar dados advindos de diversas fontes” (JESUS; CASTRO; RAMALHO, 2021, p. 2).

Nesta perspectiva, o objetivo geral deste trabalho consiste em identificar estudos que abordam a catalogação de pinturas, no cenário *Linked Data*. A pesquisa dedica-se a apresentar os conceitos e as especificidades da pintura, analisar e apresentar o estado da arte sobre a catalogação de obras de arte, com destaque para a pintura.

Desta forma, questiona-se sobre como a descrição de pinturas está sendo discutida na literatura científica da comunidade da Ciência da Informação, levando em conta que, as pinturas devidamente catalogadas, organizadas e padronizadas no contexto tecnológico vigente, em sincronia com o *Linked Data*, podem formar um rico corpo de conhecimento para as instituições como bibliotecas, arquivos e museus.

2 Fundamentação conceitual

O acúmulo de informações diante das tecnologias, traz como consequência um volume desorganizado e caótico de dados em meios digitais, necessitando de uma organização centrada na recuperação efetiva por usuários e máquinas.

Os dados vêm se multiplicando à medida que o desenvolvimento das tecnologias permite maiores possibilidades de compartilhamento e utilização de diversos tipos de conteúdo nos ambientes digitais (TRIQUES; ARAKAKI, 2021, p. 186).

Ao verificar a necessidade de padrões e a utilização de ligação entre dados para uma melhor recuperação no futuro, percebe-se a possibilidade de *links* entre objetos e a utilização de filtros, uma estratégia muito útil para a representação de recursos em ambientes digitais.

Nesse cenário, o *Linked Data* (LD), ou dados conectados, proposto por Tim Berners-Lee (2006), tem como objetivo principal a estruturação e a publicação de dados na *Web*, via *hiperlinks*. Segundo Homburg *et al.* (2021, p. 2, tradução nossa), o LD auxilia no “compartilhamento de definições de metadados que podem ser consultados na nuvem para futuras comparações”.

Segundo Nhacuongue, Rozsa e Dutra (2018, p. 24), o LD é considerado “uma extensão da *Web* para a conexão de dados globais sobre pessoas, organizações, livros, publicações científicas, filmes, músicas, programas de televisão, ensaios clínicos, etc.”

Entretanto, existem outras denominações semelhantes com *Linked Data* na literatura científica, como é o caso do *Linked Open Data* (LOD). Silva (2013) apresentou a diferença entre LD e LOD e, conforme o autor, ao serem verificadas as necessidades de dados abertos, a utilização de *Open Data*, dados para uso livre, foi possível a criação do LOD, focado nesses dados na *Web* para uma ligação mais efetiva.

Santarém Segundo (2015, p. 225) explica que o LOD nada mais é que “[...] um conjunto de melhores práticas para publicação e conexão de dados estruturados na *Web*” formando assim um único conjunto de dados de maneira global.

Portanto, LOD pode ser considerado dados que são mais acessíveis que o LD e mais interligáveis na *Web*, utilizando como base os princípios do *Open Data*, facilitando a interoperabilidade e, também, uma maior acessibilidade por parte dos usuários, tanto os que manuseiam dados quanto aqueles que recebem os dados diante de uma busca na *Web*.

A base para uma catalogação LD requer, como uma das primeiras decisões, uma infraestrutura para publicação e dados sobre obras de arte. Dessa forma, espera-se que se obtenham dados voltados aos patrimônios de alguma base que já realizou a catalogação e a descrição desses, para que assim seja possível um reaproveitamento dos dados e, por fim um melhoramento desses (DIJKSHOORN *et al.*, 2018).

Sob esta ótica, quanto mais empregado no ambiente artístico o LD, a catalogação de obras de arte será mais eficiente, permitindo um reaproveitamento e melhoramento dos dados obtidos.

2.1 Metadados e padrões de metadados

Para Haynes (2018, p. 224, tradução nossa), os metadados são a descrição de objetos informacionais, podendo ser tanto dados brutos ou descritivos sobre determinado indivíduo e, além disso, os metadados constituem uma forma de tornar a informação mais acessível, garantindo seu acesso de maneira ampla. Portanto, “a apresentação eficaz de metadados melhora sua usabilidade”.

Ainda, descrever e padronizar os metadados é fundamental para a documentação, criação, organização e gestão, além de questões referentes à preservação digital, pois “[...] produzir uma boa documentação significa principalmente atribuir metadados de qualidade aos dados” (SAYÃO; SALES, 2015, p. 28). Entretanto, existem algumas exigências em relação ao nível de descrição, como também a atribuição de metadados (SAYÃO; SALES, 2015).

Greenberg (2017) apresenta metadados como aqueles que descrevem e dão contexto, fundamentais para interpretação. Ao tornar o metadado acessível permite

ao usuário uma interpretação entre objeto e seus dados conforme a atividade realizada.

No caso de Buckland (2017, p. 113, tradução nossa), o autor explica que, “o uso de formatos padronizados para armazenamento e exibição torna o uso de metadados mais fácil”, ou seja, usar padrões em vocabulários controlados “tem a vantagem de consistência e ajuda na compreensão”. Portanto, para o autor, ao utilizar padrões para metadados, facilita-se o uso desses, como os vocabulários controlados, que auxilia na compreensão do recurso informacional descrito (BUCKLAND, 2017, tradução nossa).

Já para Triques e Arakaki (2021, p. 188), a padronização dos metadados é uma

garantia da construção de uma linguagem comum e compartilhada entre os componentes da Web, permitindo explicitar o sentido agregado aos dados, e assim, estabelecer as ligações de significado entre eles de tal forma que as máquinas tenham maior capacidade de processar e trocar informações semanticamente.

Por fim, Silva *et al.* (2018, p. 55) argumentam que, as “pesquisas na área da anotação semântica são relevantes para solucionar problemas de busca, de localização e de recuperação da informação” e, assim, para que a padronização de metadados em conjunto com anotações semânticas sejam adequadas para gestão de acervo necessita-se: da acessibilidade por usuários, da continuidade dos dados no futuro, da construção de uma linguagem em comum na Web para os dados, da caracterização de um objeto tornando seu conteúdo acessível sem a necessidade da consulta do original e da interoperabilidade dos dados em ambiente Web (TRIQUES; ARAKAKI, 2021; LIMA; SANTOS; SANTARÉM SEGUNDO, 2016; CASTRO; SANTOS, 2014; NISO, 2004; MILSTEAD; FELDMAN, 1999).

2.2 Catalogação Descritiva

A catalogação descritiva, é “responsável pela manipulação de dados bibliográficos no ambiente informacional e como disciplina determinante pelo processo de representação documental”, isto é, a catalogação descritiva é considerada uma das bases para uma eficiente representação de recursos informacionais, podendo ser repensada para que possam ser fortalecidas de acordo com o contexto que se encontram (CASTRO, 2020, p. 108).

Portanto, a Catalogação Descritiva se aperfeiçoa no desenvolvimento de códigos, de regras de catalogação e nos padrões de descrição e, também, procura se desenvolver perante os formatos de intercâmbio e modelos conceituais, sempre visando o contexto tecnológico que está inserida (CASTRO, 2020, p. 116).

Ao catalogar, o bibliotecário responsável por essa prática na instituição em que atua, deve realizar uma leitura prévia da obra, seja um objeto inanimado, um

livro ou uma ilustração, assim “os catalogadores revisam as atribuições contraditórias e selecionam as mais confiáveis de acordo com uma série de critérios considerados confiáveis na comunidade (por exemplo, bibliografia sobre o tema, análise estilística)” (DAQUINO, 2020, p. 757, tradução nossa).

Um código de catalogação que descreve uma diversidade de materiais e possui um conjunto de regras para descrição bibliográfica e pontos de acesso, o Código de Catalogação Anglo-Americano, revisão de 2002, ou AACR2r (MEY; SILVEIRA, 2009). O AACR2r possui uma estrutura específica, na primeira parte está a descrição, na segunda parte encontram-se os pontos de acesso, títulos uniformes e remissivas e, nos apêndices podemos localizar desde o uso de maiúsculas e minúsculas até o uso dos artigos iniciais.

Infelizmente, o código AACR2r “não se apresenta adaptado aos recursos digitais e às tecnologias de informação” do cenário vigente (MODESTO, 2010, p. 1), o que o torna obsoleto para atender às exigências da descrição em ambientes digitais. Já o *Resource Description and Access* (RDA) contempla a descrição de toda tipologia de documentos e se adequa ao ambiente digital, em relação aos recursos digitais (SILVA, 2013, p. 9).

2.3 Pinturas

Para uma melhor compreensão do termo utilizado como base neste trabalho, entende-se que existe a necessidade da definição por meio de um vocabulário controlado no domínio das Artes. A pesquisa no vocabulário Getty, que fornece o Tesouro de Arte e Arquitetura Online (*Art & Architecture Thesaurus Online - ATT*), auxiliou na identificação de um termo comum para o tema “pintura”.

Identificou-se, que “painting” é o termo para o domínio das Artes. Segundo Getty (c2004), a pintura é a criação de imagens, processos e técnicas de criação de imagens, isto é, a arte e a prática de aplicação de pigmentos suspensos sobre uma superfície.

Foram selecionados 4 dicionários e 4, dos quais 2 dicionários e 2 enciclopédias são em português (Priberam, Michaelis, Wikipédia e Britannica Escola), e 2 dicionários e 2 enciclopédias são em inglês (*Cambridge University, Merriam-Webster, Wikipedia e Britannica*).

Quanto aos dicionários, mais especificamente no dicionário Priberam (c2021), a pintura é a “arte de pintar”. No dicionário Michaelis (c2021) a pintura é “ato ou efeito de pintar(-se)”. Já no *Cambridge Dictionary* (2021) pintura, é a cobertura de algo com tinta. Quanto ao dicionário *Merriam-Webster* (c2021), a pintura nada mais é que a aplicação de “cor, pigmento ou tinta”.

Para a Wikipédia (c2021), em português, a pintura é uma técnica, em que se aplica pigmento, em suas diferentes

formas, com o objetivo de colorir. Já para Britannica Escola (c2021) a própria pintura é a arte, de uma forma em que se criam imagens com cores, formas, texturas. No caso da *Wikipedia* (c2021), em inglês, a pintura é a prática de aplicar tintas, pigmentos, cores, sobre superfície sólida. Já na *Britannica* (c2021), em inglês, a pintura é a “expressão de ideias e emoções” com qualidades estéticas.

A pintura possui uma pluralidade de conceitos, desde uma generalização de termos, como “pintar algo”, até uma especificação sobre “aplicação de tintas/pigmentos”. Considerando tais resultados, foi definido como base para este estudo a pintura como aplicação de pigmento a uma superfície.

Marcondes (1998), em seu Dicionário de Termos Artísticos, clássico no domínio das Artes, traz sua definição a partir do termo “pintura” como uma técnica ou processo na aplicação de tintas sobre superfície. Para o autor, a pintura é “capaz de produzir uma imagem” e se diferencia assim então das outras artes plásticas. Por fim, para ele, a pintura depende especificamente do “emprego da cor” (MARCONDES, 1998, p. 229).

Outro autor que apresenta o conceito do termo “pintura” é Wölfflin (2006). De acordo com o autor, a pintura se baseia numa “denominação tradicional” de um único atributo: a partir dos seus pigmentos, essa tem como “objetivo básico a criação de superfícies”, o que para ele é o que a distingue do desenho. Logo, Wölfflin (2006, p. 57) argumenta que obras de arte são organismos e, como tal, são inevitáveis, projetando exatamente como são, como é o caso das pinturas, por exemplo.

Costa (2009), em seu texto “O que é ‘arte’?”, apresenta o termo pintura como aquela que “imita a natureza”, ou seja, de acordo com o autor, a pintura representa as práticas humanas no meio ambiente. Já Maximiano (2013), afirma que a pintura possui determinadas características e, uma delas é a “aplicação de tintas de diversos tipos sobre uma determinada superfície”, isto é, a tinta aplicada em algum suporte com o intuito de “transmitir uma informação, um sentimento, uma experiência”, é mais do que apenas o uso de instrumentos (MAXIMIANO, 2013, p. 29).

Em relação à conceituação do termo na área da Ciência da Informação, essa é explicada como a representação “para designar cenas, pessoas e objetos tridimensionais em um espaço bidimensional” (AGUSTIN LACRUZ, 2006 *apud* MAIMONE; TÁLAMO, 2008, p. 5).

Segundo o AACR2r (2002), no Apêndice D-9, original de arte (*art original*), é

obra de arte original, em duas ou três dimensões, (que não seja gravura, q.v., ou fotografia) criada pelo artista (p.ex. uma pintura, um desenho ou uma escultura em contraposição à reprodução de sua pintura, desenho ou escultura).

Já para o RDA (2022, tradução nossa), que possui uma definição expressa sobre pintura, essa é considerada uma unidade de extensão que consiste em um item no

qual as imagens são formadas principalmente pela aplicação direta de pigmentos suspensos em um meio, dispostos em massas de cores sobre uma superfície geralmente bidimensional. Este termo se aplica a imagens estáticas.

Ainda, a pintura é uma forma de manifestação, técnica, habilidade que os seres humanos praticam e que essas estão ligadas à estética, por meio da “expressão de idéias, percepções, sentimentos etc., e provocam sentimentos na pessoa que afrui” e, tradicionalmente, “estes produtos são as esculturas e as pinturas” (RODRIGUES; CRIPPA, 2009, p. 3).

Como forma de representação da realidade, de acordo com Klenczon e Rygiel (2014), a pintura há muito tempo é adotada como uma manifestação artística das pessoas e suas histórias, culturas, tradições, interesses, costumes, entre outros; é, também, produzida a partir de pontos de vistas específicos, podendo ser interpretada de formas variadas, por diferentes grupos de indivíduos, com suas necessidades e competências.

A informação transmitida a partir de uma pintura depende também, como explicam Navarrete e Owen (2016), da leitura pelo observador, com regras e métodos de interpretação. Assim, “uma pintura pode ser ‘lida’ de maneira diferente por um pintor (observando a cor e a pincelada), um historiador da arte (determinando o valor cultural e histórico) e um químico (inspeccionando a composição mineral)” (NAVARRETE; OWEN, 2016, p. 112, tradução nossa).

Costa e Almeida (2010) afirmam que, ao descrever corretamente obras de arte, isto influenciará em grande parte na qualidade da pesquisa e na compreensão de coleções. Já Klenczon e Rygiel (2014) abordam as características de uma imagem e consideram que essas podem até contribuir para a compreensão da obra num todo. Alguns recursos, como título, autor, aparência, podem ser facilmente identificados e usados na descrição como pontos de acesso, mas ainda assim não a definem, pois, seu significado é subjetivo para cada indivíduo, exigindo muitas vezes de uma abordagem diferente.

Ao garantir a digitalização das obras, no caso de pinturas, para que assim possam ser reutilizadas, seguindo modelos de catalogação, indexação e classificação, permite-se possível o acesso eficiente às pinturas, à informação e ao conhecimento que elas representam (KLENCZON; RYGIEL, 2014, p. 43).

3 Procedimentos metodológicos

Este trabalho caracteriza-se por uma pesquisa descritiva e de natureza exploratória, com a finalidade de

proporcionar a construção de um referencial teórico e metodológico sobre a catalogação de pinturas no cenário *Linked Data*.

Optou-se pela Revisão Sistemática de Literatura (RSL), a fim de identificar o estado da arte sobre a catalogação de pinturas e sua relação com o *Linked Data*, por ser um método relevante, que garante a consistência dos resultados e sua capacidade de reprodução (SAMPAIO; MANCINI, 2007) [1].

Para Galvão e Ricarte (2019, p. 57), a Revisão Sistemática “é uma modalidade de pesquisa, que segue protocolos específicos e busca dar alguma logicidade a um grande corpus documental”; além de possuir “etapas e critérios rigorosos” que garantirão a “qualidade e a reprodutibilidade dos documentos” (CERRAO; CASTRO; JESUS, 2018, p. 109).

Primeiramente, na pesquisa exploratória e empírica foram analisados os *sites* de projetos e de iniciativas de catalogação de pinturas e sua relação com o *Linked Data*, em busca de documentos e modelos de descrição que apresentam um ambiente informacional de obras de arte/pinturas com características *Linked Data*.

Em março de 2022 foram realizadas as buscas nas bases de dados, com 25 documentos recuperados. Na etapa de seleção verificou-se a existência de 2 documentos duplicados (8%), 10 documentos aceitos (40%) e 13 documentos rejeitados (52%). Já na etapa de execução, após as leituras dos documentos recuperados, foram excluídos 3 documentos (30%) e foram aceitos 7 documentos (70%) para análise.

Assim, para possibilitar a análise dos resultados da Revisão Sistemática foram estabelecidas 3 categorias de análise, que atuaram como um roteiro para a análise individual dos documentos, auxiliando na extração de dados: enfoque dos documentos; exemplos e características de projetos, de iniciativas e de modelos de descrição de obras de arte à luz do *Linked Data*; e, padrões de descrição utilizados na catalogação de pinturas.

Após a análise individual dos artigos recuperados, identificou-se, que grande parte dos documentos aceitos, após a etapa de execução, trazem estudos sobre modelos de ontologias para a descrição de patrimônios culturais. Durante a análise, como um dos campos de extração de dados, foram identificados os modelos de descrição utilizados para a catalogação de obras de arte que serão apresentados a seguir.

4 Resultados

O Código de Catalogação Anglo-Americano, revisão de 2002, ou AACR2r, no Brasil, é considerado, como comentado anteriormente, um código de catalogação que descreve uma diversidade de materiais e possui um

conjunto de regras para descrição bibliográfica e pontos de acesso.

Entretanto, o AACR2r, por mais que tenha como objetivo ser aplicado a toda tipologia de materiais, este não atende plenamente aos recursos digitais (SILVA, 2013) e, para atender às necessidades de descrição no ambiente digital, surge o RDA, como um possível sucessor do código de catalogação AACR2r, desenvolvido pelo *Committee for Development of RDA*.

O *Resource Description and Access (RDA)*, segundo a *International Federation of Library Associations and Institutions (IFLA, 2009)*, vem como uma forma de atender às necessidades da descrição de recursos de todas as tipologias, isso porque o RDA possui diretrizes mais eficientes, além de compatíveis com outros padrões e modelos internacionais que se adequam aos ambientes informacionais digitais contemporaneamente.

Assim, o RDA, conforme observado, é de suma importância para descrição de recursos informacionais no cenário atual e apresenta-se compatível ao *Linked Data*, tanto pela sua versatilidade, quanto por trazer a continuação de documentos passados, tornando-os compatíveis com as tecnologias vigentes.

Já o VRA Core, por exemplo, de acordo com seu *website* oficial, é “um padrão de dados para a descrição de obras de cultura visual, bem como das imagens que as documentam”, em que é hospedado por “*Network Development* e pelo *MARC Standards Office da Library of Congress (LC)*” (VRA CORE, 2022).

Em um de seus tutoriais, na introdução a respeito do padrão, é apresentada a finalidade e funções do VRA Core. Explica que a descrição de obras de cultura visual inclui: “objetos ou eventos como pinturas, desenhos, escultura, arquitetura, fotografias, assim como livros, artes decorativas e performáticas” (VRA CORE Introduction, 2014).

Compreende-se também que, ao seguir as recomendações e os protocolos internacionais, torna-se mais fácil a interoperabilidade entre máquinas e sistemas, podendo, assim, corroborar com o cenário *Linked Data* e a catalogação de pinturas, por exemplo, como no Modelo de Referência Conceitual CIDOC (CIDOC-CRM).

Já o CIDOC-CRM foi criado a partir do grupo CIDOC *Documentation Standards Group*, vinculado ao Comitê Internacional de Documentação do Conselho Internacional de Museus. Até 1994, segundo StringFixer ([20--]), o CIDOC era focado apenas em informações presentes em museus e, a partir de 1996, trouxe algumas mudanças metodológicas, o que resultou no “primeiro Modelo de Referência Conceitual CIDOC (CRM)” em 1999 (STRINGFIXER, [20--]).

Segundo Reis, Castro e Fujita (2023, p. 2),

[...] em execução, o modelo conceitual abarca as diferentes tipologias de itens museológicos e suas especificidades, abordando domínios artístico-culturais, arqueológicos e históricos, e podendo integrar o contexto do objeto na esfera do simbólico.

Para Dijkshoorn *et al.* (2018, tradução nossa), o *Europeana Data Model*, além de ser um modelo, é considerado uma ontologia para instituições de patrimônio que consigam estruturar dados, desenvolvido para utilizar outros modelos que trabalhem com dados mais exclusivos de instituições e unidades de informação, como é o caso de bibliotecas, arquivos e museus.

Os conceitos que são utilizados como base para descrição são agrupados de forma lógica e, “os grupos têm a ver com as noções de participação, estrutura, localização, avaliação e identificação, finalidade, motivação, uso, e assim por diante (EDM DEFINITION, 2017, tradução nossa).

Em relação às diretrizes para a descrição de obras de arte, tem-se o *Categories for the Description of Works of Art* (CDWA), esse se apresenta como “um conjunto de diretrizes para a descrição de arte, arquitetura e outras obras culturais”, ou seja, possui o foco totalmente em descrição de arte e cultura, mais específico que o CIDOC e Europeana, que são focados no patrimônio cultural como um todo (GETTY, 2022, tradução nossa).

Como explica o *site* do CDWA, que é mantido pelo Getty (2022) e, portanto, no CDWA um pequeno subconjunto de categorias é centralizado para representar informações que são consideradas mínimas e essenciais para descrever e individualizar uma obra, incluindo discussões e diretrizes para catalogação mínimas, além de exemplos (GETTY, 2022, tradução nossa).

Outro conjunto de diretrizes para a catalogação de objetos culturais foi criado com foco em padrões de dados de conteúdo para a catalogação descritiva, na sua base os metadados descritivos e o controle de autoridade, o *Cataloging Cultural Objects* (CCO), que abrange muitos tipos de obras culturais, incluindo arquitetura, pinturas, escultura, gravuras, manuscritos, fotografias e outras mídias visuais, performance arte, sítios arqueológicos e artefatos, e vários objetos funcionais do reino da cultura material. (ALA, 2007, p. XIII, tradução nossa).

Além disso, o manual auxilia na descrição, na documentação e na catalogação de artefatos e mídias visuais, projetado para coleções de museus, coleções de recursos, arquivos e bibliotecas com ênfase primária em arte e arquitetura. Quanto ao *Resource Description Framework* (RDF) e, de acordo com a W3C (2014, tradução nossa), este é um “modelo padrão para intercâmbio de dados na Web”.

Observa-se que o RDF, foi criado com o objetivo de ser uma “arquitetura genérica de metadados” para

representação de recursos na Web, como por exemplo título, autor, data de atualização, mas também é considerado uma linguagem de marcação, isto é, para projetar declarações em triplas, que possuem propriedades e valores (LIMA; CARVALHO, 2005, p. 2).

Ainda, a linguagem RDF estende uma estrutura de *links Web*, para nomear relacionamento entre dados, de forma que sua sintaxe é baseada em XML, registrando e tornando interoperáveis os dados processados por máquinas com o uso de URI (LIMA; CARVALHO, 2005). Já o esquema XML, um protocolo de internet, recomendado pelo W3C, para descrição de classes e propriedades, possui um conjunto de classes para catalogação de objetos culturais (CARRASCO, 2019).

O protocolo SPARQL (*Protocol and RDF Query Language*), consiste em “uma linguagem de consulta para bases de dados”, o qual permite que dados possam ser consultados nessas bases para o recolhimento de informações, que são armazenadas no formato RDF (UTA, 2023).

Algumas iniciativas e modelos foram também recuperados pela RSL, dentre elas, o *International Committee for Documentation* (CIDOC) - Conceptual Reference Model (CRM) (CIDOC-CRM) que é o mais citado na literatura científica, segundo os resultados da RSL. Quanto ao Europeana Data Model (EDM), um modelo de dados da Europeana, esse é utilizado como base apenas para 1 documento dentre os aceitos.

Dentre outros exemplos de iniciativas, o *Rijksmuseum* (RIJKSMUSEUM, [20--], tradução nossa), museu da Holanda que, de acordo com seu *site*, conta a história de 800 anos holandesa, desde 1200 até os dias de hoje e, também, organizam “várias exposições por ano a partir da própria coleção e com empréstimos (inter)nacionais”.

Já o *ARIADNE Reference Model* (ARIADNE..., c2012, tradução nossa) teve como maior desafio “desenvolver um esquema global e extensível na forma de uma ontologia formal que permita a integração sem perda de significado. CIDOC CRM foi escolhido como a espinha dorsal do Modelo de Referência ARIADNE e um conjunto de extensões foi desenvolvido para lidar com a complexidade da integração de dados arqueológicos”.

Quanto ao *CrossCult Knowledge Base* (CROSSCULT..., 2020, tradução nossa), projeto europeu, foi planejado para possibilitar para “a reutilização do patrimônio cultural digital em cortes transversais sensíveis ao contexto da história europeia” e visa estimular uma mudança na forma como os cidadãos europeus avaliam a história” e “promoverá a reinterpretação do que os cidadãos podem ter aprendido à luz das interconexões transfronteiriças entre peças do patrimônio cultural, pontos de vista de outros cidadãos e espaços físicos”.

O caso do *Google Images*, da empresa Google, de acordo com a Wikipédia (GOOGLE, 2022) “é um serviço de pesquisa de propriedade do Google que permite aos usuários pesquisar por conteúdo de imagem na *World Wide Web*. Foi introduzido em 12 de julho de 2001, devido a uma demanda por fotos do vestido verde Versace de Jennifer Lopez que a pesquisa regular do Google não conseguiu lidar”.

O *Knowledge Graphs* (KG), também utilizado pela empresa Google, é comentado com frequência pelos autores e, para eles os “KGs podem ser definidos como uma combinação de ontologias que modelam o domínio de interesse e dados materializados publicados no formato de dados abertos conectados (LOD)” (MONACO *et al.*, 2022, p. 127, tradução nossa).

No caso do *Raphael Research Resource* (RAPHAEL..., [20--], tradução nossa), a partir de um estudo das pinturas do pintor Raphael, foi reunido neste projeto uma “coleção excepcional de material ao longo dos 500 anos desde sua criação no século XVI, abrangendo sua história e proveniência, bem como os materiais e técnicas utilizados em sua confecção”.

O Sistema de Museus (*The Museum System - TMS*), de acordo com o site oficial (PAQUA; NULLMAN; CURRAN, c2018, tradução nossa), “é um aplicativo de banco de dados relacional criado especificamente para gerenciamento de coleções de museus. O TMS foi projetado para ser eficaz com uma ampla variedade de tipos de museus, bem como com coleções individuais heterogêneas”.

The World of Art on the Semantic Web (CLAROS) - O Mundo da Arte na Web Semântica, “é uma federação internacional de pesquisa interdisciplinar que utiliza os mais recentes desenvolvimentos em Tecnologias de Informação e Comunicação para levar a arte do mundo a todos. É liderado pela Universidade de Oxford e sediado em seu centro de pesquisa eletrônica, OeRC.” (OXFORD, 2011, tradução nossa).

Por fim, o *Villa I Tatti – Berenson Library*, é um dos casos pouco comentados. Igualmente como o projeto *Raphael Research Resource*, esta iniciativa não foi aprofundada por Daquino (2020), apenas há uma breve apresentação sobre a *Villa I Tatti – Berenson Library*, por ser um dos 3 acervos fotográficos históricos mais representativos da arte.

5 Conclusão

Para responder à questão de pesquisa, isto é, como a descrição de pinturas está sendo discutida na literatura científica da comunidade da Ciência da Informação, foi realizada uma Revisão Sistemática de Literatura (RSL) com o intuito de verificar o estado da arte da catalogação de pinturas no cenário *Linked Data*.

Foi possível a construção de um estado da arte sobre o tema, no repasse de um referencial teórico e metodológico rigoroso e explícito sobre a catalogação de pinturas no cenário *Linked Data*.

Quanto aos objetivos específicos, foram apresentados os conceitos e as especificidades da pintura, considerando o ponto de vista do domínio das Artes e da Ciência da Informação.

Quanto ao impacto científico em questão, o tema é de certa maneira novo para a área da Ciência da Informação, ao identificar uma lacuna no estudo mais aprofundado da temática em apreço. A utilização de metadados e dos padrões de metadados, pode auxiliar quanto a questões de descrição e interoperabilidade entre ambientes informacionais heterogêneos, garantindo o acesso e a preservação de pinturas futuramente, além do uso e reuso.

Percebe-se também, que é preciso levar em conta a utilidade, a qualidade, a confiabilidade, a integridade, a precisão, a consistência e os vocabulários controlados, pois estes aumentam a potencialidade na interoperabilidade, tanto na catalogação descritiva quanto na produção e na aplicação de metadados.

Tendo isso em vista, percebe-se, que ao utilizar padrões de metadados e *links* entre objetos, além de filtros, possibilita-se uma melhora na organização, na recuperação e na descrição de dados em ambientes *Web*, portanto o *Linked Data* facilita o reuso de dados, evitando retrabalho ao tornar dados legíveis para usuários e máquinas, permitindo o compartilhamento e a conexão entre recursos.

A não aplicação do *Linked Data* resulta em algumas consequências, por exemplo, a falta de identificação e de recuperação de recursos de maneira mais precisa. Quanto mais o emprego do *Linked Data* no ambiente artístico, a catalogação de pinturas será mais adequada e completa, podendo fazer um reaproveitamento e melhoramento dos dados obtidos.

Assim, o RDA, é de suma importância para descrição de recursos informacionais no cenário atual, pois apresenta compatibilidade ao *Linked Data* por causa de sua versatilidade, permitindo a reutilização de dados e metadados, acompanhando as tecnologias vigentes.

Portanto, essa pesquisa se apresenta como um ponto de partida para os estudos que levem em consideração a descrição de obras de arte, em especial, as pinturas, a serem representadas utilizando-se das tecnologias *Linked Data*, contribuindo também para os *sites* e catálogos de museus, galerias, artistas e sistemas de informação no domínio das Artes e permitindo novos olhares para pesquisas inovadoras no campo da Ciência da Informação.

Notas

[1] Os dados de pesquisa, assim como o protocolo de busca da RSL estão disponíveis no Repositório Institucional da Universidade Federal de São Carlos (UFSCar).

Agradecemos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro recebido para o desenvolvimento dessa pesquisa.

Referências

- AMERICAN LIBRARY ASSOCIATION (ALA). **Introduction to cataloging cultural objects**. 2007. Online. Disponível em: <http://www.ala.org/alacts/events/ac/2006/ccbootcamp>. Acesso em: 8 mar. 2023.
- ARIADNE REFERENCE MODEL. [European Union]: European Commission's 7th Framework Programme, c2012. Disponível em: <http://legacy.ariadne-infrastructure.eu/resources-2/ariadne-reference-model/>. Acesso em: 8 mar. 2023.
- BERNERS-LEE, Tim. **Linked Data: design issues**. Online. 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 8 mar. 2023.
- BUCKLAND, Michael. **Information and society**. London: MIT Press, 2017. 234 p. Disponível em: <https://muse.jhu.edu/book/51614>. Acesso em: 8 mar. 2023.
- CARRASCO, Laís Barbudo. **Integração de conteúdos culturais heterogêneos em ambientes digitais do patrimônio cultural: harmonização de modelos conceituais**. Tese (doutorado) – Universidade Estadual Paulista (UNESP), Faculdade de Filosofia e Ciências, Marília, São Paulo. Disponível em: <http://hdl.handle.net/11449/182113>. Acesso em: 29 maio 2023.
- CASTRO, Fabiano Ferreira de. **Catálogo Descritiva: necessidade de revisão em suas bases teórico-conceituais? Informação & Informação**, Londrina, v. 25, p. 107-134, jul./set. 2020. Disponível em: https://www.researchgate.net/publication/347878552_Catalogo_Descritiva_necessidade_de_revisao_em_suas_bases_teorico-conceituais. Acesso em: 8 mar. 2023.
- CASTRO, Fabiano Ferreira de; SANTOS, Plácida L. V. A. C. **Elementos de interoperabilidade na perspectiva da catalogação descritiva. Informação e Sociedade: Estudos**, Campina Grande, v. 24, n. 3, p. 13-25, 2014. Disponível em: <http://hdl.handle.net/11449/130230>. Acesso em: 29 maio 2023.
- CERRAO, Natalia Gallo; CASTRO, Fabiano Ferreira de; JESUS, Ananda Fernanda de. **O método de revisão sistemática da literatura (RS) na área da Ciência da Informação no Brasil: análise de dados de pesquisa. Informação & Tecnologia (ITEC)**, Marília/João Pessoa, v. 5, n. 1, jan./jun. 2018.
- CÓDIGO DE CATALOGAÇÃO ANGLO-AMERICANO (AACR2r). **Original de arte / art original**. Preparado sob a direção do Joint Steering Committee for Revision of AACR. 2. ed. rev. 2002. Tradução para a língua portuguesa sob a responsabilidade da FEBAB. São Paulo: FEBAB/Imprensa Oficial do Estado de São Paulo, 2004. p. 664.
- COSTA, Claudio F. **O que é 'arte'? Dossiê Walter Benjmin**, [S. l.], v. 4, n. 6, 2009. Disponível em: <https://periodicos.ufop.br/raf/article/view/706>. Acesso em: 8 mar. 2023.
- COSTA, Ivani Di Grazia; ALMEIDA, Maria Christina Barbosa de. **Vocabulário de arte: ferramentas fundamentais no trabalho cooperativo em bibliotecas, museus e arquivos. In: SEMINÁRIO SERVIÇOS DE INFORMAÇÃO EM MUSEUS, 1., 2010, São Paulo. Anais [...]** São Paulo: 2010. p. 89-101. Disponível em: <http://biblioteca.pinacoteca.org.br:9090/bases/biblioteca/322829.pdf>. Acesso em: 8 mar. 2023.
- CROSSCULT: Where History meets IT. **European Union: European Union's Horizon, 2020**. Disponível em: <https://www.crosscult.lu/>. Acesso em: 8 mar. 2023.
- DAQUINO, Marilena. **A computational analysis of art historical Linked Data for assessing authoritativeness of attributions. Journal of the Association for Information Science and Technology**, [S.l.], v. 71, n. 7, p. 757-769, 2020. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/asi.24301>. Acesso em: 8 mar. 2023.
- DIJKSHOORN, Chris; AROYO, Lora; van OSSENBRUGGGEN, Jacco; SCHREIBER, Guus. **Modeling cultural heritage data for online publication. Applied Ontology | IOS Press**, [S. l.], v. 13, n. 4, 2018. Disponível em: <https://content.iospress.com/articles/applied-ontology/ao201>. Acesso em: 8 mar. 2023.
- EDM DEFINITION. **Definition of the Europeana Data Model, version 5.2.8**. Online. 2017. Disponível em: <https://pro.europeana.eu/page/edm-documentation>. Acesso em: 8 mar. 2023.
- GALVÃO, Maria Cristiane Barbosa; RICARTE, Ivan Luiz Marques. **Revisão sistemática da literatura: conceituação, produção e publicação. Logeion: Filosofia da Informação**, Rio de Janeiro, v. 6, n. 1, p. 57-73, 2019. Disponível em: <https://revista.ibict.br/fiinf/article/view/4835>. Acesso em: 8 mar. 2023.
- GETTY. **Categories for the Description of Works of Art (CDWA)**. Online. c2004. Disponível em: https://www.getty.edu/vow/AATFullDisplay?find=painting&logic=AND¬e=&english=N&prev_page=5&subjectid=300054216. Acesso em: 29 maio 2023.
- GETTY. **Categories for the Description of Works of Art (CDWA)**. Online. 2022. Disponível em: https://www.getty.edu/research/publications/electronic_publications/cdwa/. Acesso em: 8 mar. 2023.
- GOOGLE. **Google Imagens**. [S. l.]: Wikipédia, 2022. Disponível em: https://pt.wikipedia.org/wiki/Google_Imagens. Acesso em: 8 mar. 2023.
- GREENBERG, Jane. **Big metadata, smart metadata, and metadata capital: toward greater synergy between data science and metadata. Journal of Data and Information Science**, Philadelphia, v. 2, n. 3, p. 19-36, jul. 2017.

Disponível em: <https://sciendo.com/article/10.1515/jdis-2017-0012>. Acesso em: 8 mar. 2023.

HAYNES, David. **Metadata for information management and retrieval: understanding metadata and its use**. [S. l.]: Facet Publishing, 2018.

HOMBURG, Timo; CRAMER, Anja; RADDATZ, Laura; MARA, Hubert. Metadata schema and ontology for capturing and processing of 3D cultural heritage objects. **Herit Sci**, [S. l.], v. 9, n. 91, 2021. Disponível em: <https://heritagesciencejournal.springeropen.com/articles/10.1186/s40494-021-00561-w>. Acesso em: 8 mar. 2023.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **International implementation of RDA**. Netherlands: IFLA, 2009. Disponível em: <https://www.ifla.org/news/international-implementation-of-rda/>. Acesso em: 8 mar. 2023.

JESUS, Ananda Fernanda de; CASTRO, Fabiano Ferreira de; RAMALHO, Rogério Aparecido Sá. O papel das bibliotecas no *Linked Data*. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 26, p. 1-21, 2021. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/75909>. Acesso em: 8 mar. 2023.

KLENCZON, Wanda; RYGIEL, Paweł. Librarian cornered by images, or how to index visual resources. **Cataloging & Classification Quarterly**, Warsaw, Poland, v. 52, n. 1, p. 42-61, 2014. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01639374.2013.848123?journalCode=wccq20>. Acesso em: 8 mar. 2023.

LIMA, Fábio Rogério Batista; SANTOS, Plácida Leopoldina V. A. C.; SANTARÉM SEGUNDO, José Eduardo. Padrão de metadados no domínio museológico. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 21, n. 03, jul./set. 2016. Disponível em: <https://www.scielo.br/j/pci/a/SrN8qVHNnkvYgTBBSvmdJ/s/?lang=pt>. Acesso em: 8 mar. 2023.

LIMA, Júnior César de; CARVALHO, Cedric L. de. **Resource Description Framework (RDF)**. Goiás: INFUFG, 2005. 28 p. Relatório Técnico (RT-INF_003-05). Disponível em: https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_003-05.pdf. Acesso em: 8 mar. 2023.

MAIMONE, Giovana Deliberali; TÁLAMO, Maria de Fátima Gonçalves Moreira. Tratamento informacional de imagens artístico-pictóricas no contexto da ciência da informação. **DataGramZero**, [s. l.], v. 9, n. 2, 2008. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/6242>. Acesso em: 8 mar. 2023.

MARCONDES, Luiz Fernando. **Dicionário de termos artísticos**: com equivalências em inglês, espanhol e francês. Rio de Janeiro: Pinakothek, 1998.

MAXIMIANO, Angélica Rodrigues Marques. **A prática da pintura no ensino fundamental**: relato de uma experiência na E. E. Maria de Magalhães. 2013. 54 f. Monografia (Especialização em Ensino de Artes Visuais) – Programa de Pós-Graduação em Artes, Escola de Belas Artes, Universidade Federal de Minas Gerais, Belo Horizonte, 2013. Disponível em:

<https://repositorio.ufmg.br/handle/1843/BUOS-9J2JSX>. Acesso em: 8 mar. 2023.

MEY, Eliane S. A.; SILVEIRA, Naira C. **Catálogo no plural**. Brasília: Briquet de Lemos, 2009. Disponível em: https://www.academia.edu/43366250/Cataloga%C3%A7%C3%A3o_no_Plural_Eliane_Serr%C3%A3o_. Acesso em: 8 mar. 2023.

MILSTEAD, Jessica; FELDMAN, Susan. Metadata: Cataloging by Any Other Name. **Online**, [S. l.], v. 23, jan. 1999. Disponível em: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.195.2851&rep=rep1&type=pdf>. Acesso em: 8 mar. 2023.

MODESTO, Fernando. **O divórcio do catalogador: AACR2 ou RDA**. [S. l.]: INFOHOME, 2010. Disponível em: http://www.ofaj.com.br/colunas_conteudo.php?cod=559. Acesso em: 8 mar. 2023.

MONACO, Daniele; PELLEGRINO, Maria Angela; SCARANO, Vittorio; VICIDOMINI, Luca. Linked open data in authoring virtual exhibitions. **Journal of Cultural Heritage**, [S. l.], v. 53, p. 127-142, 2022. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1296207421001667>. Acesso em: 8 mar. 2023.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). **Understanding Metadata**. Bethesda, MD: NISO Press, 2004. 20 p. Disponível em: <https://groups.niso.org/higherlogic/ws/public/download/17446/Understanding%20Metadata.pdf>. Acesso em: 8 mar. 2023.

NAVARRETE, Trilce; OWEN, John Mackenzie. The Museum as Information Space: Metadata and Documentation. In: BOROWIECKI, Karol Jan; FORBES, Neil; FRESA, Antonella. (ed.) **Cultural Heritage in a Changing World**. [S. l.]: Springer/Cham, 2016. p. 111-123. Disponível em: https://link.springer.com/chapter/10.1007%2F978-3-319-29544-2_7. Acesso em: 8 mar. 2023.

NHACUONGUE, Januário Albino; ROZSA, Vitor; DUTRA, Moisés Lima. *Linked Data* e ciência da informação: diretrizes para a publicação de datasets institucionais abertos. **Biblios**, Peru, n. 73, p. 20-34, 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/110426>. Acesso em: 8 mar. 2023.

OXFORD. **The world of art on the semantic web (CLAROS)**. Oxford: Oxford e-Research Centre, 2011. Disponível em: <https://digital.humanities.ox.ac.uk/project/world-art-semantic-web-claros>. Acesso em: 8 mar. 2023.

PAINT. In: Merriam-Webster Dictionary- Since 1828. [s. l.]: Merriam-Webster, c2021. Disponível em: <https://www.merriam-webster.com/dictionary/paint>. Acesso em: 8 mar. 2023.

PAINTING. In: Cambridge Dictionary. Reino Unido: Cambridge University Press, c2021. Disponível em: <https://dictionary.cambridge.org/pt/dicionario/portugues-ingles/painting>. Acesso em: 8 mar. 2023.

PAINTING. In: Britannica - Encyclopedia. [s. l.]: Encyclopedia Britannica, Inc., c2021. Disponível em: <https://www.britannica.com/art/painting>. Acesso em: 8 mar. 2023.

- PAINTING. *In*: WIKIPEDIA - The Free Encyclopedia. [S. l.]: Wikimedia Foundation, Inc., c2021. Disponível em: <https://en.wikipedia.org/wiki/Painting>. Acesso em: 8 mar. 2023.
- PAQUA, Megan; NULLMAN, Samantha; CURRAN, Katie. **The Museum System (TMS)**: Description of Software. [S. l.]: Collection Management Systems, c2018. Disponível em: <https://collectionmanagementsystems.wordpress.com/the-museum-system-tms/>. Acesso em: 8 mar. 2023.
- PINTURA. *In*: Britannica - Escola. [S. l.]: Encyclopedia Britannica, Inc., c2021. Disponível em: <https://escola.britannica.com.br/artigo/pintura/482141>. Acesso em: 8 mar. 2023.
- PINTURA. *In*: MICHAELIS - Dicionário Brasileiro da Língua Portuguesa. [S. l.]: Editora Melhoramentos Ltda., c2021. Disponível em: <https://michaelis.uol.com.br/moderno-portugues/busca/portugues-brasileiro/pintura>. Acesso em: 8 mar. 2023.
- PINTURA. *In*: Priberam Dicionário. [S. l.]: Priberam Informática S. A., c2021. Disponível em: <https://dicionario.priberam.org/pintura>. Acesso em: 8 mar. 2023.
- PINTURA. *In*: Wikipédia: A enciclopédia livre. [S. l.]: Wikimedia Foundation, Inc., c2021. Disponível em: <https://pt.wikipedia.org/wiki/Pintura>. Acesso em: 8 mar. 2023.
- RAPHAEL RESEARCH RESOURCE BETA. **Welcome to the Research Resource**. [London]: The National Gallery, [20--]. Disponível em: <https://cima.ng-london.org.uk/documentation/>. Acesso em: 8 mar. 2023.
- REIS, Mateus da Silva; CASTRO, Fabiano Ferreira de; FUJITA, Mariângela Spotti Lopes. Análisis del modelo conceptual CIDOC-CRM desde los principios básicos de la catalogación descriptiva en museos. *Anales de Documentación*, v. 26, 2023. Disponível em: <https://revistas.um.es/analesdoc/article/view/536221>. Acesso em: 19 mar. 2023.
- RDA Steering Committee. **About RDA**. Online. 2020. Disponível em: <http://www.rda-rsc.org/content/about-rda>. Acesso em: 8 mar. 2023. RDA 2022?
- RIJKSMUSEUM. **Over ons**. Amsterdam: Rijksmuseum, [20--]. Disponível em: <https://www.rijksmuseum.nl/nl/over-ons>. Acesso em: 8 mar. 2023.
- RODRIGUES, Bruno César; CRIPPA, Giulia. A ciência da informação e suas relações com arte e museu de arte. *Biblionline*, [S. l.], v. 5, n. 1/2, 2009. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/16800>. Acesso em: 8 mar. 2023.
- SANTARÉM SEGUNDO, José Eduardo. Web semântica, dados ligados e dados abertos: uma visão dos desafios do Brasil frente às iniciativas. **Tendências da Pesquisa Brasileira em Ciência da Informação**, João Pessoa, v. 8, n. 2, 2015. Disponível em: <https://www.brapi.inf.br/index.php/res/v/119595>. Acesso em: 8 mar. 2023.
- SAYÃO, Luis Fernando; SALES, Luana Farias. Documento seus dados. *In*: SAYÃO, Luis Fernando; SALES, Luana Farias. **Guia de gestão de dados de pesquisa para bibliotecários e pesquisadores**. Rio de Janeiro: CNEN/IEN, 2015. p. 27-38. Disponível em: <http://www.aben.com.br/Arquivos/420/420.pdf>. Acesso em: 8 mar. 2023.
- SILVA, Oscar Américo da. **AACR2 e RDA**: mudanças e estimativas para Biblioteca Nacional. 2013. 68 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Florianópolis, 2013. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/103965/AACR2%20e%20RDA%20mudan%C3%A7as%20e%20estimativas%20para%20BN.pdf?sequence=1&isAllowed=y>. Acesso em: 8 mar. 2023.
- SILVA, Renata Eleuterio da. **As tecnologias da web semântica no domínio bibliográfico**. 2013. 134 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Filosofia e Ciências, Marília, 2013. Disponível em: <http://repositorio.unesp.br/handle/11449/93653>. Acesso em: 8 mar. 2023.
- SILVA, Walison Dias da; PARREIRAS, Fernando Silva; MAIA, Luiz Cláudio Gomes; BRANDÃO, Wladimir Cardoso. Anotação semântica automática do currículo Lattes utilizando Linked Open Data. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 4, n. 23, out./dez. 2018. Disponível em: <https://www.scielo.br/j/pci/a/Bk7tsSvGVxHRhHYTXkFfkdQ/?lang=pt#>. Acesso em: 8 mar. 2023.
- STRINGFIXER. **Modelo de referência conceitual CIDOC-CRM**. Online. [20--]. Disponível em: https://stringfixer.com/pt/CIDOC_CRM. Acesso em: 8 mar. 2023.
- TRIQUES, Maria Lígia; ARAKAKI, Ana Carolina Simionato. Representação de patrimônios culturais em plataformas digitais: o modelo de dados da Europeana. **Em questão**, Porto Alegre, v. 27, n. 1, p. 185-209, jan/mar. 2021. Disponível em: <https://www.seer.ufrgs.br/EmQuestao/article/viewFile/100265/59316>. Acesso em: 8 mar. 2023.
- UTA. **SPARQL**. Definirtec, 2023. Online. Disponível em: <https://definirtec.com/sparql/>. Acesso em: 8 mar. 2023.
- VRA CORE. Official Website. **A data standard for the description of images and works of art and culture**. 2022. Online. Disponível em: <https://www.loc.gov/standards/vracore/>. Acesso em: 8 mar. 2023.
- VRA CORE. Official Website. **An introduction to VRA Core**. 2014. Online. Disponível em: <https://www.loc.gov/standards/vracore/>. Acesso em: 8 mar. 2023.
- WÖLFFLIN, Heinrich. **Conceitos fundamentais da história da arte**. Tradução João Azenha Jr. São Paulo: Martins Fontes, 2006.
- WORLD WIDE WEB CONSORTIUM (W3C). **RDF**: Estrutura de descrição de recursos (RDF). Online. 2014. Disponível em: <https://www.w3.org/RDF/>. Acesso em: 29 maio 2023.

Ciência cidadã em Botânica: um estudo sobre a colaboração de cidadãos em herbários e jardins botânicos

Citizen Science in Botany: a study on public collaboration with herbaria and botanical gardens

Hadassa De Zen Itepan (1), Nair Yumiko Kobashi (2)

(1) Escola de Comunicações e Artes da Universidade de São Paulo (ECA-USP), Av. Prof. Lúcio Martins Rodrigues, 443 - Butantã, São Paulo - SP, 05508-020, hadassa.itepan@usp.br. (2) nairkobashi@gmail.com

Resumo

Estudo exploratório sobre as relações entre Ciência Cidadã, Botânica e Organização do Conhecimento (KO) em diferentes plataformas: herbários virtuais, bases de dados de biodiversidade e páginas de Facebook. Os projetos de Ciência Cidadã têm fundamental importância para a coleta de dados, em especial no campo da biodiversidade e as instituições de ensino e estudo da área da Botânica podem se valer de esforços coletivos, tanto para a curadoria de dados quanto para a coleta de novas informações. No entanto, verifica-se que há dificuldades para desenvolver programas e bases de dados de Ciência Cidadã, tanto pela abrangência e diversidade do público participante, como também pelas políticas de compartilhamento de dados adotadas. A Organização do Conhecimento (KO) e a Ciência da Informação podem contribuir para a compreensão e desenvolvimento de sistemas de informação melhor ajustados à necessidade de projetos de Ciência Cidadã. Para isso, devem ser implementados novos estudos, especialmente no que diz respeito à licenças de uso dos dados (*Creative Commons*), no compartilhamento de dados de biodiversidade e desenvolvimento de programas de bases de dados de biodiversidade e de gestão de coleções botânicas.

Palavras-chave: Ciência Cidadã; Botânica; Jardins Botânicos; Herbários; Organização do Conhecimento (KO); Ciência da Informação.

Abstract

Exploratory study on the relationships between Citizen Science, Botany and Knowledge Organization (KO) in different platforms: virtual herbaria, biodiversity databases and Facebook pages. Citizen Science projects are of fundamental importance for data collection, especially in the field of biodiversity, and teaching and study institutions in the field of Botany can make use of collective efforts, both for data curation and for the collection of new information. However, it appears that there are difficulties in developing Citizen Science programs and databases, both due to the scope and diversity of the participating public, as well as the data sharing policies adopted. Knowledge Organization (KO) and Information Science can contribute to the understanding and development of information systems that are better suited to the needs of Citizen Science projects. For this, new studies must be implemented, especially with regard to data use licenses (*Creative Commons*), in the sharing of biodiversity data and the development of programs for biodiversity databases and management of botanical collections.

Keywords: Citizen Science; Botany; Botanical Gardens; Herbarium; Knowledge Organization (KO); Information Science.

Resumen

Estudio exploratorio sobre las relaciones entre Ciencia Ciudadana, Botánica y Organización del Conocimiento (KO) en diferentes plataformas: herbarios virtuales, bases de datos de biodiversidad y páginas de Facebook. Los proyectos de Ciencia Ciudadana son de fundamental importancia para la recopilación de datos, especialmente en el campo de la biodiversidad, y las instituciones de enseñanza y estudio en el campo de la Botánica pueden hacer uso de los esfuerzos colectivos, tanto para la curación de datos como para la recopilación de nueva información. Sin embargo, parece que existen dificultades para desarrollar programas y bases de datos de Ciencia Ciudadana, tanto por la amplitud y diversidad del público participante, como por las políticas de intercambio de datos adoptadas. La Organización del Conocimiento (KO) y la Ciencia de la Información pueden contribuir a la comprensión y desarrollo de sistemas de información que se adapten mejor a las necesidades de los proyectos de Ciencia Ciudadana. Para ello, se deben implementar nuevos estudios, especialmente en lo que respecta a las licencias de uso de datos (*Creative Commons*), en el intercambio de datos de biodiversidad y el desarrollo de programas para bases de datos de biodiversidad y gestión de colecciones botânicas.

Palabras clave: Ciencia Ciudadana; Botánica; Jardines botânicos; Herbario; Organización del Conocimiento (KO); Ciencias de la Información.

1 Introdução

De acordo com Luís (2022), à medida que a ciência ganha maior destaque perante a sociedade, maior é o envolvimento e diálogo entre ciência e sociedade. Diante deste contexto, diversas iniciativas de ciência aberta e mais especificamente de ciência cidadã têm surgido ao longo da última década.

A Ciência aberta é um fenômeno que tem causado muitas mudanças socioculturais e tecnológicas na comunicação da comunidade científica. As mudanças estão orientadas para a abertura dos procedimentos e resultados de pesquisas científicas, que abrangem desde o desenvolvimento de softwares abertos, até o livro acesso aos dados de pesquisa. (CARVALHO; LEITE, 2021).

A Ciência Cidadã é um movimento que se enquadra nas iniciativas de Ciência Aberta já que estimula a participação da comunidade não acadêmica na pesquisa científica, de modo a maximizar a quantidade de dados coletados e a expandir a compreensão do público sobre a Ciência (ADAMS, 2020). De acordo com Cohn (2008), a Ciência Cidadã tem impactos positivos nos processos de produção, comunicação e uso da informação científica. Com efeito, a colaboração entre pesquisadores e cidadãos voluntários pode ajudar a aprimorar técnicas de coleta de dados e expandir os resultados obtidos.

Diante do movimento de Ciência Cidadã, nos perguntamos quais são as possibilidades de um público voluntário contribuir para a coleta e curadoria de dados botânicos e de biodiversidade de Herbários e Jardins Botânicos, que são instituições de importância essencial para a Botânica. Os herbários são importantes centros de estudo, pesquisa e documentação da flora (PEIXOTO; MAIA, 2013); os Jardins Botânicos, por sua vez, detêm coleções vivas de plantas e sementes que são utilizadas para estudos e programas educativos, bem como para desenvolver ações de conservação da fauna. (FARAJI; KARIMI, 2020).

Neste estudo, foram selecionados para análise os projetos de Ciência Cidadã, na área de Botânica, em diferentes plataformas: herbários virtuais, bases de dados de biodiversidade e páginas de Facebook, com a finalidade discutir as possíveis contribuições da Organização do Conhecimento (KO) para a compreensão e proposta de sistemas de informação melhor ajustados à necessidade de projetos de Ciência Cidadã.

1.1 Objetivos do estudo:

Objetivo principal: analisar o movimento da Ciência Cidadã e sua contribuição para as instituições de ensino e estudo da Botânica, à luz das bases teóricas da Organização do Conhecimento (KO) e da Ciência da Informação.

Objetivos específicos: a) analisar contribuições específicas da Ciência Cidadã para a Botânica em bases de dados de biodiversidade, herbários virtuais e páginas de Facebook; b) compreender as características do seu público, as formas de compartilhamento de dados, bem como seu potencial para a educação científica.

1.2 Metodologia

Este estudo é exploratório, considerando a ausência de publicações sobre as relações entre Ciência Cidadã, Botânica e Organização do Conhecimento (KO). Inicialmente, foi realizado um levantamento bibliográfico sobre Ciência Cidadã, público participante, qualidade de dados e disponibilização de dados de pesquisa.

Em seguida, realizou-se um segundo levantamento sobre projetos de Ciência Cidadã, com foco em Botânica. Deu-se preferência para a apresentação de projetos desenvolvidos em meios digitais diferentes, que são usuais para o desenvolvimento de pesquisas e projetos: um herbário virtual, uma base de dados de biodiversidade e uma página de Facebook.

E, seguidamente, realizou-se uma pesquisa sobre teorias da Organização do Conhecimento (KO), com foco em instituições voltadas ao estudo da Botânica, como os Jardins Botânicos, Herbários e Museus de História Natural. Procurou-se entender as relações entre a Ciência Cidadã e os atos de coleta de coleções biológicas, em especial no contexto de instituições possuidoras de uma ampla rede de coleta de informações, muitas vezes coletadas por um público não envolvido diretamente em atividades científicas. Por fim, discute-se como a Organização da Informação (KO) e a Ciência da Informação podem contribuir para o desenvolvimento de sistemas de informação para projetos de Ciência Cidadã.

2 Conceitos essenciais da Ciência Aberta e Ciência Cidadã.

A comunidade científica sempre estabeleceu meios de comunicação para fins de construção, debate e divulgação dos resultados de suas pesquisas. Contudo, o meio utilizado para tais diálogos se alterou ao longo do tempo, especialmente pelas mudanças tecnológicas e dos meios de comunicação.

Entretanto, o crescimento exponencial da literatura especializada fez surgir desafios próprios, como a dificuldade para a seleção de material de qualidade pertinente às áreas de estudo.

Outra importante revolução na comunicação científica se deu com a internet, que possibilitou o compartilhamento de forma mais fácil e rápida. Nesse processo, os periódicos impressos perderam espaço para as versões digitais, que passaram a ser incluídos em

bases de dados virtuais. Contudo, o modelo de negócios de publicação de periódicos, conhecido como modelo de subscrição, tornou-se fonte de muita insatisfação e discussão na comunidade científica pelos entraves que impõe à livre circulação do conhecimento científico. O processo de publicação é atrelado a uma editora comercial, que recebe no mínimo duas taxas: do autor, ou da instituição, e a de acesso à obra já publicada, por meio da assinatura dos periódicos e das bases de dados. A editora não remunera o pesquisador ou avaliador por sua contribuição (SILVA; SILVEIRA, 2019). Tal modelo, baseado em assinaturas, tornou-se o “combustível” para o surgimento do Acesso Livre.

Os movimentos de Acesso Livre surgiram em 1999 como reação dos pesquisadores aos altos preços de assinatura das revistas científicas. O crescimento e desenvolvimento de redes digitais impulsionaram a crescente conscientização do impacto, necessidade e importância de conteúdos científicos disponibilizados em acesso livre, principalmente aqueles produzidos com financiamento público.

O conceito de Acesso Livre veio a consagrar-se baseado nos termos da Declaração de Budapeste (BUDAPEST OPEN ACCESS INITIATIVE, 2002), da Declaração de Bethesda sobre publicação em acesso aberto (BROWN, P. O. *et al.*, 2003) e na Declaração de Berlim sobre acesso livre ao Conhecimento nas ciências e Humanidades (MAX PLANCK SOCIETY, 2003) - sendo documentado o “acesso livre” como conceito base para o movimento de acordo com a seguinte definição:

‘Acesso livre’ à literatura científica revisada por pares significa a disponibilidade livre na Internet, permitindo a qualquer usuário ler, fazer download, copiar, distribuir, imprimir, pesquisar ou referenciar o texto integral desses artigos, recolhê-los para indexação, introduzi-los como dados em software, ou usá-los para outro qualquer fim legal, sem barreiras financeiras, legais ou técnicas que não sejam inseparáveis ao próprio acesso a uma conexão à Internet. As únicas restrições de reprodução ou distribuição e o único papel para o direito autoral neste domínio é dar aos autores o controle sobre a integridade do seu trabalho e o direito de ser devidamente reconhecido e citado. (BUDAPEST OPEN ACCESS INITIATIVE, 2002).

O movimento de Acesso Livre está diretamente ligado ao movimento de Ciência Aberta. Esta última, incentiva a transparência da pesquisa científica e o uso de softwares abertos. Busca também elaborar metodologias de gestão de dados, para que estes sejam acessíveis e possam ser reutilizados, tanto pela comunidade científica, quanto pelo público geral interessado. (SILVA; SILVEIRA, 2019).

O termo Ciência Aberta refere-se a diversos tipos de movimentos que têm o objetivo de construir uma ciência mais aberta. Estão incluídos no movimento de Ciência Aberta: o Código Aberto (*Open Source*), Acesso Aberto (*Open Access*); Dados abertos (*Open Data*); Cadernos Abertos de Laboratório (*Open Notebooks*); Recursos

Educacionais Abertos (*Open Educational Resources*); Revisão por Pares Aberta; Ciência Cidadã (*Citizen Science*) (MENESES; MORENO, 2019).

Entre todos estes movimentos, este trabalho se concentra na Ciência Cidadã, que consiste no envolvimento voluntário de cidadãos para dar respostas a perguntas científicas. Ele envolve desde a definição e desenho dos projetos, até a disseminação dos principais resultados e conclusões, passando pela coleta, discussão e interpretação dos resultados. (MENESES; MORENO, 2019).

A Ciência Cidadã sobrevém como um tópico de discussão relevante no tocante aos processos inerentes à produção, comunicação e uso da informação científica. As colaborações entre cientistas e cidadãos voluntários em pesquisas podem aprimorar as técnicas utilizadas na coleta de dados. Além disso, a Ciência Cidadã estimula a participação da comunidade não acadêmica na pesquisa científica, de modo a maximizar a quantidade de dados coletados e a expandir a compreensão desse público sobre a Ciência. (CARVALHO; LEITE, 2021).

De acordo com Gura (2013), há indicativos de que nos últimos anos a Ciência Cidadã tem se tornado mais popular entre os pesquisadores, e mais passível de se transformar em projetos legítimos e publicáveis. Isto porque os estudos feitos através da Ciência Cidadã oferecem a possibilidade de lidar com um volume de dados substancialmente maior do que o que um pesquisador ou grupo de pesquisadores poderiam coletar isoladamente. Isso é especialmente verdade para projetos que envolvem dados em Biodiversidade, em que a abrangência geográfica e temporal de espécies é de valiosa importância.

Contudo, o recrutamento do público geral pode trazer complicações, tal como encontrar as ferramentas técnicas e parceiros para organizar, desenvolver e manter os projetos. Além disso, a garantia da qualidade dos dados e a gestão do enorme volume de informações são preocupações essenciais para a integridade do desenvolvimento das pesquisas derivadas de projetos de Ciência Cidadã.

Dificuldades adicionais estão relacionadas ao treinamento do público e à sua motivação. Os participantes podem ser amadores com olhos destreinados ou até professores de Botânica aposentados. Os cientistas devem então elaborar seus projetos com abertura para um público muito amplo, além de enfrentar o desafio de recrutar voluntários e mantê-los engajados (GURA, 2013).

3 História da Ciência Cidadã aliada às coleções biológicas

A contagem de pássaros de Natal, "*Christmas Bird Count*", da Sociedade Nacional de Audubon, é usualmente citada como um dos primeiros casos de

Ciência Cidadã. Realizado desde 1900, é tido como o levantamento mais antigo e de maior abrangência geográfica sobre aves do hemisfério ocidental. A contagem inicial era feita em 25 localidades diferentes, mas, a cada ano, mais 20 a 30 locais eram adicionados. Atualmente, mais de 50.000 observadores participam da contagem, e existem cerca de 2.000 grupos de contagem, espalhados pelos EUA, Canadá e América Latina. Os dados foram usados em centenas de publicações em uma ampla gama de tópicos, como distribuição biogeográfica, mudanças morfológicas em aves e estudos populacionais. (DUNN *et al.*, 2005).

Contudo, há contestações de que o “*Christmas Bird Count*” seja de fato a primeira ocorrência na tradição de Ciência Cidadã. De acordo com Luís (2022), uma das áreas de maior destaque para a Ciência Cidadã é a área de Biodiversidade. Por séculos as pessoas demonstraram interesse pelo mundo natural e identificaram e documentaram animais e plantas. Há uma longa tradição de naturalistas amadores, pessoas fora da comunidade científica, que coletaram e publicaram em periódicos sobre suas descobertas naturais, desde o século XVII. (MILLER-RUSHING; PRIMACK; BONNEY, 2012).

Outras fontes demonstram tradições históricas de envolvimento da população na coleta e registro de informações acerca de fenômenos naturais. Na China, por exemplo, tanto funcionários quanto cidadãos registram surtos de gafanhotos há 3.500 anos, ao menos (TIAN *et al.*, 2011). No Japão, em Kyoto, foram encontrados registros sobre as mudanças em flores de cerejeiras, em diários e crônicas que remontam a mais de 1.000 anos (AONO; KAZUI, 2008).

Em épocas de grandes navegações e ocupações colonialistas, a coleta de informações sobre a História Natural dos países ocupados era de fundamental importância para a economia e manutenção de poder dos grandes centros. Portugal é um exemplo da importância do público instruído, mas também do público geral, para a coleta de informações, em especial para fins de depósito em coleções biológicas de caráter nacional, como os Museus Nacionais.

Em 1862, José Vicente Barbosa du Bocage, diretor do Museu de Lisboa, que mais tarde originou o Museu Nacional de História Natural e Ciência, publicou “As Instruções Práticas Sobre o Modo de Colligir, Preparar e Remetter Productos Zoologicos para o Museu de Lisboa.” Nesta publicação o diretor do museu apela à colaboração de indivíduos para ajudar a expandir suas coleções zoológicas (LUÍS, 2022).

Portugal é hoje o menos conhecido e explorado de todos os países da Europa; da sua Fauna apenas se conhecem poucos e raros fragmentos; nos museus mais ricos e completos, nas melhores colleções de particulares mal se avista um ou outro specimen colhido no nosso solo; mesmo o nosso antigo museu era n’este ponto um dos menos favorecidos. É tempo, cremos nós, de fazer cessar esta

vergonha, que denuncia mais do que tudo aos estrangeiros o nosso atraso e obscurantismo é tempo de estudar por nós mesmos o que é nosso, e de coligir pela forma que a sciencia prescreve os documentos que devem servir de base á historia das produções naturaes do nosso paiz. (BOCAGE, 1862, p. 8 *apud* LUÍS, 2022, p. 34).

De acordo com Luís (2022), cientistas e funcionários portugueses criaram e sustentaram uma rede de informação que permitiu ao Estado setecentista conhecer de forma aprofundada os seus domínios na Europa, Ásia, África e, sobretudo, na América. As informações coletadas pela rede permitiam reconhecer os limites físicos do Império Português, bem como as potencialidades econômicas do território administrado.

O processo de coleta de descrição de espécies e recursos naturais também contou com os povos originários, apesar de não terem sua contribuição reconhecida. Eram os povos originários que conheciam a fauna e flora da região e orientavam na utilização e aplicação de produtos, seja na área de agricultura, comércio ou farmacêutica. (LUÍS, 2022).

As coletas de espécimes realizadas durante este período eram depositadas em coleções de Museus de História Natural e Herbários. Muitas coleções existem até os dias de hoje, sendo conhecidas como coleções biológicas.

As coleções biológicas podem ser entendidas como coleções de espécimes biológicos, organismos ou parte de organismos usados como representativos de uma espécie, preservados fora de seu ambiente natural. Essas coleções são de grande importância para o estudo e documentação da biodiversidade, uma vez que os espécimes da coleção são preparados e organizados de forma a apresentar informações essenciais, como a identificação taxonômica, data e geolocalização, o que confere a elas um caráter científico. Desse modo, permite aos pesquisadores traçar um panorama temporal e geográfico da diversidade biológica depositada na coleção e desenvolver estudos que orientem de forma mais acertada as políticas de conservação ambiental. (ARANDA, 2014).

Os herbários são importantes centros de estudo botânico e documentação da flora, que abrigam coleções de plantas secas. Além dos herbários, os Jardins Botânicos também desempenham importante papel no estudo e conservação botânica. Já os Jardins Botânicos, mantêm coleções de plantas vivas, nativas e exóticas, cultivadas para pesquisa, ensino, extensão, produção, venda e conservação de mudas. Também são utilizadas para educação ambiental, estando abertas para visitação pública. (FONSECA; VIEIRA, 2015).

De acordo com a definição da União Internacional para Conservação da Natureza (IUCN), um jardim botânico contém coleções ordenadas de acordo com critérios científicos de classificação, onde as plantas possuem etiquetas com informações essenciais, disponíveis para

visitas do público para fins de recreação, educação e pesquisa (FARAJI; KARIMI, 2020).

A partir da compreensão da história da Ciência Cidadã e sua importância para a coleta e documentação de informações sobre a biodiversidade; bem como a importância de herbários e Jardins Botânicos para estudos botânicos, analisaremos as potencialidades de projetos de Ciência Cidadã dentro dessas instituições.

4 Projetos de Ciência Cidadã na área de Botânica

A digitalização dos espécimes depositados em coleções biológicas associado a informações sobre os espécimes são atualmente projetos de alta prioridade em diversas instituições do mundo. Herbários, assim como Jardins Botânicos que também possuem coleções biológicas, estão desenvolvendo Herbários Virtuais ou integrando seus dados a Bancos de Dados de Biodiversidade. (JAMES *et al.*, 2018).

De acordo com Silva *et al.* (2017), Bancos de dados de biodiversidade são cada vez mais consultados para pesquisas e tomada de decisões por parte de gestores públicos. Esses dados podem ser utilizados para a elaboração de ações de conservação, produção de listas de espécies ameaçadas, modelagem de distribuição de espécies, entre outros.

O Jardim Botânico do Rio de Janeiro (JBRJ) figura como uma das instituições botânicas brasileiras mais relevantes. (FORZZA *et al.*, 2015) Fundada em 1980, a instituição abriga, entre coleções de herbário e plantas vivas do jardim botânico, mais de 665 mil espécimes, que contribuem de forma valiosa para o conhecimento da flora neotropical (FORZZA *et al.*, 2016).

Considerando a importância da coleção e a necessidade de tornar as informações associadas a ela mais acessíveis, teve início, em 2005, uma iniciativa que buscou capturar em uma base de dados todas as informações contidas nas amostras botânicas da instituição.

O software desenvolvido pelo próprio JBRJ para armazenar e publicar dados e imagens das amostras online foi chamado de Jabot - Sistema de Gerenciamento de Coleções Botânicas. O software auxilia na curadoria dos acervos e cria uma rede de compartilhamento de informações (SILVA, 2017).

Algumas informações sobre os espécimes são consideradas essenciais para que estes sejam utilizados para pesquisas científicas, tal como a identificação taxonômica, data, geolocalização e coletor do espécime (PEIXOTO; MAIA, 2013) (BRIDSON; FORMAN, 1998).

Contudo, alguns exemplares depositados em herbários são extremamente antigos, datando de épocas coloniais, e a leitura das informações atrelados ao espécime pode

ser difícil de interpretar, tanto pela caligrafia quanto pela localização. Muitos espécimes eram atribuídos a localizações muito vagas, algo compreensível para naturalistas que estavam viajando por longos períodos dentro de florestas tropicais. Além disso, muitas cidades mudaram de nome ao longo do tempo. Dessa forma, erros acerca da identificação taxonômica, georreferenciamento e atribuição de coletas são comuns em coleções botânicas. (SILVA *et al.*, 2017).

Além desses fatores, os curadores que usualmente têm uma equipe reduzida de funcionários, enfrentam um grande desafio para avaliar e corrigir os erros de sua coleção. Baseado nessa questão e se inspirando na Ciência Cidadã, o Jardim Botânico do Rio de Janeiro desenvolveu um projeto para ajudar a corrigir os erros da coleção. Nesse projeto, os dados da coleção podem ser corrigidos de acordo com a qualificação dos voluntários, sendo eles divididos em estudantes de Botânica, de Geografia e público geral.

Os estudantes de Botânica podem auxiliar na correção de transcrições taxonômicas erradas e incompletas, ou até mesmo com o registro de uma nova determinação taxonômica. Os estudos de Geografia podem ajudar com a correção de dados geográficos informados erroneamente e indicação de possíveis localizações. Por fim, o público geral pode auxiliar com a identificação do nome do coletor principal e data de coleta informada (SILVA; OLIVEIRA, 2018).

O incentivo proposto à participação dos estudantes é a atribuição de atividades complementares, de acordo com as horas trabalhadas no projeto, que pode ser valioso para os alunos que necessitam de créditos extras para sua graduação. O projeto ainda possui valor pedagógico importante, pois pode contribuir para a formação de estudantes de áreas diversas e ainda apresentar o valor e a variedade das coleções botânicas.

Uma questão a ser destacada para a Ciência Cidadã é o envolvimento do público e sua educação científica. A compreensão das metodologias, instituições e questões científicas podem ser grandes instigadores de interesse pela Botânica.

De acordo com Martellos *et al.* (2016), os Jardins Botânicos são centros de enorme potencial para a educação científica do público e engajamento em projetos de Ciência Cidadã. Centros dedicados à ciência cidadã localizados em Jardins Botânicos poderiam fornecer treinamento para cientistas cidadãos e líderes de projetos, assim como para apoiar atividades de investigação de base comunitária e fornecer apoio e treinamento para os muitos projetos de observatórios cidadãos.

O engajamento de Jardins Botânicos forneceria um novo papel social e político para essas instituições e aumentaria sua visibilidade, bem como proporcionaria a oportunidade de desenvolver a consciência social sobre

as questões relacionadas à perda de biodiversidade no mundo atual e a necessidade de enfrentamento do problema.

De forma inesperada, o Facebook também mostra ser uma ferramenta importante para a promoção de projetos de Ciência Cidadã. A possibilidade de criação de grupos nessa mídia promove a troca de informações sobre assuntos específicos, como ocorreu com o grupo “Cogumelos do Brasil”.

O Grupo Cogumelos do Brasil foi criado em por uma micóloga, com o intuito de divulgar conhecimento sobre a funga brasileira. Nos primeiros anos, o grupo era fechado e invisível para potenciais interessados, mas a partir de 2015 tornou-se público, o que alavancou o crescimento do grupo. Em 2022, o grupo ultrapassou 18 mil membros.

A autora do grupo relata em um artigo (TRIERVEILER-PEREIRA *et al.*, 2022) a experiência com a mediação do grupo e quais as vantagens e problemáticas. De acordo com a autora, os projetos de Ciência Cidadã são importantes especialmente para macrofungos, pois estes possuem estruturas reprodutivas efêmeras e de distribuição muito restrita, o que dificulta o avistamento e identificação.

Através do grupo, se torna possível obter informações sobre espécies até então desconhecidas, ou então sobre distribuição de espécies em geral, identificação de espécies invasoras, entre outros. A troca de informações ajuda a alavancar o conhecimento sobre as espécies brasileiras e os ambientes onde ocorrem.

No entanto, grupos de Facebook requerem a curadoria dos membros para não ser poluído com anúncios ou conteúdos impróprios. A autora relata também que há pessoas com pouco conhecimento teórico-prático que se aventuram a identificar espécies de cogumelos comestíveis e/ou tóxicas, fato que pode ser extremamente perigoso. A curadoria também envolve a correção de identificações incorretas.

Outra questão referente ao grupo é acerca da acessibilidade dos dados. Para fazer o estudo sobre o grupo, foi necessário criar planilhas de alimentação de dados sobre o perfil dos participantes e conteúdos postados em termos de espécies fotografadas.

A metodologia parece custosa em termos de tempo e mão de obra para os autores, e os dados são de difícil aquisição e manipulação para outros grupos de cientistas que não fazem parte do grupo. Se os esforços estivessem integrados a um banco de dados de biodiversidade ou a algum software de identificação e discussão de espécies botânicas e fúngicas, a recuperação poderia ser melhor.

Contudo, a integração a uma base de dados de biodiversidade não é garantia de acesso livre aos dados. A GBIF - *Global Biodiversity Information Facility* é uma rede internacional de compartilhamento

de dados de biodiversidade. Os dados são providos de forma voluntária, de instituições como Herbários, Jardins Botânicos, Museus de História Natural, ou então de agências privadas e voluntários de projetos de Ciência Cidadã. Devido a essa natureza, muitos assumem que esses dados estão disponíveis para pesquisa científica.

Contudo, Groom, Weatherdon e Geijzendorffer (2017), questionaram se os dados de biodiversidade da GBIF eram de fato de Acesso aberto e analisaram as licenças de dados da base, de acordo com os tipos de contribuintes (voluntários, instituições e agências privadas).

Apesar de reafirmarem a importante contribuição dos cientistas cidadãos para a pesquisa em biodiversidade, as questões relativas à acessibilidade dos dados mostraram que são mais preocupantes. Os conjuntos de dados de biodiversidade no GBIF derivados de projetos de ciência cidadã foram frequentemente associados a licenças mais restritivas do que outros tipos de dados, e frequentemente, restringiam o uso de dados para organizações comerciais.

De acordo com os autores, isso pode ocorrer por vários motivos. Por exemplo, os detentores de dados podem usar como argumento a conservação de espécies, evitando divulgar a localização de espécies predadas e ameaçadas de extinção, para que caçadores e comerciantes não possam se aproveitar desses dados. Cientistas também podem sentir-se relutantes em compartilhar seus dados, uma vez que competem por financiamento e apoio institucional, e os dados coletados resultantes de suas pesquisas são valiosos no processo.

A variedade de licenças *Creative Commons* dentro do GBIF também pode ser problemática, confundindo os participantes nas suas escolhas e implicações no compartilhamento de dados coletados. (GROOM; WEATHERDON; GEIJZENDORFFER, 2017).

Políticas mais claras de restrições de uso comercial, e uma simplificação na escolha de licenças *Creative Commons* podem auxiliar na questão. Contudo, devemos nos perguntar: os dados de biodiversidade são de fato acessíveis? Podemos contá-los como parte do movimento de Ciência Aberta? E se não, como podemos resolver essas questões e promover o acesso a esses dados?

5 Interfaces de projetos de Ciência Cidadã com a Organização do Conhecimento (KO) e a Ciência da Informação

Algumas interfaces podem ser traçadas com a Organização do Conhecimento e a Ciência da Informação.

Em primeiro lugar, a Ciência Cidadã compreende a participação de um público muito variado, constituído por amadores entusiastas ou professores/pesquisadores botânicos. Qualquer sistema de informação que seja desenvolvido para esses projetos deve levar em conta essa amplitude de públicos e antecipar as formas diferenciadas de comunicação e busca por informação.

Em segundo lugar, a Organização do Conhecimento (KO) pode fornecer ferramentas para a compreensão da história da Ciência Botânica e sua inserção no cenário atual de herbários virtuais e bases de dados de biodiversidade.

De acordo com Hjørland (2008), a Organização do Conhecimento (KO) pode ser entendida, em um sentido mais restrito, como a representação e recuperação de documentos através de sistemas de organização do conhecimento (SOC), como classificações e linguagens documentárias. Contudo, para o autor, desenvolver sistemas eficientes de representação e recuperação da informação, é necessário primeiro entender como a informação circula e as dinâmicas de emergência e consolidação dinâmica de cada domínio científico, o que implica aspectos históricos, mas também pensar nos desenvolvimentos futuros de uma disciplina.

O movimento de Ciência Cidadã em herbários e Jardins Botânicos é de certa forma uma continuação de uma tradição antiga de colaboração de naturalistas e povos indígenas para a coleta de informações naturais. Contudo, a inserção de dados em bases de dados e a coleta em larga escala geram desafios que devem ser examinados mais detalhadamente. Certamente, a Organização do Conhecimento pode dar sua contribuição para solucionar esses problemas.

6 Conclusão

A Ciência Cidadã é um movimento compreendido dentro da Ciência Aberta e que pressupõe a contribuição do público, tanto no desenvolvimento e implementação, quanto no levantamento de dados que ajudam a responder perguntas científicas.

Os projetos de Ciência Cidadã podem ser de extrema importância, especialmente para a área de biodiversidade, por permitir o levantamento de dados sobre espécies em uma abrangência temporal e geográfica de difícil realização apenas por grupos isolados de pesquisadores.

O movimento pode trazer benefícios para instituições de divulgação e pesquisa da área de Botânica, como Herbários e Jardins Botânicos. Essas instituições também podem dar um retorno ao público e desenvolver projetos de educação científica.

Contudo, a participação de um público muito amplo nesses projetos pode resultar em dificuldades para

desenvolver sistemas de informação eficazes, ou então para lidar com o volume de dados e questões de interoperabilidade e acessibilidade.

Acreditamos que a Organização do Conhecimento (KO), bem como a Ciência da Informação, de forma mais ampla, pode trazer subsídios para entender o movimento de Ciência Cidadã e ajudar a desenvolver sistemas de informação melhor ajustados às necessidades dos pesquisadores e voluntários.

Nesse sentido, mais estudos sobre a área precisam ser desenvolvidos, em especial no que diz respeito às licenças de uso dos dados (*Creative Commons*) no compartilhamento de dados de biodiversidade; assim como sobre o desenvolvimento de programas para bases de dados de biodiversidade e de gestão de coleções botânicas.

Referências

- AONO, Y.; KAZUI, K. Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, [S. l.], v. 28, n. 7, p. 905-914, 2008.
- ADAMS, A. L. Citizen science. *Public Services Quarterly*, [S. l.], v. 16, n.1, p. 20-26, 2020.
- ARANDA, A. T. Coleções Biológicas: Conceitos básicos, curadoria e gestão, interface com a biodiversidade e saúde pública. In: SIMPÓSIO SOBRE A BIODIVERSIDADE DA MATA ATLÂNTICA, 3., 2014. *Anais [...]*. [S. l.]: 2014, p. 45-56.
- BROWN, P. O. *et al.* **Bethesda statement on open access publishing**. 2003. Disponível em: <<https://ecopersia.modares.ac.ir/page/128/Open-Access-Statement>>. Acesso em: 28 maio 2023.
- BRIDSON, B.; FORMAN, L. **The Herbarium handbook**. 3. ed. Kew, England: Royal Botanic Gardens, 1998.
- BUDAPEST OPEN ACCESS INITIATIVE. Budapest Open Access Initiative. 2002. Disponível em: <https://www.budapestopenaccessinitiative.org/read/>. Acesso em: 10 maio 2023.
- CARVALHO, E. R. S.; LEITE, F. C. L. Análise do atual cenário da pesquisa científica sobre a Ciência cidadã no campo da Ciência da Informação. **Páginas a&b: arquivos e bibliotecas**, Porto, p. 26-32, 2021.
- COHN, J. P. Citizen Science: can volunteers do the real research? *BioScience*, [S. l.], v. 58, n. 3, p. 192-197, 2008.
- DUNN, E. H.; FRANCIS, C. M.; BLANCHER, P. J.; DRENNAN, S. R.; HOWE, M. A.; LEPAGE, D.; ROBBINS, C. S.; ROSENBERG, K. V.; SAUER, J. R.; SMITH, K. G. Enhancing the scientific value of the Christmas Bird Count. *The Auk*, [S. l.], v. 122, n. 1, p. 338-346, 2005.

- FARAJI, L.; KARIMI, M. Botanical gardens as valuable resources in plant sciences. **Biodiversity and Conservation**, *[S. l.]*, p. 1-22, 2020.
- FONSECA, R. S.; VIEIRA, M. F. **Coleções botânicas com enfoque em herbário**. Viçosa, MG: Ed. UFV, 2015.
- FORZZA, R. C.; CARVALHO, A.; ANDRADE, A. C. S.; FRANCO, L.; NADRUZ, M.; TAMAIO, N. T.; FONSECA-KRUEL, V. S. Herbário do Jardim Botânico do Rio de Janeiro, Rio de Janeiro (RB). **Unisanta BioScience**, v. 4, n. 6, p. 393-398, 2015.
- FORZZA, R. C.; ANDRADE, C. S.; FRANCO, L.; SILVA, L. A. E.; FONSECA-KRUEL, V. S.; ZAPPI, D.; NADRUZ, M. A. C.; TAMAIO, N. Coleções biológicas do Jardim Botânico do Rio de Janeiro À luz das metas da GSPC/CDB: onde estamos em 2020?. **Museologia & Interdisciplinaridade**, v. 5, n. 9, p. 135-159, 2017.
- GROOM, Q.; WEATHERDON, L.; GEIJZENDORFFER, I. R. Is citizen science an open science in the case of biodiversity observations?. **Journal of Applied Ecology**, *[S. l.]*, v. 54, n. 2, p. 612-617, 2017.
- GURA, T. Citizen science: amateur experts. **Nature**, *[S. l.]*, v. 496, n. 7444, p. 259-261, 2013.
- HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, *[S. l.]*, v. 35, n. 2/3, p. 86-101, 2008.
- JAMES, S. A.; SOLTIS, P. S.; BELBIN, L.; CHAPMAN, A. D.; NELSON, G.; PAUL, D. L.; COLLINS, M. Herbarium data: global biodiversity and societal botanical needs for novel research. **Applications in Plant Sciences**, *[S. l.]*, v. 6, n. 2, 2018.
- LUÍS, C. A Ciência Cidadã: Passado, Presente e Futuro do Envolvimento Público na Investigação Científica. **Revista Lusófona de Estudos Culturais**, *[S. l.]*, v. 9, n. 2, p. 29-42, 2022.
- MARTELLOS, S.; LAGANIS, J.; BACARO, G.; BONACQUISTI, S.; DEREWNICKA, L.; ATTORRE, F. Botanical gardens and citizen science: an (as yet) under-exploited potential. **Plant Biosystems-An International Journal Dealing with all Aspects of Plant Biology**, *[S. l.]*, v. 150, n. 3, p. 381-383, 2016.
- MENESES, R. V.; MORENO, F. P. Estudo da literatura sobre ciência aberta na ciência da informação. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, *[S. l.]*, v. 14, n. 2, 2019.
- MILLER-RUSHING, A.; PRIMACK, R.; BONNEY, R. The history of public participation in ecological research. **Frontiers in Ecology and the Environment**, *[S. l.]*, v. 10, n. 6, p. 285-290, 2012.
- PEIXOTO, A. L.; MAIA, L. C. **Manual de Procedimentos para Herbários**. Editora Universitária UFPE: Recife, 2013.
- SILVA, L. A. E.; OLIVEIRA, F. A. Aplicação da Ciência Cidadã em Banco de Dados de Herbários. In: WORKSHOP DE COMPUTAÇÃO APLICADA À GESTÃO DO MEIO AMBIENTE E RECURSOS NATURAIS, 9., 2018. **Anais [...]**. *[S. l.]*: Sociedade Brasileira de Computação, 2018.
- SILVA, F. C. C.; SILVEIRA, L. O ecossistema da Ciência Aberta. **Transinformação**, Campinas, v. 31, 2019.
- SILVA, L. A. E. da; FRAGA, C. N. de; ALMEIDA, T. M. H. de; GONZALEZ, M.; LIMA, R. O.; ROCHA, M. S. da; BELLON, E.; RIBEIRO, R. da S.; OLIVEIRA, F. A. de; CLEMENTE, L. da S.; MAGDALENA, U. R.; MEDEIROS, E. von S.; FORZZA, R. C. Jabot -Sistema de Gerenciamento de Coleções Botânicas: a experiência de uma década de desenvolvimento e avanços. **Rodriguésia**, *[S. l.]*, v. 68, p. 391-410, 2017.
- TIAN, H.; STIGE, L. C.; CAZELLES, B.; KAUSRUD, K. L.; SVARVERUD, R.; STENSETH, N. C.; ZHANG, Z. Reconstruction of a 1,910-y-long locust series reveals consistent associations with climate fluctuations in China. **Proceedings of the National Academy of Sciences**, *[S. l.]*, v. 108, n. 35, p. 14521-14526, 2011.
- TRIERVEILER-PEREIRA, L.; CARDOSO, J. S.; PRADO-ELIAS, A.; NEVES, M. A.; KARSTEDT, F. Cogumelos do Brasil e a Ciência Cidadã na divulgação da funga brasileira. **Journal of Education Science and Health**, *[S. l.]*, v. 2, n. 3, p. 1-16, 2022.

Computational Archival Science: uma nova abordagem para a organização do conhecimento arquivístico

Computational Archival Science: a new approach to archival knowledge organization

Raquel Torrão Valentim (1), Natália Bolfarini Tognoli (2), Thiago Henrique Bragato Barros (3)

(1) Universidade Federal Fluminense, raquel_valentim@id.uff.br (2) Universidade Federal Fluminense, nataliatognoli@id.uff.br (3) Universidade Federal do Rio Grande do Sul, bragato.barros@ufrgs.br

Resumo

Objetivo: Apresentar a *Computational Archival Science* (CAS) enquanto uma nova abordagem na Organização do Conhecimento Arquivístico (OCA) a partir de um olhar complementar entre as áreas. **Metodologia:** trata-se de um estudo exploratório, de natureza qualitativa a partir de revisão bibliográfica sobre os referenciais teóricos da CAS e OCA. **Resultados:** Tendo em vista o pensamento computacional como seu principal fundamento teórico, podemos englobar nos estudos da CAS as atividades de organização e representação do conhecimento e vislumbrá-las em dimensões ainda pouco exploradas, tais como a automação de atividades como a identificação, recuperação, indexação e estruturação de conceitos. Apontamos que a CAS é capaz de abranger questões até então pouco exploradas nos estudos da Organização do Conhecimento Arquivístico, como a indexação, o relacionamento com *big data* e o uso de ontologias. **Conclusões:** A CAS retrata uma oportunidade inédita para estreitar os estudos entre a Computação e a Arquivologia e Organização do Conhecimento, implementando estudos que possibilitem avanços no desempenho das atividades de gestão e curadoria de documentos arquivísticos

Palavras-chave: *Computational Archival Science*. Organização do Conhecimento Arquivístico. Ontologias. Sistemas de Organização do Conhecimento. Análise de Domínio

Abstract

Objective: To introduce Computational Archival Science (CAS) as a new approach in the Archival Knowledge Organization (AKO) from a complementary perspective between the areas. **Methodology:** This is an exploratory, qualitative study based on a literature review of the theoretical references of CAS and AKO. **Results:** Considering computational thinking as its main theoretical foundation, we can encompass in the studies of CAS the activities of organization and representation of knowledge and glimpse them in dimensions still little explored, such as the automation of activities like identification, retrieval, indexing, and structuring of concepts. We point out that the CAS can cover issues hitherto little explored in studies of Archival Knowledge Organization, such as indexing, the relationship with big data, and the use of ontologies. **Conclusions:** The CAS represents an unprecedented opportunity to strengthen studies between Computer Science and Archival Science and Knowledge Organization, implementing studies that enable advances in addressing management activities and curating archival records.

Keywords: Computational Archival Science. Archival Knowledge Organization. Ontologies. Knowledge Organization Systems. Domain Analysis.

Resumen

Objetivo: Presentar la *Computational Archival Science* (CAS) como un nuevo enfoque en la Organización del Conocimiento Arquivístico (OCA) desde una mirada complementaria entre las áreas. **Metodología:** Se trata de un estudio exploratorio, de carácter cualitativo a partir de una revisión bibliográfica sobre los marcos teóricos de la CAS y la OCA. **Resultados:** Considerando el pensamiento computacional como su principal fundamento teórico, podemos englobar en los estudios de CAS las actividades de organización y representación del conocimiento y vislumbrarlas en dimensiones aún poco exploradas, como la automatización de actividades como identificación, recuperación, indexación y estructuración de conceptos. Destacamos que la CAS es capaz de abarcar temas hasta ahora poco explorados en los estudios de Organización del Conocimiento Arquivístico, como la indexación, la relación con *big data* y el uso de ontologías. **Conclusiones:** La CAS representa una oportunidad sin precedentes para fortalecer los estudios entre Computación y Archivística y Organización del Conocimiento, implementando estudios que permitan avances en la realización de actividades de gestión y curaduría.

Palabras clave: *Computational Archival Science*. Organización del Conocimiento Arquivístico. Ontologías. Sistemas de Organización del Conocimiento. Análisis de Dominio.

1 Introdução

O aporte das tecnologias digitais nas atividades arquivísticas e no tratamento de documentos arquivísticos digitais vem sendo objeto de estudos na comunidade acadêmica e profissional da Arquivologia e da Ciência da Informação (CI). Dentre esses, destacamos, aqui, o desenvolvimento da *Computational Archival Science* (CAS), que se configura enquanto um domínio emergente resultado da integração dos pensamentos arquivístico e computacional, com ênfase no último (VALENTIM, 2023). Dentre suas propostas, está a utilização de ferramentas e recursos computacionais nas atividades arquivísticas, como a Inteligência Artificial, abrangendo desde a produção documental até sua difusão e armazenamento.

Movida pela necessidade de expandir os conhecimentos sobre os usos e benefícios das tecnologias digitais na gestão e curadoria de documentos de arquivo, a CAS emerge enquanto um novo domínio destinado a reunir conhecimentos teóricos e metodológicos sobre o tema e aplicá-los em benefício das atividades profissionais.

A partir do uso de recursos e ferramentas computacionais, seus pesquisadores vêm desenvolvendo projetos que possibilitam a automação de atividades arquivísticas. É possível, então, pensar a automação de atividades como a identificação, recuperação, indexação e estruturação de conceitos. Nesse sentido, podemos vislumbrar uma aproximação da CAS com os estudos da Organização do Conhecimento (OC), levando-nos a considerá-la como uma nova abordagem para a Organização do Conhecimento Arquivístico (OCA).

Portanto, discutimos a CAS enquanto um domínio emergente, caracterizado previamente por meio de uma análise de domínio (VALENTIM, 2023), conforme as abordagens epistemológicas, histórica – de modo não exaustivo –, bibliométrica e terminológica (HJØRLAND, 2002), no escopo das atividades da OCA a partir de uma análise comparativa entre os campos.

Buscamos neste estudo apresentar a CAS enquanto uma nova abordagem para a OCA. Para tanto, partiremos de uma caracterização da CAS enquanto domínio emergente, com base nos estudos de Valentim (2023), para, em seguida, apresentar suas contribuições para a Organização do Conhecimento Arquivístico, com ênfase na construção de sistemas de organização do conhecimento (SOCs).

Esse estudo apresenta parte dos resultados de um esforço de pesquisa conjunto sobre as novas abordagens para a Organização do Conhecimento arquivístico empreendido no âmbito dos Grupos de Pesquisa Arquivos, Bibliotecas e Organização do

Conhecimento (GPABOC), da Universidade Federal Fluminense, e Organização e Representação do Conhecimento: abordagens linguísticas em arquivos e bibliotecas (ORCALAB), da Universidade Federal do Rio Grande do Sul. Trata-se, portanto, de uma colaboração entre instituições e entre grupos de pesquisa certificados pelo CNPq.

Como resultado, apontamos que por ser um domínio em construção, emergindo a partir de uma relação entre disciplinas – em nível ainda não determinado – a *Computational Archival Science* é capaz de englobar questões até então pouco exploradas nos estudos da Organização do Conhecimento Arquivístico, como a indexação, as relações com a Inteligência Artificial, as Humanidades Digitais e o uso de ontologias, por exemplo. Em especial no que toca a essa última, ressaltamos que, como o pensamento computacional é o principal fundamento teórico da CAS, seus estudos representam uma oportunidade para os pesquisadores da OC no contexto dos arquivos de desenvolverem métodos mais eficientes para construí-la.

2 A Organização do Conhecimento Arquivístico (OCA)

Nos últimos anos a relação entre a Arquivologia e a Organização do Conhecimento ganhou contornos mais nítidos, notadamente a partir da inserção das discussões sobre os processos de organização e descrição de documentos de arquivo na literatura científica da *International Society for Knowledge Organization* – ISKO, em seus congressos internacionais e em seus capítulos regionais. A partir da década de 1990, ficou claro para os autores de ambas as áreas que a OC deveria incorporar os estudos dos processos e sistemas voltados aos documentos de arquivo em suas práticas, consagrando-se o termo Organização do Conhecimento Arquivístico (OCA) enquanto um conjunto de atividades mediadoras entre a produção e o uso dos documentos de arquivo, que levam em conta a natureza especializada e contextualizada desses documentos para sua organização e representação.

O contexto é, portanto, o elemento determinante das circunstâncias de produção e uso dos documentos e, por esse motivo, um pressuposto para os processos, instrumentos e produtos da OCA que, em sua essência, são os responsáveis pelo estabelecimento de “pontes” entre essas duas circunstâncias – produção e uso (Tognoli et al. 2022, tradução nossa). Assim, o conhecimento do contexto é fundamental para a compreensão da organização do conhecimento arquivístico.

Nesse sentido, esse conhecimento contextualizado e especializado caracteriza-se por três facetas indivisíveis, conforme definido por Tognoli et al. (2019, grifo nosso), quais sejam: o *conceito de*

proveniência - sobre o qual incidirão todos os processos de organização - , *o conhecimento da forma documental* - adquirido por meio da análise diplomática - e *o conhecimento do contexto de produção* - por meio da metodologia da identificação arquivística e por meio da análise funcional.

Logo, a organização e representação desse conhecimento requer especial atenção ao elemento contextual que caracteriza e evidencia as relações entre o produtor, as atividades e os documentos produzidos. Para tanto, os processos de classificação e descrição do conhecimento arquivístico têm contado com ferramentas metodológicas e conceituais, como a identificação arquivística e a análise funcional, consideradas métodos para a identificação do órgão produtor e dos elementos centrais do documento de arquivo, cujos conteúdos - como tipo documental e assunto - podem ser extraídos para a construção de sistemas de organização do conhecimento e, a partir deles, sua representação.

Tradicionalmente, a Arquivologia tem se valido dos planos de classificação, quadros de arranjo e normas de descrição para organizar e representar o conhecimento arquivístico. Esses instrumentos, além evidenciarem as relações entre o produtor, as atividades e os documentos produzidos por eles, são instrumentos para garantia da transparência das ações e autenticidade dos documentos. No entanto, sua construção até recentemente permaneceu calcada apenas na capacidade de análise e síntese dos profissionais envolvidos no processo, o que significa que esses instrumentos, em sua maioria, apresentam poucas possibilidades de replicação e reuso.

A fim de garantir sua efetividade, especificamente no que diz respeito à organização e representação nos SOCs, para além dos instrumentos tradicionais, o contexto tecnológico implica a aproximação com outras disciplinas e áreas do conhecimento (BARROS, 2020; 2022), para a construção de instrumentos de controle conceitual no âmbito orgânico- funcional, por exemplo.

3 Computational Archival Science e Organização do Conhecimento Arquivístico: primeiras aproximações

Os usos e benefícios da aplicação de novas tecnologias digitais nas atividades arquivísticas ocupam discussões recorrentes na literatura e são refletidos no desenvolvimento de pesquisas teóricas e práticas. Apesar da produção de documentos de arquivo em ambiente digital desencadear estudos sobre seu tratamento e preservação, resultando em avanços teóricos e metodológicos sobre o tema, a constante evolução tecnológica segue demandando pesquisas contínuas sobre o assunto.

A noção do uso de ferramentas tecnológicas digitais como resposta aos desafios enfrentados pelos arquivistas, vem ganhando espaço nas pautas teóricas do campo e sendo verificadas por pesquisadores e profissionais em projetos práticos.

É o caso, por exemplo, dos estudos desenvolvidos no escopo da *Computational Archival Science*, uma iniciativa que ganhou espaço nos debates arquivísticos por propor reunir conhecimentos já existentes sobre o tratamento de documentos arquivísticos digitais e o uso de ferramentas e recursos tecnológicos nas atividades arquivísticas, bem como prosseguir com novos estudos a fim de promover seu ensino. Seus pesquisadores defendem a criação de um campo transdisciplinar, fundamentado no conhecimento teórico e metodológico da Arquivologia e da Ciência da Computação, que compile esses conhecimentos e avance em direção ao ensino nas graduações (MARCIANO *et al.*, 2018a).

A ideia de reunir esses conhecimentos em uma nova abordagem ganhou forças a partir do desenvolvimento de projetos experimentais, sobretudo na Universidade de Maryland, junto aos estudantes, promovendo o ensino teórico e aplicado do uso de recursos computacionais nas atividades arquivísticas. Ressalta-se que a maioria dos projetos em curso atualmente são liderados por Richard Marciano, pesquisador na Universidade e um dos co-fundadores da CAS, cujas contribuições são atingem um número relevante de projetos e artigos sobre o tema.

Podemos apontar que o diferencial da CAS em relação a outras iniciativas reside no empenho em tornar o ensino dessas ferramentas e recursos realidade nos cursos de graduação e os esforços de seus pesquisadores em criarem veículos de comunicação para disseminação do conhecimento em canais abertos de comunicação, como é o caso das apostilas *Computational Archival Science Education module*, ou CASE module, disponíveis em um repositório *on-line* chamado *Computational Archival Research Educational System (CASES)*[1], a promoção de *workshops* – sendo o primeiro organizado em 2016 pela *IEEE Big Data 2016 Conference* (MARCIANO *et al.*, 2019a) – e o *website* colaborativo de seus pesquisadores, o *Advanced Information Collaboratory*, onde estão disponíveis textos fundamentais para a compreensão dos projetos desenvolvidos, informações sobre seus pesquisadores e eventos, além de apresentar os objetivos dessa iniciativa, que consistem, sobretudo em contribuir para o desenvolvimento teórico da nova transdisciplina, projetar bases educacionais e fornecer treinamentos para atuação profissional e desenvolver um laboratório físico e virtual a fim de testar os avanços desenvolvidos em um ambiente colaborativo (ADVANCED INFORMATION COLLABORATORY, 2023, *on-line*, tradução nossa).

Sua institucionalização remete a dois marcos importantes. O primeiro foi o Simpósio “*Finding New Knowledge: Archival Records in the Age of Big Data*” em abril de 2016 promovido pela Universidade de Maryland em colaboração com a *King’s College London*, que reuniu estudantes, pesquisadores e profissionais de diferentes países com o objetivo de defini-la (MARCIANO *et al.*, 2018a). O segundo foi a publicação do capítulo de livro intitulado *Archival Records and Training in the Age of Big Data* em 2018, de autoria de seus cofundadores, com exceção de Gregory Jansen, onde apresentaram-na enquanto um campo de pesquisa destinado a promover a integração do pensamento arquivístico com o pensamento computacional (LEE *et al.*, 2017; MARCIANO *et al.*, 2018a).

Não é surpresa, portanto, que na Universidade de Maryland (Estados Unidos) a CAS já seja reconhecida enquanto uma área de especialização que explora a conjunção de métodos computacionais emergentes com a prática arquivística e o *big data*, junto a outras 13 áreas, conforme disponível em seu *website* (COLLEGE OF INFORMATION STUDIES, 2023, on-line). Apesar de a maioria dos pesquisadores ser vinculada à Universidade de Maryland, a CAS estende-se para além dessas fronteiras, tendo pesquisadores também na *King’s College London* (Reino Unido), *University of British Columbia* (Canadá), *Texas Advanced Computing Center* (Estados Unidos) e *Former NARA* (Estados Unidos).

Como defendido por Valentim (2023), a CAS configura um domínio emergente (SMIRAGLIA, 2012; BARITÉ, 2020), caracterizado por meio de uma análise de domínio, sob as abordagens epistemológica, histórica, bibliométrica e terminológica, a partir da análise de sua comunidade discursiva, através de suas escolhas terminológicas e seu referencial teórico.

A partir desse estudo, verificou-se que a CAS não apenas configura um domínio emergente, cuja principal base teórica é o pensamento computacional, isto é, um conjunto de práticas baseadas no poder e nos limites dos processos computacionais que inclui o uso de diferentes ferramentas e pensa na preservação, proteção e recuperação no âmbito computacional a partir de ações de redundância, contenção de dados e correção de erros (WEINTROP *et al.*, 2016; WING, 2006), integrado ao pensamento arquivístico, como também possui uma comunidade discursiva ativa com ideias convergentes e conceitos compartilhados. Concluiu-se também que ainda é preciso expandir os horizontes da CAS para além de seu pólo centralizador – que atualmente é a Universidade de Maryland – promovendo parcerias e ampliando sua comunidade discursiva e seu repertório teórico, além da necessidade de compilar seus conhecimentos em um material teórico.

Ressalta-se que essa integração dos pensamentos computacional e arquivístico, assim como a constituição da CAS são motivadas por oito estudos de caso temáticos que apresentam exemplos de esforços interdisciplinares para abordar o contexto de mudança de *recordkeeping* e o caráter dos documentos de arquivo (MARCIANO *et al.*, 2018a; UNDERWOOD *et al.*, 2017). Disponíveis e explicados no capítulo publicado por seus cofundadores os estudos de caso temáticos consistem em: Prototipagem evolutiva e linguística computacional; Análise de gráficos, humanidades digitais e representação arquivística; Auxílios de pesquisa computacional; Curadoria digital; Engajamento público com conteúdo (de arquivo); Autenticidade; Confluências entre teoria arquivística e métodos computacionais - ciberinfraestrutura e o *Records Continuum*; e Análise espacial e temporal (MARCIANO *et al.*, 2018a; UNDERWOOD *et al.*, 2017).

A ênfase da computação nos estudos da CAS direciona os debates da comunidade acadêmica para novas formas de utilizar as tecnologias em favor das atividades profissionais. Como consequência, a aplicação da integração dos estudos arquivísticos com o pensamento computacional vem sendo objeto de projetos teóricos e práticos no escopo do domínio, com ênfase no uso da Inteligência Artificial e da *Machine Learning* nas etapas de estruturação e processamento de textos através de ferramentas digitais. Além disso, destaca-se o avanço no desenvolvimento das atividades de indexação e na produção de ontologias.

Entendemos uma ontologia enquanto uma representação formal de um conjunto de conceitos e as relações entre eles. Essa representação é usada para descrever um domínio de conhecimento de forma que possa ser entendido tanto por humanos quanto por computadores.

Nesse contexto, a CAS utiliza-se de técnicas como a mineração de dados e análise de texto para a construção de ontologias. No primeiro caso, há a descoberta de padrões em grandes conjuntos de dados, que podem ser utilizados para identificação de conceitos e relacionamentos dentro do domínio conceituado. Para tanto, a CAS identifica padrões e usos de termos por meio da análise de texto, o que pode contribuir por exemplo, para a identificação de tipos documentais e elementos contextuais, enriquecendo os dados de proveniência, tão ricos à Organização do Conhecimento Arquivístico, como já mencionado na seção 2 deste trabalho.

No que toca às ontologias, por se tratar de ferramentas semânticas, há uma grande possibilidade de reuso em outros domínios, podendo ser utilizadas para a recuperação das informações contextuais em um sistema de recuperação. Por exemplo, os elementos constituintes da classificação arquivística são o

conjunto de ações desempenhadas pelo produtor dos documentos, assim a construção da ontologia se dá a partir do próprio desenvolvimento da classificação. Isso significa que à medida que se constituem informações sobre o órgão produtor, no contexto da categorização taxonômica realizada, se constroem as instâncias e relacionamentos na ontologia, facilitando o processo de busca e recuperação da informação em sistemas informacionais para além dos termos exclusivamente hierárquicos ou funcionais (BARROS, BASTOS & SANTOS, 2022).

Outra contribuição da CAS para a elaboração de SOCs é seu aporte para a indexação e recuperação das informações nos documentos, ao garantir consistência e precisão dos conceitos e termos, criando uma estrutura hierárquica e associativa para os termos relevantes àquele contexto específico de produção, organização e uso. Para tanto, a CAS faz uso da Inteligência Artificial, por meio das técnicas de processamento de linguagem natural (PLN), extraindo e analisando os termos frequentes e relevantes no conjunto de documentos a ser organizado, ferramentas que podem ser utilizadas nos mais variados contextos.

Exemplos da aplicabilidade da CAS na Organização do Conhecimento Arquivístico podem ser verificados nos estudos de Cox et al (2018), onde os pesquisadores aplicaram o pensamento computacional por meio de processos de análise e engenharia de dados, automatizando atividades através de um sistema de softwares, banco de dados e ferramentas de visualização para organizar o conhecimento de centenas de milhares de documentos a serem transcritos, possibilitando a otimização do tempo disposto para as atividades.

Marciano et al (2018b) também relatam o uso das técnicas da CAS com vistas ao equilíbrio entre a garantia da privacidade e o acesso à informação na curadoria, processamento e análise de documentos de arquivo da série de cartões Campo de Encarceramento de Japoneses-Americanos da Segunda Guerra Mundial, localizada no *National Archives and Records Administration* (NARA) em Washington D.C (EUA). Segundo o autor, foi formada uma equipe multidisciplinar com a missão de detectar as informações pessoais e garantir a privacidade dos indivíduos. A equipe desenvolveu e aplicou um algoritmo para automatização da tarefa em cerca de 10.000 cartões da série, alcançando o êxito desejado (MARCIANO *et al.*, 2018b).

Nota-se, portanto, que o processamento de grandes volumes de documentos pode ser facilitado com o uso da Inteligência Artificial e de *Machine Learning*, permitindo otimizar o tempo de trabalho de arquivistas e outros profissionais atuando no âmbito da organização e representação do conhecimento e da curadoria digital. Contudo, salienta-se que a

Arquivologia ainda carece de argumentos que sustentem as teorias e efetivem as propostas em meios viáveis de aplicação (JAILLANT, 2022; MARCIANO, 2022), sobretudo ao considerarmos o dever de preservar os princípios arquivísticos e as características dos documentos de arquivo. Ou seja, os estudos e aplicações práticas desenvolvidas no âmbito da CAS, ainda requerem revisões e aprofundamento teórico. A automação das atividades de organização e representação do conhecimento arquivístico ainda não foi alcançada de forma completa com suas ferramentas e abordagens atuais, ainda que existam estudos que buscam viabilizar o reuso da classificação arquivística, por meio de ontologias como desenvolvido por Barros e Gomes (2018).

Nesse contexto, acreditamos que a aproximação da CAS aos estudos da OCA é um caminho viável não apenas para a automação da identificação de conceitos em documentos de arquivos e sua estruturação em Sistemas de Organização do Conhecimento, mas também, para a caracterização da CAS enquanto uma nova abordagem para a organização do conhecimento arquivístico.

A partir do que vem sendo desenvolvido pelos pesquisadores da CAS, podemos vislumbrar o uso de ferramentas e recursos da Computação no desenvolvimento de atividades como a identificação, recuperação, indexação e estruturação de conceitos, contribuindo para sua automação e possibilitando a otimização do tempo dos processos. Essas projeções alcançam, não apenas a Organização do Conhecimento Arquivístico, mas se efetivadas, podem se estender a outros ambientes de curadoria digital.

4 Conclusão

Observamos ao longo deste trabalho, que as relações entre a Arquivologia e a Organização do Conhecimento ganham contornos mais robustos nos últimos dez anos. Essas relações perpassam desde a identificação dos processos de classificação e descrição arquivística enquanto pertencentes à OC, até a definição do conceito de conhecimento arquivístico, por meio da análise, composição e estrutura do conceito a partir de estudos complementares entre OC e Arquivologia.

Tal aproximação permite à Arquivologia uma expansão de suas fronteiras metodológicas e conceituais, abrindo caminho para uma interdisciplinaridade (ou até mesmo transdisciplinaridade) com outras áreas para além daquelas até então estabelecidas, como a História, a Diplomática e a Administração. Isso se deve, em partes, ao rápido avanço tecnológico que leva a disciplina a um repensar de suas bases teóricas, metodológicas e relacionais.

Nesse contexto, novas abordagens para a Organização do Conhecimento Arquivístico emergem como resposta

aos desafios deste século, dentre as quais destacamos a CAS.

Neste trabalho, partimos dos estudos de Valentim (2023) para definir a CAS enquanto um domínio emergente com elementos que a configuram, também, enquanto uma nova abordagem para a Organização do Conhecimento Arquivístico. Por se tratar de um domínio ainda em construção, sua sistematização carece de estudos mais aprofundados, especialmente no que toca à sua epistemologia. No entanto, podemos já identificar uma fundamentação teórica e metodológica baseada nos conhecimentos da Arquivologia e da Computação que abre possibilidades inovadoras para o tratamento técnico de acervos digitais.

Destacamos que o empenho de seus pesquisadores em promover o ensino sobre o uso de ferramentas computacionais nas atividades de gestão e curadoria possibilita sua incorporação nos estudos teóricos da OCA. Os projetos desenvolvidos no escopo da CAS demonstram aproximações entre as áreas e como a CAS pode ser utilizada como uma abordagem para a automação das atividades de organização e representação do conhecimento arquivístico.

Retomamos, por fim, as discussões desenvolvidas nas seções anteriores para definirmos a CAS enquanto uma abordagem inovadora para a Organização do Conhecimento Arquivístico, uma vez que é capaz de combinar diversas técnicas avançadas, como Processamento de Linguagem Natural (PLN), modelagem de domínios por meio de ontologias terminológicas, análise de dados e mineração de texto. Ao propor uma integração entre Ciência da Computação e Arquivologia, a CAS estabelece um enredamento entre a teoria da primeira e as ferramentas da segunda, possibilitando o uso de técnica que, quando aplicadas aos conjuntos documentais, permitem a extração de elementos fundamentais para a construção de sistemas de organização do conhecimento arquivístico, como planos de classificação, ontologias, índices, sistemas de recuperação e automatização de processos.

Com a CAS, a Arquivologia se beneficia do avanço tecnológico e das competências da Ciência da Computação, permitindo uma evolução significativa na forma como os processos arquivísticos são concebidos e implementados. A integração de técnicas como PLN e mineração de texto, por exemplo, possibilita a extração de informações valiosas dos documentos arquivísticos, facilitando a construção de sistemas de organização mais eficientes e precisos.

Essa abordagem híbrida traz uma perspectiva promissora para o campo da Arquivologia, combinando o conhecimento teórico tradicional com as ferramentas e técnicas da era digital. A CAS abre caminho para a criação de soluções inovadoras permitindo um melhor desempenho das atividades de gestão e curadoria dos

documentos, impulsionando a eficiência, a precisão e a automação nos processos de organização do conhecimento arquivístico.

Notas

[1] Disponível em: <http://cases.umd.edu/>.

[2] Este trabalho apresenta resultados parciais de uma pesquisa financiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Referências

ADVANCED INFORMATION COLLABORATORY. 2023. Disponível em: <https://ai-collaboratory.net/cas/>. Acesso em: 1 jan. 2023.

BARITÉ, M. Terminología de urgencia y garantías para la representación temática: elementos para el análisis de dominios de emergencia súbita (des). **Informação & Informação**, v. 25, n. 3, p. 239-265, 2020.

BARROS, T. H. B. Touching from a distance: concept theory and archival hierarchical classification. In In: LYKKE, M. *et al.* (ed). Knowledge Organization at the Interface: proceedings of the sixteenth international Isko Conference. **Proceedings** [...]. Würzburg: Baden-Baden: Ergon, 2020, v.18,p. 465-469.

BARROS, T. H. B. Archival Classification as Knowledge Organization. In: LYKKE, M. *et al.* (Orgs). Knowledge Organization across Disciplines, Domains, Services, and Technologies: proceedings of the seventeenth international Isko Conference. **Proceedings** [...]. Würzburg: Baden-Baden: Ergon, 2022, v.19,p. 23-36..

BARROS, T. H. B.; GOMES, D. L. Classification and Knowledge Organization Systems: ontologies and archival classification. In: RIBEIRO, F.; CERVEIRA, M. E. (Orgs.). Challenges and opportunities for knowledge organization in the digital age: proceedings of the fifteenth international Isko Conference. **Proceedings** [...]. Würzburg: Ergon-Verlag, 2018, v. 16, p. 103-111.

BARROS, T. H. B.; BASTOS, C. M. C.; DOS SANTOS, A. C. R. Sistemas de organização do conhecimento no contexto da arquivologia: aportes metodológicos para seu desenvolvimento. **Acervo**, [S. l.], v. 35, n. 2, p. 1–20, 2022.

COLLEGE OF INFORMATION STUDIES. 2023. Disponível em: <https://ischool.umd.edu/research/expertise-areas/> Acesso em: 18 dez. 2022.

COX, R. *et al.* A case study in creating transparency in using cultural Big Data: the legacy of slavery project. In: IEEE BIG DATA CONFERENCE, 10-13 dec. 2018, Seattle, WA. **Proceedings** [...]. [S.l.]: IEEE, 2018.

HJØRLAND, B. Domain Analysis in information science: eleven approaches— traditional as well as innovative. **Journal of Documentation**, v. 58, p. 422-62, 2002.

JAILLANT, L. **Archives, access and artificial intelligence: working with born-digital and digitized archival collections**. Bielefeld: Bielefeld University Press, 2022.

- LEE, M. *et al.* Heuristics for assessing Computational Archival Science (CAS) research: the case of the human face of Big Data project. *In: IEEE BIG DATA CONFERENCE*, 11-14 dec. 2017, Boston, MA. **Proceedings** [...]. Boston, MA: IEEE, 2017.
- MARCIANO, R. Afterword: Towards a new discipline of Computational Archival Science (CAS). *In: JAILLANT, L. Archives, access and artificial intelligence: working with born-digital and digitized archival collections.* Bielefeld: Bielefeld University Press, 2022.
- MARCIANO, R. *et al.* Archival records and training in the age of Big Data. *In: PERCELL, J. et al. (eds.). Re-envisioning the MLS: perspectives on the future of library and information science education.* Bingley, UK: Emerald, 2018a, p. 179-199. (Advances in Librarianship, v. 44B).
- MARCIANO, R. *et al.* Automating the detection of personally identifiable information (PII) in Japanese-American WWII Incarceration Camp Records. *In: INTERNATIONAL CONFERENCE ON BIG DATA*, 10-13 dec. 2018, Seattle, WA. **Proceedings** [...]. Seattle, WA: IEEE, 2018b, p. 2725-2732.
- MARCIANO, R. *et al.* Establishing an international computational network for librarians and archivists. *In: iCONFERENCE*, 31 mar.-3 apr. 2019, Washington, DC. **Proceedings** [...]. Washington, DC.: Springer, 2019a.
- SMIRAGLIA, R. P. Epistemology of domain analysis. *In: SMIRAGLIA, R. P.; LEE, H.-L. (Orgs.). Cultural frames of knowledge.* Würzburg: Ergon, 2012, p. 111-124.
- TOGNOLI, N. B.; RODRIGUES, A. C.; GUIMARÃES, J. A. C. Definindo o conhecimento arquivístico: estruturas conceituais. **Informação & Informação**, v. 24, n. 2, p. 58-75, 2019.
- TOGNOLI, N.B; SCHMIDT, C. M. S.; GUIMARÃES, J.A.C. Context as a Core Concept in Archival Knowledge Organization *In: LYKKE, M. et al. (ed). Knowledge Organization across Disciplines, Domains, Services, and Technologies: proceedings of the seventeenth international Isko Conference. Proceedings* [...]. Würzburg:Baden-Baden: Ergon, 2022, v.19, p.. 273-284.
- VALENTIM, R. T. **Computational Archival Science: caracterização de um domínio emergente.** 2023. Dissertação (Mestrado acadêmico em Ciência da Informação) – Instituto de Arte e Comunicação Social, Universidade Federal Fluminense, Niterói, 2023. No prelo.
- WEINTROP, D. *et al.* Defining computational thinking for mathematics and science classrooms. **Journal of Science Education and Technology**, v. 25, n. 1, p. 127-147, 2016.
- WING, J. Computational thinking. **Communications of the ACM**, v. 49, n. 3, p. 33-35, 2006.

As comunidades virtuais de aprendizagem e o uso das tecnologias digitais na formação do gestor de compras e licitações no IFPB

The virtual learning communities and the use of digital technologies in the training of the purchasing and bidding manager at the IFPB

Állysson Albuquerque Andrade (1), Josilene Aires Moreira (2), Roberto Vilmar Satur (3)

(1) Universidade Federal da Paraíba (UFPB), Centro de Educação – UFPB – Campus I – Castelo Branco – João Pessoa/PB – CEP 58051-900, allysson3@gmail.com, (2) josilene@ci.ufpb.br, (3) roberto.satur@academico.ufpb.br

Resumo

O objetivo da pesquisa é o de analisar o papel das comunidades virtuais de aprendizagem e do uso das tecnologias digitais da informação e da comunicação (TDICs) na capacitação dos gestores de compras e licitações, através do estudo de caso Escola Virtual do Governo (EV.G), realizado durante a pandemia da Covid-19, no Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB). Esse é um estudo de caso exploratório e descritivo que teve (i) revisão bibliográfica, compreendendo a contextualização e legislação aplicável ao gestor de compras e licitações, comunidades virtuais de aprendizagem, tecnologia digital aplicada à capacitação, e (ii) dados de pesquisa coletados através de questionário estruturado com os gestores responsáveis pelos processos de compras públicas no IFPB, analisados de modo quantitativo (através de análise estatística descritiva). Os resultados indicam que há uma boa recepção do uso da EV.G por parte dos gestores, embora existam limitadores internos a serem superados. Ficaram evidenciadas as seguintes necessidades: (a) Capacitar o servidor público antes de inseri-lo no setor de compras e licitação; (b) Incentivar a participação dos gestores nos cursos ofertados pela EV.G; (c) Apoiar e acompanhar a atualização contínua dos gestores e; (d) Fomentar uma cultura de compartilhamento do aprendizado adquirido. Conclui-se que as TDICs desempenham importante papel no desenvolvimento das organizações e na formação do profissional de compras e licitações, que devem buscar se adaptar a esses tempos de cultura digital e de aprendizagem por toda a vida.

Palavras-chave: Comunidades virtuais de aprendizagem. Aprendizagem contínua. Gestor de Compras. IFPB.

Abstract

The objective of the research is to analyze the role of virtual learning communities and the use of digital information and communication technologies (DICT) in the training of purchasing and bidding managers, through the case study Virtual School of Government (EV.G), carried out during the Covid-19 pandemic, at the Federal Institute of Education, Science and Technology of Paraíba (IFPB). This is an exploratory and descriptive case study that had (i) bibliographical review, including the context and legislation applicable to the purchasing and bidding manager, virtual learning communities, digital technology applied to training, and (ii) research data collected through structured questionnaire with managers responsible for public procurement processes at the IFPB, analyzed quantitatively (through descriptive statistical analysis). The results indicate that there is a good reception of the use of the EV.G by the managers, although there are internal limitations to be overcome. The following needs were evident: (a) Train the public servant before inserting it in the purchasing and bidding sector; (b) Encourage the participation of managers in courses offered by EV.G; (c) Support and monitor the continuous updating of managers and; (d) Foster a culture of sharing acquired learning. It is concluded that DICT play an important role in the development of organizations and in the training of purchasing and bidding professionals, who must seek to adapt to these times of digital culture and lifelong learning.

Keywords: Virtual learning communities. Lifelong learning. Purchasing manager. IFPB.

Resumen

El objetivo de la investigación es analizar el papel de las comunidades virtuales de aprendizaje y el uso de las tecnologías digitales de la información y la comunicación (TDICs) en la formación de gerentes de compras y licitaciones, a través del estudio de caso Escuela Virtual del Gobierno (EV.G), realizado durante la pandemia de Covid-19, en el Instituto Federal de Educación, Ciencia y Tecnología de Paraíba (IFPB). Se trata de un estudio de caso exploratorio y descriptivo que contó con (i) revisión bibliográfica, incluyendo el contexto y legislación aplicable al gestor de compras y licitaciones, comunidades virtuales de aprendizaje, tecnología digital aplicada a la formación, y (ii) datos de investigación recolectados a través de un cuestionario estructurado con directivos responsable de los procesos de contratación pública en la IFPB, analizados cuantitativamente (mediante análisis estadístico descriptivo). Los resultados indican que existe una buena acogida del uso de la EV.G por parte de los controladores, aunque existen limitaciones internas a superar. Se evidenciaron las siguientes necesidades: (a) Capacitar al servidor público antes de insertarlo en el sector de compras y licitaciones; (b) Fomentar la participación de los gerentes en los cursos ofrecidos por EV.G; (c) Apoyar y monitorear la actualización continua de los gerentes y; (d) Fomentar una cultura de compartir el aprendizaje adquirido. Se concluye que las TDIC juegan un papel importante

en el desarrollo de las organizaciones y en la formación de profesionales de compras y licitaciones, quienes deben buscar adaptarse a estos tiempos de cultura digital y aprendizaje permanente.

Palabras clave: Comunidades virtuales de aprendizaje. Aprendizaje continuo. Gerente de compras. IFPB.

1 Introdução

O IFPB (Instituto Federal de Educação, Ciência e Tecnologia da Paraíba) é uma autarquia federal vinculada ao Ministério da Educação – MEC que conta com 21 unidades estrategicamente distribuídas em todo o estado da Paraíba, desempenhando importantes tarefas em prol do desenvolvimento de diversas cidades e regiões do estado, uma vez que além de ofertar cursos que promovem o desenvolvimento profissional, tecnológico e humanístico de forma ética e sustentável, também coordena projetos diversos de pesquisa, extensão e tecnologia capazes de transformar socialmente, economicamente e ambientalmente a sociedade.

Diante da relevância dos serviços ofertados, é necessário manter um bom gerenciamento dos processos de contratações do instituto, uma vez que a eficácia desses processos irão garantir a continuidade dos serviços prestados, assim como servir como parâmetros para o planejamento das contratações futuras. Nesse caso, o Gestor de Compras e Licitação é a figura responsável pela instrução processual de compras nos diversos órgãos e entidades da Administração Pública.

De acordo com a Lei nº 14.133, de 01 de abril de 2021, em seu Capítulo IV:

Art. 7º Caberá à autoridade máxima do órgão ou da entidade, ou a quem as normas de organização administrativa indicarem, promover gestão por competências e designar agentes públicos para o desempenho das funções essenciais à execução desta Lei que preencham os seguintes requisitos:

I - sejam, preferencialmente, servidor efetivo ou empregado público dos quadros permanentes da Administração Pública;

II - tenham atribuições relacionadas a licitações e contratos ou possuam formação compatível ou qualificação atestada por certificação profissional emitida por escola de governo criada e mantida pelo poder público. (BRASIL, 2021).

Até o primeiro trimestre do ano de 2020, os processos de compras públicas no IFPB eram tramitados, quase que em sua totalidade, em meio físico, mas com o advento da pandemia Covid-19, foi publicado a Instrução Normativa nº 21, na qual estabeleceu orientações aos órgãos e entidades do Sistema de Pessoal Civil da Administração Pública Federal, quanto às medidas de proteção para enfrentamento da emergência de saúde pública, entre as quais o *home office*, que na língua portuguesa se denomina teletrabalho. A partir dessa nova situação emergencial vários unidades administrativas passaram a editar

normativos sobre a nova forma de trabalho (*home office*), prática que ainda era desconhecida para muitos.

Uma definição interessante sobre *home office* (ou teletrabalho) quem traz é o autor Pedreira (2000), quando diz que consiste em atividade que é desenvolvida total ou parcialmente em locais distantes da empresa, através de serviços informatizados.

Nesse novo contexto que todos passaram a viver (pandemia e isolamento) o trabalho de compras públicas também foi afetado e praticamente todos os processos passaram a acontecer na forma remota, por meio virtual, ainda com pouquíssima regulamentação (e muitas dúvidas) que orientasse os responsáveis sobre práticas de tramitação e migração de processos físicos para digitais, sobre uso de ferramentas de produtividade, assim como o uso das demais tecnologias de informação e comunicação e compartilhamento de procedimentos, processos e conhecimento nesse novo ambiente com novas prática híbridas ainda desconhecidas ou pouco conhecidas de muitos dos agentes que precisavam agora com ela atuarem. Assim, a tendência que se observou foi a utilização das tecnologias como forma de romper com o isolamento físico, sem transgredir as regras de distanciamento entre as pessoas. (SILVA; LIMA; COUTO, 2020, p. 1505).

Além da crise causada pela pandemia Covid-19, que modificou a forma de trabalhar, o gestor de compras e licitações públicas deveria estar preparado para enfrentar as diversas alterações legislativas que, caso não fossem devidamente compreendidas, poderiam gerar vários problemas, tanto para o usuário do serviço público como para o próprio gestor. Por isso, a capacitação contínua dos servidores que atuam nessa atividade se revela mais do que apenas uma política de desenvolvimento individual e organizacional, mas um meio para fornecer a segurança necessária para execução das diversas atividades administrativas da instituição, inclusive como uma forma de maior segurança jurídica para o profissional que executa tais ações. Assim, se capacitar continuamente não é apenas uma questão de desejo pessoal, mas uma necessidade profissional, e portanto, “torna-se necessário refletir sobre o valor que o servidor público atribui à capacitação no seu ambiente de trabalho”. (APPUGLIESE, 2010, p. 1).

Nesse cenário, surgiu com grande intensidade a demanda pelas capacitações advindas da cultura digital, bem típicas desse século XXI que tem por base a tecnologia digital da informação, da comunicação e da aprendizagem. Na visão de Duarte *et al.* (2015, p. 2) “a evolução dos estudos ocasionada pelas mudanças

tecnológicas e pelas necessidades infocomunicacionais das pessoas provoca uma revolução no saber-fazer no ambiente das organizações que sinaliza a Era do Conhecimento”. Nesse sentido, são formadas as comunidades virtuais de aprendizagem que se consolidaram de forma eficaz e acessível de práticas de educação à distância.

De fato, a construção de espaços *on-line* de formação, capacitação, treinamento e atualização compreende um desafio que não se trata apenas de migrar a forma de aprendizagem presencial para um ambiente virtual, mas na consolidação de uma metodologia de aprendizagem e conhecimento que permita uma formação adequada, dentro de um contexto organizacional que demanda gestão de tempo, readequação de rotinas, perfil do discente e a dinâmica tutor/facilitar/professor com o participante do curso.

Assim, a criação da comunidade de aprendizagem *on-line* pressupõe que todos os membros envolvidos detenham conhecimentos tecnológicos necessários à participação naquela sala ou ambiente de trocas e compartilhamento. Isso exige que os tutores, facilitadores e/ou professores, estejam envolvidos com seus aprendizes e/ou participantes do ambiente numa dinâmica de colaboração, compartilhamento e construção conjunta do conhecimento estudado. A interação e a troca de experiências vividas é uma das formas de aprendizado nesse ambiente, em que todos aprendem com todos, tendo liberdade de expressar naquela sala virtual suas dúvidas e tê-las sanadas com a interação de quem sabe como resolver tal questão. Essa interação orientativa pode ser feita pelo tutor, facilitador ou professor mas também pelos demais membros que já passaram por aquela situação e aprenderam com ela.

Como essas comunidades de aprendizagem que usam as tecnologias digitais são utilizadas na capacitação dos gestores de compras e licitações do IFPB, então o objetivo desse artigo foi o de analisar o papel das comunidades virtuais de aprendizagem e do uso das tecnologias digitais na capacitação dos gestores de compras e licitações do IFPB.

1.1 Metodologia

A pesquisa é um “procedimento reflexivo sistemático, controlado e crítico, que permite descobrir novos fatos ou dados, relações ou leis, em qualquer campo do conhecimento” (LAKATOS e MARCONI, 2020, p. 169). E para que a pesquisa aconteça é preciso que se tenha um método. Sendo assim, a metodologia proposta para este estudo, quanto aos meios, será bibliográfica e de campo, com pesquisa estruturada (questionário). Quanto aos fins, a pesquisa tem caráter exploratório e descritivo onde Cervo & Bervian (1996, p. 49) chamam de descritiva a investigação que “observa, registra, analisa e correlaciona fatos ou fenômenos sem manipulá-los”.

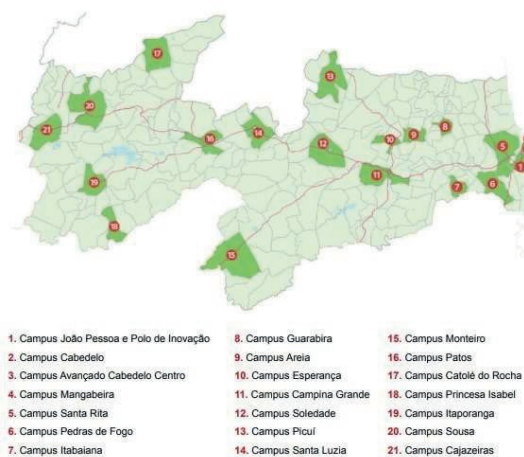
Tratou-se de um estudo de caso referente ao IFPB e seus profissionais responsáveis pela instrução processual e operacionalização das compras públicas. Dessa forma, esse estudo de caso envolveu pesquisa exploratória e descritiva, apresentando:

(i) Revisão bibliográfica, compreendendo a contextualização e legislação aplicável ao gestor de compras e licitações, comunidades virtuais de aprendizagem, tecnologia digital aplicada à capacitação, na qual são obtidas informações de livros, artigos, legislação, manuais, revistas, jornais e internet.

Esse tipo de pesquisa se caracteriza pelo levantamento de toda a bibliografia relevante já publicada e que tenha relação com o tema em estudo, com a finalidade de colocar o pesquisador em contato direto com tudo aquilo que foi escrito sobre determinado assunto em determinado recorte do tempo. No caso o recorte foi o que se tinha de mais contemporâneo sobre o assunto.

(ii) Os dados de pesquisa foram coletados junto aos setores de compras e licitação, no período de 03 de dezembro de 2021 a 09 de dezembro de 2021, em todos os *Campi* do IFPB (21 unidades) distribuídos nas cidades de Areia, Cabedelo, Cajazeiras, Campina Grande, Catolé do Rocha, Esperança, Guarabira, Itabaiana, Itaporanga, João Pessoa, Monteiro, Patos, Pedras de Fogo, Picuí, Princesa Isabel, Santa Rita, Santa Luzia, Soledade e Sousa (IFPB, 2021).

Figura 1. Área de abrangência do Instituto Federal da Paraíba (IFPB)



Fonte: IFPB-PDI (2020).

O universo de pesquisa foi de 22 gestores responsáveis pelos processos de compras públicas no IFPB (população-alvo da pesquisa), entre Diretores e Coordenadores de Compras e Licitação, dos quais conseguiu-se uma coleta amostral de 20 respondentes, próxima à totalidade da população. O instrumento de coleta foi o questionário estruturado, constituído por uma série ordenada de perguntas, que foram analisadas de modo quantitativo (através de análise estatística descritiva), com a garantia de não divulgação de suas identidades.

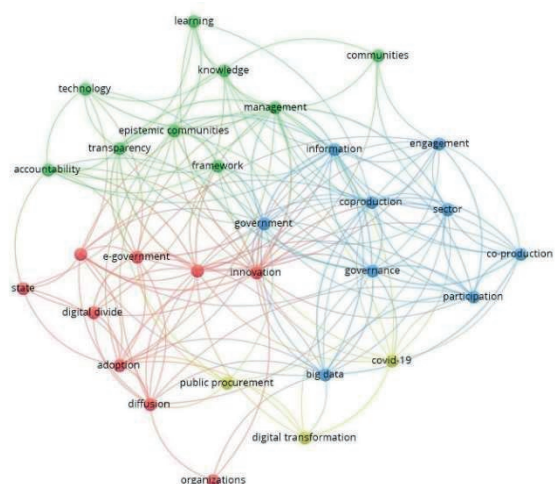
Com a finalidade de justificar a relevância desse estudo para a literatura nacional sobre o tema, assim como corroborar com a observação do pesquisador de que ainda existe uma lacuna do conhecimento de comunidades de aprendizagem na área das compras públicas e do uso das tecnologias digitais na capacitação de gestores de compras e licitações, buscou-se realizar a pesquisa na base de dados *web of science*, pelo alto impacto das publicações científicas que lá estão.

A pesquisa foi iniciada utilizando-se os operadores booleanos “OR” e “AND”, e as palavras-chave comunidades de aprendizagem, tecnologias digitais, compras públicas e aprendizagem contínua relacionadas à Administração Pública. A partir de alguns refinamentos realizados, com o intuito de aumentar a confiabilidade do estudo, tomando por base a produção científica dos últimos dez anos, selecionando-se apenas artigos, publicados em inglês e português, assim como definindo as categorias dos temas e área de pesquisa (Administração Pública), o resultado da busca indicou 167 publicações.

Diante disso, observou-se uma nítida evolução ao longo dos anos (de 2012 a 2021), com evidente crescimento a partir do ano de 2016, chegando-se a 32 publicações no ano de 2021. Os principais países que estão pesquisando sobre esses temas são Inglaterra, Estados Unidos, Holanda, Brasil e Austrália. O Brasil ocupa a 4ª posição do ranking, com pouco mais de 14 publicações no período. Esse fato é preponderante para a relevância do presente estudo, uma vez que possa contribuir efetivamente para construção de uma literatura própria nacional.

O software *VOSviewer* foi utilizado para a criação de um mapa baseado em coocorrência de palavras-chave, mostrando os *clusters* de pesquisa e as relações entre esses termos.

Figura 2. Mapa baseado em coocorrência de palavras-chave



Fonte: VOSviewer utilizando dados da pesquisa (2021).

Observando a figura pode-se afirmar que:

No *cluster 1* (vermelho), existe o termo com a maior força de ligação entre os demais, que é o termo inovação. Esse cluster identifica bem as publicações que trazem termos importantes para a área na qual o estudo é desenvolvido.

O *cluster 2* (verde) está ligado às teorias envolvidas na pesquisa, a partir de termos nos quais a pesquisa se desenvolverá, quais sejam, tecnologia, comunidades, bem como conhecimento e aprendizagem.

O *cluster 3* (azul) está ligado ao ambiente e ações dos atores da pesquisa, como governo, participação, coprodução e governança.

O *cluster 4* (amarelo) está relacionado aos impactos da relação entre o *cluster 1* e 3, abordando termos como transformação digital e compras públicas.

A bibliometria serviu como forma de mensurar a contribuição do conhecimento científico e tecnológico das publicações na área pesquisada, com o objetivo de fundamentar a justificativa da importância da pesquisa para a literatura nacional.

2 As comunidades virtuais de aprendizagem mediadas pelas tecnologias digitais

As universidades, escolas de governo, cursos de treinamentos e preparatórios e tantos outros se constituem em centros de aprendizagem. Com o advento da Cibercultura - conceito que ganhou repercussão a partir dos estudos de Pierre Lévy (1999) - e mais ainda agora após a crise sanitária mundial do coronavírus, as comunidades de aprendizagem buscaram se organizar utilizando de forma mais intensa a internet e as TDICs para alcançarem suas finalidades.

A Cibercultura foi definida por Lévy (1999, p. 92) como um “conjunto de técnicas (materiais e intelectuais), de práticas, de atitudes, de modos de pensamento e de valores que se desenvolvem juntamente com o crescimento do ciberespaço” e este ambiente virtual, consiste no meio de comunicação que surge da interconexão mundial dos computadores, seja pela infraestrutura material da comunicação digital ou pelos seres humanos. É o que Martino (2015, p. 11) vai chamar de “espaço de interação criado no fluxo de dados digitais em redes de computadores” e Assis (2010, p. 7) irá conceituar como “vasto e interconectado mundo de mundos que visitamos sempre que mantemos contato com a web”. Nesse espaço, a internet,

É o tecido de nossas vidas. Se a tecnologia da informação é hoje o que a eletricidade foi na Era Industrial, em nossa época a Internet poderia ser equiparada tanto a uma rede elétrica quanto ao motor elétrico, em razão de sua capacidade de distribuir a força da informação por todo o domínio da atividade humana. (CASTELLS, 2003, p. 7).

A opção do uso do ensino à distância (EaD), bem como outras formas de ensino remoto ou híbridos síncronos ou assíncronos, especialmente no seu formato *on-line*, se tornou a opção viável na crise da pandemia por manter todos separados fisicamente, mas ao mesmo tempo podendo estar presentes em um mesmo ambiente numa sala ou plataforma digital. Além de ser o formato mais viável diante das condições expostas, despertou especial atenção também por ter baixo custo. A partir disso também se fortalecem as bibliotecas digitais, surgem novos formatos de ambientes de aprendizagem e ensino, gravam-se e disponibilizam-se novas vídeo aulas, entre outros. O fortalecimento das comunidades de aprendizagem virtuais também tem muita relação com esse novo momento virtual de tempos pandêmico e pós-pandêmico. Elas se caracterizam, na definição de Cross (1998, p. 1, tradução nossa), como “grupos de pessoas envolvidas em interação intelectual com o objetivo de aprender” de forma remota, com o uso de tecnologias digitais, independente do local físico e geográfico onde estiverem, desde que estejam dispostas a se conectar por TDICs pela rede mundial de computadores na plataforma proposta.

As TDICs e demais ferramentas de colaboração garantiram que o teletrabalho fosse executado durante a mais recente pandemia. Entretanto, a maioria dos servidores públicos, habituada com uma rotina presencial, não estava preparada para uma mudança brusca em sua rotina de trabalho, como a que ocorreu em 2020. De repente, os servidores públicos se viram frente a novas rotinas e demandas que, em muitos casos destes, nem sabiam como proceder: (a) fazer a tramitação de todos os processos agora no formato eletrônico; (b) participar de reuniões *on-line* por vídeo, precisando para tal baixar aplicativos, acessar plataformas, fazer cadastros e saber acessar e interagir por elas; (c) passar a fazer praticamente todos os contatos por meio eletrônico através do uso de aplicativos de mensagem, tendo que aprender a usá-los, saber baixar, enviar ou receber documentos pelos mesmos; (d) na formalização de documentos eletrônicos, construção de relatórios e textos de forma conjunta e compartilhada no formato *on-line* em drives “nas nuvens”, mediante o uso de ferramentas de colaboração, desconhecida, até então, de muitos servidores; (e) necessidade de assinatura digital, tendo que para isso descobrir que plataforma segura permitia fazer isso e como se poderia fazer tal ação; (f) e outras tantas ações do “mundo digital” que se tornou necessária “do dia para a noite” e que gerou insegurança, ansiedade e até desespero de muitos que se viram obrigados ou até forçados a aderir a uma inovação tecnológica que não pediram e nem desejavam para participar.

A inovação exige uma postura de aprendizagem constante como condição necessária para que os processos e demais atividades na organização possam ser melhorados e executados de maneira satisfatória e,

boa parte do sucesso nas atividades organizacionais pode ser atribuída à necessária e contínua capacitação dos servidores públicos envolvidos nos inúmeros processos desenvolvidos. A essa nova tendência que foi submetido a grande parte do mundo do trabalho pela cultura digital, Satur e Silva (2021) dizem ser o novo desafio da necessidade de aprender por toda vida, da aprendizagem que nunca cessa (*lifelong learning*).

Por isso entende-se a visão de Sabbag (2007, p. 336-7 *apud* SATUR; SILVA, 2020), que a tecnologia digital deverá ser voltada à aprendizagem e alinhada às demandas, aos interesses e aos desafios da organização, dos colaboradores e dos que fazem uso dos serviços públicos. A tecnologia digital aplicada à capacitação deve ser capaz de contribuir com o conhecimento gerado e compartilhado, fornecer métricas para melhorar a aprendizagem, assim como otimizar o processo do ensinar-aprender.

Cabe destacar os avanços tecnológicos mais recentes, com a construção da nova Sociedade 5.0, que se utiliza dos conceitos de internet das coisas, inteligência artificial, robótica e *blockchain*, que servirão, num futuro próximo, para promover inclusive novas transformações digitais nas comunidades virtuais de aprendizagem. (MOREIRA *et al.*, 2020).

Então, sobretudo a partir do *home office*, que passou do contexto de mais uma opção de trabalho (antes da pandemia) para o cenário de imposição (durante a pandemia da Covid-19), como modo para a continuidade da prestação de serviços e manutenção das atividades, outra preocupação dos órgãos públicos foi a de continuar capacitando os seus servidores no pós-pandemia pois, pelo que se observa, essa nova condição de trabalho veio para ficar em muitas atividades. Em vista disso, no Brasil, de acordo com Oliveira, R. (2020, p. 2), a Escola Nacional de Administração Pública (ENAP) toma maior importância no cenário nacional, “como instituição pública de educação corporativa, focada na capacitação de servidores, que implementa estratégias necessárias para garantir a realização contínua de suas ações com finalidade educacional”.

Nesse cenário houve grande expansão da Escola Virtual de Governo (EV.G), (OLIVEIRA, Ivan, 2020, p. 3) que constitui-se em comunidade de aprendizagem virtual, uma vez que reúne as características de comunidade de aprendizagem e atinge a sua finalidade através do uso de tecnologias digitais aplicadas à capacitação, quando centraliza a divulgação, hospedagem, gestão acadêmica, emissão de certificados; atualização constante dos cursos ofertados; facilidade no acesso e plataforma intuitiva. O entendimento de seus objetivos e funcionalidades podem ser visualizados na figura a seguir.

Figura 3. *Benefícios associados à Escola Virtual do Governo*



Fonte: Teixeira e Pontes (2017).

Além de gestão acadêmica, a EV.G oferece serviços de indicadores e de ambiente virtual, através das plataformas MOOC, com cursos de curta e média duração, abertos e gratuitos, e SPOC, voltado para grupos menores e personalizados, disponibilizando recursos e tecnologia virtual com engajamento pessoal entre professores e alunos.

Os cursos ofertados pela EV.G (2021) na área de logística e compras pública são: a) Formação de pregoeiros – teoria, b) Formação de pregoeiros – prática, c) Curso básico de licitações - enfrentando (e vencendo) tabus, d) Prevenção e detecção de cartéis em licitações, e) Fundamentos da gestão da logística pública, f) Elaboração de termos de referência para contratação de bens e serviços, g) Noções introdutórias de licitação e contratos administrativos.

3 Análise do papel das comunidades virtuais de aprendizagem e do uso das TDICs na capacitação dos gestores de compras e licitações do IFPB: Resultados da pesquisa

Do total dos respondentes da pesquisa feita com os gestores de compras e licitação no Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, 40% são mulheres e 60% são homens, sendo que 50% têm idade entre 26 e 35 anos e 50% entre 36 e 50 anos. Na amostra não foi observada a presença de pessoas com 25 anos ou menos e nem pessoas com mais de 50 anos.

Quanto ao maior grau de escolaridade dos gestores de compras e licitação do IFPB, 25% possuem mestrado, 55% possuem alguma pós-graduação em nível de especialização (*lato sensu*), 10% possuem curso superior e 10% possuem o ensino médio.

O que ajuda a explicar os números consideráveis de servidores com mestrado e especialização é o plano de cargos, carreira e remuneração, que prevê o incentivo remunerado à qualificação aos servidores que buscarem grau de escolaridade superior ao exigido para o cargo,

nos moldes da Lei nº 11.091, de 12 de janeiro de 2005 (BRASIL, 2005).

Com relação ao tempo de atuação como gestor de compras, 30% atuam de 1 a 3 anos na função, 45% atuam de 4 a 7 anos, 15% de 8 a 10 anos e 10% há mais de 10 anos no cargo. Os dados da pesquisa revelam ainda que 35% dos gestores (7 deles) não foram encaminhados, quando da designação para o setor de compras das suas unidades de lotação, para que realizassem cursos específicos para o desempenho das atividades de compras e licitações. Isso indica uma fragilidade no planejamento de capacitação dos servidores do Instituto, pois esses cursos deveriam ser concluídos ou ao menos iniciados antes mesmo do servidor assumir as atribuições de gestor de compras e licitações.

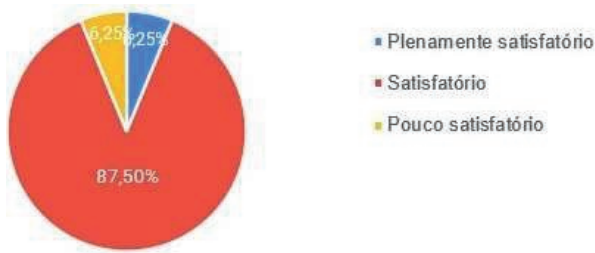
3.1 Sobre a importância da comunidade de aprendizagem EV.G

Aos pesquisados foi perguntado sobre a participação deles em curso de formação na área de logística e compras públicas ofertados pela Escola Virtual do Governo (EV.G) da ENAP. O resultado foi que 80% (16 gestores) afirmaram já ter participado de cursos na área de logística e compras públicas. A expansão da Escola Virtual de Governo (EV.G), durante o período pandêmico, foi justamente para que a ENAP continuasse ofertando cursos de capacitação aos servidores públicos, tendo em vista os ótimos resultados obtidos na educação à distância e com a facilidade de acesso e gerenciamento dos estudos permitidas pelas tecnologias digitais aplicadas à capacitação. Entretanto, 4 gestores (20%) informaram nunca ter participado de curso algum oferecido pela EV.G/ENAP, indicando que ainda não existe, no âmbito do IFPB, uma adesão total aos cursos ofertados pela EV.G por parte dos gestores de compras e licitações.

Embora haja uma política de treinamento e de aprendizagem, inclusive com a promoção da capacitação contínua, prevista no Plano de Desenvolvimento de Pessoas do IFPB, a pesquisa revelou que os diretores, coordenadores e demais chefes imediatos não estão completamente atentos ao compromisso e necessidade de capacitar os servidores continuamente antes de colocá-los à frente de uma atividade no instituto, especialmente nesse período pandêmico em que as atividades migraram para o digital. Observou-se que apenas 65% (13 gestores) tiveram incentivo das chefias imediatas para a participação nos cursos da EV.G/ENAP.

Com relação ao impacto no desempenho no setor de compras e licitações por causa do conhecimento adquirido, 87,5% estavam satisfeitos, conforme resultado apresentado no gráfico 1.

Gráfico 1. Avaliação dos conhecimentos adquiridos e do seu uso nas atividades do setor



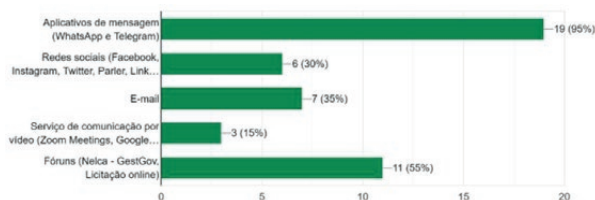
Fonte: Dados da pesquisa.

É fundamental que além de treinados e capacitados, os gestores saibam aplicar bem o conhecimento adquirido, assim como possam disseminar as boas práticas aprendidas entre os servidores que participam e colaboram na instrução processual de compras e licitações do IFPB. Na verdade, é importante não apenas ter acesso ao conhecimento, mas conseguir interpretá-lo e aplicá-lo da melhor forma, melhorar ações, transformar processos, tornando-os mais eficientes.

3.2 Sobre os cursos e o uso das tecnologias digitais inserida na EV.G/ENAP

Perguntado aos respondentes qual tecnologia foi mais utilizada durante a realização dos cursos à distância promovidos pela EV.G/ENAP, 56,25% responderam que utilizaram o desktop como tecnologia principal, enquanto 43,75% deles fizeram uso de notebook. Já quando questionados sobre quais meios virtuais eram utilizados para a aquisição de conhecimento 95% afirmaram que faziam uso dos aplicativos de mensagens (especialmente WhatsApp e Telegram). Também, 55% responderam utilizar os fóruns de discussão, como Nelca – GestGov e o Licitação Online; enquanto que 35% afirmaram que utilizam também o e-mail e apenas 15% dos participantes da pesquisa responderam que tiram dúvidas e adquirem conhecimento fazendo chamadas de vídeo com outros gestores, conforme gráfico 2.

Gráfico 2. Meios virtuais utilizados pelos gestores de compras do IFPB

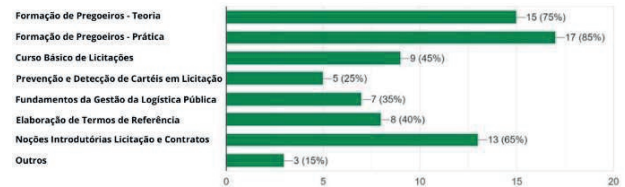


Fonte: Dados da pesquisa.

Quanto à realização de cursos desde que assumiram as funções de gestores de compras e licitação no IFPB, 43,75% responderam que realizaram entre 1 a 3 cursos, 43,75% entre 4 a 7 cursos e apenas 12,50% participaram de 8 a 10 cursos. Quando foi solicitada a informação de

quais cursos eles achavam mais importantes para as atribuições desenvolvidas no setor, as respostas foram variadas, conforme o gráfico 3.

Gráfico 3. Cursos da EV.G/ENAP mais relevantes da área de compras



Fonte: Dados da pesquisa.

Os cursos “Formação de pregoeiros – teoria” (75%), de “Formação de pregoeiros – prática” (85%) e o de “Noções introdutórias de licitação e contratos administrativos” (65%), foram os que registraram o maior interesse por parte dos gestores de compras e licitações. A preferência maior pelos cursos de pregoeiro – teoria e prática – se justifica tendo em conta ser esta a modalidade de licitação predominante nas contratações públicas no âmbito do IFPB.

A última questão abordada foi quanto ao compartilhamento dos conhecimentos adquiridos, nos diversos cursos de capacitação, dentro do IFPB, entre os gestores. A pesquisa revelou que 75% gestores tentavam compartilhar o conhecimento adquirido e apenas 25% se acomodavam, indicando que não havia compartilhamento entre os servidores do IFPB, o que alerta para a necessidade de investigação complementar sobre como o conhecimento efetivamente se propaga e quais são os fatores dificultadores que impedem uma maior disseminação dos conhecimentos adquiridos, assim como buscar promover uma gestão participativa e incentivar a cultura de compartilhamento, até porque compartilhar conhecimento consiste em um dos três processos necessários à gestão do conhecimento nas organizações.

4 Conclusão

Os gestores entrevistados representam o perfil do agente de compras do IFPB, servidores que ingressaram no IFPB no período de expansão dos *campi* e participaram ativamente dos processos relativos à instrução processual, planejamento da contratação, licitação, compra direta e até mesmo formalização de contratos e seus aditivos.

Os dados da pesquisa permitiram observar que mais de 80% desses gestores realizaram de 1 até 7 cursos, utilizando-se da comunidade virtual de aprendizagem da ENAP.

Foi observado durante as entrevistas que esses gestores reúnem a experiência de trabalhar desde o início em suas

unidades e desenvolveram a consciência da necessidade de capacitação contínua, sendo confirmado pela quantidade de cursos realizados na EV.G/ENAP. Entretanto, isso está muito presente apenas no plano individual, não existindo ainda uma cultura organizacional de aprendizagem e compartilhamento de informações, tanto que se percebeu que não há uma adesão total ao uso das comunidades virtuais de aprendizagem no âmbito do IFPB, assim como há o problema do compartilhamento do conhecimento adquirido dentro da instituição, visto por 25% dos gestores como algo deficiente ou insatisfatório.

Uma solução para esse entrave poderia ser o de investir na associação da EV.G com comunidades de prática do IFPB, que poderiam ser criadas justamente para proporcionar colaboração à distância, padronização e melhoria de documento, desenho e definição de fluxos processuais, compartilhamento de conhecimento construído e formação de uma organização aprendente.

De acordo com Costa e Gouvinhas (2005) as comunidades de prática possuem a capacidade de oferecer um ambiente de aprendizado forte, podendo abranger um mesmo setor, áreas distintas de uma organização ou até mesmo organizações diferentes. Mas, é necessário lembrar que a participação nessas comunidades normalmente é voluntária.

Não é possível afirmar que há uma quantidade exata de cursos que irá garantir o bom desempenho das funções do gestor de compras e licitações do IFPB, no entanto as respostas dos gestores revelaram as seguintes necessidades:

- 1) Capacitar o servidor antes de inseri-lo no setor de compras e licitação;
- 2) Incentivar a participação dos gestores de compras nos cursos ofertados pela EV.G;
- 3) Apoiar e acompanhar a atualização contínua dos gestores de compras;
- 4) Fomentar uma cultura de compartilhamento do aprendizado adquirido.

A realidade de vários *campi* do IFPB é a de apenas um gestor de compras e licitação para o *campus* inteiro, sendo acumuladas as atividades de planejamento das contratações, pesquisa e formação de preços, compra direta, licitação, formalização de contratos, termos aditivos e apostilamentos, publicação de extratos de contratos e suas alterações e até mesmo com a fiscalização de alguns desses contratos.

Diante desse contexto, as tecnologias digitais favorecem, pois, permitem que haja contato com outros gestores de compras e licitação, o que acaba sendo uma das alternativas mais rápida e eficiente de elucidar dúvidas, compartilhar experiências e comparar as práticas utilizadas visando ao aperfeiçoamento.

Sobre o uso das tecnologias digitais (TDICs) também é possível aferir a importância das ferramentas tecnológicas aplicadas à aprendizagem. Então, o que deve ser trabalhado é o papel pedagógico e de integração permitido por causa da tecnologia, entre os diversos agentes que fazem parte das comunidades virtuais de aprendizagem, evitando que se limite apenas a um mero aparato de comunicação (fios, tela, rede, vídeo, som, imagem).

Os resultados revelaram a boa recepção dos gestores de compras e licitação do IFPB à Escola Virtual de Governo da ENAP, some-se ainda o fato interessante que, mesmo sem o incentivo da chefia imediata, 30% dos entrevistados informaram ter participado de cursos na área de compras ofertados pela EV.G/ENAP. Assim, mais do que estabelecer um número exato de cursos a serem realizados, faz-se necessário investir em capacitação do servidor, antes de assumir funções de compras e licitação, e de forma contínua, assim como fomentar uma cultura de compartilhamento do aprendizado adquirido.

Por fim, observa-se que vivemos tempos de cultura digital e tempos líquidos em que o aqui e o agora imperam, que não do que se aprende é definitivo e para sempre e que, portanto, a atualização e a aprendizagem precisam ser promovidas o tempo todo, por toda a vida. A organização e o profissional que não se atentarem para essa realidade ficarão defasados e não conseguirão ter desempenho satisfatório nesse novo contexto cibercultura.

Referências

APUGLIESE, Marcilene Maria Enes. **Capacitação dos servidores visando eficácia dos serviços públicos: um estudo de caso do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Cubatão (IFSP)**. 2010. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2010. Disponível em: <https://www.lume.ufrgs.br/handle/10183/40407>. Acesso em: 23 dez. 2021.

ASSIS, Emanuel Cesar Pires de. Ciberespaço e Pós-Modernidade em Neurocomancer de William Gibson. *In: ENCONTRO DE ESTUDOS MULTIDISCIPLINARES EM CULTURA (ENECULT)*, 6., 2010, Salvador. **Anais [...]** Salvador: UFBA, 2010. Disponível em: <http://www.cult.ufba.br/wordpress/24841.pdf>. Acesso: 20 dez. 2021.

BRASIL. **Instrução Normativa nº 21, de 16 de março de 2020**. Estabelece orientações aos órgãos e entidades do Sistema de Pessoal Civil da Administração Pública Federal - SIPEC, quanto às medidas de proteção para enfrentamento da emergência de saúde pública de importância internacional decorrente do coronavírus (COVID-19). Disponível em: <https://www.in.gov.br/en/web/dou/-/instrucao-normativa-n-21-de-16-de-marco-de-2020-248328867>. Acesso: 21 dez. 2021.

BRASIL. **Lei nº 11.091, de 12 de janeiro de 2005.** Dispõe sobre a estruturação do Plano de Carreira dos Cargos Técnico-Administrativos em Educação, no âmbito das Instituições Federais de Ensino vinculadas ao Ministério da Educação, e dá outras providências. Brasília, DF, 13 jan. 2005. Disponível em:

http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/lei/111091.htm. Acesso: 12 jan. 2022.

BRASIL. **Lei nº 14.133, de 1º de abril de 2021.** Lei de Licitações e Contratos Administrativos. Diário Oficial da União, Poder Legislativo, Brasília, DF, Seção 1, 01 abr. 2021, p. 1. Disponível em:

http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14133.htm. Acesso: 21 dez. 2021.

CASTELLS, Manuel. **A Galáxia da Internet: reflexões sobre a internet, os negócios e a sociedade.** Tradução: Maria Luíza X. de A. Borges. Rio de Janeiro: Zahar, 2003.

CERVO, Amado Luiz; BERVIAN, Pedro Alcino. **Metodologia científica.** São Paulo: MAKRON Books, 1996.

COSTA, Paulo Eduardo de C.; GOUVINHAS, Reidson Pereira. **Proposta de uma ferramenta de gestão do conhecimento como uma comunidade de prática: o caso SEBRAE/RN.** In: ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO (ENEGEP), 25., 2005, Porto Alegre. **Anais [...].** Porto Alegre, 2005.

CROSS, Kathryn Patricia. Why learning communities? Why now? **About Campus**, [S.l.], v. 3, n. 3, 1998.

DUARTE, E. N.; SATUR, R. V.; LIRA, S. L.; SILVA, N. B. X.; LIMA, E. S. Práxis de gestão do conhecimento no ambiente das organizações no escopo da Ciência da Informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 16., 2015, João Pessoa. **Anais [...].** João Pessoa: 2015.

EV.G. Escola Virtual de Governo. **Catálogo de cursos.** 2021. Disponível em: <https://www.escolavirtual.gov.br/catalogo>. Acesso: 08 jan. 2022.

IFPB. Instituto Federal de Educação, Ciência e Tecnologia da Paraíba. **Sobre o IFPB.** 2021. Disponível em: <https://www.ifpb.edu.br/institucional/sobre-o-ifpb>. Acesso: 14 dez. 2021.

IFPB-PDI. **Plano de Desenvolvimento Institucional IFPB 2020-2024.** 2020. Sobre o IFPB. Disponível em: https://www.ifpb.edu.br/transparencia/documentos-institucionais/documentos/pdi_ifpb_2020-2024.pdf/view. Acesso: 16 dez. 2021.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos de metodologia científica.** 8. ed. São Paulo: Atlas, 2020.

LEVY, Pierre. **Cibercultura.** São Paulo. Editora 34, 1999.

MARTINO, Luís Mauro Sá. **Teoria das mídias digitais: linguagens, ambientes e redes.** 2. ed. Petrópolis, RJ: Vozes, 2015.

MOREIRA, Gabriel; SILVA, Ricardo; MATTOS, Giorgia; AIRES, Josilene. Ensino de informática: habilidades requeridas para a Sociedade 5.0. In: CONGRESSO SOBRE TECNOLOGIAS NA EDUCAÇÃO (CTRL+E), 5., 2020, Evento online. **Anais [...].** Porto Alegre: Sociedade Brasileira de Computação, 2020. p. 648-654.

OLIVEIRA, Ivan Lucas Alves. **O uso de ferramenta de autoria na implementação de conteúdo de cursos a distância da Escola Virtual Governo (EV.G).** Brasília: ENAP, 2020.

OLIVEIRA, Rafael Carvalho Rezende. **Licitações e contratos administrativos: teoria e prática.** 9. ed. Rio de Janeiro: Forense; São Paulo: Método, 2020.

PEDREIRA, Pinho. O teletrabalho. **Revista LTr**, São Paulo, v. 64, n.5, 2000.

SATUR, Roberto Vilmar; SILVA, Armando Malheiro da. Aprendizagem visando a competência em informação na sociedade em tempos de infoesfera. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 10, p. 1-22, 2020. Disponível em:

<https://periodicos.ufpb.br/index.php/pgc/article/view/48660/29748>. Acesso: 22 dez. 2021.

SILVA, P.; LIMA, D. M.; COUTO, E. S. Lives de festas nos tempos da Covid-19: arranjos, vínculos e performances. **Revista Brasileira de Pesquisa (Auto) Biográfica**, Salvador, v. 5, n. 16, p. 1503-1517, 2020.

TEIXEIRA, Natalia; PONTES, Bruno. Escola Virtual do Governo Federal: proposta de solução para a capacitação continuada a distância no serviço público federal. In: CONGRESO INTERNACIONAL DEL CLAD SOBRE LA REFORMA DEL ESTADO Y DE LA ADMINISTRACIÓN PÚBLICA (ENAP), 22., 2017. **Anais [...].** ENAP: 2017. Disponível em: <https://repositorio.enap.gov.br/handle/1/3080>. Acesso em: 12 jan. 2022.

O conceito *Linked Data* no Brasil

The Linked Data concept in Brazil

Ana Carolina Simionato Arakaki (1), Felipe Augusto Arakaki (2), Paula Regina Ventura Amorim Gonçalves (3)

(1) Universidade Federal de São Carlos (UFSCar), São Carlos, acsimionato@ufscar.br. (2) Universidade de Brasília (UnB), Brasília, felipe.arakaki@unb.br. (3) Universidade Federal do Espírito Santo (UFES), Vitória, paula.goncalvez@ufes.br

Resumo

Objetivo: Analisar o conceito utilizado nos trabalhos publicados desde 2012 indexados na Base de dados em Ciência da Informação-BRAPCI. **Metodologia:** A pesquisa é quantitativa, com caráter descritivo e exploratório, a partir da análise de 138 artigos publicados entre 2012 e 2022, na Base de Dados de Periódicos em Ciência da Informação (BRAPCI) faz o mapeamento da forma conceitual empregada para o termo *Linked Data*, tendo como base os fundamentos da Teoria do Conceito de Hjørland (1997 e 2009), buscando a compreensão como a adoção do termo foi empregada no Brasil. Para a comparação e a sistematização dos dados foi utilizado o *software* IRAMuTeQ. **Resultados:** A análise consistiu na verificação de 118 textos, separados em 251 segmentos de texto, com aproveitamento de 151 segmentos (60,16%). Dessa quantidade, foram identificadas 7.752 ocorrências (palavras, formas ou vocábulos), com 1.480 palavras distintas. Assim, foi gerada a categorização pelo *software* IRAMuTeQ a partir do termo *Linked Data* e dos termos relacionados. Identificou-se dois eixos: 1) fundamentação teórica e 2) aplicação e instrumentos. **Conclusões:** É apresentado um mapeamento do emprego do conceito de *Linked Data* nos artigos publicados nos últimos 11 anos no Brasil. O uso do *software* IRAMuTeQ possibilitou a identificação dos trabalhos nas categorias: princípios *Linked Data* e na caracterização do conceito e instrumentos, indicando que os autores estão alinhados com as definições de Bizer, Heath e Berners-Lee.

Palavras-chave: *Linked data*. Teoria do conceito. *Web* Semântica.

Abstract

Objective: To analyze the concept used in works published since 2012 indexed in the Information Science Database - BRAPCI. **Methodology:** The research is quantitative, with a descriptive and exploratory character, based on the analysis of 138 articles published between 2012 and 2022, in the Database of Periodicals in Information Science (BRAPCI) mapping the conceptual form used for the term *Linked Data*, based on the foundations of Hjørland's Theory of Concept (1997 and 2009), seeking to understand how the term was adopted in Brazil. For data comparison and systematization, the IRAMuTeQ software was used. **Results:** The analysis consisted of checking 118 texts, separated into 251 text segments, with use of 151 segments (60.16%). Of this amount, 7,752 occurrences were identified (words, forms or words), with 1,480 different words. Thus, the categorization was generated by the IRAMuTeQ software from the term *Linked Data* and related terms. Two axes were identified: 1) theoretical foundation and 2) application and tools. **Conclusions:** A mapping of the use of the *Linked Data* concept in articles published in the last 11 years in Brazil is presented. The use of the IRAMuTeQ software enabled the identification of works in the categories: *Linked Data* principles and in the characterization of the concept and instruments, indicating that the authors are aligned with the definitions of Bizer, Heath and Berners-Lee.

Keywords: *Linked data*. Concept theory. Semantic Web.

Resumen

Objetivo: Analizar el concepto utilizado en trabajos publicados a partir de 2012 indexados en el Banco de Datos de Ciencias de la Información - BRAPCI. **Metodología:** La investigación es cuantitativa, de carácter descriptivo y exploratorio, a partir del análisis de 138 artículos publicados entre 2012 y 2022, en la Base de Datos de Publicaciones Periódicas en Ciencias de la Información (BRAPCI) mapeando la forma conceptual utilizada para el término datos vinculados, con base en sobre los fundamentos de la Teoría del Concepto de Hjørland (1997 y 2009), buscando comprender cómo fue adoptado el término en Brasil. Para la comparación y sistematización de datos se utilizó el *software* IRAMuTeQ. **Resultados:** El análisis consistió en la revisión de 118 textos, separados en 251 segmentos de texto, con uso de 151 segmentos (60,16%). De esta cantidad, se identificaron 7.752 ocurrencias (palabras, formas o palabras), con 1.480 palabras diferentes. Así, la categorización fue generada por el *software* IRAMuTeQ a partir del término Datos Vinculados y términos relacionados. Se identificaron dos ejes: 1) fundamentación teórica y 2) aplicación e instrumentos. **Conclusiones:** Se presenta un mapeo del uso del concepto de Datos Vinculados en artículos publicados en los últimos 11 años en Brasil. El uso del *software* IRAMuTeQ posibilitó la identificación de obras en las categorías: principios de Datos Vinculados y en la caracterización del concepto e instrumentos, indicando que los autores están alineados con las definiciones de Bizer, Heath y Berners-Lee.

Palabras clave: Datos vinculados. Teoría del concepto. *Web* Semántica.

1 Introdução

Tornar acessíveis catálogos e dados sempre foi um desafio para profissionais que tenham como seu objeto de trabalho a informação, como os bibliotecários, os arquivistas e os museólogos.

Com a crescente quantidade de informações produzidas e disponibilizadas no ambiente *Web*, para que elas sejam acessadas, utilizadas e reutilizadas para satisfazer necessidades informacionais e/ou para a tomada de decisões, a interoperabilidade torna-se condição *sine qua non* para a troca de informações entre sistemas de forma eficaz e eficiente.

Souza e Alvarenga (2004, p. 139) definem interoperabilidade como: “[...] a possibilidade de sistemas diferentes conversarem entre si”.

Nesse contexto, Sayão e Marcondes (2008, p.134) afirmam que os conteúdos disponibilizados na Internet “[...] só serão usados e reusados se estiverem ancorados por sistemas que permitam um alto grau de Interoperabilidade”.

Em frente à nova demanda, no início do século XXI, pesquisas e publicações a respeito da *Web Semântica* sinalizam estudos e avanços nessa área.

Berners-Lee, Hendler e Lassila (2001), ao publicarem o artigo *The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities*, estabelecem os princípios para *Web Semântica*. Assim, questões de interoperabilidade encontram respaldo nos recursos da *Web*. Nesse artigo, os autores afirmam que a *Web Semântica* seria uma *Web* na qual os computadores poderiam entender o contexto das pessoas para interpretar o significado da informação. Ainda complementam que:

A *Web Semântica* não é uma *Web* separada, mas uma extensão da atual, em que a informação tem um significado bem definido, permitindo que os computadores e as pessoas trabalhem em cooperação (BERNERS-LEE; HENDLER; LASSILA, 2001, não paginado, tradução nossa).

Em 2006, Shadbolt, Berners-Lee e Hall publicam *The Semantic Web Revisited*. A partir de pesquisas, publicações e com o desenvolvimento da *Web Semântica*, surgem conceitos de dados conectados. O conceito *Linked Data* é proposto nesse mesmo ano por Tim Berners-Lee, o inventor da *World Wide Web*, em um artigo intitulado *Linked Data*, veiculado na revista científica *Scientific American*. O objetivo de Berners-Lee era tornar a *Web* mais inteligente, permitindo que os dados fossem conectados e acessados de forma eficiente por humanos e máquinas.

Berners-Lee (2006, não paginado, tradução nossa) é categórico ao declarar que:

[...] A *Web Semântica* não é apenas sobre a colocação de dados na *web*. É sobre como fazer ligações, de modo que uma pessoa ou máquina possa explorar a *web* de dados. Com o *Linked Data*, quando você tem um pouco dele, você pode encontrar outros dados relacionados. (BERNERS-LEE, 2006, não paginado, tradução nossa).

Desde então, o conceito de *Linked Data* ganhou popularidade entre os pesquisadores e desenvolvedores de tecnologia da informação. Vários projetos e iniciativas surgiram em todo o mundo para aplicar os princípios do *Linked Data* a diferentes áreas, como ciência, cultura, governo, saúde e comércio.

Em 2007, a organização sem fins lucrativos *Open Data Consortium* foi fundada com o objetivo de promover a publicação de dados abertos e conectados na *Web*. Desde então, várias outras organizações e iniciativas surgiram para promover e incentivar a adoção do *Linked Data*, incluindo a *World Wide Web Foundation* e o *Linked Data Research Centre*.

No ano de 2009, Bizer, Heath e Berners-Lee usaram o termo *Linked Open Data* (Dados Abertos Conectados) definido como um conjunto de boas práticas para publicar, de forma aberta, e interconectar conjuntos de dados estruturados na *Web* com o intuito de criar uma *Web* de Dados.

A partir do uso dos termos *Linked Data* e *Linked Open Data* para sinalizar a associação dos dados no ambiente *Web*, pesquisadores da Ciência da Informação, desde 2012, têm publicado o resultado de suas pesquisas no Brasil.

Nesse cenário, este estudo tem como objetivo mapear o conceito utilizado para o termo *Linked Data*, verificando quais autores são citados, ou seja, qual a base teórica utilizada para a conceituação desses termos nos artigos publicados na Ciência da Informação. Nossa pesquisa está respaldada em Dias e Naves (2007, p. 63-64), que definem o termo conceito como o “[...] conjunto de características, que são os elementos dos conceitos, e traduzem os atributos das coisas designada [...]”, e também em Barros (2016, p. 26), que afirma que o conceito é a ideia bem delineada, “[...] evocada a partir de uma palavra ou expressão verbal que passa, desde então, a ser operacionalizada sistematicamente em um campo do saber”.

Ressaltamos a relevância deste trabalho pelo fato de que as tecnologias utilizadas para o desprendimento e uso do *Linked Data* envolvem a criação de conjuntos de dados conectados, que permitem uma recuperação e uso mais eficiente das informações, o que requer um entendimento tanto dos aspectos tecnológicos quanto dos aspectos teóricos e práticos da organização e representação da informação.

2 Referencial teórico

Considerando os desafios interdisciplinares já mencionados em relação ao conceito de *Linked Data*, é necessário abordá-lo a partir de uma perspectiva que contemple as diversas áreas do conhecimento envolvidas, como a Ciência da Computação, a Ciência da Informação, a Biblioteconomia, entre outras.

No âmbito da Biblioteconomia e Ciência da Informação, há diversos autores que tratam e refletem sobre o tema dos conceitos, dentre os quais destacam-se, mas não se limitam a estes: Ingetraud Dahlberg, Brian Vickery, Maria Antonieta R. Piedade, Maria Luiza Almeida Campos, Hagar Espanha Gomes, Nair Yumiko Kobashi, Mário Barité e Birger Hjørland. (MACULAN; LIMA, 2017, não paginado).

Nesse contexto, Birger Hjørland (1997), de forma pragmática, propôs a teoria do conceito, que tem como objetivo analisar e explicar como os conceitos são criados, definidos e utilizados na comunicação científica.

Nessa abordagem busca-se o significado que há no uso dos termos em um universo específico do discurso. O conhecimento é analisado em sua dimensão social, considerando a dinâmica contextual e instável do processo de comunicação entre indivíduos. A importância da abordagem pragmática consiste na possibilidade de construir uma estrutura contextual e flexível para a representação do conhecimento. (MELO; BRÄSCHER, 2014, p. 78).

De acordo com Hjørland (1997), um conceito é uma unidade de pensamento que representa uma ideia abstrata ou uma categoria de objetos, eventos ou fenômenos. Ele acredita que os conceitos são criados pela interação entre os indivíduos e a sociedade, e sua definição e uso são influenciados pelo contexto cultural e histórico em que são utilizados.

Em complemento à sua publicação de 1997, Hjørland, em 2009, esclarece que os conceitos são construções significativas dinâmicas e negociadas que categorizam o mundo de acordo com as teorias e os interesses envolvidos. O autor argumenta que os conceitos não podem ser entendidos isoladamente dos interesses e teorias que motivaram sua criação, já que há, geralmente, concepções e conceitos concorrentes em todos os domínios do conhecimento.

Hjørland (2009) também enfatiza a importância da análise conceitual na pesquisa científica, afirmando que a compreensão clara e precisa dos conceitos é fundamental para o avanço do conhecimento em uma determinada área. Propõe uma abordagem crítica à análise conceitual, que envolve examinar cuidadosamente as definições existentes e as práticas de uso de conceitos, avaliando sua adequação para representar e explicar os fenômenos em questão.

Em resumo, a teoria do conceito de Hjørland (1997) destaca a importância dos conceitos na construção do conhecimento científico, bem como a influência do contexto cultural e histórico na sua criação, definição e uso. Sugere uma abordagem crítica e reflexiva à análise conceitual, tendo em vista aprimorar a compreensão e a aplicação dos conceitos na pesquisa científica.

De acordo com Hjørland (2009), a principal função dos conceitos é representar aspectos da realidade em sistemas simbólicos, como o pensamento e a linguagem, a fim de facilitar sua comunicação e reflexão. No entanto, essa função estabilizadora não impede que os conceitos possam mudar ao longo do tempo. Por essa razão, é necessário considerar o contexto temporal ao estudar os conceitos e reconhecer que o significado de um termo não é fixo e imutável, mas sim dinâmico e evolui conforme a construção social do conhecimento.

Hjørland (2007) entende que os relacionamentos entre conceitos derivam das teorias e epistemologias que afetam de maneira variada todos os campos do conhecimento. Por isso o autor rejeita a ideia de que relações semânticas possam ser estabelecidas por meio da aplicação de pressupostos universais e da descrição objetiva de características dos conceitos. Hjørland (2007) afirma que diferentes abordagens para a organização do conhecimento resultam em diferentes perspectivas sobre os relacionamentos semânticos entre os conceitos.

Hjørland (2007) lembra também que as relações semânticas estão relacionadas com uma tarefa ou situação específica e nem todos os usuários de um conjunto de relações semânticas compartilharão a mesma visão sobre a equivalência dos termos. No entanto, argumenta que não é possível projetar sistemas para mais de um usuário ou situação com base em uma visão individualista ou idiossincrática de conceito e semântica. É necessária a criação de princípios mais estáveis que determinem as relações semânticas. Para isso é preciso desenvolver uma teoria semântica sobre o significado de palavras como forma de práticas tipificadas. O conhecimento sobre a semântica das práticas tipificadas pode, então, ser utilizado por pesquisadores da informação para incluir ou excluir determinados documentos.

Nesse cenário, foi analisado o termo *Linked Data*, que é um conjunto de práticas recomendadas para publicar e conectar dados na *Web* de maneira estruturada e padronizada (BIZER; HEATH, 2011; ISOTANI; BITTENCOURT, 2015), tornando-os mais acessíveis e úteis para humanos e máquinas. Essas práticas são baseadas em tecnologias da *Web Semântica*, como o *Resource Description Framework* (RDF), que permite que os dados sejam expressos em formatos legíveis por máquina e interconectados com outros dados na *Web*.

Diversos autores já discutiram e apresentaram o conceito de *Linked Data*, como Berners-Lee (2006); Berners-Lee (2009); Bizer, Heath e Berners-lee (2009); Heath (2011); Baker et al. (2011), Heath e Bizer (2011); Marcondes (2012); Hooland e Verborgh (2014); Santarém Segundo (2015); Arakaki (2016); Rautenberg (2019); Serra e Santarém Segundo (2021); Jesus, Castro e Ramalho (2021); e Arakaki e Triques (2022).

Para que diferentes conjuntos de dados possam ser associados e combinados para gerar novos conhecimentos, é necessária a criação de identificadores únicos, as URIs, para cada recurso de dados, e a utilização de padrões e vocabulários comuns para descrever e conectar esses recursos.

Algumas aplicações do *Linked Data* incluem a criação de bancos de dados abertos e conectados em áreas, como saúde, ciência, governo, cultura e educação, além de melhorar a pesquisa e a análise de dados na *Web*, nas redes sociais e no *Big Data*.

3 Procedimentos metodológicos

O estudo consiste em uma pesquisa de natureza quantitativa, com caráter descritivo e exploratório. A partir dos objetivos propostos, buscou-se mapear, de forma conceitual, o termo *Linked Data*, com base nos fundamentos da Teoria do Conceito de Hjørland (1997, 2009), procurando compreender como a tradução do termo foi adotada no Brasil pela Ciência da Informação e pela comunicação científica da área.

Durante a primeira etapa do trabalho, os procedimentos metodológicos constituíram-se na identificação de artigos que continham o termo *Linked Data*. A base de dados escolhida foi a Base de Dados de Periódicos em Ciência da Informação (BRAPCI). A coleta dos dados foi realizada no dia 10 de março de 2023.

Na coleta de dados, foram utilizadas as possíveis traduções para o termo *Linked Data*. Assim, foram recuperadas ocorrências dos termos: dato vinculado, dado vinculado, *Linked Open Data*, *Linked Data*, dado ligado, dado aberto ligado, dado conectado e dado aberto conectado. Para os termos dado entrelaçado e dado interconectado, não houve nenhuma ocorrência.

Nesse cenário, foram recuperados 240 artigos. Desses artigos recuperados, foram excluídos os duplicados e também os documentos que não apresentavam o termo *Linked Data* ou suas variações de tradução no título nas palavras-chave ou no resumo.

Dessa forma, ao final da aplicação desses critérios, foram identificados 138 artigos. Após uma segunda análise, foram excluídos trabalhos em eventos e textos que foram recuperados em outro idioma, totalizando 118 artigos para a composição do corpus de análise deste estudo.

Os dados coletados durante as avaliações foram compilados em planilhas para facilitar a análise dos resultados, além da preservação para futuros reusos. Com isso foi possível realizar comparações e sistematizações utilizando o *software* IRaMuTeQ.

O IRaMuTeQ, um acrônimo para *Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*, é um *software* livre ligado ao pacote estatístico R para análises de conteúdo, lexicometria e análise do discurso. Foi lançado em 2008 e desenvolvido pelo *Laboratoire d'Études et de Recherches Appliquées en Sciences Sociales* (LERASS) da Universidade de Toulouse, na França.

Entre as funcionalidades do *software* IRaMuTeQ estão: busca de especificidade a partir da segmentação de texto definida; definição de classes conforme o agrupamento de segmentos de texto; estatísticas clássicas de texto; análise de similaridade nas formas completas de um corpus recortado em segmentos de texto.

O algoritmo do IRaMuTeQ automatiza o método de classificação hierárquica descendente, método proposto por Reinert (1983). Esta classificação é realizada de três maneiras: 1) classificação simples no texto: os textos permanecem na íntegra e a classificação permite o agrupamento pelos textos mais próximos, 2) classificação simples em segmento de texto: esta maneira refere-se ao agrupamento pelos segmentos de texto e 3) classificação dupla para os grupos de segmento de texto: esta classificação é realizada em duas tabelas nas quais as linhas não são mais segmentos de texto, mas grupos de segmentos de texto. O mesmo processamento é assim feito duas vezes, mas alterando o número de formulários ativos por RST.

Desse modo, foi utilizado segmentos de texto para a análise do conceito do termo *Linked data*. Com o uso do *software*, foi extraído o conceito apresentado do termo *Linked Data* nos artigos selecionados. Contudo, antes da execução pela ferramenta, foi necessária uma padronização dos dados (ou higienização dos dados), isto é, alguns termos compostos tiveram que ser adicionados com o uso de um travessão para o programa entender que se referem a um único termo. Quanto às definições que estavam em outro idioma, foi feita uma tradução livre para que os dados pudessem ser comparados. Alguns caracteres tiveram que ser excluídos, como aspas simples, aspas compostas, hífen, colchetes, chaves, asteriscos, ponto de interrogação e exclamação. O arquivo foi salvo em *txt*. Após a curadoria, os dados foram preparados para análise no *software* IRaMuTeQ.

4 Análise dos dados e resultados

A análise consistiu na verificação de 118 textos, separados em 251 segmentos de texto, com

aproveitamento de 151 segmentos (60, 16%). Dessa quantidade, foram identificadas 7.752 ocorrências (palavras, formas ou vocábulos), com 1.480 palavras distintas.

A Figura 1 (Apêndice A) apresenta a categorização gerada pelo IRaMuTeQ a partir do termo *Linked Data* e dos termos relacionados. As palavras foram divididas pelo IRaMuTeQ em dois eixos. Para melhor compreensão, foram identificadas as principais temáticas abordadas pelo recorte do programa e, assim, os eixos foram denominados como: Eixo 1 (fundamentação teórica) e Eixo 2 (aplicação e instrumentos).

Os dois eixos foram divididos em categorias e subcategorias. Pelo IRaMuTeQ, as divisões são denominadas como classes, contudo adotamos, terminologicamente, categoria, em virtude da coerência com os trabalhos de Hjørland.

O primeiro eixo apresenta as categorias 1 e 2, subordinadas às categorias 2, 2.1 e 2.2. O segundo eixo constitui a categoria 3. Da mesma forma, foram identificados os trabalhos que o programa utilizou para realizar a classificação e, assim, as categorias e subcategorias foram denominadas conforme o esquema a seguir:

1) *Fundamentação teórica*

categoria 1 - princípios *Linked Data* (classe 1)

categoria 2 - caracterização

subcategoria 2.1 - definição (classe 2)

subcategoria 2.2 - benefícios (classe 3)

2) *Aplicação e instrumentos*

categoria 3 - instrumentos e tecnologias (classe 4)

Na Figura 1, observa-se, também, que a categoria 1, denominada princípios *Linked Data*, teve 23,2% das palavras. Foram destacadas: Bizer, Heath, global, fonte, diferente, espaço, *links*, item, grafo, *software*, sistema, formar, conexão, organização, possível, estruturado.

Diante da primeira categoria gerada, foram identificados os autores Bizer e Heath entre as principais ocorrências, o que mostra que esse conjunto de termos está apoiado nesses autores e em seus estudos (BIZER; HEATH; BERNERS-LEE, 2009). Assim, as palavras elencadas estabelecem noções gerais do que é *Linked Data*, ou seja, uma fundamentação do conceito do termo.

Na categoria 2, é evidenciada a teoria de Hjørland (2009), considerando as associações do conceito. A categoria 2, identificada por caracterização, é subdividida. A subcategoria 2.1 — definição — teve 37,1% de ocorrência. Os principais termos que compuseram essa categoria foram: semântico, conceito, disponibilizar, *Linked Data*, proposta, dado, grande,

contexto, apresentado, semântica, criar, conjunto, denominar, serviço, pesquisa e ferramenta.

Entende-se que a categoria gerada e as palavras evidenciadas repercutem na concepção do conceito do termo *Linked Data*. Observa-se, ainda, um detalhamento mais aprofundado da fundamentação, relacionando o *Linked Data* com aspectos semânticos. Essa é uma proposta, associada a um serviço, pesquisa e ferramenta.

Na classe 3 da Figura 1, foram registrados 17,2% de ocorrências. Os termos que tiveram destaque foram: apresentar, facilitar, processo, acesso, interoperabilidade, autor, humano, criação não base, busca, agente, coleção, descobrir, significado, representação. Essas duas classes possuem forte relação com a classe 1 e apresentam termos que ajudam a entender melhor o que é o *Linked Data*.

A partir das palavras citadas relacionadas com a subcategoria 2.2 — benefícios — (classe 3) acima citadas, entende-se que esta categoria trata da avaliação e do emprego do ambiente com dados conectados, isto é, os benefícios, as vantagens, as possibilidades e a democratização de uso e reúso do *Linked Data*.

O eixo 2 e categoria 3, denominada como instrumentos e tecnologias, correspondeu a 22,5% dos termos encontrados. As palavras foram: URI, identificador, Isotani, rdf, *description*, *identifier*, *World Wide Consortium*, Bittencourt, termo, triplo, *framework*. Destacaram-se autores como Isotani, Bittencourt e o W3C Consórcio. Esse conjunto de termos está associado aos aspectos tecnológicos para publicação de dados em *Linked Data*.

A categoria 3 registra os autores Isotani e Bittencourt (2015) e também o emprego das tecnologias desenvolvidas para a representação e modelagem das triplas e conexão dos dados, ou seja, as ferramentas e tecnologias necessárias para publicação dos dados em *Linked Data*.

A partir da coleta realizada, foi possível identificar a separação entre as categorias das palavras geradas, como observado na Figura 2 (Apêndice B), que apresenta uma associação e categorização dos termos da Figura 1 que foram divididos em quatro quadrantes: A, B, C e D.

Observa-se, na organização do sistema, que os termos estabelecidos na categoria 2 estão alocados ao centro do quadrante (palavras em verde). Isso revela a importância dessa categoria, que apresenta o cerne do conceito estudado. Pode-se perceber ainda com essa divisão a forte relação das subcategorias 2.1 e 2.2 que estão predominantemente estabelecidas nos quadrantes A e B. No quadrante C, o foco está nos termos da categoria 1 e o quadrante D apresenta os termos da categoria 3.

As relações semânticas elucidadas por Hjørland (2009) são ilustradas pelos grandes agrupamentos da Figura 3,

como consta no Apêndice C. Na figura, é possível verificar três grandes agrupamentos. O primeiro corresponde aos termos em volta dos termos dado e *Linked Data*. Destaca-se que o primeiro agrupamento apresenta termos que estão relacionados com o termo *Linked Data*. O segundo agrupamento está associado ao termo *Web* e mostra vocábulos que envolvem questões de ambiente de aplicação, como a *Web*, e dos recursos tecnológicos do *Linked Data* de uma forma mais geral. Por fim, o terceiro agrupamento apresenta-se em torno do termo URI, ou seja, algumas tecnologias necessárias para publicação dos dados em *Linked Data*.

De forma isolada, a figura ressalta a sistematização adotada para os trabalhos que não conceituam o *Linked Data* ou ainda que não citam autoridade sobre o assunto. A sistemática não define o termo, porém demonstra todos os trabalhos que não apresentaram o conceito de *Linked Data*.

5 Considerações finais

Como resultado deste estudo, tem-se o mapeamento do emprego do conceito de *Linked Data* nos artigos publicados nos últimos 11 anos no Brasil. O uso do software IRAMuTeQ possibilitou a identificação dos trabalhos nas categorias: princípios *Linked Data* e na caracterização do conceito e instrumentos, indicando que os autores estão alinhados com as definições de Bizer, Heath e Berners-Lee. Há também um entendimento da importância do *Linked Data*, quando se verifica o uso dos termos: facilitar, acesso, interoperabilidade, representação para o entendimento do conceito, vantagens, possibilidades, democratização do uso e reúso quando é utilizado o *Linked Data*.

O estudo também apontou que 21 trabalhos não definem o termo *Linked Data* e 39 artigos não citam os autores, na definição dos conceitos, de forma direta nem indireta, o que causa preocupação.

A relação de palavras com os termos aponta o entendimento do conceito do *Linked Data* entre os pesquisadores da área, bem como o ambiente de aplicação e as tecnologias necessárias.

Além desses autores, foram identificados nomes de pesquisadores brasileiros que foram citados mais de três vezes por seus estudos sobre *Linked Data* como: Arakaki (2016, 2017) e Santarém Segundo (2015, 2016, 2017 e 2018).

Por fim, observa-se a importância dos estudos sobre os conceitos utilizados para o termo *Linked Data* para fundamentar sua utilização, visto que esse entendimento é relevante para a criação dos conjuntos de dados e, conseqüentemente, para a recuperação, uso e reúso da informação.

Para trabalhos futuros, sugerem-se: pesquisas sobre o estabelecimento de características para o termo *Linked Open Data*; observação da frequência do uso de outros termos, como dados ligados e/ou dados vinculados; e a verificação, ao longo dos anos, se houve aproximação de um consenso na literatura sobre a definição do termo e principalmente, propiciar um panorama internacional com o uso de fontes internacionais para coleta de dados.

Referências

- ARAKAKI, A. C. A.; TRIQUES, M. L. A convergência da curadoria digital com os princípios Linked data: o exemplo Linked Logainm. **Fronteiras da Representação do Conhecimento**, [s. l.], v. 1, p. 56, 2022.
- ARAKAKI, A. C. S. Família FR, BIBFRAME e Linked Data: integração de dados bibliográficos. In: ENCONTRO DE REPRESENTAÇÃO DOCUMENTAL, 1., 2017, São Carlos. **Anais [...]**. São Carlos: UFSCar, 2017.
- ARAKAKI, F. A. **Linked Data**: ligação de dados bibliográficos. 2016. 144 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista ‘Júlio de Mesquita Filho’, Marília/SP, 2016. Disponível em: <https://repositorio.unesp.br/handle/11449/147979>. Acesso em: 29 mar. 2022.
- BAKER, T.; BERMÈS, E.; COYLE, K.; DUNSIRE, G.; ISAAC, A.; MURRAY, P.; MURRAY, M.; SCHNEIDER, J.; SINGER, R.; SUMMERS, E.; WAITES, W.; YOUNG, J.; ZENG, M. **Library Linked Data Incubator Group Final Report**. [s. l.]: W3C Incubator Group Report, 2011. Disponível em: <http://www.w3.org/2005/Incubator/ld/XGR-ld-20111025/>. Acesso em: 29 mar. 2022.
- BARROS, J. D. **Os conceitos: seus usos nas ciências humanas**. Petropolis: Vozes, 2016.
- BERNERS-LEE, T. **Linked Data**: Design Issues. [s. l.]: W3C, 2006. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 25 mar. 2023.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web. **Scientific American**, [s. l.], May 2001, p. 29-37.
- BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data: the story so far. **International Journal on Semantic Web and Information Systems**, [s. l.], v. 5, n. 3, p. 1-22, 2009. Disponível em: <http://tomheath.com/papers/bizer-heath-berners-lee-ijswis-linked-data.pdf>. Acesso em: 25 mar. 2023.
- DIAS, E. W.; NAVES, M. L. **Análise de assunto**: teoria e prática. Brasília: Thesaurus, 2007.
- HEATH, T. **Linked Data**: connect distributed data across the Web. [s. l.]: Linked Data community, 2011. Disponível: <http://linkeddata.org/home>. Acesso em: 29 mar. 2022.
- HEATH, T. T.; BIZER, C. **Linked data**: evolving the web into a global data space. Berlim: Morgan & Claypool, 2011. 136 p. Disponível em: <http://linkeddatabook.com/editions/1.0/#htoc8> Acesso em: 15 jul. 2020.

HJØRLAND, B. Concept theory. **Journal of the American Society for Information Science and Technology**, [s. l.], v. 60, n. 8, p. 1519-1536, 2009. Disponível em: <https://doi.org/10.1002/asi.21082>.

HJØRLAND, B. **Information seeking and subject representation: an activity-theoretical approach to information science**. Westport: Greenwood Press, 1997.

HJØRLAND, B. Semantics and knowledge organization. **Information Science and Technology**, [s. l.], v. 41, n. 1, p. 367-405, 2007. Disponível em: <https://doi.org/10.1002/aris.2007.1440410115>.

HOOLAND S. van; VERBORGH, R. **Linked Data for Libraries, Archives and Museums: How to Clean, Link and Publish your Metadata**. Chicago (USA): American Library Association Editions, 2014.

ISOTANI, S., BITTENCOURT, I. I. **Dados abertos conectados**. [s. l.]: Editora Novatec, 2015.

JESUS, A. F.; CASTRO, F. F.; RAMALHO, R. A. S. O papel das bibliotecas no Linked Data. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [s. l.], v. 26, p. 01-21, 4 jan. 2021. Disponível em: <http://dx.doi.org/10.5007/1518-2924.2021.e75909>. Acesso em: 23 mar. 2023.

MACULAN, B. C. M. dos S.; LIMA, G. A. B. de O. Buscando uma definição para o conceito de “conceito”. **Perspectivas em Ciência da Informação**, [s. l.], v. 22, p. 54-87, 2017. Disponível em: <https://doi.org/10.1590/1981-5344/2963>. Acesso em: 23 mar. 2023.

MARCONDES, C. H. “Linked data” – dados interligados - e interoperabilidade entre arquivos, bibliotecas e museus na web. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [s. l.], v. 17, n. 34, p. 171-192, 9 ago. 2012. Universidade Federal de Santa Catarina (UFSC). Disponível em: <http://dx.doi.org/10.5007/1518-2924.2012v17n34p171>. Acesso em: 23 mar 2023.

MELO, M. A. F.; BRÄSCHER, M. Termo, conceito e relações conceituais: um estudo das propostas de Dahlberg e Hjørland. **Ciência da Informação**, [s. l.], v. 43, n. 1, 2014. Disponível em: <https://doi.org/10.18225/ci.inf.v43i1.1419>.

RAUTENBERG, S. Web Semântica e Bibliografia: um estudo para a publicação de registros bibliográficos como dados abertos conectados. **Em Questão**, [s. l.], p. 313-330, 17 out. 2019. Faculdade de Biblioteconomia Comunicação. Disponível em: <http://dx.doi.org/10.19132/1808-5245250.313-330>. Acesso em: 23 mar 2023.

REINERT, A. Une méthode de classification descendante hiérarchique: application à l'analyse lexicale par contexte. **Cahiers de l'Analyse des Données**, [s. l.], v. 8, n. 2, p. 187-198, 1983. Disponível em: http://www.numdam.org/item/CAD_1983__8_2_187_0.pdf. Acesso em: 15 maio 2023.

SANTARÉM SEGUNDO, J. E. O uso de elementos semânticos no processo de recuperação da informação em ambientes digitais. **Texto Digital**, Florianópolis, v. 13, n. 2, p. 93-111, jul./dez. 2017.

SANTARÉM SEGUNDO, J. E. Web semantica, dados ligados e dados abertos: uma visão dos desafios do brasil frente às iniciativas. **Tendências da Pesquisa Brasileira em Ciência da Informação**, [s. l.], v. 8, n. 2, 2015. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/119595>. Acesso em: 23 mar. 2023.

SANTARÉM SEGUNDO, J. E. Web Semântica: fluxo para publicação de dados abertos e ligados. **Informação em Pauta**, Fortaleza, v. 3, n. especial, p. 117-140, 26 nov. 2018.

SANTARÉM SEGUNDO, J. E.; SIMIONATO, A. C. Uma abordagem sobre a estrutura do geonames e suas contribuições para o linking open data. **Informação & Tecnologia**, [s. l.], v. 3, n. 1, p. 117-137, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/40979>. Acesso em: 23 mar. 2023.

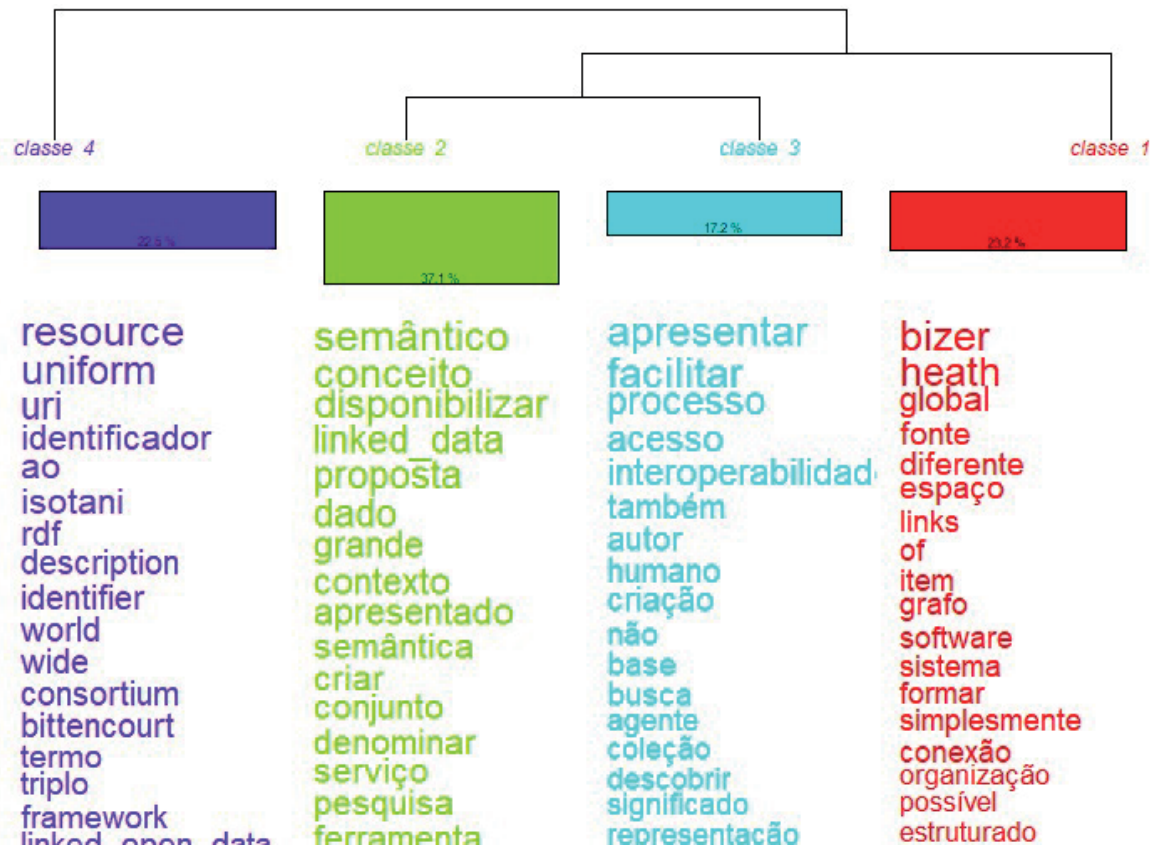
SAYÃO, L. F.; MARCONDES, C. H. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **Transinformação**, Campinas, v. 20, n. 2, p. 133-148, 2008. Disponível em: <https://periodicos.puc-campinas.edu.br/transinfo/article/view/6254>. Acesso em: 20 mar. 2023.

SERRA, L. G.; SANTAREM SEGUNDO, J. E. Dos silos de dados à Web dos dados: bibliotecas e o linked data. **Informação & Informação**, Londrina, v. 26, n. 2, p. 625, 8 jul. 2021. Disponível em: <http://dx.doi.org/10.5433/1981-8920.2021v26n2p625>. Acesso em: 23 mar 2023.

SHADBOLT, N.; BERNERS-LEE, T.; HALL, W. The Semantic Web Revisited, **IEEE Intelligent Systems**, [s. l.], v. 21, n. 3, p. 96-101, 2006.

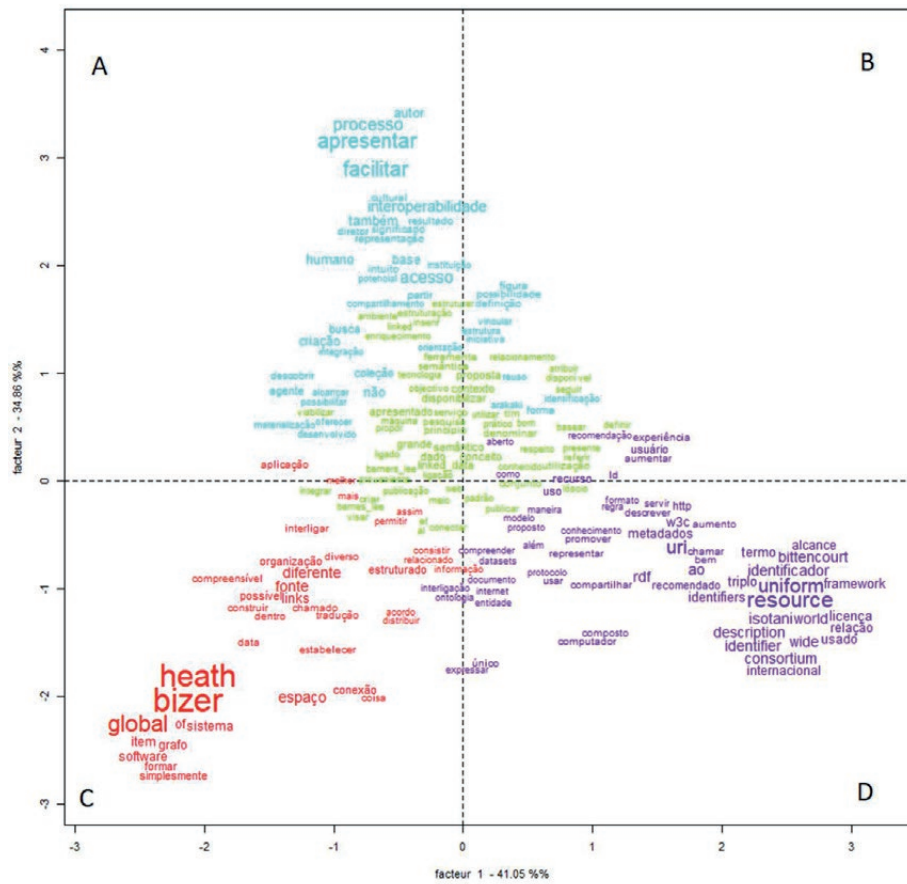
SOUZA, R. R.; ALVARENGA, L. A web semântica e suas contribuições para a ciência da informação. **Ciência da Informação**, [s. l.], v. 33, n. 1, 2004. Disponível em: <https://revista.ibict.br/ciinf/article/view/1077>. Acesso em: 20 mar. 2023.

Apêndice A – Figura 1: Categorias geradas pelo IRaMuTeQ



Fonte: Dados de pesquisa.

Apêndice B – Figura 2: Dispersão das palavras geradas pelo IRaMuTeQ



Fonte: Dados de pesquisa.

Construção de modelo para enriquecimento semântico de imagem biomédica: metodologia para extração de conceitos

Construction of a model for the semantic enrichment of biomedical images: methodology for extracting concepts

Jóice Cleide Cardoso Ennes de Souza (1), Sergio de Castro Martins (2), Elan Cardozo Paes de Almeida (3), Matheus Souza da Silva (4)

(1) Universidade Federal Fluminense (UFF), Rua Lara Vilela 126, São Domingos, Niterói/RJ, joicecardoso@id.uff.br, (2) Universidade Federal do Rio de Janeiro (UFRJ), Av. Pasteur 250, Urca, Rio de Janeiro/RJ, sergiomartins@facc.ufrj.br. (3) Universidade Federal Fluminense (UFF), Rua Dr. Silvio Henrique Braune 22, Centro, Nova Friburgo/RJ, elancardozo@id.uff.br. (4) Universidade Federal Fluminense (UFF), Rua Lara Vilela 126, São Domingos, Niterói/RJ, matheuss@id.uff.br

Resumo

Objetivo: Este trabalho tem como objetivo desenvolver uma metodologia para embasamento de modelo de ontologia aplicada no âmbito da área biomédica, através de processamento de imagens de lâminas histopatológicas, de modo a extrair categorias de conceitos e entidades que irão compor a ontologia. **Metodologia:** Revisão bibliográfica nacional e internacional sobre aplicação de ontologias na área biomédica, e com o aporte teórico da Teoria do Conceito de I. Dahlberg, dentre outras teorias para construção de ontologias, como Moreiro González, foram analisadas lâminas e laudos histopatológicos de modo a se extrair os conceitos necessários para o estabelecimento de classes e entidades que irão compor um modelo esquemático para ontologia aplicada. As análises tiveram como parâmetro as descrições macroscópicas dos objetos selecionados. **Resultados:** Os conceitos extraídos mediante análise foram categorizados nas seguintes classes: dimensões, cor, consistência, localização, formato e lesão. Estas classes compreendem todas as ocorrências histopatológicas observadas nos laudos, que irão se constituir em valores ou atributos destas classes. **Conclusão:** A partir da metodologia de extração de conceitos, foi possível identificar aqueles presentes nos laudos anatomopatológicos e determinar as classes com base na literatura da Patologia, mais precisamente com os parâmetros adotados pelos patologistas na descrição macroscópica de lâminas histopatológicas. Uma vez definidas as classes e conceitos, será possível a realização de análises para o estabelecimento de relações entre atributos e classes que permitirão o desenvolvimento de uma ontologia aplicada de relação causal para auxílio no diagnóstico, seja no âmbito de pesquisa quanto no âmbito pedagógico.

Palavras-chave: Tecnologias Semânticas. Ontologia Aplicada. Histopatologia. Teoria do Conceito.

Abstract

Objective: The objective of this work is to develop a methodology to base an ontology model applied in the biomedical area, through image processing of histopathological slides, in order to extract categories of concepts and entities that will compose the ontology. **Methodology:** After a national and international bibliographic review on the application of ontologies in the biomedical area, and with the theoretical support of the Theory of Concept by I. Dahlberg, among other theories for the construction of ontologies, such as M. González, slides and reports were analyzed histopathology in order to extract the necessary concepts for the establishment of classes and entities that will compose a schematic model for applied ontology. The analysis had as a parameter the gross descriptions of the selected objects. **Results:** The concepts extracted through analysis were categorized into the following classes: dimensions, color, consistency, location, shape and lesion. These classes include all histopathological occurrences observed in the reports, which will constitute values or attributes of these classes. **Conclusion:** With the concept extraction methodology, it was possible to identify those present in the anatomopathological reports and determine the classes based on the pathology literature, more precisely with the parameters adopted by pathologists in the gross description of histopathological slides. Once defined the classes and concepts, it will be possible to establish relationships between attributes and classes and to develop an applied ontology of causal relationship to aid in diagnosis, as in the pedagogical context.

Keywords: Semantic Technologies. Ontologies. Histopathology. Theory of Concept.

Resumen

Objetivo: El objetivo de este trabajo es desarrollar una metodología para fundamentar un modelo de ontología aplicado en el área biomédica, a través del procesamiento de imágenes de láminas histopatológicas, con el fin de extraer categorías de conceptos y entidades que compondrán la ontología. **Metodología:** Luego de una revisión bibliográfica nacional e internacional sobre la aplicación de ontologías en el área biomédica, y con el apoyo teórico de la teoría de conceptos de I. Dahlberg, entre otras teorías para la construcción de ontologías, como la de M. González, diapositivas Se analizaron informes e histopatología con el fin de extraer los

conceptos necesarios para el establecimiento de clases y entidades que compondrán un modelo esquemático para la ontología aplicada. Los análisis tuvieron como parámetro las descripciones macroscópicas de los objetos seleccionados. **Resultados:** Los conceptos extraídos a través del análisis fueron categorizados en las siguientes clases: dimensiones, color, consistencia, ubicación, formato y lesión. Estas clases incluyen todas las ocurrencias histopatológicas observadas en los informes, que constituirán valores o atributos de estas clases. **Conclusión:** A partir de la metodología de extracción de conceptos, fue posible identificar aquellos presentes en los laudos anatomopatológicos y determinar las clases con base en la literatura de la Patología, más precisamente con los parámetros adoptados por los patólogos en la descripción macroscópica de láminas histopatológicas. Una vez definidas las clases y conceptos, será posible la realización de análisis para el establecimiento de relaciones entre atributos y clases permitirán el desarrollo de una ontología aplicada de relación causal para ayudar en el diagnóstico, sea en el ámbito de investigación como en el ámbito pedagógico.

Palabras clave: Tecnologías Semánticas. Ontología aplicada. Histopatología. Teoría del Concepto.

1 Introdução

Com o avanço das tecnologias, uma grande quantidade de dados e informações vem sendo gerada diariamente, e a heterogeneidade dos dados produzidos representam novos desafios no que diz respeito ao tratamento e recuperação da informação. Na área biomédica, a quantidade de dados produzidos vem aumentando progressivamente e acompanha o desenvolvimento de tecnologias para a produção e captura de imagens que auxiliam no diagnóstico de doenças e são também utilizadas no processo de ensino-aprendizagem desse domínio do conhecimento.

O desenvolvimento de sistemas, artefatos e modelos que visam ao enriquecimento semântico, e o uso de ferramentas que potencializam o armazenamento, organização e recuperação de informações de origem imagéticas, são tecnologias fundamentais para a Ciência da Informação. Podemos citar a Web Semântica, o uso de ontologias e o enriquecimento semântico como fundamentos que visam à organização da quantidade de dados que vêm sendo disponibilizados em diversos domínios do conhecimento.

Novas tecnologias são utilizadas objetivando o enriquecimento semântico dos dados, potencializando os processos de organização e recuperação da informação, principalmente de motores de busca, em que a aplicação de ontologias é usada para este fim. O enriquecimento semântico tem em vista dotar as imagens objeto deste estudo de variados elementos que possam representar suas relações com outros objetos, como laudos, estudos, relatórios, artigos etc.

Consultando os bancos de imagens médicas, os profissionais médicos acessam e interpretam as imagens para fundamentar diagnósticos clínicos. Portanto, as imagens devem oferecer metadados e descrição para serem recuperadas. A especialidade médica e o órgão foco do exame podem gerar um número diferenciado de imagens que exige um tratamento informacional específico, fazendo-nos refletir sobre novos procedimentos para a gestão do acervo de imagens biomédicas. Segundo Wong e Hoo Jr. (2002), a imagem biomédica possui uma

particularidade por representar informações funcionais de partes do corpo humano, além da estrutura e conteúdo anatômico.

Diante do grande número de imagens biomédicas produzidas no âmbito dos exames de imagem, nos perguntamos: como aplicar tecnologias semânticas para permitir a descoberta de conhecimento nos repositórios de imagens biomédicas? Em que medida o enriquecimento semântico pode proporcionar a melhor representação dos metadados de imagens biomédicas, de forma a aumentar a recuperação?

Diante dessas questões, nosso objetivo geral é estabelecer uma metodologia para extração de classes e conceitos necessários à construção de uma ontologia aplicada visando ao enriquecimento semântico de imagens biomédicas, utilizando um conjunto de imagens de lâminas histopatológicas, de modo a aumentar a recuperação da informação.

O oferecimento de bancos de imagens biomédicas na *web* consiste em um desafio para a organização e a recuperação deste tipo de documento. No caso de acervos que reúnem lâminas histopatológicas, a descrição das alterações morfológicas é decisiva para o diagnóstico clínico do paciente. Uma vez que os bancos de imagens biomédicas são acervos muito povoados, a *biomedical imaging informatics* (BMII), informática de imagens biomédicas, surge e assume um protagonismo por concentrar

[...] na melhoria dos resultados dos pacientes através do uso efetivo de imagens e informações obtidas a partir das imagens derivadas de pesquisa e de cuidados clínicos. Os desenvolvimentos no campo têm implicações no diagnóstico de doenças, na otimização do tratamento, no acompanhamento da resposta a doenças e na previsão de resultados (AMERICAN MEDICAL INFORMATICS ASSOCIATION, 2020, tradução nossa).

Cabe à informática de imagens biomédicas o processamento, otimização e uso de qualquer tipo de imagem biomédica (radiologia, lâminas histopatológicas, tomografia computadorizada, ressonância magnética, entre outras), bem como o estudo de sua aplicação na pesquisa científica.

2 Enriquecimento semântico

O enriquecimento semântico refere-se à potencialização dos metadados das imagens mediante a inserção de vocabulários e ontologias, de modo a dotá-las de variados elementos que possam representar suas relações com outros objetos, como laudos, estudos, relatórios, artigos etc. O enriquecimento semântico tem como fim proporcionar uma exponencial taxa de recuperação das imagens e informações correlatas, possibilitando aspectos como inferência e interoperabilidade, fatores que caracterizam uma representação e recuperação de informações mais ricas, úteis e com adição de elementos de inteligência autônoma de processamento de informações por parte do sistema. De acordo com Zeng,

O termo "enriquecimento" pode referir-se ao processo da aplicação de uma ferramenta de enriquecimento aos seus resultados, como novos metadados criados no final do processo. Na literatura, vários termos podem fazer referência a métodos como reconciliação, mapeamento, alinhamento, correspondência, massagem, fusão, interligação, etc. O resultado geral é claramente o enriquecimento dos metadados existentes, com significados mais contextualizados (ZENG, 2019, p. 7).

Gruber considera Ontologia “[...] como sendo uma especificação explícita de uma conceptualização [...]” (GRUBER, 1993, p. 1, tradução livre), sendo que a conceptualização é algo essencialmente abstrato, tratando-se da representação simplificada de algo ao qual desejamos simbolizar no mundo. Também Dahlberg (1978a) afirma que o homem passou a utilizar palavras (conjunto de símbolos) para designar os objetos em sua volta. Sendo assim, Dahlberg afirma que “[...] podemos então dizer que a linguagem constitui a capacidade do homem designar os objetos que o circundam assim como de comunicar-se com os seus semelhantes [...]” (DAHLBERG, 1978a, p. 101). A linguagem permitiu ao homem se relacionar com os objetos e assim atribuir enunciados a eles, formando então os conceitos que nos aproximam da ideia de conceptualização apresentada por Gruber (1993). Para Dahlberg, “Podemos agora definir a formação dos conceitos como a reunião e compilação de enunciados verdadeiros a respeito de determinado objeto” (DAHLBERG, 1978a, p. 102).

Dessa forma, entendemos que uma Ontologia é constituída inicialmente a partir de conceitos, na qual a conceptualização apresentada por Gruber (1993) nos remete às ideias apresentadas por Dahlberg (1978a) na Teoria do Conceito: “É possível definir, então, o conceito como a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico” (DAHLBERG, 1978a, p. 102). Ainda segundo Gruber (1993), a representação formal

de um determinado domínio do conhecimento constitui um conjunto de objetos ao qual se atribui um nome do universo de discurso, pois, em uma Ontologia, as definições dos termos do vocabulário daquele domínio associam os nomes das entidades e conceitos no universo de discurso (classes, relações, funções, objetos) com a descrição de textos legíveis pelo ser humano. As definições, assim, são descritas em axiomas formais, logicamente verdadeiras, cujo objetivo é restringir as interpretações do termo.

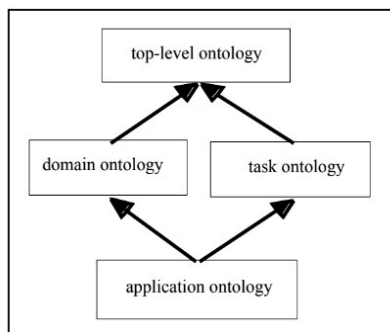
Visando ao suporte ao processo de ensino aprendizagem e a descoberta do conhecimento, o uso de ontologias se mostra como assertivo, pois potencializa o processo de organização, representação e recuperação da informação e do conhecimento. Conforme afirma Moreiro González,

As ontologias supõem um salto qualitativo: do uso de motores de busca (com palavras-chave), passa-se a usar agentes de software inteligentes (com conceitos). Possibilitam passar da mera recuperação de informação para a obtenção de respostas precisas a consultas concretas (recuperar conhecimento) (MOREIRO GONZÁLEZ, 2011, p. 76).

De maneira adicional, Moreiro González (2011) apresenta Ontologia como “[...] a soma de uma série de conceitos relevantes que representam o conhecimento compartilhado pelos membros de um determinado domínio” (MOREIRO GONZÁLEZ, 2011, p. 76), corroborando a importância da Teoria do Conceito, suas características e relações entre termos, para compreendermos a base inicial do design de uma ontologia. Morais e Ambrósio (2007) destacam cinco tipos principais de ontologias, que diferem segundo seus propósitos. Neste sentido, há ontologias genéricas, ontologias de domínio, ontologias de tarefas, ontologias de aplicação e ontologias de representação.

Guarino (1998) também aponta diferentes tipos de ontologia considerando o nível de generalidade: ontologia de alto nível, de tarefa, de domínio e de aplicação. A ontologia de alto nível descreve conceitos gerais como espaço, tempo, matéria, objeto, evento, ação, entre outros, que independem de um problema ou domínio em particular, sendo destinada a grandes comunidades de usuários. A ontologia de domínio descreve termos de um determinado domínio de conhecimento, como Medicina. Já a ontologia de tarefa representa o vocabulário de tarefas ou atividades, de certa forma especificando os termos apresentados na ontologia de alto nível. A ontologia de aplicação é uma especialização dos termos da ontologia de domínio ou de tarefa, podendo representar papéis desempenhados. Tais ontologias podem ser representadas esquematicamente conforme a Figura 1:

Figura 1. Taxonomia de ontologias segundo a generalidade.



Fonte: GUARINO, 1998.

Na área biomédica, as ontologias são utilizadas para facilitar o processo de busca e recuperação de informações e, conseqüentemente, possibilitam o compartilhamento do conhecimento médico ao enriquecer a pesquisa e potencializar os resultados das buscas em bases de dados. De acordo com Bodenreider (2009),

As ontologias tornaram-se recursos importantes para pesquisas biomédicas e os pesquisadores passaram a contar com ontologias como a International Classification of Diseases e a Gene Ontology em uma grande variedade de aplicações, o que garante a sua existência (BODENREIDER, 2009, p. 9, tradução livre).

O desenvolvimento e aplicação de ontologias na área biomédica, ademais, permite o encurtamento do ciclo pesquisa-experimento-teste-aplicação de soluções médicas, otimizando o tempo entre o início de pesquisas, sua fase de testes e a sua aplicação efetiva. Nos últimos anos, no cenário de pandemia mundial de COVID-19, o desenvolvimento de vacinas em tempo recorde exemplifica o ganho de conhecimento pelo encurtamento deste ciclo devido a modelos mais dinâmicos e efetivos de representação, publicação e recuperação de informações.

Segundo esta visão, uma ontologia desenhada a partir de uma metodologia eficiente pode oferecer recursos valiosos para os processos de organização, representação e recuperação do conhecimento. Ultimamente, novas metodologias de *design* de ontologias vêm sendo desenvolvidas para a construção e manutenção de ontologias em diferentes domínios do conhecimento, sobretudo os domínios das áreas de saúde.

3 Material e Método

No contexto da pesquisa, foi estudado um tipo de imagem biomédica: imagens de lâminas histopatológicas produzidas como material didático no âmbito da disciplina de Patologia Geral, do Departamento de Ciências Básicas (FCB) do Campus

de Nova Friburgo da Universidade Federal Fluminense (UFF). A disciplina de Patologia Geral é ministrada aos alunos dos cursos de Biomedicina, Fonoaudiologia e Odontologia e seu ensino apresenta papel fundamental na formação de profissionais da área médica, por ser a principal conexão entre as ciências básicas e a prática clínica, assim como referência para a pesquisa básica.

Além das imagens de lâminas histopatológicas, também utilizamos laudos anatomopatológicos elaborados pelo professor da disciplina de Patologia Geral. A Patologia é uma disciplina fundamental à formação dos profissionais da área médica e é dividida em Patologia Geral e Anatomia Patológica (patologia sistêmica).

Segundo Brasileiro Filho (2006, n.p.), Patologia é “a ciência que estuda as causas das doenças, os mecanismos que as produzem, os locais onde ocorrem e as alterações moleculares, morfológicas e funcionais que apresentam”. A Patologia apresenta um papel relevante para a compreensão das doenças por oferecer aspectos para identificar os elementos necessários para prevenção, manifestações clínicas, diagnóstico, tratamento, evolução e prognóstico.

A Patologia Geral une e articula a ciência básica à prática clínica pelo estudo das alterações estruturais e funcionais nas células e tecidos doentes, introduzindo conceitos de processos de doença e lesão, e reação tecidual (PEREIRA *et al*, 2018).

As imagens utilizadas no âmbito da disciplina de Patologia Geral são de origem veterinária, não ferindo as normas do Comitê de Ética e Pesquisa de uso de material humano. Os tecidos animais e dos humanos possuem muitas similaridades na morfologia e também nas doenças, permitindo seu uso no ensino. Segundo a Comissão de Ética no Uso de Animais da UFF, “O uso de animais para fins didáticos-científicos vêm ao longo dos tempos contribuindo para o avanço dos conhecimentos na área da saúde do homem e dos animais” (UNIVERSIDADE, 2023).

Os laudos anatomopatológicos informam as alterações patológicas observadas no material em análise, sendo as alterações interpretadas para a conclusão de um diagnóstico para um caso específico. O diagnóstico final é baseado na avaliação crítica e no reconhecimento de padrões, e posterior confirmação da informação na literatura médica, pois o registro de experiências anteriores ou similares pode ser consultado para melhor definição do diagnóstico (STROMBERG *et al*, 2019).

Por sugestão do especialista em Patologia, fizemos o levantamento dos conceitos a partir da descrição macroscópica das imagens constantes nos laudos analisados. Tal decisão justificou-se pela complexidade do domínio e de sua modelagem conceitual. A imagem macroscópica é aquela cuja análise é feita a olho nu,

não sendo necessária a utilização de microscópio para observar as alterações, conforme exemplo ilustrado na Figura 2:

Figura 2. *Imagem macroscópica de fígado normal*



Fonte: UNIVERSIDADE (2013).

A partir dessa deliberação, iniciamos a modelagem conceitual com a análise diplomática dos 14 laudos, identificando sua organização textual e de conteúdo. Ao observarmos o laudo (Figura 3, em apêndice), verificamos que o documento possui uma apresentação estruturada, dividindo-se em quatro partes: identificação do paciente, natureza do material e histórico clínico; descrição macroscópica; descrição microscópica e conclusões.

4 Análise e resultados

A partir da leitura dos laudos, identificamos os conceitos apresentados na descrição macroscópica com base nos princípios da Teoria do Conceito. O processo de determinação do conceito, para Dahlberg (1978), se dá no momento em que é selecionado um item de referência (referente) e analisado dentro de um determinado Universo. A partir daí, atribuem-se características ao referente, selecionado as mais relevantes. Essas características auxiliam no processo de denominação (forma verbal) que denota o conceito.

Foram identificados 207 conceitos, sendo estes colocados em ordem alfabética para identificar as repetições. Foram apurados 67 termos repetidos, totalizando 140 termos únicos. Com base na definição de Patologia, em um processo indutivo, iniciou-se a definição das categorias a partir da análise das características dos conceitos.

A definição de Patologia auxiliou na identificação das categorias na modelagem, aliado à consulta aos aspectos que devem ser observados pelo patologista ao fazer a descrição macroscópica: localização (tecido ou órgão, localização anatômica no órgão), distribuição (focal, multifocal ou múltipla, multifocal a coalescente, difuso, simétrica, assimétrica), cor (branco, amarelo, verde, vermelho, preto, cinza, azul), forma e

demarcação (redonda, oval, retangular, triangular, irregular, plano elevado, plano deprimido), tamanho (em 2 ou 3 planos), consistência (firme, firme-elástica, elástica, macia, rangente, pétrea, superfície de corte compacta, cística, policística) e características especiais (odor) (STROMBERG *et al*, 2019).

Com base no exposto, chegamos às seguintes categorias e suas definições:

- Dimensões - Tamanho dos órgãos
- Cor - Cor da alteração do tecido ou órgão.
- Consistência - Qualidade da amostra analisada obtida através do toque.
- Localização - Órgão ou tecido afetado.
- Formato - A forma como se apresenta a alteração observada.
- Lesão - Alterações moleculares, morfológicas e funcionais.

Como primeiro exercício, aplicamos os conceitos identificados às categorias definidas, conforme exemplo apresentado no Quadro 1:

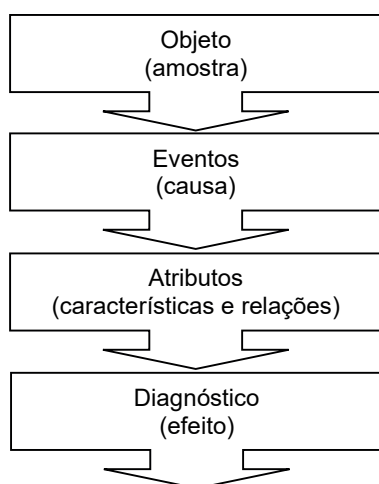
Quadro 1: *Exemplo de categorização dos conceitos nas categorias.*

CATEGORIAS	DADOS DAS AMOSTRAS
Dimensões	- 0,5cm diâmetro
	- 0,7cm x 0,6cm
	- 0,7cm x 0,6cm x 0,3 cm
	- 0,8cm x 0,4 cm
	- 1,0cm x 1,0cm x 0,3 cm
Cor	- Amostra acinzentada
	- Amostra enegrecida
	- Cápsula de gordura brancacenta
	- Cornos uterinos creme
Consistência	- Pelos <u>negros</u>
	- Consistência elástica
	- Consistência firme
	- Consistência firme-elástica
	- Consistência macia
Localização	- Tecido adiposo macio
	- Articulações
	- Brônquios
	- Cornos uterinos
	- Epidídimos
Formato	- Orofaringe
	- Amostra irregular
	- Base elíptica
	- Aspecto nodular
Lesão	- Face cirúrgica lobada
	- Área ulcerada
	- Áreas císticas
	- Lesão ulcerada
	- Nódulo
	- Nódulo exófitico

Fonte: Elaborado pelos autores

Cada categoria identificada representa um aspecto a ser considerado na descrição macroscópica e na definição do diagnóstico morfológico que vem a ser o produto da descrição morfológica, indicando a suspeita clínica. Desta forma, elas irão compor as entidades que farão parte da modelagem semântica. A cada análise de lâminas, novos elementos podem ser incorporados a uma destas categorias e, assim, tendo como base os diagnósticos, relações dinâmicas causais podem ser determinadas entre o objeto, os eventos e o diagnóstico, conforme a Figura 4:

Figura 4. *Relação dinâmica causal.*



Fonte: Elaborado pelos autores.

Uma vez mapeadas e descritas as dinâmicas de interações entre eventos e características incidentes no objeto, relações causais poderão ser estabelecidas, obtendo-se relações semânticas possíveis a serem apontadas pela ontologia. A título de exemplo, numa amostra, um tecido de cor acinzentada, de tamanho entre 2,0 e 3,0 cm e consistência firme, causalmente pode não apontar para um diagnóstico preocupante; por outro lado, num mesmo tecido, ao possuir cor acinzentada e em dimensões > 5 cm e formato irregular, esta amostra pode possuir elementos potencialmente causais entre um evento (lesão, ulceração) e um diagnóstico preocupante.

5 Conclusão

Na concepção da modelagem conceitual, as etapas apontadas por Moreira Gonzalez (2011) nortearam as atividades desenvolvidas. A partir de Guarino (1998), cogitou-se trabalharmos com o domínio Patologia, mas o material analisado (imagens macro e microscópicas) apresenta interseções com outras áreas da saúde, indicando complexidade. Com base no conjunto dos termos e das categorias estabelecidas na modelagem conceitual, verificamos que a ontologia do nosso estudo é de aplicação por trabalharmos com termos específicos da Patologia (domínio) e com tarefas ou

atividades, representadas pelas alterações sofridas pelo material em análise.

Em nossa pesquisa, optamos por construir a ontologia a partir da identificação de conceitos constantes nos laudos anatomopatológicos e determinar as classes com base na literatura da Patologia, mais precisamente com os parâmetros adotados pelos patologistas na descrição macroscópica. A modelagem realizada tem como objetivo principal permitir a recuperação de itens a serem usados como material didático exemplificando as lesões e diagnósticos morfológicos para os pesquisadores poderem comparar com suas pesquisas ou casos clínicos. A modelagem de imagens macroscópicas nos permitiu entender um pouco mais sobre o domínio, além de possibilitar uma interação mais efetiva com a professora da disciplina de Patologia, que é a especialista.

Como continuidade da pesquisa, definiremos os vocabulários para o enriquecimento semântico e iniciaremos os experimentos semânticos, quando a ontologia será escrita em *Ontology Web Language* (OWL) no software Protégé.

O presente trabalho foi realizado com apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Referências

- AMERICAN MEDICAL INFORMATICS ASSOCIATION. **Working Group Biomedical Imaging Informatics**. 2020. Disponível em: <https://www.amia.org/programs/working-groups/biomedical-imaging-informatics>. Acesso em: 4 de out. 2020
- BRASILEIRO FILHO, G. **Bogliolo Patologia**. 7. ed. Rio de Janeiro: Guanabara Koogan, 2006.
- DAHLBERG, I. A Referent-oriented analytical concept theory for Interconcept. **International Classification**, v. 5, n. 3, p. 142-150, 1978.
- GUARINO, N. (ed.), *Formal Ontology in Information Systems. Proceedings of FOIS'98*, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15.
- MOREIRO GONZÁLEZ, José Antônio. **Linguagens documentárias e vocabulários semânticos para a web: elementos conceituais**. Salvador : EDUFBA, 2011.
- PEREIRA, Patricia Fonseca *et al.* O ensino da patologia e sua influência na atuação de patologistas e infectologistas no Rio de Janeiro. **Revista Brasileira de Educação Médica**, v. 42, p. 216-225, 2018.
- STROMBERG, Paul C. *et al.* **Opening Pandora's Box: Gross description and interpretation in anatomic veterinary pathology**. Illinois, USA: Davis-Thompson Foundation, 2019.
- UNIVERSIDADE FEDERAL FLUMINENSE. **Comissão de Ética no Uso de Animais (CEUA)**. Niterói, 2023. Disponível em: <http://www.ceua.uff.br/>. Acesso 1 mar 2023.

UNIVERSIDADE FEDERAL FLUMINENSE. Foto Macroscopia. **Patologia Geral**. Nova Friburgo, 2013.

Disponível em:

<https://patogeralpunf.wixsite.com/punfuff/macro>. Acesso em: 2 mar. 2023.

WONG, S.; HOO Jr, K. S. (2002). Medical imagery. In Castelli, V. & Bergman, L. D. (editors) **Image Databases:**

Search and Retrieval of Digital Imagery (pp 83-103). John Wiley and Sons, New York.

ZENG, M. (2019). Semantic enrichment for enhancing LAM data and supporting digital humanities. Review article. **El profesional de la información**, v. 28, n. 1.

Apêndice – Figura 3. *Fac-símile de laudo anatomopatológico*

EXAME ANATOMOPATOLÓGICO			Exame Histopatológico Data da entrada: 16/11/2022
Número: DV 11.740			
Proprietário: Dra. XXXXXXXX			
Veterinário solicitante: XXXXXXXX – (Hospital Veterinário XXXXXXXX) Nome: XXXXXXXXX			
Espécie: Canina	Sexo: Macho	Porte: -	
Raça: Fox Paulistinha	Idade: 06 anos	Pelagem: -	
Natureza do material: Peça cirúrgica dígitos.			
Histórico clínico: Tumoração macia, infiltrativa em dígito. Local da lesão: Dígitos. Local da biópsia: Local da lesão. Suspeita clínica: Mastocitoma.			
Macroscopia:			
Foi recebida peça cirúrgica de dois dígitos, de formato irregular, medindo 3,5 x 2,5 x 1,5 cm; superfície externa revestida por pele tricotomizada creme, com respectivos coxins; região interdigital com área ulcerada de 0,5 cm de diâmetro; face cirúrgica composta por músculos, articulações e ossos, revestidos por pele. Aos cortes, na região interdigital, consistência firme-elástica, superfície de corte compacta e brancacenta.			
Microscopia:			
Foi observado, na região interdigital, nódulo circunscrito, não encapsulado, ulcerado com necrose e infiltrado neutrofílico. Estava composto por acentuado infiltrado inflamatório crônico granulomatoso constituído por material amorfo a filiforme, compatível com fragmentos de pelos e queratina, circundados por células gigantes multinucleadas, células epitelioides e macrófagos, por vezes fagocitando o material, ao redor intensa proliferação de fibroblastos. Havia um foco com infiltrado de neutrófilos, pelos fragmentados, circundado por células gigantes multinucleadas e fibroblastos, neovascularização e focos hemorrágicos. Margem da amostra com epiderme queratinizada, folículos pilosos e anexos cutâneos normais; coxim normal.			
Conclusão: PROCESSO CRÔNICO GRANULOMATOSO POR CORPO ESTRANHO.			
Obs.: Devido histórico prévio de mastocitoma no local, foi solicitada coloração especial para investigação de possíveis células neoplásicas.			

A convergência da Organização do Conhecimento e do Design para apresentação de informação acessível: a criação de um modelo de tesouros acessível

The convergence of Knowledge Organization and Design for accessible information presentation: foundations for creating an accessible thesaurus model

Natalia Nakano (1), Mirele Carolina Souza Ferreira Costa (2), Milton Shintaku (3)

(1) Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Brasília, natalianakano@ibict.br
(2) mirelecosta@ibict.br (3) shintaku@ibict.br

Resumo

Objetivo: A Organização do Conhecimento (KO) e o Design são áreas que envolvem o projeto de sistemas com parte integrante das suas atividades, e, portanto, é razoável pensar que as áreas têm muitos outros aspectos em comum. Assim, tendo como premissa o contexto da KO e do Design, e nele o Design da Informação, o Design Universal e a Encontrabilidade da informação, é relevante discutir como essas disciplinas podem, conjuntamente com a KO cooperar na criação de ambientes digitais informacionais acessíveis, mais especificamente, no caso deste estudo, na apresentação da informação acessível em um website de tesouros acessível. **Metodologia:** Trata-se de pesquisa bibliográfica, documental e aplicada. Apresentamos os conceitos e princípios do Design da Informação que informaram este estudo, do Design Universal para sistemas informacionais e da Encontrabilidade da Informação. **Resultados:** Os resultados apresentam o website de tesouros acessível. **Conclusão:** Conclui-se que relaciona os conceitos aplicados à criação do website bem como o conhecimento adquirido durante o estudo.

Palavras-chave: Documentos acessíveis. Design da Informação. Linguagem simples. Libras. Acessibilidade.

Abstract

Objective: Knowledge Organization (KO) and Design are areas that involve systems design as an integral part of their activities, and therefore it is reasonable to think that the areas have many other aspects in common. Thus, having as a premise the context of KO and Design, and within it Information Design, Universal Design and Findability, this study finds discusses how these areas can, together with KO, cooperate in the creation of accessible informational digital environments, more specifically, in the case of this study, in the presentation of accessible information on an accessible thesaurus website. **Methodology:** It is a bibliographical, documental and applied research. We present the concepts and principles of Information Design that informed this study, Universal Design for informational systems and Findability. **Results:** The results present the accessible thesaurus website page. **Conclusions:** The conclusion relates the concepts applied to the creation of the website as well as the knowledge built during the study.

Keywords: Accessible documents. Information Design. Plain English. Sign language. Accessibility.

Resumen

Objetivo: La Organización del Conocimiento (KO) y el Diseño son áreas que comprenden el diseño de sistemas como parte integral de sus actividades, por lo que es razonable pensar que las áreas tienen muchos otros aspectos en común. Así, teniendo como premisa el contexto del KO y el Diseño, y dentro de él el Diseño de la Información, el Diseño Universal y la Encontrabilidad, este estudio analiza cómo estas áreas pueden, junto con el KO, cooperar en la creación de sitios digitales informativos accesibles, más específicamente, en el caso de este estudio, en la presentación de información accesible en un sitio web de tesoro accesible. **Metodología:** Es una investigación bibliográfica, documental y aplicada. Presentamos los conceptos y principios del Diseño de la Información que informaron este estudio, del Diseño Universal para sistemas de información y Encontrabilidad de la información. **Resultados:** Los resultados presentan la página web del tesoro accesible. **Conclusiones:** La conclusión relaciona los conceptos aplicados a la creación del sitio web así como el conocimiento construido durante el estudio.

Palabras clave: Documentos accesibles. Diseño de la Información. Portugués sencillo. Lengua de señas. Accesibilidad.

1 Introdução

A Organização do Conhecimento, tradução de *Knowledge Organization* (KO) é definido por Smiraglia (2012, p. 225) como

[...] o domínio no qual a ordem do conhecimento é tanto o paradigma primário para a investigação científica quanto a aplicação primária no desenvolvimento de sistemas.

Por isso, a KO torna-se amparo conceitual de estudos, como uma área de investigação, da mesma forma em que atua no apoio ao desenvolvimento de sistemas de informação.

Martínez-Ávila, San Segundo e Zurian (2014) e Martínez-Ávila (2015), por sua vez, relacionam o KO com o design de sistema, aumentando a atuação deste campo de pesquisa. Com isso, a junção do KO e design oferta desafios e oportunidades de estudos, na medida em que aliam pontos complementares.

Como o design ou projeto de sistemas é parte integrante da KO, é razoável pensar que o Design da Informação pode ter muitos outros aspectos em comum com a KO. Nesse sentido, de acordo com Orna e Stevens (1991, p. 197), o Design da Informação pode ser entendido de forma ampla como “[...] tudo o que fazemos para tornar as ideias visíveis para que outros possam torná-las suas e usá-las para seus próprios propósitos.” Esse é, de forma simplificada, um processo que envolve signos que representam conhecimento.

A ciência do Design, e nele o Design da Informação, é uma relevante fonte de subsídios para a criação de sistemas informacionais digitais acessíveis. A ciência do Design e suas metodologias estão centradas no ser humano e as suas variações, barreiras ao acesso, desafios e outros, da mesma forma que que agrupam habilidades que nos caracterizam, como a intuitividade, o reconhecimento de padrões, a construção de ideias com significado emocional, e a autoexpressão.

A Organização do Conhecimento, similarmente, se concentra nos princípios e métodos de organização, categorização e representação da informação e do conhecimento, habilidades também centrais e inerentes dos seres humanos. A KO tem como desafio tornar os recursos de informação mais acessíveis e utilizáveis, estruturando-os de forma que reflitam seu conteúdo intelectual e seus relacionamentos. Portanto, preocupe-se com o desenvolvimento e gerenciamento de sistemas, ferramentas e técnicas para organizar recursos de informação de forma a facilitar sua descoberta, recuperação e uso por diferentes tipos de cidadãos.

Os estudos voltados para a organização e o design da informação devem contar que, se estão voltados para o ser humano, a humanidade não é uniforme na forma de percepção da informação. Inúmeras diferenças nos componentes do público-alvo de um sistema revelam a complexidade desses estudos, no qual a acessibilidade

para portadores de deficiências se apresenta como uma das afirmativas.

Assim, tendo como premissa o contexto da KO e do Design, e nele o Design da Informação, o Design Universal e a Encontrabilidade da informação, é relevante discutir como essas disciplinas podem, conjuntamente com a KO cooperar na criação de ambientes digitais informacionais acessíveis, mais especificamente, no caso deste estudo, na apresentação da informação acessível em um website de tesouros acessível.

2 Metodologia

Trata-se de pesquisa bibliográfica, documental e aplicada que utiliza como fonte documental o Modelo de Acessibilidade em Governo Eletrônico (e-MAG) (BRASIL, 2014), as Diretrizes de Acessibilidade para Conteúdo Web (WCAG) e as recomendações da ABNT 17060. Como fontes bibliográficas, utilizamos a Tese de doutorado intitulada “Princípios do design da informação na curadoria digital de ambientes virtuais de aprendizagem sob a perspectiva da ciência da informação [...]”, de Nakano (2019), o livro “A emergência do Design da Informação na contemporaneidade da Ciência da Informação”, 2ª edição, de Jorente, Nakano e Padua (2020) e o livro “Encontrabilidade da Informação”, de Vechiato e Vidotti (2014).

Considerando que o Design da Informação ainda não está consolidado como interdisciplinar à Ciência da Informação e à Organização do Conhecimento, embora as áreas tenham em comum muito mais do que as questões humanas, o objeto informação e apresentação da informação para sua eficiente recuperação (NAKANO *et al.*, 2018), julgamos necessário apresentar os conceitos que fundamentaram a construção do website do tesouros acessível. Assim, as seções seguintes apresentam o conceito e alguns princípios do Design da Informação que informaram este estudo, o Design Universal para sistemas informacionais e a Encontrabilidade da Informação. Os resultados apresentam um exemplo da página do website do tesouros acessível e a conclusão relaciona os conceitos aplicados à criação do website bem como o conhecimento adquirido durante o estudo.

3 Como o Design da Informação contribui para sistemas informacionais acessíveis

O principal objetivo do Design da Informação (DI), para Rune Pettersson (2013, p. 77), consiste na clareza da comunicação de mensagens. Para isso, mensagens bem desenhadas devem caracterizar-se pela sua precisão. Sua produção e distribuição devem mirar a correta interpretação e compreensão por aqueles que com ela

interagem, independentemente do meio - físico ou digital - ou das habilidades ou deficiências. Para o autor, o objetivo do DI consiste em subsidiar a criação de materiais informacionais, incluindo mensagens, que possibilitem aos receptores o desempenho de tarefas específicas.

Design da Informação é um conceito reconhecidamente amplo e, portanto, complexo (JORENTE; NAKANO; PADUA, 2020). Isso porque os estudos em Informação, Design, Mídia e Comunicação constituem áreas amplas e interdisciplinares que se conectam com diversas outras áreas também preocupadas com a apresentação, a representação e a codificação da informação para a transmissão de mensagens.

Horn (1999, sem paginação, tradução nossa), define o DI como “[...] a arte e ciência de preparar informação para que possa ser usada por seres humanos com eficiência e eficácia.” Para o autor, o que diferencia o DI de outras disciplinas do Design são as noções de eficiência e eficácia do objetivo comunicacional. Assim, os objetivos do DI incluem, para Horn (1999):

- desenvolver documentos que sejam compreensíveis, rápida e precisamente recuperáveis, e fáceis de serem traduzidos em ação efetiva;
- desenhar interações com equipamentos que sejam fáceis, naturais e o mais agradável possível, o que envolve resolver problemas de interfaces humano-computador;
- assegurar que as pessoas possam se mover em espaços tridimensionais com conforto, especialmente espaços urbanos, mas também no espaço virtual.

Outra definição, mencionada por Pettersson (2013, p. 87, tradução nossa), foi elaborada por Zwaga, Boersma e Hoonhour (1999). Para os autores, o DI é:

[...] primeiramente o desenvolvimento efetivo de organização de dados para transformar esse dado em informação, e então o desenvolvimento de um instrumento (geralmente um produto gráfico) para transferir a informação de tal forma que contribua para a base de conhecimento do usuário ou oriente a realização de uma tarefa de maneira efetiva e convincente.

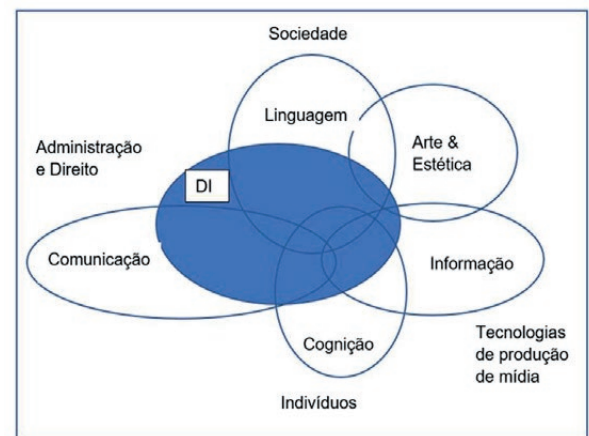
Uma definição mais recente foi estabelecida pelo *Information Design Journal* (IDJ), indexado na área *Library and Information Sciences* pela *Scimago Journal & Country Rank*: o DI, aqui, é definido como a aplicação de processos de design (planejamento) para a comunicação da informação (conteúdo, linguagem e forma) (PETTERSON, 2013).

O DI frequentemente é relacionado e até representado por outros conceitos, como Infografia, Comunicação Visual, Visualização da Informação, Design de Comunicação, etc. Na definição de Pettersson (2013), o

DI objetiva satisfazer as necessidades informacionais dos receptores. Para isso, “[...] o DI compreende a análise, planejamento, apresentação e entendimento de uma mensagem – seu conteúdo, linguagem e forma” (PETTERSSON, 2013, p. 112, tradução nossa). Para o autor, um material com bom desenho, independente do meio, atende a requisitos estéticos, econômicos, ergonômicos e subjetivos.

O DI, para Pettersson (2013), é uma área interdisciplinar que sofre influências e influencia mais de 50 áreas de pesquisa. Conforme modelo teórico apresentado na figura 1, as principais áreas interdisciplinares são as disciplinas de linguagem, arte e estética, disciplinas que investigam a informação, a comunicação, o comportamento, a cognição, a gestão e o direito, e as tecnologias de produção de mídia.

Figura 1. Principais áreas interdisciplinares ao DI



Fonte: Adaptada de Pettersson (2013).

Caracteriza-se pela interdisciplinaridade, por derivar e relacionar-se a outras áreas, e por possuir tanto um componente científico puro quanto um componente de ciência aplicada, que desenvolve serviços e produtos. No atributo referente ao desenvolvimento de serviços e produtos observa-se o seu caráter aplicado, voltado, por exemplo, à criação de sistemas e ambientes digitais informacionais, e por que não, websites de tesouros.

Outra característica do DI, ainda de acordo com Pettersson (2012), é a inexistência de regras claras, dados os muitos fatores que o influenciam, como o contexto, a tarefa, o público-alvo, etc. Por isso, o autor considera o DI tanto como uma arte quanto uma ciência. Porém, foram estabelecidos princípios considerados fundamentais pelos estudiosos do DI e áreas convergentes. As diretrizes baseadas nesses princípios subsidiaram o desenho de materiais apropriados para o público-alvo. Pettersson (2012) agrupou em seis categorias os princípios funcionais do DI: Definição do Problema, Estrutura, Clareza, Simplicidade, Ênfase e Unidade.

3.1 Definição do Problema

No processo de definição do problema, é necessário considerar: quem é o produtor da mensagem; para quem a mensagem é direcionada; e o meio, ou seja, o ambiente digital. É preciso avaliar, também, que a mensagem contextualizada no meio resulta na representação (mensagem + meio = representação). Nesse sentido, a representação pode ser um material impresso (texto ou imagem), apresentações audiovisuais ou uma combinação de mídias, por exemplo. Em relação ao meio, é necessário levar em consideração o contexto em que a mensagem está inserida, ou seja, o contexto digital deve ser transparente e familiar a fim de possibilitar a realização da tarefa proposta por quem interage com o produto informacional.

3.2 Estrutura

Quando bem desenhada, a estrutura facilita a percepção, a interpretação, o entendimento, e a memória da informação. Portanto, as informações devem ser organizadas da mais importante para a menos importante. A estrutura deve, também, limitar o conteúdo e seus elementos, além de agrupar a informação relacionada de forma que sua relação esteja explicitada. Os elementos precisam estar alinhados de forma a facilitar a navegação entre eles: detalhes informacionais em excesso devem ser evitados; os níveis textuais não devem ultrapassar três ou quatro níveis; o layout e a variação tipográfica devem ser estruturados para direcionar o olhar e a leitura da interface, com cabeçalhos em destaque, por exemplo; as cores devem ser consistentes e esclarecer a estrutura do texto - com cores de fundo diferenciadas, por exemplo; os cabeçalhos devem salientar a importância e a hierarquia para melhorar a compreensão do texto.

No contexto digital, os princípios de estrutura devem ser aplicados tanto na interface quanto no conteúdo da mensagem (NAKANO, 2019). Para Pettersson (2012, p. 67), as mensagens verbais e visuais podem ser escolhidas da seguinte maneira:

As mensagens verbais funcionam bem quando o conteúdo da mensagem é analítico, detalhado, lógico, narrativo, teórico e sequencial. Mensagens visuais são preferidas quando o conteúdo é emocional, holístico, imediato, espacial e visual. Para mensagens complexas, representações verbais e visuais combinadas podem ser a melhor escolha.

3.3 Clareza

Para garantir a clareza, a mensagem deve ser clara, simples, inequívoca e transparente. Para isso, o uso de tipos de letras incomuns e fontes que demasiado grandes ou pequenas deve ser evitado, pois o texto é lido como um todo, não palavra por palavra ou letra por letra.

A legibilidade do texto na tela do computador – isto é, a facilidade com que o leitor consegue ler, ver e distinguir

as partes do texto, considerando cor e tipografia – deve ser buscada por meio de contraste adequado entre texto e fundo (PETTERSSON, 2012). Em relação a fontes, é recomendável empregar tipos apropriados para telas, como Trebuchet, Verdana e Georgia. Fontes maiores devem ser usadas. Por fim, o espaçamento duplo entre linhas de um texto contínuo é recomendável para alcançar máxima legibilidade em tela de computador.

A regra de contraste aplica-se, também, à legibilidade de figuras: elas devem conter somente informação essencial e um contraste adequado entre fundo e figura para uma ótima visualização.

Questões de contraste também são relevantes quanto às cores do material informacional. Texto em preto ou marrom escuro em fundo branco ou amarelo proporcionam melhores contrastes para leitura. Fundo claro ou escuro contrastando com conteúdo escuro ou claro são recomendados na tela do computador. Embora mais chamativas, outras combinações de cores são menos legíveis, e, portanto, devem ser evitadas a fim de garantir a legibilidade (JORENTE; NAKANO; PADUA, 2020).

As questões de clareza relacionadas ao DI mencionadas por Pettersson e as outras autoras convergem com algumas das recomendações para acessibilidade para evitar que se configurem em barreiras para pessoas com baixa visão, e estão alinhadas às recomendações do WCAG e e-MAG.

3.4 Simplicidade

Considerado por Pettersson (2012) um dos princípios mais importantes do DI, a simplicidade objetiva facilitar ao máximo a percepção, o processamento e a memória relacionada a uma mensagem. Portanto, o princípio da simplicidade deve ser prioritário em um projeto de DI.

A simplicidade de um material informacional na Web pode ser enriquecida se condizente com estudos sobre legibilidade textual podem nortear a criação de materiais informacionais na Web. Segundo Pettersson (2012, p. 103),

[...] um texto facilmente compreensível é caracterizado por sentenças curtas, palavras curtas, e estrutura gramatical simples. O texto então deve ser conciso, consistente e preciso.

A escolha do vocabulário constitui outra variável relevante para esse objetivo: em relação ao grau de abstração exigido, que deve ser baixo, por exemplo; quanto ao número de sílabas nas palavras, que acrescenta complexidade ao texto; e a familiaridade do leitor com as palavras empregadas. Outros fatores que devem ser considerados: a escolha do assunto, a subdivisão em parágrafos, a extensão do parágrafo, cabeçalhos e sub cabeçalhos, extensão das linhas, a distância entrelinhas, ilustrações, o tamanho das fontes, a relevância do texto para o leitor, o tamanho da página

(PETTERSSON, 2012). Similarmente, essas recomendações coincidem com as recomendações de acessibilidade para pessoas com baixa escolaridade ou pessoas com deficiência cognitiva, e corroboram as recomendações de acessibilidade levantadas na literatura (NAKANO, 2021).

Outro fator relevante para a legibilidade são as cores. Quando empregadas para destacar ou ocultar algo ou exacerbar diferenças e similaridades, as cores auxiliam os leitores a lembrar informação e a encontrar o que buscam. As pessoas costumam preferir o uso de no máximo três ou quatro cores na mesma página, tela ou slide, uma diretriz a ser considerada no desenho desses materiais (JORENTE; NAKANO; PADUA, 2020).

3.5 Ênfase

O princípio da ênfase é empregado para atrair, direcionar e manter a atenção. Tal objetivo é alcançado por meio do uso de contraste e elementos como “[...] complexidade, direcionalidade, características exageradas, humor, isolamento ou movimento” (PETTERSSON, 2012, p. 121). A maioria das pessoas realiza a leitura de um material informacional selecionando o que está mais enfatizado, e não completamente do início até o fim do documento. De acordo com Pettersson (2012), as pessoas tendem a combinar, em seu comportamento leitor, a leitura transversal com a leitura de cabeçalhos, o olhar direcionado a ilustrações e suas legendas e a leitura mais ou menos cuidadosa de determinadas partes do texto. Por isso, elementos como cabeçalhos, fotos, desenhos e elementos gráficos são o ponto de entrada de uma página pelo seu poder de atrair a atenção. Porém, é necessário considerar, também, a influência que o tamanho e a colocação desses elementos exercem sobre o leitor em sua decisão de ler a página – figuras demasiado grandes podem ser ignoradas e, por isso, devem estar em tamanho e posição adequadas. Aqui, complementamos que, deve-se evitar elementos gráficos ou imagéticos meramente decorativos. Esses elementos, se indispensáveis, devem vir acompanhados de legendas ou alternativas em texto para leitores de tela, ou ainda, no CSS para que sejam ignorados por leitores de tela, e assim, facilitar a leitura por pessoas que utilizam o recurso (NAKANO, 2022).

3.6 Unidade

Em conjunto com o princípio da harmonia, o princípio da unidade refere-se à necessidade de que as interfaces apresentem coerência e união no geral, já que as inconsistências confundem os internautas. A unidade consiste, assim, no emprego consistente de estilos e terminologias, layout e tipografia, e técnicas de ênfase. Isso se traduz, por exemplo, em posicionar as figuras próximas da parte do texto ao qual estão relacionadas. O princípio da unidade está relacionado, também, ao princípio da proporção estética.

3.7 Harmonia

O princípio da harmonia é atendido quando os elementos estão bem agrupados e formam relações harmoniosas, equilibradas. Isso porque o ser humano possui um senso intuitivo de equilíbrio que, nesse caso, pode ser formal ou informal.

Enquanto o equilíbrio formal consiste na apresentação de simetria, que resulta na percepção de estabilidade, o equilíbrio informal resulta na percepção de dinamismo. Em consonância com Bradshaw (1996, 2003), Pettersson (2012) defende que o aprendizado, em materiais informacionais, é reduzido se há desequilíbrio e uso inconsistente de cores, gráficos ou tipografia.

3.8 Proporção Estética

Apesar de o termo proporção remeter a algo objetivo e matemático, a proporção estética é um conceito subjetivo, relativo ao receptor da mensagem no que tange ao que avalia belo ou feio. Por isso, no contexto informacional, é necessário assegurar que os elementos gráficos, como figuras ou ilustrações, não sejam utilizados para mera decoração, ou seja, seu uso deve ser limitado a uma função informativa. Nesse sentido, a simplicidade e a hierarquia devem ser priorizadas.

4 Convergência do Design da Informação e do Design Universal para sistemas informacionais acessíveis

O projeto de Design da Informação de um sistema informacional acessível deve preparar a informação para que essa seja recuperada, utilizada e transformada em conhecimento por todos, independentemente das suas habilidades ou deficiências, e, portanto, o DI deve conduzir e regular o uso de linguagens multimodais. As linguagens ativas nesses ambientes podem ser textuais, imagéticas ou audiovisuais que convergem e interoperam em interfaces para interação humano-informação. Nas palavras de Moore e Kearsley (2007, p. 77) “[...] quanto mais combinações adequadas de conteúdo, mídias e tecnologias são oferecidas, maior a abrangência e a acessibilidade do conteúdo criado”.

O DI converge com o Design Universal (DU), uma vez que os Princípios do Design Universal buscaram simplificar e sistematizar diretrizes para criação de produtos, comunicações e ambientes acessíveis. Esses princípios foram criados por um grupo de designers, arquitetos e engenheiros liderados por Ron Mace, no *Center for Universal Design* na *North Carolina State University*. Nas palavras dos criadores do DU

A intenção do Design Universal é simplificar a vida para todas as pessoas, tornando os produtos, as comunicações e o ambiente construído mais utilizável pelo maior número de pessoas possível, com pouco ou nenhum custo extra. O Design Universal beneficia pessoas de todas as idades e habilidades (Connell *et al.*, 1997, p. 58).

De acordo com Horton (2006), o DU não deseja que conteúdo específico ou direcionado para um determinado público seja construído para suprir uma deficiência, mas que se permita que indivíduos com deficiência possam ter acesso à mesma informação.

Assim, esses produtos (informações comunicadas, ambientes digitais, imagens, objetos digitais, etc.) devem contemplar usabilidade [1] para todas as pessoas. Nesse sentido, uma vez que o projeto de DI de sistemas informacionais é desenvolvido com a mentalidade do Design Universal, nenhuma adaptação ao ambiente ou ao conteúdo são necessários (ALVES, 2012). Destaca-se que os princípios de design aqui mencionados devem ser considerados pelas pessoas envolvidas no momento da concepção do ambiente digital acessível; a equipe responsável deve estar ciente das recomendações de acessibilidade antes da criação e implementação do ambiente.

A partir de nossa experiência em ambientes digitais no levantamento conceitual do trabalho da tese de doutorado (NAKANO, 2019), da intersecção do Design da Informação com o Design Universal, elaboramos os exemplos que seguem dos sete princípios do Design Universal em ambientes digitais:

1. Uso equitativo: é o princípio que norteia todos os outros; preza que o designer deve ter em mente que o sistema/conteúdo deve contemplar uma diversidade de pessoas, e não um público específico. Por exemplo, o uso de fontes com tamanho suficientemente grande e contraste de cores permite não apenas que pessoas com baixa visão consigam realizar a leitura da informação, mas também facilita a leitura para pessoas sem esse marcador.
2. Uso flexível: dita que o design deve acomodar pessoas com diferentes preferências e habilidades. Ambientes que permitem customização de interface e mudanças de cores e tamanhos de fontes são um exemplo.
3. Uso simples e intuitivo: o design deve ser de fácil entendimento independente da experiência, conhecimento ou habilidades do internauta, como por exemplo o campo de busca do Google ou o design do aparelho de celular do iPhone, ou até mesmo uma maçaneta, em que o formato informa o sujeito se deve ser alavancada ou rodada.
4. Informação perceptível: o design deve comunicar a informação de forma eficiente, independentemente das condições do ambiente ou das habilidades sensoriais da pessoa: um conteúdo comunicado por vídeo deve também trazer legendas e tradução para libras, por exemplo.
5. Tolerância ao erro: o design deve minimizar as consequências de erros acidentais cometidos na

realização de tarefas ou até mesmo ações não intencionais. Por exemplo, mais de uma possibilidade de voltar à página anterior deve estar disponível na interface.

6. Baixo esforço físico: o design deve prever baixo esforço físico ou baixo uso cognitivo. Em ambientes digitais, um bom exemplo seria a utilização de menos clicks possíveis para executar uma tarefa.
7. Tamanho e espaço para aproximação e uso: esse princípio estabelece que espaço e tamanho apropriados estejam disponíveis para que o sujeito consiga alcançar, manipular e utilizar o conteúdo, independentemente do tamanho do corpo, postura ou mobilidade. Embora este princípio objetive contemplar ambientes físicos, em ambiente digital, podemos pensar em projetos de conteúdos que sejam interoperáveis com software de tecnologia assistiva, como por exemplo de reconhecimento de voz, e possam ser acionados por diferentes recursos, como por exemplo ponteira de cabeça, acionadores, etc.

Nesse contexto, consideramos que os princípios do Design da Informação convergidos e associados aos princípios do Design Universal e complementares à Encontrabilidade da Informação podem guiar a criação de sistemas informacionais acessíveis eficientes e inclusivos que contemplem a diversidade de pessoas que necessitam interagir e se informar nesses ambientes, em especial, no ambiente do tesaurus acessível do projeto Documentos Digitais Governamentais Acessíveis desenvolvido no âmbito do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

5 A Encontrabilidade da Informação para sistemas informacionais acessíveis

Termo cunhado por Peter Morville, em 2005, no seu livro *Ambient Findability, findability* tem sua tradução para o português já consolidada como encontrabilidade. No nosso projeto, utilizamos o termo Encontrabilidade da Informação.

No âmbito dos ambientes informacionais, a Encontrabilidade da Informação é uma característica do sistema ou ambiente e está diretamente relacionada à navegação e à busca da informação. Pode-se conceituar a Encontrabilidade da Informação como:

- a. A qualidade de ser localizável ou navegável; b. O grau no qual um determinado objeto é facilmente descoberto ou localizado; c. O grau no qual um sistema ou ambiente suporta a navegação e recuperação. (MORVILLE, 2005 *apud* VECHIATO; VIDOTTI, 2014, p. 110).

A Encontrabilidade da Informação envolve não apenas o efetivo acesso e o uso da informação, mas a garantia de que a informação seja encontrada de modo facilitado

com um sistema de navegação e recuperação eficiente, e que deva incluir as necessidades dos sujeitos e suas características particulares. Enfatizamos aqui que a Encontrabilidade da Informação promove a eliminação de barreiras de acesso à informação quando determina que os sujeitos e suas necessidades devem ser considerados.

Nessa linha, Vechiato e Vidotti (2014, p. 165) propõem o conceito operatório para a Encontrabilidade da Informação

A encontrabilidade da informação sustenta-se fundamentalmente na interseção entre as funcionalidades de um ambiente informacional e as características dos sujeitos informacionais.

De acordo com Vechiato e Vidotti (2014), a encontrabilidade é um desafio, uma vez que a equipe multidisciplinar envolvida no desenvolvimento de sistemas e websites, e no caso deste estudo, tesouros acessível, necessita colaborar e reconhecer a diversidade de internautas para os quais a equipe desenha e conhecê-los a partir da concepção do projeto.

Os autores enfatizam ainda a importância dos estudos em “Intencionalidade” para a encontrabilidade em sistemas informacionais digitais. A Intencionalidade, doutrina da fenomenologia, explica que cada ato de consciência é intencional, é consciência ou experiência de algo ou outrem.

[...] a Intencionalidade do sujeito é carregada de experiências, necessidades e competências (tanto as informacionais quanto as tecnológicas), entendimento, cognição e satisfação, fornecendo, inclusive, subsídios para a estruturação de sistemas e ambientes informacionais. (VECHIATO; VIDOTTI, 2014, p. 114).

Os autores mencionam ainda a folksonomia como exemplo de intencionalidade do internauta. O ambiente permite que os internautas atribuam tags aos objetos informacionais para sua recuperação. A folksonomia decorre de uma característica da evolução da Web 1.0 para a Web 2.0, em que os internautas não apenas “usam” a informação, mas especialmente produzem e compartilham informação. A intencionalidade, então, direciona para a encontrabilidade da informação.

Vechiato e Vidotti (2014) propõem uma lista de recomendações de encontrabilidade da informação que pode ser seguida em qualquer ambiente digital. O Quadro 1 a seguir apresenta as recomendações:

Quadro 1. *Recomendações de encontrabilidade da informação para qualquer ambiente digital*

Recomendações	
1	Utilizar instrumentos para organização da informação como taxonomias navegacionais e instrumentos de controle terminológico como tesouros e ontologias, quando viável.

2	Implementar recursos de classificação social (folksonomia) e de navegação por meio de tags atribuídas pelos sujeitos.
3	Representar os recursos informacionais por metadados.
4	Investir na mediação infocomunicacional dos sujeitos.
5	Ampliar as possibilidades dos sujeitos de encontrar a informação por meio de diferentes ambientes e dispositivos.
6	Investigar os comportamentos, as competências, as experiências e a Intencionalidade dos sujeitos.
7	Aplicar recomendações e avaliações de acessibilidade e de usabilidade
8	Investir na utilização de affordances para orientar os sujeitos no espaço (wayfinding) no decorrer da navegação.
9	Investir na utilização de affordances para a query e os resultados de busca.
10	Proporcionar busca pragmática

Fonte: Vechiato e Vidotti (2014, p. 173).

A partir dos conceitos apresentados e como resultado da pesquisa, apresentamos a proposta para a apresentação da informação no website do tesouros do projeto Documentos Governamentais Digitais Acessíveis. O objetivo é propor a apresentação da informação de acordo com as recomendações da literatura do Design da Informação, do Design Universal e da Encontrabilidade da Informação, além da pesquisa realizada nos documentos elaborados pelos órgãos nacionais e internacionais sobre acessibilidade digital.

6 Resultados

Proposta para apresentação da informação no website do tesouros [2].

Os resultados apresentados foram aplicados em um tesouro implementado com o software livre Tematres, pela flexibilidade e facilidade em ajustes necessários. Com isso, os elementos levantados foram aplicados em um website.

6.1 Página Inicial

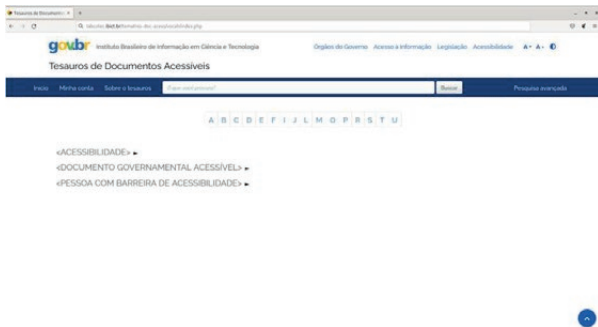
Apresentação da informação de forma linear, em uma coluna. Utiliza-se o mínimo possível de elementos textuais e contempla-se um design minimalista, simples, claro e objetivo, agradável e de fácil navegabilidade.

Para internautas sem barreiras de visão, o campo de busca simples, tipo google, está centralizado e faz uso de um vocabulário controlado, sendo familiar à maioria dos internautas. Por estar bem destacado, facilita a busca e a encontrabilidade da informação.

Abaixo do menu principal, está disposta a lista dos termos em ordem alfabética, e abaixo os meta-termos. A organização dos termos do tesouro em ordem alfabética logo abaixo do campo de busca se configura em estratégia simples e eficiente, harmoniosa ao campo de busca.

Na página inicial, a barra com menu principal horizontal e caixa de busca se destaca e permanece na navegação por todas as páginas. Logo abaixo, um menu alfabético possibilita listar todos os termos do tesouros de forma ordenada. A lista de metatermos possibilita navegar conforme a estrutura hierárquica dos termos (figura 2).

Figura 2. *Página inicial do Tesouros de Documentos Acessíveis*



Fonte: IBICT (2023).

O Tesouros oferece ao internauta mais de uma possibilidade de encontrar a informação: por meio do campo de busca simples (do tipo google), familiar aos internautas; a lista em ordem alfabética, e navegando por meio do meta-termos. Dessa maneira, a encontrabilidade da informação está contemplada na página inicial.

5.2 *Página sobre o Tesouros*

Possivelmente, a página que melhor representa as funcionalidades de acessibilidade implementadas no tesouros é a “sobre o tesouros”. Nesta página, textos explicativos são apresentados em vários formatos, cada qual voltado ao atendimento de internautas com necessidades específicas.

A apresentação dos elementos da página “sobre o tesouros” em uma coluna segue a ordem: título, botões de atalho, recurso de áudio, recurso de vídeo em libras, descrição textual do website em português, descrição textual do website em português simplificado, participantes da pesquisa separados por uma discreta linha cinza. Os recursos de áudio e vídeo estão centralizados.

Considerando a estrutura hierárquica da apresentação da informação, os elementos devem estar organizados de forma a explicitar a sua relação. Assim, todos os elementos referentes a um termo ou assunto devem estar na mesma página, ao invés de abas diferentes, por exemplo. Podem estar separados em caixas distintas para explicitar aos internautas videntes que se trata de elemento distinto. Assim, por exemplo, como a definição em português simplificado espelha exatamente a nota bibliográfica, elas devem estar apresentadas na mesma página.

Para internautas sem barreiras visuais, a organização dos elementos em caixas distintas, separadas por uma discreta linha cinza revela que se tratam de conteúdos distintos – aspecto positivo da organização em caixas (ou blocos) –, e a exibição das caixas em uma mesma página denota sua estreita relação.

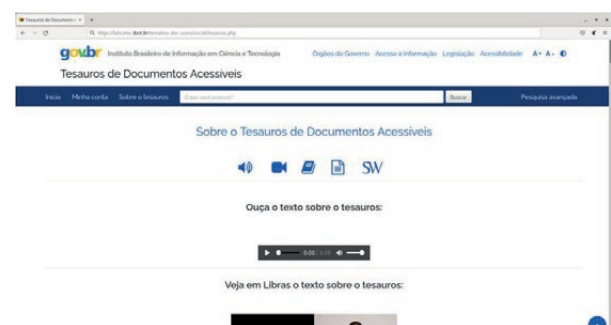
Em se tratando da acessibilidade, os leitores de tela ignoram as linhas que separam os elementos (em cinza claro) e realizam a leitura de forma linear da página, assim, a informação destacada em caixas não interferirá na leitura do conteúdo.

De forma a suplantiar os diversos tipos de barreiras, a página “sobre o tesouros” oferta a informação em vários formatos:

- Textual original: explicação convencional voltado para internautas interagindo no tesouros;
- Textual simplificado: para internautas com baixa escolaridade ou algum tipo de deficiência cognitiva;
- Áudio: para internautas com deficiência visual ou iletrados;
- Vídeo em Libras para internautas surdos
- Registrado em SignWriting, forma gráfica de representação da Libras.

A figura 3 mostra as teclas de atalho logo abaixo do título que representa o conteúdo da página (sobre o tesouros de documentos acessíveis). Vale destacar que as teclas de atalho trazem a descrição da sua função quando o mouse é apontado sobre o ícone, além de ser lido por leitores de tela. Os internautas videntes podem então se guiar pelos ícones ou pela descrição do botão, ou ambos. Ainda na figura 3, é possível observar o recurso de áudio logo abaixo das teclas de atalho para contemplar familiaridade no sistema.

Figura 3. *Teclas de atalho e recurso de áudio*



Fonte: Brasil (2023).

A informação também é apresentada em Libras, por meio de vídeo (figura 4), atendendo a usuários que possuem maior familiaridade com essa língua. A Lei da Libras - Lei nº 10.436/2002 (BRASIL, 2002), reconhece a Libras como a forma de comunicação e expressão dos surdos, por isso a importância da inclusão deste formato.

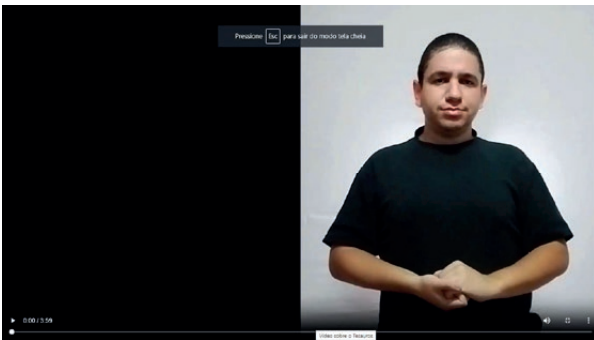
Figura 4. Vídeo com descrição em Libras



Fonte: Brasil (2023).

O vídeo oferta o recurso de apresentação em tela cheia, de forma a dar maior ênfase aos sinais apresentados, diante das características da língua (figura 5). Da mesma forma melhora a visualização das legendas, para que possa relacionar a escrita com os sinais.

Figura 5. Vídeo em modo tela cheia



Fonte: Brasil (2023).

A descrição simplificada é um recurso utilizado para atender a usuários com pouca escolaridade ou com problemas com a leitura de textos mais complexos. Este recurso tem semelhança aos que os anglofalantes utilizam no chamado *Plain English*, simplificando a escrita de textos.

Figura 6. Descrição em linguagem simples



Fonte: Brasil (2023).

A fim de incluir o maior número de pessoas possível, independentemente das suas habilidades ou deficiências, o tesaurus traz um recurso que, embora previsto nas recomendações nacionais e internacionais de acessibilidade e nos princípios de design para eficiente comunicação da mensagem, encontrou-se de difícil caracterização e estudos na literatura, qual seja o

recurso de descrição em linguagem simples. A descrição em linguagem simples segue a recomendação do eMAG com a determinação de que as linguagens utilizadas em sistemas digitais governamentais não devam exigir mais do que o nível de ensino fundamental. Assim, um estudo foi realizado no sentido de definir o léxico específico e, a partir dele, a descrição do termo em linguagem simples para incluir pessoas de baixa escolaridade, pessoas com algum déficit cognitivo, entre outras.

7 Conclusão

Considerado dentro da grande área do Design, o Design da Informação (DI) se preocupa com o processo de criação de espaços informacionais centrados nas necessidades dos sujeitos que interagem nesses ambientes como protagonistas, colaboradores e criadores de informação e conhecimento. É uma área do conhecimento focada na apresentação otimizada da informação, para que aqueles que interajam com o ambiente tenham acesso com facilidade e realizem as tarefas com eficiência (HORN, 1999).

A proposta da apresentação da informação no website do tesaurus do projeto Documentos Governamentais Acessíveis contempla os princípios de estrutura, clareza, simplicidade, ênfase e unidade do Design da Informação (PETTERSSON, 2012). Considera, ainda, os princípios do Design Universal e as recomendações de encontrabilidade da informação (VECHIATO; VIDOTTI, 2014).

Com relação à **Estrutura** do website, uma estrutura bem desenhada facilita a percepção, a interpretação, o entendimento, a aprendizagem e a memorização da informação. A informação deve, assim, ser organizada de forma hierárquica, e a organização da estrutura, o conteúdo e seus elementos devem ser limitados e agrupar informações relacionadas de tal forma que a relação fique explicitada.

Os elementos de uma estrutura devem estar alinhados de maneira que o internauta consiga navegar por eles: o layout e a variação tipográfica devem se estruturar de maneira a direcionar o olhar e a leitura na interface, por exemplo, destacar cabeçalhos, utilizar linhas e caixas.

Com relação à **Clareza**: o texto deve ser fácil de ser lido e o leitor deve conseguir distinguir suas partes, mas o mais importante é criar bom contraste entre texto e fundo.

A **Simplicidade**, por outro lado, foi prioridade no projeto, tanto com relação à interface quanto à mensagem textual. O uso das cores deve ser ponderado, não mais do que três ou quatro cores na mesma página.

A **Ênfase** atrai, direciona a leitura e mantém a atenção. Vale destacar que certas partes do texto são lidas com mais ou menos atenção e, outras, evitadas

completamente. Assim, elementos como cabeçalhos, fotos, desenhos e gráficos atraem a atenção das pessoas e, portanto, devem ser os primeiros pontos a serem observados em uma página.

A **Unidade e Harmonia** é observada no uso consistente de estilos, termos, layout e tipografia. Os elementos do website estão equilibrados e simétricos.

Os recursos de acessibilidade foram contemplados com a observação de utilização de multilinguagens. A inserção de recurso de áudio, vídeo com tradução em libras e a descrição do termo ou tópico em linguagem simples.

Vale destacar a atenção especial dada à encontrabilidade da informação, especialmente no planejamento dos sistemas de navegação e busca possibilitados no website.

Notas

[1] Usabilidade entendida como um produto ou ambiente fácil de usar por qualquer pessoa.

[2] Website ainda não publicado.

Referências

- ALVES, A. V. **Informação e inclusão**: um estudo da aplicação de acessibilidade em portais legislativos estaduais. 122f. (Dissertação de Mestrado). Programa de Pós-Graduação em Ciência da Informação. Escola de Ciência da Informação da Universidade Federal de Minas Gerais, 2012.
- BRASIL. **eMAG: Modelo de Acessibilidade em Governo Eletrônico**. Departamento de Governo Eletrônico, 2014. Disponível em: <http://emag.governoeletronico.gov.br>. Acesso em: 19 mar. 2023.
- BRASIL. Lei nº 10.436, de 24 de abril de 2002. Dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências. **Diário Oficial da União**: Seção 1, Brasília, DF, p. 23, 25 abr. 2002.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). Tesouros de Documentos Acessíveis. Brasília, DF: IBICT, 2023. (em fase de elaboração).
- CONNELL, B.; JONES, M.; MACE, R.; MUELLER, J.; MULLICK, A.; OSTROFF, E.; SANFORD, J.; STEINFELD, E.; STORY, M.; VANDERHEIDEN, G. **The Principles of Universal Design**: version 2.0. [s. l.]: Raleigh: Center for Universal Design, North Carolina State University, 1997.
- HORN, R. E. Information Design: emergence of a new profession. In: JACOBSON, R. (ed.). **Information Design**. Cambridge: MIT Press, 1999.
- HORTON, S. **Access by Design**: a guide to Universal Usability for Web Designers. Berkeley: New Riders; Pearson Education, 2006.
- JORENTE, M. J. V.; NAKANO, N.; PADUA, M. C. A **emergência do Design da Informação na contemporaneidade da Ciência da Informação**. 2. ed. São Paulo: Cultura Acadêmica, 2020.
- MARTÍNEZ-ÁVILA, D. Knowledge Organization in the Intersection with Information Technologies. **Knowledge Organization**, [s. l.], v. 42, n. 7, p. 486-498, 2015.
- MARTÍNEZ-ÁVILA, D.; SAN SEGUNDO, R.; ZURIAN, F. A. Retos y oportunidades en organización del conocimiento en la intersección con las tecnologías de la información. **Revista Española de Documentación Científica**, [s. l.], v. 37, n. 3, p. e053, 2014.
- MOORE, M. G.; KEARSLEY, G. **Educação a Distância**: uma visão integrada. São Paulo: Thomson Learning, 2007.
- NAKANO, N. Acessibilidade na Ciência da Informação: uma revisão sistemática da literatura brasileira. **Atoz**: novas práticas em informação e conhecimento, [s. l.], v. 11, p. 1-13, 2022.
- NAKANO, N. **Princípios do Design da Informação na curadoria digital de ambientes virtuais de aprendizagem sob a perspectiva da Ciência da Informação**. 2019. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho”, Marília, 2019.
- NAKANO, N.; MARTÍNEZ-ÁVILA, D.; JORENTE, M. J. V.; CANTISANI, M. Information Design, Information Science, and Knowledge Organization: a domain analysis from the perspective of complexity. **Scire**, [s. l.], v. 24, n. 1, p. 67-75, 2018.
- ORNA, E.; STEVENS, G. Information design and information science: a new alliance? **Journal of Information Science**, [s. l.], v. 17, n. 4, p. 197-208, 1991.
- PETTERSSON, R. **Information Design 1**: message design. Wien: International Institute for Information Design (IIID), 2013.
- PETTERSSON, R. **It Depends**: ID Principles and guidelines. 4. ed. Tullinge: Institute for Infology, 2012.
- SMIRAGLIA, R. P. Knowledge Organization: some Trends in an Emergent Domain. **El Profesional de la Información**, [s. l.], v. 21, n. 3, p. 225-227, 2012.
- VECHIATO, F. L.; VIDOTTI, S. A. B. G. **Encontrabilidade da informação**. São Paulo: Cultura Acadêmica, 2014.

Correlações entre a intencionalidade dos sujeitos informacionais e o conceito de polirrepresentação

Correlations between the intentionality of informational subjects and the concept of polyrepresentation

**Fernanda Alves Sanchez (1), Silvana Ap. B. G. Vidotti (2), Fernando Luiz Vechiato (3),
Sandra Milena Roa-Martínez (4)**

(1) Universidade Estadual Paulista (UNESP), Marília, SP, fernanda.a.sanchez@unesp.br, (2) São Paulo, SP, silvana.vidotti@unesp.br, (3) Universidade Federal do Rio Grande do Norte (UFRN), Natal, RN, vechiato2008@gmail.com, (4) Universidad del Cauca, Colômbia, smroa@unicauca.edu.co

Resumo

A Intencionalidade dos sujeitos informacionais é um conceito formalizado por representações, sejam elas dentro do domínio da linguística ou das percepções visuais. O conceito de Polirrepresentação visa reunir diversas formas de representações de objetos e/ou processos informacionais, a fim de aplicar diferentes tipos de recursos e potencializar os processos de Recuperação da Informação. Nesse sentido, o objetivo do estudo foi identificar as correlações entre os conceitos de Intencionalidade dos sujeitos informacionais e a Polirrepresentação, com o intuito de apresentar contribuições acerca da construção de ambientes informacionais digitais. Trata-se de uma pesquisa que se apresenta como bibliográfica e exploratória de abordagem qualitativa, realizada por meio das bases de dados que trazem representatividade e relevância qualitativa e quantitativa na comunidade científica, sendo elas: o Portal de Periódicos da Capes, a Base de Dados em Ciência da Informação (BRAPCI) e o Repositório Institucional da Unesp. Dentre os resultados obtidos com o levantamento bibliográfico, foi possível estabelecer a correlação das temáticas da Intencionalidade dos sujeitos informacionais com a Polirrepresentação por meio das representações. Por fim, fica claro o potencial de ambos os conceitos em contribuir para a qualidade dos ambientes informacionais digitais, especialmente, no que concerne em melhorias para os sistemas de representação e recuperação, buscando propiciar a Encontrabilidade da Informação.

Palavras-chave: Intencionalidade. Sujeitos Informacionais. Polirrepresentação. Comportamento de Busca.

Abstract

The Intentionality of informational subjects is a concept formalized by representations, whether within the domain of linguistics or visual representations. The concept of Polyrepresentation aims to bring together different forms of representation of objects and/or informational processes, in order to apply different types of resources and enhance the Information Retrieval processes. In this sense, the objective of the study was to identify the correlations between the concepts of Intentionality of informational subjects and Polyrepresentation, with the aim of presenting contributions regarding the construction of digital informational environments. This is a research that presents itself as a bibliographical and exploratory qualitative approach carried out through databases that bring representativeness and qualitative and quantitative conversion in the scientific community, namely: the Portal of Periodicals da Capes, the Database in Information Science (BRAPCI) and Institutional Repository of Unesp.. Among the results obtained with the bibliographical survey, it was possible to establish the correlation of the themes of the Intentionality of the informational subjects with the Polyrepresentation through the representations. Finally, the potential of both concepts to contribute to the quality of digital information environments is clear, especially with regard to improvements to representation and retrieval systems, seeking to provide Information Findability.

Keywords: Intentionality. Informational Subjects. Polyrepresentation. Information Behavior.

Resumen

La Intencionalidad de los sujetos informacionales es un concepto formalizado por representaciones, ya sea dentro del dominio de la lingüística o de las percepciones visuales. El concepto de Polirrepresentación pretende reunir diferentes formas de representación de objetos y/o procesos informacionales, con el fin de aplicar diferentes tipos de recursos y potenciar los procesos de Recuperación de Información. En ese sentido, el objetivo del estudio fue identificar las correlaciones entre los conceptos de Intencionalidad de los sujetos informacionales y la Polirrepresentación, con la intención de presentar aportes sobre la construcción de ambientes informacionales digitales. Esta es una investigación que se presenta como un enfoque cualitativo bibliográfico y exploratorio, realizado a través de bases de datos que aportan representatividad y relevancia cualitativa y cuantitativa en la comunidad científica, a saber: el Portal de Periódicos da Capes, la Base de Datos en Ciencias de la Información (BRAPCI) y Repositorio Institucional de la Unesp. Entre los resultados obtenidos con el levantamiento bibliográfico, fue posible establecer la correlación de los temas de la

Intencionalidad de los sujetos informacionales con la Polirrepresentación a través de las representaciones. Finalmente, es claro el potencial de ambos conceptos para contribuir a la calidad de los entornos de información digital, especialmente en lo que se refiere a mejoras en los sistemas de representación y recuperación, buscando proporcionar Information Findability.

Palabras clave: Intencionalidad. Sujetos informacionales. Polirrepresentación. Comportamiento de Búsqueda.

1 Introdução

No âmbito da Ciência da Informação (CI), entende-se que a Informação é o objeto no qual o sujeito informacional aplica a sua Intencionalidade, dentro desse contexto, entende-se que a Intencionalidade é um conceito inerente ao conceito de Representação tratando-se de um processo anterior. A Intencionalidade é formalizada por meio das representações que podem ser consideradas o: “[...] ato de utilizar elementos simbólicos – palavras, figuras, imagens, desenhos, mímicas, esquemas, entre outros – para substituir um objeto, uma ideia ou um fato” (LIMA; ALVARES, 2012, p. 21).

Nesse sentido, diante das constantes evoluções no cenário das Tecnologias de Informação e Comunicação (TIC) é possível destacar as mudanças no papel dos sujeitos informacionais, que são vistos como sujeitos que consomem, mas também produzem informação constantemente. Nesse contexto, eles passam a contribuir de forma direta com a representação e organização dos dados e informações em ambientes informacionais, desse modo, entende-se que: “[...] ao acessarmos uma informação em qualquer prática cotidiana, estamos acessando uma representação sobre aquilo que é informado” (MASSONI; LUIS, 2018, p. 76).

O conceito de Intencionalidade dos sujeitos informacionais é advindo da Teoria da Intencionalidade e essa está essencialmente ligada à bagagem cognitiva dos sujeitos. A teoria segue a escola do pensamento da Fenomenologia que enxerga a Intencionalidade como sua doutrina nuclear e a define como: “[...] cada ato de consciência que nós realizamos, cada experiência que nós temos, é intencional: é essencialmente "consciência de" ou uma "experiência de" algo ou de outrem” (SOKOLOWSKI, 2004, p. 17).

Em suma, o conceito de Polirrepresentação, se trata de múltiplas e diversas representações de objetos ou processos de informação que visam trazer qualidade ao processo de Recuperação da Informação (LIOMA; LARSEN; INGWERSEN, 2012). Seu princípio abrange todas as sinergias, sejam elas, físicas, cognitivas e/ou pragmáticas, o que amplia e aprimora a capacidade dos ambientes informacionais digitais de fornecer um processo de recuperação eficiente (ROA-MARTÍNEZ, 2019).

Diante dessas explanações, a questão que norteia o desenvolvimento deste estudo, é: Quais são as contribuições do conceito de Polirrepresentação em

ambientes informacionais digitais a partir das correlações com o conceito de Intencionalidade dos sujeitos informacionais?

Desse modo, o objetivo desse estudo busca identificar as correlações entre os conceitos de Intencionalidade dos sujeitos informacionais e a Polirrepresentação, a fim de apresentar contribuições acerca da construção de ambientes informacionais digitais, especialmente, no que condiz a organização, representação, recuperação e encontrabilidade de dados e informação para geração de conhecimento.

2 Procedimentos Metodológicos

Como procedimentos metodológicos a realização desta pesquisa que se apresenta como bibliográfica e exploratória de abordagem qualitativa, foi realizado um levantamento nas seguintes bases de dados: o Portal de Periódicos da CAPES, na Base de Dados em Ciência da Informação (BRAPCI) e no Repositório Institucional da Unesp, no segundo semestre do ano de 2022.

As palavras-chave utilizadas para realização do levantamento bibliográfico foram: Teoria da Intencionalidade, Intencionalidade dos sujeitos informacionais, Polirrepresentação e Encontrabilidade da Informação. As estratégias de busca foram: a) os termos deveriam aparecer nos títulos dos materiais; b) os textos deveriam estar inseridos na área da Ciência da Informação; c) as buscas dos termos foram feitas em três idiomas (português, inglês e espanhol) e d) não delimitar os anos das publicações.

O Quadro 1, que segue, apresenta os resultados obtidos com as pesquisas nas bases de dados mencionadas.

Quadro 1. *Dados do levantamento bibliográfico*

<i>Textos recuperados:</i>	Teoria da Intencionalidade; Intencionalidade dos sujeitos informacionais	36
	Encontrabilidade da Informação	68
	Polirrepresentação	7
<i>Total</i>		111
<i>Textos selecionados:</i>	Teoria da Intencionalidade; Intencionalidade dos sujeitos informacionais	8
	Encontrabilidade da Informação	3
	Polirrepresentação	8
<i>Total</i>		19

Fonte: elaborado pelos autores (2023).

É importante ressaltar que os números mencionados na primeira parte do quadro refletem a quantidade total dos resultados recuperados a partir dos termos mencionados. Na segunda parte, são apresentados os números referentes aos textos utilizados para a construção desse texto. Após o levantamento foi necessário uma análise dos materiais para o enquadramento na pesquisa, do total de dezenove textos, três deles são referências conhecidas a partir de pesquisas anteriores visando enriquecer as discussões presentes nesse trabalho.

3 Intencionalidade dos sujeitos informacionais

O conceito de Intencionalidade dos sujeitos informacionais é advindo da Teoria da Intencionalidade que tem como um dos pioneiros o filósofo John Searle. Ao falar de Intencionalidade se faz a associação com a intenção no sentido de ação, propósito, desejo etc., entretanto, Searle (2002, p. 4), afirma que: “[...] Intencionalidade é direcionalidade; ter a intenção de fazer algo é apenas uma forma de Intencionalidade entre outras”.

Os primeiros estudos sobre Intencionalidade que foram desenvolvidos no Brasil, derivam-se da tese de Miranda (2010), que devolve o estudo sobre Intencionalidade no contexto da *Findability* e posteriormente na tese e no livro sobre Encontrabilidade da Informação de Vechiato (2013) e Vechiato e Vidotti (2014), onde a Intencionalidade dos sujeitos informacionais é apresentado como um dos atributos dentro dos aportes operacionais e conceituais desse conceito.

Para Miranda (2010, p. 273) a definição do conceito de Intencionalidade para a *Findability*, significa:

[...] *direcionalidade* de informação e se funda na experiência de cada sujeito (*user experience* para controle na produção, organização e partilha de informação); a informação que é produzida é sempre *acerca de e dirigida a*, isto é, um sujeito, com a sua experiência, cria informação acerca de e dirigida a para atingir seus objetivos. E é nesse sentido, da capacidade da experiência do usuário, da consciência Intencional, que se baseia a web da inovação, do paradigma atual (MIRANDA, 2010, p. 273, grifo do autor).

A definição do atributo de Intencionalidade dos sujeitos informacionais para Vechiato e Vidotti (2014), é definido como um elemento essencial no que diz respeito aos ambientes informacionais digitais, pois ele é carregado de:

[...] experiências, necessidades e competências (tanto as informacionais quanto as tecnológicas), entendimento, cognição e satisfação, fornecendo, inclusive, subsídios para a estruturação de sistemas e ambientes informacionais (VECHIATO; VIDOTTI, 2014, p. 113).

Searle (2002), na capa de seu livro coloca a seguinte frase: “[...] A representação de uma sentença deriva da Intencionalidade da mente”. A sentença é considerada

pelo autor, estados intencionais, nelas estão estados intrínsecos da mente humana, como as crenças e os desejos, essas sentenças podem ser “[...] os sons emitidos pela boca ou os sinais gráficos que se fixam no papel” (SEARLE, 2002, p. 8).

Com base nessa afirmação, é possível identificar dois grandes pilares que fundamentam o desenvolvimento dos estudos sobre Intencionalidade, sendo eles: a linguagem e as percepções visuais. Nesse texto, esses alicerces serão chamados de domínios, Searle (2002), os identificam como: 1) a Intencionalidade linguística ou Intencionalidade da linguagem, que está ligada aos atos de fala e 2) a Intencionalidade da percepção direcionada, especialmente, às percepções visuais.

Nesse contexto, dentre os estudos que John Searle desenvolveu ao longo da vida sobre a Teoria da Intencionalidade, são encontrados frequentemente alguns elementos nos textos do filósofo que fundamentam o conceito de Intencionalidade, como por exemplo: os estados mentais, os estados intencionais, a Rede, o *Background*, a direção de ajuste, as condições de satisfação e direcionalidade etc.

É de suma importância entender cada um desses para posteriormente compreender como eles se enquadram dentro dos domínios. O Quadro 2, apresenta de forma sucinta esses elementos e suas definições.

O Quadro 2, apresenta a definição de alguns dos elementos mais pronunciados por Searle para compor as discussões sobre a Intencionalidade, é importante ressaltar que a teoria não se resume a isso. Desse modo, o primeiro domínio se trata da Intencionalidade linguística ou Intencionalidade da linguagem e está diretamente ligado ao que Searle chama de atos de fala, na CI temos os estudos sobre semântica, linguística e metadados que podem se enquadrar nesse domínio. Vale ressaltar que Searle (2002, p. 7), afirma que:

Ao explicar a Intencionalidade em termos de linguagem, não pretendo sugerir que a Intencionalidade é essencial e necessariamente linguística. Ao contrário, parece-me obvio que os recém-nascidos e muitos animais que, em um sentido ordinário, não possuem uma linguagem nem realizam atos de fala apresentam, mesmo assim, estados Intencionais.

Sendo assim, os atos de fala são considerados as representações dos estados intencionais dos sujeitos, para Miranda (2010, p. 157), é possível a partir desses atos identificar a Intencionalidade dos sujeitos informacionais:

[...] porque, por meio deles, são expressas as crenças, desejos, temores, dúvidas; são compostos por um conteúdo proposicional, aquilo em que se acredita, se deseja, teme, duvida (MIRANDA, 2010, p. 157).

Os atos de fala são compostos por três (3) pontos importantes sendo eles: I) a direção de ajustamento, a II) sinceridade ao expressar um ato de fala com conteúdo

proposicional e III) as condições de satisfação (MIRANDA, 2018).

Quadro 2. *Elementos da Teoria da Intencionalidade*

<i>Elementos</i>	<i>Definições</i>
Estados mentais	Estados mentais são estados intrínsecos do ser humano, ligados às operações do cérebro, que podem ser conscientes e inconscientes. Segundo Searle (2002) diz que os estados mentais são causados por fenômenos biológicos que acabam causando outros fenômenos biológicos. Nem todos os estados mentais serão intencionais, como, por exemplo: a ansiedade, exaltação e melancolia (SEARLE, 2002; MIRANDA, 2018).
Estados Intencionais	Estados intencionais podem ser causados por meio desses fenômenos biológicos, entretanto, estão ligados à direcionalidade desses estados mentais, além das condições de satisfação dessa intenção (intenção no sentido literal de ação). Alguns exemplos de estados intencionais, podem ser: admiração, respeito, amor, vergonha etc. (SEARLE, 2002).
Rede	Carrega a junção dos estados: mentais e intencionais, e conseqüentemente influencia nas condições de satisfação dos estados intencionais, como por exemplo: particularidades, interpretações, significados etc. (SEARLE, 2002; CARVALHO, 2021)
Background	Um “[...] conjunto de capacidades mentais não-representacionais que permite a ocorrência de toda representação (SEARLE, 2002, p. 198). É dividido em dois (2) momentos, sendo o primeiro: O <i>Background</i> de base ou profundo: aquele relacionado a constituição biológica do ser humano (habilidades, competências, comportamentos); O <i>Background</i> local: aquelas experiências adquiridas culturalmente em uma vida em sociedade, ou seja, essas capacidades são adquiridas e podem sofrer influências diante do contexto (social, cultural, político, tecnológico etc.) (SEARLE, 2002; CARVALHO, 2021).
Direção de ajuste	A direção de ajuste está ligada ao conteúdo intencional sob diferentes perspectivas, como por exemplo, as de mente-mundo e mundo-mente. Dessa forma, os estados intencionais com direção de ajuste são considerados representações das condições de satisfação de um sujeito. Sendo assim, para compreender se um estado intencional tem uma direção de ajuste e se essa direção de ajuste irá atender as condições de satisfação do estado intencional, Searle (2002, p. 2) afirma que: “[...] se um estado E é Intencional, deve haver uma resposta para perguntas como: A que se refere E? Em que consiste E? O que é um E tal que?.”

Fonte: Elaborado pelos autores (2023).

O primeiro (I) ponto, está relacionado ao objetivo do ato de fala, sendo caracterizados como: enunciados, descrições e asserções (atos de fala assertivos) e as ordens, comandos e solicitações (atos de falas diretivos). O segundo (II), é sobre o conteúdo do ato de fala se caracterizar como um conteúdo proposicional (desejos, dúvidas, crenças) ou ao seu conteúdo representativo, ou seja, um ato de fala pode ser insincero, irônico ou até mesmo uma mentira – nesse caso, não há uma direção de ajuste e o conteúdo proposicional passa a ser pressuposto pelo outro. O terceiro (III), está ligado a satisfação ou sucesso do ato de fala, ou seja, do atendimento ou entendimento ao fazer um pronunciamento ou uma ordem, essas condições são internas aos estados intencionais (MIRANDA, 2018).

Isto posto, o segundo domínio trata-se da Intencionalidade da percepção, especialmente, as percepções visuais, este domínio pode estar relacionado diretamente às imagens (analógicas e digitais), ícones e símbolos tratados na CI. Dentro desse contexto, alguns pontos estão ligados a esse domínio, sendo eles: a diferença entre apresentação e representação, bem como os aspectos ligados ao conceito de significados que estão imbricados a discussões como particularidades, perspectivas e condições de satisfação diferentes sob o mesmo objeto para cada sujeito informacional.

Dentre os elementos da Intencionalidade apresentados anteriormente, a Rede e o *Background*, são especialmente responsáveis por afetar e influenciar a percepção dos sujeitos Searle (2002, p. 76), aponta que:

[...] muitas de nossas experiências visuais sequer são possíveis sem o domínio de certas capacidades de *Background*, entre as quais figuram com destaque as linguísticas.

Podemos observar que as representações são formadas pelos dois grandes domínios atribuídos pela Teoria da Intencionalidade, como a Intencionalidade da linguagem e a Intencionalidade da percepção.

Na Teoria da Representação, que influenciou diretamente o que conhecemos no âmbito da CI como a área de Representação da Informação, temos a: Representação da palavra e a Representação do objeto. Ambas as teorias se relacionam e contribuem diretamente para a CI quando se trata da construção de ambientes informacionais.

Constatando o que Searle aponta sobre as capacidades de *Background* e a linguística, Andrade (2016, p. 282, tradução nossa), diz que: “[...] quando a conexão da palavra com a representação do objeto é estabelecida, o aparelho de linguagem articula representações que produzem um sentido”, ou seja, essa afirmativa esclarece a importância do visual e da linguagem e como elas são imbricadas uma na outra.

No que diz respeito aos pontos apresentados na Intencionalidade da percepção, há discussões sobre os

diferentes conceitos de apresentação e representação. Resumidamente, é possível compreender que o conceito de apresentação está ligado a ações, ao modo como são apresentadas as coisas ao mundo e aos sujeitos, já o conceito de representação está direcionada à bagagem que essa coisa carrega e suas significações para o mundo e para os sujeitos, ainda que muito parecidas são não idênticas, assim como os estados mentais e intencionais.

Dentro do contexto da Intencionalidade da percepção, onde os conceitos de apresentação e representação estão ligados a experiência visual, é necessário compreender que a experiência visual são apresentações das condições de satisfação e são elas que determinam essa experiência, ou seja, a experiência visual é carregada de apresentações e tornam-se representações com significados a partir, por exemplo, da crença e da lembrança de cada sujeito.

Um exemplo utilizado pelo filósofo John Searle é a de uma flor, desse modo: dentro da percepção visual temos a experiência visual de ver a flor, ou seja, trata-se de uma apresentação, conseqüentemente nossa bagagem cognitiva associa a flor a uma lembrança, desse modo, a flor passa a ser uma representação, pois dentro dessa lembrança há experiências e associações a momentos em que o sujeito viu flores ao longo da vida, e por fim, as condições de satisfação para esse sujeito será determinada por meio dessa experiência. (SEARLE, 2002).

Nesse contexto, Arnao (2008, p. 200), atribui que a:

[...] representação é algo que está no lugar de outra coisa, e por conseguinte é intencional e, ao mesmo tempo tanto as representações de palavra, que não remetem a uma única coisa, nem as representações de coisa a um único traço mnêmico ou objeto indiferente à maneira de objetivo, então o próprio conceito de intencionalidade se torna mais complexo e, com ele, o de significado.

Sob o ponto de vista do significado Searle (2002), declara que há dois aspectos nas intenções de significação: sendo a intenção de representar e a intenção de comunicar. Ainda que parte do que comunicamos como sujeitos informacionais sejam as nossas próprias representações, é possível ter a intenção de representar, sem ter a intenção de comunicar, mas não o contrário.

Para Carvalho (2021, p. 5), o significado na versão searleana é um tipo de intencionalidade que por: “[...] “intencionalidade” entende-se a capacidade que a mente possui de se dirigir ou representar fatos no mundo”.

Desse modo, ao visitar partes dos aportes teóricos da Teoria da Intencionalidade, é possível enfim definir que o termo Intencionalidade dos sujeitos informacionais é adequado diante dos elementos e domínios que à compõem estarem imbricados puramente a essência e ao potencial que apenas os seres humanos apresentam em sua bagagem cognitiva.

Além disso, é possível afirmar que se trata de uma temática necessária para pertencer e agregar aos estudos na área da Ciência da Informação, visto que desde os primórdios tem como foco os sujeitos informacionais. Atualmente a CI vem compreendendo a importância e o papel dos sujeitos informacionais não apenas como consumidor da informação e sim, como um grande aliado para gerar, organizar e representar informação, especialmente, no âmbito digital.

Desse modo, a Intencionalidade dos sujeitos informacionais está diretamente relacionada ao conceito de Representação da Informação. A Intencionalidade é um processo anterior à construção das representações, assim que definidas, elas formalizam a Intencionalidade dos sujeitos o que acabam por indicar, por exemplo, as suas necessidades informacionais.

Por fim, é possível dizer que dentro do conceito de Intencionalidade fica claro que a ideia de representação não é caracterizada apenas como linguagem ou como visual, esses conceitos são complementares.

4 O conceito de Polirrepresentação

Entende -se a Polirrepresentação como um conceito que prioriza:

[...] além de diversas representações dos tipos de conteúdo e informações dos recursos informacionais, a polirrepresentação considera a representação da necessidade informacional do usuário, buscando tornar o Sistema de Recuperação da Informação uma plataforma contextual enriquecida (ROA-MARTÍNEZ, 2019, p. 24).

Dentre as áreas do conhecimento que complementam o objetivo geral da Polirrepresentação e estão inseridas no âmbito da CI temos: a Representação da Informação, a Recuperação da Informação, a Web Semântica, a Organização do conhecimento, o Comportamento de Busca e a Encontrabilidade da Informação.

Os estudos da Polirrepresentação foram iniciados por Peter Ingwersen em 1996, desde então o conceito vêm sendo ampliado, principalmente, no que condiz às suas funções operacionais, para o autor a Polirrepresentação:

[...] procura representar a necessidade de informação corrente do usuário, estados de problema e conhecimento, domínio da tarefa de trabalho ou interesse na forma de estruturas contextuais da causalidade. Ao mesmo tempo, implica que devemos aplicar diferentes métodos de representação e uma variedade de técnicas de RI [Recuperação da Informação], de diferente origem cognitiva e funcional, aos objetos de informação, no espaço de informação. Os objetivos são melhorar o acesso intelectual às fontes de informação e, ao mesmo tempo, fornecer ao SRI [Sistema de Recuperação da Informação] uma plataforma contextual enriquecida, que possa sustentar a busca de informação do usuário (INGWERSEN, 1996, p. 4, tradução nossa).

Dando continuidade aos estudos de Ingwersen (1996), Lioma *et al.*, (2010) propõe um modelo matemático tratável e flexível denominado *Subjective Logic* – Lógica Subjetiva, que visa expressar a Polirrepresentação das necessidades de informação do usuário e seu contexto, os autores entendem que a Polirrepresentação abrange em seu conceito dois tipos de representação, sendo eles:

As representações cognitivamente diferentes: que foram derivadas das interpretações de diferentes atores e;

As representações funcionalmente diferentes: que foram derivadas do mesmo ator, como o autor gerando estruturas de texto, recursos de imagem, legendas de diagramas e referências (LIOMA *et al.*, 2010, p. 2-3).

Seguindo o princípio da Polirrepresentação, os autores acreditam que quanto mais representações, sejam elas cognitivas ou funcionais, maiores serão a qualidade da recuperação das informações dos objetos digitais. O modelo de Lioma *et al.*, (2010), declara a importância da estruturação do sistema de recuperação da informação, pois são eles que fornecem os principais elementos da Polirrepresentação, tanto por parte do usuário, quanto pela perspectiva do sistema.

Dentre os elementos advindos por meio das bases de entrada do sistema são encontrados: as relevâncias e as percepções implícitas e explícitas de informação do usuário, os *feedbacks* do usuário, as classificações, as recomendações, as tags, as citações etc., essas permitem que o sistema de recuperação seja personalizado, tanto por uma ótica social (comunidade), quanto individual. Já nos elementos de saída do sistema de recuperação, temos as seguintes opções: mudanças nos mecanismos de busca, algoritmos, erros, vies, recursos das coleções e recuperação etc.

Para Lioma *et al.*, (2010, p. 9, tradução nossa), é importante que existam possibilidades para modificar a forma clássica de consulta do usuário, isso pode ser realizado:

[...] através de listas de termos sugeridos, trechos, estruturas conceituais, representações de ontologias etc. Todos esses podem ser vistos como diferentes representações feitas a partir de diferentes perspectivas e interpretações (LIOMA *et al.*, 2010, p. 9, tradução nossa).

O estudo de Lioma *et al.*, (2010), por meio de uma formalização matemática, considerada tratável e flexível visa a extração das representações (cognitivas e funcionais), dentro das áreas da Ciência da Informação e da Lógica Subjetiva, a fim de, potencializar principalmente a Recuperação da Informação.

No Brasil, o conceito de Polirrepresentação é apresentado por Roa-Martínez (2019) em sua tese de doutorado na área da CI no contexto das imagens digitais. Para Roa-Martínez (2019, p. 172), a Polirrepresentação:

[...] se torna relevante quando se pretende fornecer e aprimorar a Recuperação da Informação desde suas diversas óticas, sejam estas, a física, a cognitiva e a pragmática, pois não só centra a atenção na descrição do recurso informacional, nas necessidades do usuário ou nas técnicas usadas pelo sistema para recuperar, mas sim, às abrange todas com suas sinergias (ROA-MARTÍNEZ, 2019, p. 172).

No âmbito das imagens digitais, o conceito de Polirrepresentação:

[...] permitirá ao sistema oferecer diversas opções e mecanismos de recuperação que quando enriquecida pelas tecnologias da Web Semântica, como a ontologia *OntoImage*, ampliarão os espaços de busca que podem surgir durante a interação do usuário com o sistema. [...] A polirrepresentação constitui-se no elemento integrador da representação do recurso imagético e das necessidades do usuário que contempla seu contexto, as informações e conteúdos visuais e não visuais (ROA-MARTÍNEZ, 2019, p. 213-218).

A ontologia *OntoImage*, foi criada no estudo de Roa-Martínez (2019), a partir da necessidade de uma ontologia para imagens digitais. Ela abrange os conteúdos visuais, informações não visuais e de contexto, propiciando a recuperação e a encontrabilidade dos recursos imagéticos, a ampliação das fontes de consulta, a interoperabilidade, dentre outras vantagens (ROA-MARTÍNEZ, 2019).

Vale ressaltar que a Polirrepresentação de Roa-Martínez (2019), foi construída com foco na melhoria de encontrabilidade de imagens digitais, seu modelo foi baseado nos estudos anteriores sobre Polirrepresentação e tecnologias da Web Semântica, além das temáticas da: Recuperação da Informação e Comportamento de Busca da Informação.

Ainda que o direcionamento do seu estudo está em imagens digitais, os componentes práticos e conceituais apresentados sobre a Polirrepresentação podem ser adaptados a todo e qualquer tipo de ambiente informacional digital, pois se trata de recursos e tecnologias que enriquecem a capacidade do sistema, a qualidade da encontrabilidade da informação e consequentemente a experiência do sujeito.

5 Discussão e Resultados

Apresentados os referenciais teóricos, nessa seção serão realizadas algumas discussões sobre as correlações das temáticas da Intencionalidade dos sujeitos informacionais e da Polirrepresentação dentro do âmbito da Ciência da Informação.

Primeiramente, para correlacionar o conceito de Polirrepresentação a Intencionalidade dos sujeitos informacionais é importante esclarecer alguns pontos conceituais sobre a Encontrabilidade da Informação.

Na tese de Roa-Martínez (2019), a autora utiliza o termo em inglês do conceito de Encontrabilidade da Informação, no caso, a *Information Findability* – nele são posicionados alguns dos elementos da Encontrabilidade da Informação, como por exemplo, o fato de que a Encontrabilidade é transversal à Recuperação da Informação e ao Comportamento de Busca, além de deixar mais claro que a Recuperação da Informação está relacionada com o sistema (máquina), já a Encontrabilidade da informação tem ligação com os sujeitos informacionais, pois ela só poderá ser determinada a partir das condições de satisfação e pertinência desses.

A autora apresenta uma definição para a *Information Findability*, como a:

[...] capacidade do usuário de determinar o encontro dos recursos informacionais pertinentes em um ambiente informacional a partir dos resultados entregues pelo Sistema de Recuperação da Informação; e para que a *Information Findability* aconteça se pressupõe a existência das habilidades, os conhecimentos, as experiências e outros elementos contemplados pelo Comportamento de Busca da Informação próprios do contexto do usuário (ROA-MARTÍNEZ, 2019, p. 102).

Dito isso, os autores Sanchez, Vidotti, Vechiato e Roa-Martínez (2022) concordam que o termo em inglês *Information Findability* deve ser utilizado apenas como tradução para as publicações realizadas sobre o conceito de Encontrabilidade da Informação, evitando confusões já que o termo era utilizado com essa finalidade em alguns estudos anteriores. Desse modo, é mantida a trajetória do conceito consolidado em âmbito brasileiro desde 2013 e são apresentados novos posicionamentos e contribuições para o conceito de Encontrabilidade. Sendo assim, ao ler *Information Findability* deve-se compreender que se trata do conceito de Encontrabilidade da Informação.

Por conseguinte, a Polirrepresentação de Roa-Martínez (2019), aborda o conceito de Comportamento de busca do usuário e posiciona o usuário nos processos que antecedem e sucedem a Recuperação da Informação.

Assim, sobre o processo que antecede a RI entende-se que o usuário influencia a entrada das informações e o sistema deve apresentar os resultados da melhor forma possível a partir de um determinado contexto. Sobre o processo posterior a RI interpreta-se que o usuário irá determinar se a informação realmente foi encontrada e em qual nível suas condições de satisfação foram atendidas.

É nesse momento, quando Roa-Martínez (2019) atribui toda essa relevância ao usuário e a sua capacidade de encontrabilidade é que podemos fazer a relação com a Intencionalidade dos sujeitos informacionais.

Após a investigações sobre o conceito de Intencionalidade é possível dizer que o Comportamento

de busca do usuário faz parte do conceito de Intencionalidade, ou seja, a Intencionalidade tem o comportamento como um dos seus fundamentos.

Dessa maneira, entende-se que é possível analisar o comportamento do sujeito, a fim de, extrair dados que possam contribuir para o ambiente, entretanto, a partir da literatura investigada, defende-se a ideia de que devemos ir para além do comportamento, buscando compreender e coletar elementos advindos da Intencionalidade dos sujeitos informacionais. Vale ressaltar que a Intencionalidade é vista em:

[...] todos os “tipos” de sujeitos informacionais ligados aos projetos dos ambientes (informáticos, profissionais da informação e usuário) e esses manifestam suas habilidades, competências e comportamentos nos níveis que lhe são cabíveis dentro do ambiente gerando assim, representações que afetam positivamente ou negativamente os sistemas de recuperação da informação, tanto nos mecanismos de busca, quanto na navegação por meio da interface. Desse modo, esses processos acarretam a encontrabilidade que só pode ser determinada a partir das condições de satisfação do sujeito (SANCHEZ *et al.*, 2022, p. 13).

Nesse sentido, a Intencionalidade dos sujeitos informacionais é formada por tudo o que os antecede, o que acarreta a formação de sua bagagem cognitiva e é essa bagagem a responsável por influenciar todas as suas ações.

A Intencionalidade, como visto, não se dá de maneira isolada, mas é dependente da relação. Sendo assim, vários personagens devem ser considerados no âmbito da intencionalidade: o sujeito informacional, os profissionais da informação, o produtor da informação, o espaço informacional, o tipo de informação, a linguagem, o material em que a informação se plasma, o contexto, entre outros (SANCHEZ *et al.*, 2022, p. 10).

Em seguimento às discussões sobre a correlação das temáticas de Intencionalidade dos sujeitos informacionais e a Polirrepresentação, a seguir serão apresentadas discussão sob um olhar mais técnico e prático a partir de estudos na área da CI.

Os autores Guo, Li e Shao (2012), mencionam estudos envolvendo tecnologias que possibilitem uma extração da Intencionalidade, nesse caso, a intencionalidade cognitiva direcionada a níveis de discurso. Segundo os autores há dois desafios nesse tipo de extração que são:

[...] (1) alcançar resultados de extração com alta precisão e (2) elevar a tecnologia de processamento de níveis lexicais e sintáticos para níveis de discurso mais elevados (GUO; LI; SHAO, 2012, p. 36, tradução nossa).

Esse modelo, elaborado a partir de bases matemáticas, denominado *Pragmatic-Tree* (P-Tree) é considerado: “[...] uma ferramenta heurística para produzir uma abordagem da funcionalidade do discurso” (GUO; LI; SHAO, 2012, p. 37, tradução nossa).

O estudo supracitado, pode ser correlacionado ao estudo de Lioma, Larsen e Ingwersen (2012), pois buscam a

partir de uma formalização matemática extrair representações usando as crenças e proposições dos sujeitos para compor a Polirrepresentação que irá potencializar os sistemas de Recuperação da Informação.

Entende-se que esses estudos estão ligados à Intencionalidade, pois ambos buscam extrair representações resultadas dos atos dos sujeitos informacionais e como foi dito anteriormente, as representações são a formalização da Intencionalidade. Entretanto, é importante afirmar que a extração que Guo, Li e Shao (2012) propõem pode ser caracterizada como uma extração de elementos encontrados no conceito de Comportamento de Busca da Informação e não da Intencionalidade em si.

Desse modo, diante da complexidade e da ambiguidade em que as representações podem apresentar por serem passíveis de interpretação dentro da bagagem cognitiva de cada sujeito, a Polirrepresentação se apresenta como uma vantagem, para que humanos e máquinas tenham opções para recuperar e encontrar a informação desejada. Além disso, a Polirrepresentação irá potencializar os recursos inseridos em diversas partes dos projetos e implementação de ambientes informacionais digitais, sendo a Polirrepresentação o conceito que traz a formalização da Intencionalidade dos sujeitos informacionais por meio de diferentes formas de representação.

Além disso, Ferreira (2018) afirma que quanto mais recursos foram implementados em uma interface, melhor será a experiência do sujeito ao interagir com o ambiente. Nesse sentido, encontramos também na Polirrepresentação o mesmo princípio direcionando nessa funcionalidade, especialmente, para os motores de busca nos ambientes, o que irá auxiliar diretamente na Representação e Recuperação da Informação e consequentemente na Encontrabilidade da Informação.

6 Considerações Finais

Diante do objetivo proposto, é possível afirmar que é por meio das representações que a correlação entre a Polirrepresentação e a Intencionalidade dos sujeitos informacionais é estabelecida.

A Polirrepresentação por ser um conceito que busca agregar diversas formas de representação pode ser visto como um aspecto prático da formalização da Intencionalidade. Além disso, é importante posicionar a Intencionalidade dos sujeitos informacionais como um conceito que abarca o Comportamento de Busca da Informação e assim, amplia e potencializa o conceito da Encontrabilidade da Informação.

Nesse sentido, se considera importante a interdisciplinaridade com áreas do conhecimento como a Ciência da Computação e os Sistemas de Informação, que desenvolvem estudos envolvendo recursos

tecnológicos da Web Semântica e da Inteligência Artificial, a fim de, fortificar o conceito de Polirrepresentação e a qualidade com que ocorrem a recuperação de objetos digitais, sejam eles textuais ou imagéticos.

A Intencionalidade dos sujeitos informacionais deve ser vista como um fundamento teórico forte, mas que apresenta características práticas que devem estar atrelados aos conceitos ligados à construção de ambientes informacionais digitais. Isso acontece, pelo fato de que todo trabalho humano desenvolvido em um projeto é carregado de Intencionalidade, dessa forma, ela estará presente direta ou indiretamente desde o primeiro momento.

Portanto, é credível que a inclusão de diversas formas e recursos representacionais realizada a partir dos princípios da Polirrepresentação e atreladas aos aspectos extraídos da bagagem cognitiva dos sujeitos informacionais (comportamentos, competências, habilidades e experiências) podem propiciar qualidade na eficiência e eficácia dos espaços digitais.

Referências

- ANDRADE, C. B. The Specificity Of Language In Psychoanalysis. **Ágora: Estudos em Teoria Psicanalítica**, v. 19, p. 279-294, 2016. Disponível em: <https://doi.org/10.1590/S1516-14982016002009>. Acesso em: 25 jan. 2023.
- ARNAO, M. A distinção entre representação de palavra e representação de coisa na obra freudiana: mudanças teóricas e desdobramentos filosóficos. **Ágora: Estudos em Teoria Psicanalítica**, v. 11, n. 2, p. 187-201, 2008. Disponível em: <https://doi.org/10.1590/S1516-14982008000200002>. Acesso em: 19 jan. 2023.
- CARVALHO, J. M. Searle e os desafios da inteligência artificial (IA) forte. **Revista de Filosofia Reflexões**, v. 2021, n. 18, p. 1, 2021. Disponível em: https://revistareflexoes.com.br/wp-content/uploads/2021/03/1.2-NB-Joelma-Carvalho_Publica%C3%A7%C3%A3o.pdf. Acesso em: 20 jan. 2023.
- FERREIRA, A. M. J. F. C. **Contribuições da experiência do usuário para a arquitetura da informação**. 165 f. Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2018. Disponível em: <http://hdl.handle.net/11449/157487>. Acesso em: 27 jan. 2023.
- GUO, Y; LI, Y; SHAO, Z. Cognitive intentionality extraction from discourse with pragmatic-tree construction and analysis. **Information Sciences**, v. 214, p. 35-55, 2012. Disponível em:

<https://doi.org/10.1016/j.ins.2012.06.002>. Acesso em: 26 jan. 2023.

INGWERSEN, P. Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of documentation*, v. 52, n. 1, p. 3-50, 1996.

LIMA, J. L. O.; ALVARES, L. Organização e representação da informação e do conhecimento. **Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações**. São Paulo: B4 Editores, v. 248, p. 21-48, 2012.

LIOMA, C *et al.* A subjective logic formalisation of the principle of polyrepresentation for information needs. In: **Proceedings of the third symposium on Information interaction in context**. 2010. p. 125-134. Disponível em: <https://arxiv.org/pdf/1704.01610.pdf>. Acesso em: 05 jan. 2023.

LIOMA, C; LARSEN, B; INGWERSEN, P. Preliminary experiments using subjective logic for the polyrepresentation of information needs. In: **Proceedings of the 4th Information Interaction in Context Symposium**. 2012. p. 174-183. Disponível em: <https://arxiv.org/pdf/1704.01603.pdf>. Acesso em: 05 jan. 2023.

MASSONI, F. H; LUIS, V. J. M. Ética e teoria das representações sociais: uma discussão a partir da ciência da informação. **Logeion: filosofia da informação**, v. 4, n. 1, p. 73-85, 2018. Disponível em: <https://brapci.inf.br/index.php/res/download/43385>. Acesso em: 15 jan. 2023.

MIRANDA, M. K. F. de O. **O acesso à informação no paradigma pós-custodial: da aplicação da intencionalidade para findability**. 353f. Tese (Doutorado em Informação e Comunicação em Plataformas Digitais) – Faculdade de Letras, Universidade do Porto. Disponível em: <https://hdl.handle.net/10216/50422>. Acesso em: 20 jan. 2023.

MIRANDA, M. K. F de O. **A Teoria da Intencionalidade e a Informação: definições e propriedades para o fenômeno**. In: ROJAS, M. Á. R. (Ed.). **La intencionalidad en la Ciencia de la**

Información Documental. UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información, 2018b. 172p. Disponível em: http://ru.iibi.unam.mx/jspui/handle/IIBI_UNAM/L156. Acesso em: 08 jan. 2023.

ROA-MARTÍNEZ, S. M. **Da information findability à image findability: aportes da polirrepresentação, recuperação e comportamento de busca**. 235f. Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2019. Disponível em: <http://hdl.handle.net/11449/182465>. Acesso em: 10 jan. 2023.

SANCHEZ, F. A. *et al.* Intencionalidade e mediação da informação no contexto dos ambientes informacionais digitais. **Em Questão**, Porto Alegre, v. 28, n. 2, p. 111998, 2022. Disponível em: <https://doi.org/10.19132/1808-5245282.111998>. Acesso em: 20 fev. 2023.

SANCHEZ, F. A. *et al.* Encontrabilidade da Informação e *Information Findability*: relações com a Intencionalidade. In: **Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), XXII, 2022**, Porto Alegre. ISSN 2177-3688. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxii/enancib/paper/viewFile/1119/793>. Acesso em: 08 jan. 2023.

SEARLE, J. R. **Intencionalidade: um ensaio de filosofia da mente**. Lisboa: Relógio D'Água Editores, 1999. 346 p.

SOKOLOWSKI, R. **Introdução à Fenomenologia**. São Paulo: Edições Loyola, 246p. 2004.

VECHIATO, F. L. **Encontrabilidade da informação: contributo para uma conceitualização no campo da Ciência da Informação**. 2013. Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2013. Disponível em: <http://hdl.handle.net/11449/103365>. Acesso em: 21 set. 2021.

VECHIATO, F. L; VIDOTTI, S. A. B. G. **Encontrabilidade da informação**. São Paulo: Cultura Acadêmica, 2014. (Coleção PROPG Digital-UNESP). Disponível em: <https://repositorio.unesp.br/handle/11449/126218>. Acesso em: 06 jan. 2023.

Dados de pesquisa em enfermagem: um estudo sobre a organização para transformação em Dados *FAIR* (*FAIRification Process*)

Research data in nursing area: a study on planning for transformation into FAIR data (FAIRification Process)

Cláudio José Silva Ribeiro (1)^{||}, Isaque Fontes Braga (2)

(1) Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Av. Pasteur, 458, Rio de Janeiro, claudio.ribeiro@unirio.br, (2) isaque.unirio@gmail.com

Resumo

As dificuldades no processo de implantação do compartilhamento de dados, em especial com a prática dos princípios *FAIR*, permeiam várias áreas do conhecimento. Este relato trata de um estudo sobre o compartilhamento, uso e reúso de dados, voltado para a área de Enfermagem. Investiga artigos que contenham conjuntos de dados para a execução de um processo de *FAIRification*, além de realizar um levantamento sobre o processo de transformação de *datasets* presentes na área de Enfermagem. Caracteriza-se como de natureza qualitativa para a análise de dados, com análise bibliográfica e documental na BRAPCI e em portais da área de saúde e no Google Acadêmico. Os resultados apontam para a literatura sobre *FAIR* e para as iniciativas em curso na área de Enfermagem, além de uma proposta para realizar as atividades de transformação dos *datasets* em *FAIR*, com possível aplicabilidade para a área. Conclui-se que, para torná-los *FAIR*, o processo de transformação de dados de pesquisa ainda é uma iniciativa difícil de ser conduzida e com algumas lacunas na área de Enfermagem.

Palavras-chave: *FAIR*. *FAIRification*. Conjuntos de dados. Enfermagem. Compartilhamento de dados.

Abstract

In order to build any approach for data sharing using the practice for Open Science, we have to overpass some difficulties, especially with the practice of the *FAIR* principles that pass through several areas of knowledge. This research deals with studies about the sharing, use and reuse of data aimed at the Nursing in Health Science field. It explores articles that have datasets for the execution of *FAIRification* workflow and comes up with more data about the process of transforming datasets present in the Nursing area. It is characterized by a qualitative approach of the data analysis, with bibliographic and documentary analysis on BRAPCI, on Health Science portals and Google Scholar. The results point to the literature on *FAIR* and ongoing initiatives in the Nursing area, in addition to a proposal to carry out dataset transformation activities in *FAIR* with a possible application in the area. It is therefore concluded that, to make them *FAIR*, the data transformation process is still a difficult initiative to be conducted and with some gaps in Nursing area.

Keywords: *FAIR*. *FAIRification*. Datasets. Nursing. Data sharing.

Resumen

Las dificultades en el proceso de implantación del uso compartido de los datos, sobre todo con la práctica de los principios *FAIR*, permean diversas áreas del conocimiento. Este informe se basa en un estudio sobre el intercambio, el uso y la reutilización de datos, vuelto hacia el área de la Enfermería. Investiga artículos que contienen conjuntos de datos para ejecutar el proceso de *FAIRification*, además de llevar a cabo una encuesta sobre el proceso de transformación de conjuntos de datos existentes en el área de la Enfermería. Está caracterizado por su naturaleza cualitativa para el análisis de datos, con la exploración bibliográfica y documental en BRAPCI, en sitios del área de la Salud y en el buscador Google Académico. Los resultados señalan la literatura sobre los principios *FAIR* y las iniciativas en marcha en el área de la Enfermería, así como una sugerencia para que se realicen las actividades de cambio de los conjuntos de datos en *FAIR* con la posibilidad de emplearlos en el área. Se concluye que, para volverlos *FAIR*, el proceso de transformación de datos de investigación es una iniciativa aún difícil de conducirse y presenta algunas lagunas en el área de la Enfermería.

Palabras clave: *FAIR*. *FAIRification*. Conjuntos de datos. Enfermería. Uso compartido de los datos.

1 Introdução

Nos dois últimos anos, vivemos sob a sombra de uma impensável pandemia, tema que só era presente no nosso cotidiano nas leituras sobre acontecimentos

históricos. O retorno à normalidade só foi possível depois da chegada de vacinas, que se desenvolveram a partir de pesquisas compartilhadas entre diferentes profissionais espalhados ao redor do globo. Pode-se depreender que o alinhamento entre os temas do

compartilhamento, da Ciência Aberta, do reúso de dados, da compreensão e da colaboração converge na obtenção de resultados mais ágeis para o desenvolvimento da Ciência.

Esta pesquisa busca esta convergência por meio da compreensão dos princípios *FAIR* e da possível contribuição que poderiam levar à área de Enfermagem, para auxiliar os pesquisadores desse campo na organização e compartilhamento dos dados de pesquisa. As dificuldades no processo de implantação do compartilhamento de dados permeiam várias áreas do conhecimento. O reúso de dados gerados por pesquisas ainda é assunto de difícil condução, pois, até o momento atual, as iniciativas de mudança no comportamento dos investigadores carecem de impulso na área de Enfermagem.

A provocação para esta investigação surgiu a partir de um trabalho de conclusão de curso idealizado por Braga (2021) na Pós-Graduação *lato sensu* Informação Científica e Tecnológica em Saúde, na Fundação Oswaldo Cruz (Fiocruz), com o título “Modelo de plano de gestão de dados de pesquisa em Enfermagem”. Como consequência, vislumbrou-se a oportunidade de disseminar e incentivar os pesquisadores a conhecerem e adotarem *FAIR*, além de instrumentalizar o seu uso para auxílio na pesquisa e gestão de dados.

Assim, este trabalho tem como objetivo geral fazer um levantamento sobre o processo de *FAIRificação* na área de Enfermagem, bem como identificar as iniciativas de compartilhamento de dados e de execução de processo de transformação.

2 Procedimentos metodológicos

Conforme os preceitos descritos por Gil (2002), este trabalho tem por base os princípios da pesquisa qualitativa para análise dos dados. Em relação à pesquisa realizada, esta foi feita por meio de análise bibliográfica e documental.

O processo de pesquisa seguiu os passos descritos nas alíneas abaixo:

- a) Pesquisa bibliográfica na Base de Dados em Ciência da Informação (BRAPCI), que tem foco em Biblioteconomia e Ciência da Informação. Neste caso, foram buscados materiais relacionados à temática *FAIR* com um intervalo de dez anos (2013-2023);
- b) Pesquisa bibliográfica no Banco de Dados em Enfermagem (BDENF), por meio do Portal da Biblioteca Virtual em Saúde (BVS, 2023);
- c) Pesquisa bibliográfica no Google Acadêmico para agregar consistência ao trabalho no que tange à temática de *FAIR* (*FAIRification*, *FAIRness* e Enfermagem);

d) Utilização do portal da Escola de Enfermagem Alfredo Pinto (EEAP), da Universidade Federal do Estado do Rio de Janeiro (UNIRIO) para verificação de literatura, no qual não foi localizado nenhum texto, mas foi possível identificar as pesquisadoras integrantes da Rede *GO FAIR* Brasil Saúde Enfermagem; e

e) Vista dos currículos Lattes de cada uma das componentes da Rede *GO FAIR*, com o intuito de verificar publicações realizadas e as atividades ligadas à temática *FAIR* de que as pesquisadoras tenham participado.

2.1 Trabalhos correlatos na BRAPCI

Os termos utilizados foram: “princípio fair”; “princípio” AND “fair”; “fair principle”; “fair” AND “principle”; “fairness”; e “plano” AND “gestão” AND “dado”. Esses termos foram adotados porque foram entendidos como os que mais se aproximavam da proposta pensada para a pesquisa.

Como estratégia de busca, foi utilizado o filtro de delimitação de data de dez anos (2013-2023), e o esquema de organização dos documentos foi do mais antigo para o mais novo. As buscas foram realizadas sem e com a expressão “AND” para testagem de resultados e, em alguns casos, a expressão gerou resultados diferentes. A última expressão de busca (“plano” AND “gestão” AND “dado”) foi empregada porque contém o potencial de encontrar documentos que abordem a temática *FAIR* ligada a planos de gestão de dados. Foi feita uma breve análise dos documentos por meio de seus títulos, resumos e texto integral, baseado nas orientações sobre o processo de leitura técnica (LANCASTER, 2004). Os resultados foram avaliados em relação à pertinência e se o texto estava disponível em formato completo ou não.

2.2 Trabalhos correlatos no Portal da BVS/BDENF e Google Acadêmico

Sabe-se que há outras possíveis bases de Enfermagem para a realização de buscas; mas, para simplificação do processo, pelo fato de ser gratuita e pela questão de ser totalmente nacional, a base escolhida foi a BDENF.

Os filtros aplicados ao pesquisar no Portal da BVS foram: base – BDENF; idioma – Português; e período – últimos cinco anos (2018-2023).

Foram utilizadas, ao todo, sete termos e expressões de busca: “fair”; “dados”; “compartilhamento” AND “dados”; “dados” AND “abertos”; “compartilhamento de dados”; “dados abertos”; e “fairification process” (tabela 3, em apêndice F).

A fim de estruturar o processo de pesquisa, optou-se pelos seguintes critérios: as palavras utilizadas na busca seriam colocadas entre aspas para limitar a busca apenas a palavras ou termos que estivessem presentes nas

estruturas dos textos localizados; e a intenção em escolher as expressões de busca listadas foi a de começar por um termo mais geral, propositadamente, para depois chegar a um certo nível de afinamento. Não foi intenção dos pesquisadores saturar expressões, isto é, elencar quantas expressões fossem possíveis, pois isso geraria quantidade significativa de dados que não caberiam ao escopo deste trabalho.

As buscas efetuadas no *Google Acadêmico* retornaram os mesmos registros obtidos nesta subseção e na anterior (2.2 e 2.1, respectivamente), corroborando a abordagem de pesquisa e seleção.

3 Os princípios FAIR

FAIR é um acrônimo para *Findable, Accessible, Interoperable e Reusable*, princípios ou “[...] elementos norteadores ao processo de gestão de dados de pesquisa” (HENNING *et al.*, 2018, p. 4). No âmbito nacional, a aplicação dos princípios *FAIR* ainda dá seus primeiros passos:

No Brasil, os debates acerca deste tema ainda estão em caráter inicial, restritos a alguns grupos de investigação específicos de universidades e institutos de pesquisa localizados na Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Universidade de São Paulo (USP), Universidade Estadual Paulista (UNESP), Universidade Federal da Paraíba (UFPB), Universidade Federal do Espírito Santo (UFES), Fundação Oswaldo Cruz (Fiocruz), Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA), Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), entre possíveis outros (HENNING *et al.*, 2018, p. 4).

Tomando aquilo que foi visto na introdução deste trabalho, vê-se que a problemática relacionada aos princípios *FAIR* e consequentemente ao compartilhamento, abertura e reuso de dados permeia outras áreas do conhecimento além da Enfermagem. Neste sentido, reflete-se que

Vivenciamos um cenário de mudança de paradigma no mundo científico, onde as práticas de produção, acesso e disseminação do conhecimento estão em franca revisão. Percebe-se que nunca existiram tantos pesquisadores produzindo ciência no mundo como nos últimos anos (HENNING *et al.*, 2019, p. 389).

Ainda, Henning e outros (2019), citando Markam (2017), relatam que há um número expressivo de pesquisadores, que gastam um grande volume de recursos, algo que gira em torno de “[...] 2,3 trilhões de dólares com pesquisas realizadas” (HENNING *et al.*, 2019, p. 340). Desta forma, com tantos recursos gastos, os quais, em certos casos, são públicos, o mais adequado para a divulgação e evolução do conhecimento seria que os dados obtidos por suas pesquisas fossem abertos, o que, como se sabe, é uma das premissas dos princípios *FAIR*. Então, conforme relato da autora supracitada, apenas cerca de 10% do volume de dados produzidos

por esses cientistas é aberto. Isso confirma o que ocorreu no breve mapeamento feito na área de Enfermagem da EEAN, pois certamente ora os pesquisadores guardam seus dados em meios pessoais, ora em lugares que não estão abertos ao público, ou mesmo têm receio de divulgarem seus dados. Corroborando essa informação, tem-se que:

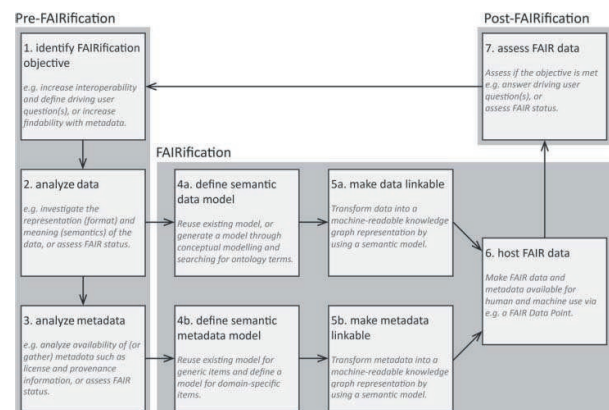
É mais frustrante ainda se pensarmos que a maior parte dessas pesquisas foi financiada com verbas públicas e que dificilmente teremos acesso ao seu resultado. Portanto, algumas ações precisavam ser conduzidas para mudar esse cenário, e é neste sentido que o acesso aberto às publicações científicas e, mais recentemente, a Ciência Aberta se consolidam como novas modalidades do fazer científico. A adoção de diferentes dinâmicas, voltadas para o trabalho colaborativo, que visam aos interesses sociais e coletivos, além do compartilhamento e reutilização dos dados e da informação, são premissas que fazem parte desta nova maneira de conduzir a ciência (HENNING *et al.*, 2019, p. 390).

Os princípios *FAIR* são a composição de 15 itens que orientam ações nas dimensões do compartilhamento: Encontráveis (*Findable*); Acessáveis (*Accessible*); Interoperáveis (*Interoperable*); e Reutilizáveis (*Reusable*) (quadro 1, em apêndice A). A sua adoção está relacionada ao processo de transformação de *datasets* segundo esses princípios (*FAIRificação*).

3.1 Processo de FAIRificação

Segundo Jacobsen *et al.* (2020), o processo de *FAIRificação* é, na verdade, um fluxo de trabalho para tornar os dados de uma pesquisa de acordo com as preconizações indicadas nos princípios *FAIR*. Assim, este processo é dividido em três fases: *pré-FAIRificação*, *FAIRificação* e *pós-FAIRificação*. Dentro destas fases, há sete passos que determinam como deve ser o processo. A figura 1 demonstra o fluxo de trabalho indicado para o processo:

Figura 1. Processo de FAIRificação



Fonte: Jacobsen *et al.* (2020, p. 58).

Alharbi *et al.* (2022) comentam que o processo de *FAIRificação* se caracteriza como um trabalho permeado por desafios, ou barreiras, as quais podem ser

financeiras, técnicas, legais e organizacionais. Destarte, esses autores refletem acerca dos desafios inerentes à proposta de *FAIRificação*:

Os desafios relacionados à questão financeira dizem respeito ao custo dos recursos necessários para a implementação do processo, posto que é importante para a manutenção de estruturas computacionais que possibilitem a gestão de dados. Tais custos também envolvem o emprego de pessoas para lidar com os dados e fazer a sua manutenção a longo prazo. [...] Os desafios técnicos estão associados à infraestrutura, ferramentas e métodos necessários para a realização do processo. Neste caso, são os serviços de identificadores persistentes, registros de metadados, serviços para o desenvolvimento e manutenção de ontologias etc. [...] Os desafios legais são aqueles que correspondem aos requerimentos, os quais dizem respeito ao processamento e compartilhamento dos dados, como por exemplo, direitos de acessibilidade e a observância de regulações concernentes à proteção de dados. Estes são aspectos diretamente ligados aos critérios preconizados – acessibilidade e reusabilidade dos dados – pelos princípios *FAIR* e, neste caso, o processo de *FAIRificação*. [...] Nos desafios organizacionais estão incluídos o oferecimento de treinamentos para aqueles que estarão diretamente ligados à implementação e à manutenção do processo de *FAIRificação*. Além disso, a adoção de *FAIR* também envolve o desenvolvimento e a premissa de manter uma cultura que promova recompensas para pesquisadores e pesquisadoras as quais praticam a essa proposta em seus dados de pesquisa. (ALHARBI *et al.*, 2022, p. 2081, tradução nossa)

Assim, transpor essas barreiras e colocar em prática o fluxo de *FAIRificação* são ações que promovem a convergência no percurso da Ciência Aberta, bem como ajudam a adotar o reúso de dados, como explorado anteriormente neste relato.

4 Resultados

Os resultados a seguir convalidam as percepções e motivações apresentadas na introdução deste trabalho, auxiliando no diagnóstico e planejamento de possíveis ações na adoção de *FAIR* para a área de Enfermagem.

4.1 Análise de publicações sobre o tema na BRAPCI

Considerando aquilo que informaram Henning *et al.* (2018), os princípios *FAIR* são discutidos na comunidade científica internacional desde o ano de 2014. Já em relação ao Brasil, as publicações de artigos de periódicos acerca da temática, conforme levantamento bibliográfico feito na BRAPCI, iniciam-se a partir do ano de 2018.

Com base nas expressões de busca apresentadas em 2.1, elaborou-se um quadro que reúne os trabalhos com abordagem voltada ao uso dos princípios *FAIR*, e que foram categorizados segundo a temática principal identificada no texto (quadro 3, em apêndice E).

A observação do quadro citado permite inferir que temáticas relacionadas aos princípios *FAIR*, de acordo com o que se tem disponível na BRAPCI, começaram a ser discutidas a partir do ano de 2018, isto é, ainda há reduzida literatura nacional que suporte a temática, em especial no contexto interdisciplinar da Ciência da Informação e nas Ciências da Saúde.

4.2 Levantamento de trabalhos no Portal da BVS/BDENF

O processo de levantamento partiu das expressões apresentadas em 2.2. A estratégia de afinamento nos levou a recortar a amostra de artigos com os seguintes critérios: a expressão “compartilhamento” AND “dados” resultou em uma melhora na questão da revocação, e a quarta expressão de busca, “dados” AND “abertos”, reduziu satisfatoriamente o espectro da busca para apenas 11 resultados. Já a quinta e penúltima expressão, “compartilhamento de dados”, retornou 22 trabalhos, e a última expressão, “dados abertos”, não trouxe nenhum resultado (tabela 3, em apêndice F).

Na tentativa de verificar se há trabalhos tratando de “*fairification process*” na área de Enfermagem, foi realizada uma busca com o termo citado e esta retornou zero resultado; portanto, deduz-se que, tendo em vista os resultados obtidos com os outros termos elencados nas tabelas e quadros, em apêndice, há uma produção limitada nesta área com relação a procedimentos não apenas ligados à “*FAIRification*” mas também ao uso de *FAIR* e compartilhamento de dados.

Trabalhou-se com os dados da quarta e da quinta expressão, 11 e 22 documentos, respectivamente. Assim, durante a análise foi constatado que os pesquisadores descrevem seus resultados ligados à temática em duas grandes vertentes:

- a) Tratamento dos dados coletados de forma quantitativa e/ou qualitativa (dependendo da proposta da pesquisa), mas sem estarem claramente alinhados aos princípios *FAIR*;
- b) Os pesquisadores expõem o tema e relatam que os dados foram armazenados em algum lugar específico de seu conhecimento, ao qual normalmente só eles têm acesso.

Essas duas alíneas podem ser tomadas como o procedimento frequente em pesquisas na área de Enfermagem, mesmo considerando-se a limitada amostragem com 33 trabalhos analisados. Como observado na alínea b, a análise dos trabalhos recuperados fez notar que inexistiu menção ao repositório no qual os dados de pesquisa estão armazenados e disponíveis. Esta afirmativa nos leva a inferir que, para tratar do alinhamento aos princípios *FAIR*, é mister a adoção de infraestrutura de repositórios para compartilhamento de dados, seguindo

recomendações no contexto de pesquisa brasileiro (IBICT, 2023).

4.3 A Rede GO FAIR Brasil Saúde Enfermagem

A Rede GO FAIR Brasil Saúde Enfermagem foi criada no país no ano de 2020. Cabe ressaltar que esse ano foi definido pela Organização Pan-Americana da Saúde (OPAS) e a Organização Mundial da Saúde (OMS) como o ano comemorativo dos profissionais de Enfermagem.

[...] a Rede GO FAIR Brasil Saúde Enfermagem está vinculada à Rede GO FAIR Brasil Saúde e tem como objetivo compartilhar dados de pesquisa em Enfermagem, apoiado nos princípios FAIR, visando o reúso de dados [e, por isso,] [...] a gestão de dados é considerada essencial para condução da investigação científica no século XXI, sendo os dados brutos o principal insumo informacional para o desenvolvimento de novas pesquisas. (UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO, 2023).

Nesse contexto, entende-se ser pertinente para o nosso objetivo de pesquisa, apresentar as participações em atividades e publicações de pesquisadoras participantes desta rede de implementação (tabela 1, em apêndice C).

Ao analisar-se o quadro com as pesquisadoras, depreende-se que há um quantitativo maior de pesquisadoras da área de Enfermagem (cinco) envolvidas na Rede GO FAIR do que da área de Ciência da Informação (duas) o que é importante, pois gera o entendimento de que há um interesse em fazer a Rede seguir uma proposta interdisciplinar. Isto também pode ser explicado em decorrência de os estudos relativos aos princípios FAIR estarem ligados diretamente às disciplinas de Biblioteconomia e CI no que tange aos aspectos de informação, dados e metadados. Considera-se debater que as pesquisadoras de Enfermagem estão na Rede para que se estabeleça um olhar para dentro de sua área, a fim de aportar os conhecimentos, premissas e propostas sobre FAIR para a área de Enfermagem, criando uma relação de sentido entre estas áreas no que diz respeito ao tratamento de dados de maneira geral e seu reúso em pesquisas.

4.4 A proposta para FAIRificação

Tomando-se por base o fluxo apresentado na figura 1, as questões práticas abordadas na disciplina desenvolvida no Programa de Pós-Graduação em Biblioteconomia da Unirio (PPGB) [2] e o atual estágio de adoção apresentado nas subseções 4.1, 4.2 e 4.3, apresentamos uma proposta com a relação de atividades a serem executadas em projetos de transformação de dados para adoção de FAIR em conjuntos de dados na área de Enfermagem.

4.4.1 Para a fase de pré-FAIRificação

- Avaliar o nível FAIR a ser atendido – ferramentas para verificação do nível de FAIRness, tais como *F-JUI Automated FAIR Data Assessment Tool*

(FAIRSFAR, 2023), *Self-Assessment Tool to Improve the FAIRness of Your Dataset – SATIFYD* (SATIFYD, 2023), *FairDataBR* (FAIRDATABR+, 2023) e o *template* elaborado pelo FAIR Metrics Group (HENNING et al., 2018).

- Caso não haja infraestrutura disponível, avaliar e pré-selecionar ambientes para hospedar os conjuntos de dados.
- Analisar os dados e metadados para fazer o recorte e seleção de conjuntos de dados.
- Se os dados e metadados estiverem disponíveis em Sistemas Gerenciadores de Banco de Dados, planejar e executar extrações para determinar as coleções de dados categorizados pelos principais temas de interesse para pesquisadores da área de Enfermagem.
- Avaliar a necessidade de transformações nos dados e metadados, incluindo mapeamento origem-destino.
- Verificar a necessidade de documentar a proveniência dos dados.
- Depurar as coleções e atribuir identificadores persistentes aos dados e metadados.

4.4.2 Para a fase de FAIRificação

- Identificar os vocabulários semânticos [3] que poderão auxiliar na descrição dos dados e metadados.
- Propor um esquema para representação semântica que apoie a definição dos dados e metadados, mas que permita que os dados e metadados sejam interligados.
- Submeter os conjuntos de dados ao processo de hospedagem no ambiente selecionado.

4.4.3 Para a fase de pós-FAIRificação

- Avaliar os conjuntos de dados e se o nível de FAIRness pretendido em 4.4.1 foi alcançado.
- Realizar a disseminação dos conjuntos com o apoio de mecanismos como o *FAIR Data Point* (SANTOS et al., 2023).

5 Considerações finais

Este estudo discutiu uma possível lacuna a ser investigada para a área de Enfermagem e que pode ser separada em dois aspectos: o primeiro é a aplicação de estudos acerca dos princípios FAIR à área e a respectiva infraestrutura para sua adoção, como tem ocorrido no contexto da Rede GO FAIR Brasil Saúde Enfermagem; o segundo seria o trabalho de conscientização e cultura

de incentivo para dados *FAIR* por meio de palestras, cartilhas, exposições, políticas públicas etc.

Uma breve verificação na BDEF deixou claro que os pesquisadores da área de Enfermagem, assim como de outros domínios do conhecimento, ainda estão afastados de um modelo mais atual para o compartilhamento do conhecimento produzido. No que tange ao processo de *FAIRificação*, este estudo tem a proposta de ser um indicador de que é possível aplicá-lo a recursos digitais produzidos por pesquisadores da área de Enfermagem, seja por meio de políticas públicas, seja por alguma espécie de incentivo organizacional, seja por um movimento de conscientização dentro da área. Sabe-se da importância que os recursos *FAIR* têm para a ciência, pois hoje o fazer científico vive uma necessidade ainda maior de compartilhamento de informações e trabalho colaborativo. Assim, cita-se como exemplo todo o esforço para o desenvolvimento das vacinas para COVID-19, de modo que foi por meio do labor colaborativo e do compartilhamento de dados, no sentido de dar continuidade a trabalhos anteriores, que a pesquisa avançou significativamente. O mesmo pode ser aplicado a outras áreas do conhecimento e, no caso aqui, à área de Enfermagem. O que se pensa é que os dados de pesquisa não deveriam ficar guardados para o esquecimento, mas compartilhados em locais adequados de armazenamento, como os repositórios de dados, para que outros pesquisadores tenham como acessá-los a fim de dar continuidade ao conhecimento antes produzido.

Mesmo com uma visão ainda inicial sobre a temática, foi possível aferir que existe uma quantidade limitada de artigos na área de Enfermagem acerca de *FAIR*.

O esforço vem sendo pavimentado pela Rede *GO FAIR* Brasil Saúde Enfermagem, mas espera-se que haja um maior engajamento de outras instituições para viabilizar a adoção dos princípios *FAIR*.

Notas

- [1] O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.
- [2] No ano de 2022, o PPGB estruturou uma disciplina para elucidar e debater o compartilhamento de dados usando os princípios *FAIR* por meio dos processos de transformação de dados e metadados com o apoio de modelos semânticos (processo de *FAIRificação*).
- [3] Vocabulários que permitem a descrição de propriedades com significado. Segundo Ribeiro, Santos e Moreira (2020), possibilitam a descrição, na Web, de metadados, registros bibliográficos, proveniência e estrutura de documentos. Dentro do escopo deste relato há vocabulários específicos sobre termos da área de Saúde. (BIOPORTAL, 2022).

Referências

ALHARBI, E.; GADIYA, Y.; HENDERSON, D.; ZALIANI, A.; DELFIN-ROSSARO, A.; CAMBON-THOMSEN, A.;

KOHLER, M.; WITT, G.; WELTER, D.; JUTY, N.; JAY, C.; ENKVIST, O.; GOBLE, C.; REILLY, D. S.; SATAGOPAM, V.; IOANNIDIS, V.; GU, W.; GRIBBON, P. Selection of datasets for *FAIRification* in drug discovery and development: which, why, and how? *Drug Discovery Today*, Oxford, v. 27, n. 8, p. 2080-2085, 2022. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S1359644622001994>. Acesso em: 10 fev. 2023.

BARROZO, V. L.; ALMEIDA, C. H. M. Curadoria e preservação digital: proposta de diretrizes para a gestão de dados científicos. *Múltiplos Olhares em Ciência da Informação*, [s. l.], n. Especial, 2021. Disponível em: <https://periodicos.ufmg.br/index.php/moci/article/view/37188>. Acesso em: 24 mar. 2023.

BIBLIOTECA VIRTUAL EM SAÚDE (BVS). *Pesquisa Portal Regional da BVS*. 2023. Disponível em: <https://pesquisa.bvsalud.org/portal/>. Acesso em: 9 fev. 2023.

BONETTI, L. G.; ARAKAKI, A. C. S. Princípios Fair e a avaliação de datasets no repositório institucional da UFScar. *Informação & Informação*, Londrina, v. 27, n. 1, p. 485-510, 2022. DOI: 10.5433/1981-8920.2022v27n1p485. Acesso em: 9 jan. 2023.

BRAGA, I. F. *Modelo de plano de gestão de dados de pesquisa em Enfermagem*. 2021. Trabalho de Conclusão de Curso (Especialização em Informação Científica e Tecnológica em Saúde) - Instituto de Comunicação e Informação Científica e Tecnológica em Saúde, Fundação Oswaldo Cruz, Rio de Janeiro, 2021. Disponível em: <https://www.arca.fiocruz.br/handle/icict/47482>. Acesso em: 7 jan. 2023.

FAIRDATABR+. FairDataBr: uma ferramenta para a avaliação de conjuntos de dados. [s. l.]: UFPB, 2023. Disponível em: <https://wrc.ufpb.br/fair>. Acesso em: 7 jan. 2023.

FONSECA, D. L. S.; MONTEIRO, E. W. Encontrabilidade da informação e os princípios Fair: meios alternativos de gestão de dados abertos. *Ciência da Informação*, [s. l.], v. 48, n. 3, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/136464>. Acesso em: 9 jan. 2023.

FOSTERING FAIR DATA PRACTICES IN EUROPE (FAIRSFAR). *F-UJI Automated FAIR Data Assessment Tool*. [s. l.: s. n.], 2022. Disponível em: <https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>. Acesso em: 15 jan. 2023.

GIL, A. C. *Como elaborar projetos de pesquisa*. 4. ed. São Paulo: Atlas, 2002.

GUANDALINI, C. A.; FURNIVAL, A. C. M.; ARAKAKI, A. C. S. Boas práticas científicas na elaboração de planos de gestão de dados. *Revista Digital de Biblioteconomia & Ciência da Informação*, [s. l.], v. 17, n. 2019, p. e019034, DOI: 10.20396/rdbci.v17i0.8655895. Acesso em: 9 jan. 2023.

HENNING, P. C.; RIBEIRO, C. J. S.; SANTOS, L. O. B. S.; DOS SANTOS, P. X. *GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da Ciência Aberta*. *Em Questão*, Porto Alegre, v. 25, n. 2, p. 389-412, 2019. Disponível em:

<https://seer.ufrgs.br/EmQuestao/article/view/84753>. Acesso em: 10 fev. 2023.

HENNING, P. C.; SILVA, L. O. B.; PIRES, L. F.; SINDEREN, M. V.; MOREIRA, J. L. R. Fairness de los planes de gestión de datos: una evaluación de algunos pgds europeos. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, [s. l.], v. 15, n. 3, 2021. DOI: 10.29397/reciis.v15i3.2270. Acesso em: 9 jan. 2023.

HENNING, P.; RIBEIRO, C. J. S.; SALES, L.; MOREIRA, J.; SANTOS, L. O. B. S. Desmistificando os princípios fair. **Tendências da Pesquisa Brasileira em Ciência da Informação**, [s. l.], v. 11, n. 1, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/151750>. Acesso em: 15 mar. 2023.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (IBICT). **Projeto PGD IBICT Dataverse**. Brasília: IBICT, 2023. Disponível em: <https://repositoriopesquisas.ibict.br>. Acesso em: 7 jan. 2023.

JACOBSEN, A.; KALIYAPERUMAL, R.; SANTOS, L. O. B. da S.; MONS, B.; SCHULTES, E.; ROOS, M.; THOMPSON, M. A Generic workflow for the data FAIRification process. **Data Intelligence**, Cambridge, v. 2, n. 1-2, p. 56-65, 2020. Disponível em: https://doi.org/10.1162/dint_a_00028. Acesso em: 10 fev. 2023.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília, DF: Briquet de Lemos/Livros, 2004.

MONTEIRO, E. C. S. A.; SANTANA, R. C. G. Repositórios de dados científicos na infraestrutura de pesquisa: adoção dos princípios Fair. **Ciência da Informação**, [s. l.], v. 48, n. 3, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/136407>. Acesso em: 9 jan. 2023.

NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY (BIOPORTAL). [s.l.: s. n.], 2022. Disponível em: <https://bioportal.bioontology.org/ontologies/>. Acesso em: 20 jan. 2023.

RIBEIRO, C. J. S.; SANTOS, L. O. B. S.; MOREIRA, J. L. R. Princípios FAIR e interoperabilidade: reprodutibilidade com o uso de modelagem conceitual orientada a ontologias. **Informação & Informação**, Londrina, v. 25, n. 3, p. 28-51, 2020. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/41874>. Acesso em: 24 mar. 2023.

RODRIGUES, M. M.; DIAS, G. A.; LOURENÇO, C. A. Repositórios científicos na América do Sul: uma análise da conformidade com os princípios Fair. **Em Questão**, Porto

Alegre, v. 28, n. 2, p. 113057, 2022.

<https://doi.org/10.19132/1808-5245282.113057>. Acesso em: 27 mar. 2023.

SANTOS, L. O. B. da S.; BURGER, K.; KALIYAPERUMAL, R.; WILKINSON, M. D. FAIR Data Point: A FAIR-oriented approach for metadata publication. **Data Intelligence**, [s. l.], v. 5, n.1, 2023; p. 163-183. DOI: https://doi.org/10.1162/dint_a_00160. Acesso em: 11 jan. 2023.

SATIFYD. **Self-Assessment Tool to Improve the FAIRness of Your Dataset**. [s. l.], DANS, 2023. Disponível em: <https://satifyd.dans.know.nl/>.

SILVA, L. C.; SEGUNDO, J. E. S.; SILVA, M. F. Princípios Fair e melhores práticas do *linked data* na publicação de dados de pesquisa. **Informação & Tecnologia**, [s. l.], v. 5, n. 2, p. 81-103, 2018. DOI: 10.22478/ufpb.2358-3908.2018v5n2.44812. Acesso em: 9 jan. 2023.

UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO. **GO FAIR Brasil Saúde Enfermagem – Escola de Enfermagem Alfredo Pinto**, [s. l.]: UNIRIO, 2023. Disponível em: <http://www.unirio.br/ccbs/ceap/go-fair-brasil-saude-enfermagem>. Acesso em: 6 fev. 2023.

VEIGA, V. S. O.; HENNING, P. C.; DIB, S.; PENEDO, E.; SILVA, L. O. B.; PIRES, L. F. Plano de gestão de dados fair: uma proposta para a Fiocruz - Fair data management plan: a proposal for Fiocruz. **Liinc em Revista**, [s. l.], v. 15, n. 2, 2019. DOI: 10.18617/liinc.v15i2.5030. Acesso em: 9 jan. 2023.

WILKINSON, M. D.; DUMONTIER, M.; AALBERSBERG, I. J.; APPLETON, M. A.; BAAK, A. BLOMBERG, N.; BOITEN, J-W.; SANTOS, L. B. S.; BOURNE, P. E.; BOUWMAN, J.; BROOKES, A. J.; CLARK, T.; CCROSAS, M.; DILLO, I.; DUMON, O.; EDMUNDS, S.; EVELO, C. T.; FINKERS, R.; GONZALEZ-BELTRAN, A.; GRAY, A. J. G.; GROTH, P.; GOBLE, C.; GRETHE, J. S.; HERINGA, J.; HOEN P. A. C 'T.; HOOFT, R.; KUHN, T.; KOK, R.; KOK, J.; LUSHER, S. J.; MARTONE, M. E.; MONS, A.; PACKER, A. L.; PERSSON, B.; ROCCA-SERRA, P.; ROOS, M.; SCHAIK, R. V.; SANSONE, S-A.; SCHULTES, E.; SENGSTAG, T.; SLATER, T.; STRAWN, G.; SWERTZ, M. A.; THOMPSON, M.; LEI, MULLIGEN, J.; JAN VELTEROP, E. V.; WAAGMEESTER, A.; WITTENBURG, P.; WOLSTENCROFT, K.; ZHAO, J.; MONS, B. The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, London, v. 3, n. 1, p. 9, 2016. Disponível em: <https://www.nature.com/articles/sdata201618>. Acesso em: 10 fev. 2023.

Apêndice A

Quadro 1. *As 15 premissas FAIR*

Encontráveis	F1	É atribuído aos dados e metadados um identificador persistente único global
	F2	Os dados são descritos por metadados ricos (em acordo com R1)
	F3	Os metadados incluem de forma clara e explícita o identificador dos dados que descrevem
	F4	Os dados e metadados são registrados ou indexados em um recurso pesquisável
Acessíveis	A1	Os dados e metadados são recuperáveis por seu identificador que utiliza um protocolo de comunicação padronizado
	A1.1	O protocolo é aberto, gratuito e universalmente implementável
	A1.2	O protocolo permite um procedimento de autenticação e autorização, quando necessário
	A2	Metadados são acessíveis, mesmo quando os dados não estão mais disponíveis
Interoperáveis	I1	Os dados e metadados usam uma linguagem formal, acessível, compartilhada e amplamente aplicável para representação do conhecimento
	I2	Os dados e metadados usam vocabulários que seguem os princípios FAIR
	I3	Os dados e metadados incluem referências qualificadas a outros dados e metadados
Reutilizáveis	R1	Os dados e metadados são ricamente descritos com uma pluralidade de atributos precisos e relevantes
	R1.1	Os dados e metadados são liberados com uma licença de uso de dados clara e acessível
	R1.2	Os dados e metadados são associados com proveniência detalhada
	R1.3	Os dados e metadados atendem aos padrões da comunidade relevantes para o domínio

Fonte: Elaborado pelos autores com base em Wilkinson *et al.* (2016).

Apêndice B

Quadro 2. *Pesquisadoras da Rede GO FAIR Brasil Saúde Enfermagem*

<i>PESQUISADORA</i>	<i>ÁREA</i>	<i>LATTES</i>
Eliza Cristina Macedo	Enfermagem	http://lattes.cnpq.br/2871683085576383
Patrícia Corrêa Henning	Relações Internacionais/Ciência da Informação	http://lattes.cnpq.br/0970010723997242
Maria Simone de Menezes Alencar	Engenharia Química/Ciência da Informação	http://lattes.cnpq.br/1933365741854449
Sônia Regina de Souza	Enfermagem	http://lattes.cnpq.br/8543534166970830
Danielle Galdino de Paula	Enfermagem	http://lattes.cnpq.br/5726695441745144
Tais Veronica Cardoso Vernaglia	Enfermagem	http://lattes.cnpq.br/0328468656344207
Inês Maria Meneses dos Santos	Enfermagem	http://lattes.cnpq.br/0622001848223954

Fonte: Elaborado pelos autores.

Apêndice C

Tabela 1. *Atividades ligadas aos princípios FAIR*

PESQUISADORA	PROJETO DE PESQUISA	ARTIGOS, LIVROS, TRABALHOS OU CAPÍTULOS DE LIVROS PUBLICADOS	OUTRAS ATIVIDADES
Eliza Cristina Macedo	1	-	2
Patrícia Corrêa Henning	2	13	22
Maria Simone de Menezes Alencar	2	3	11
Sônia Regina de Souza	1	1	2
Danielle Galdino de Paula	1	1	2
Tais Veronica Cardoso Vernaglia	-	-	1
Inês Maria Meneses dos Santos	0	-	-

Fonte: Elaborado pelos autores.

Apêndice D

Tabela 2. *Resultados dos termos pesquisados na BRAPCI*

	PRINCIPIO FAIR	PRINCIPIO AND FAIR	FAIR PRINCIPLE	FAIR AND PRINCIPLE	FAIRNESS	PLANO AND GESTÃO AND DADO
Possíveis	7	2	1	1	4	2
Excluídos	4	4	3	3	0	8

Fonte: Elaborado pelos autores.

Apêndice E

Quadro 3. *Lista de artigos encontrados na BRAPCI com suas respectivas categorias*

ARTIGO	TERMO DE BUSCA
Henning, P. C.; Ribeiro, C. J. S.; Sales, L.; Moreira, J.; Santos, L. O. B. S. (2018)	Planejamento FAIR
Henning, P. C.; Ribeiro, C. J. S.; Santos, L. O. B. S.; Santos, P. X. D. (2019)	Introdução ao assunto
Silva, L. C.; Segundo, J. E. S.; Silva, M. F. (2018)	Elucidação dos princípios
Monteiro, E. C. S. A.; Santana, R. C. G. (2019)	Elucidação dos princípios
Fonseca, D. L. S.; Monteiro, E. W. (2019)	Planejamento FAIR
Barrozo, V. L. D.; Marcondes, C. H. (2021)	Planejamento FAIR
Bonetti, L. G.; Arakaki, A. C. S. (2022)	Elucidação dos princípios
Rodrigues, M. M.; Dias, G. A.; Lourenço, C. A. (2022)	Elucidação dos princípios
Henning, P. C.; Silva, L. O. B.; Pires, L. F.; Sinderen, M. V.; Moreira, J. L. R. (2021)	Planejamento FAIR
Veiga, V. S. O.; Henning, P. C.; Dib, S.; Penedo, E.; Silva, L. O. B.; Pires, L. F. (2019)	Planejamento FAIR
Guandalini, C. A.; Furnival, A. C. M.; Arakaki, A. C. S. (2019)	Planejamento FAIR

Fonte: Elaborado pelos autores.

Apêndice F

Tabela 3. *Termos e resultados obtidos na busca no Portal BVS/BDENF*

TERMOS	RESULTADOS
“fair”	63
“dados”	7604
“compartilhamento” AND “dados”	207
“dados” AND “abertos”	11
“compartilhamento de dados”	22
“dados abertos”	0
“fairification process”	0

Fonte: Elaborado pelos autores.

O *Data Quality Vocabulary* (DQV): um Sistema de Organização do Conhecimento?

Data Quality Vocabulary (DQV): a Knowledge Organization System?

Ananda Fernanda de Jesus (1), José Eduardo Santarem Segundo (2)

(1) Universidade Estadual Paulista (UNESP), Av. Hygino Muzzi Filho, 737 - Mirante - Marília/SP, af.jesus@unesp.br.

(2) Universidade Estadual Paulista (UNESP) / Universidade de São Paulo (USP), Av. Bandeirantes, 3900 - Vila Monte Alegre, Ribeirão Preto, santarem@usp.br.

Resumo

O presente contexto é marcado pela presença massiva de dados, especialmente na Web, atuando como insumo em diversos setores, o que torna relevante a descrição formal e a avaliação da qualidade dos dados visando facilitar o emprego de agentes computacionais em seu processo de recuperação e seleção. Os Sistemas de Organização do Conhecimento (SOCs) atuaram tradicionalmente nos processos de representação e recuperação de recurso informacionais, tendo sido ampliados para atuarem também na descrição formal de conjuntos de dados. O *Data Quality Vocabulary* (DQV) é um vocabulário semântico criado para a descrição formal da qualidade de dados. **Objetivo:** discutir o DQV enquanto um Sistema de Organização do Conhecimento, buscando-se compreender como esse vocabulário é definido pela literatura e como tem sido aplicado. **Metodologia:** análise exploratória, levantamento documental e bibliográfico, pautado no preenchimento de protocolo de busca. **Resultados:** Ao definir o DQV, destacam-se os seus propósitos descritivos, na explicitação de limitações de qualidade a ser realizada pelos publicadores e na agregação de resultados de avaliação de qualidade, que podem ser realizados por usuários e instituições de certificação. Destaca-se também o seu potencial uso na seleção de fontes. Ele tem sido aplicado principalmente como parte integrante de ferramentas de avaliação de qualidade, permitindo o compartilhamento formal de resultados do processo de avaliação. Também atua na criação de novos SOCs, como vocabulários semânticos e modelos de qualidade de dados. **Conclusão:** Entende-se que o DQV é um SOC, especificamente um vocabulário semântico do tipo padrão de estrutura de dados, que viabiliza descrição dos conceitos e relações derivados do processo de avaliação de qualidade. Conclui-se ainda que seu uso pode facilitar os processos de recuperação, seleção e reuso dos dados disponíveis na Web.

Palavras-chave: Sistemas de Organização do Conhecimento. Vocabulários Semânticos. Qualidade de Dados. Data Quality Vocabulary.

Abstract

The present context is marked by the massive presence of data, especially on the Web, acting as an input in several sectors, which makes the formal description and evaluation of data quality relevant, aiming to facilitate the use of computational agents in their recovery process and selection. Knowledge Organization Systems (KOSs) have traditionally acted in the processes of representation and retrieval of informational resources, having been expanded to also act in the formal description of data sets. The Data Quality Vocabulary (DQV) is a semantic vocabulary created for the formal description of data quality. Objective: to discuss the DQV as a Knowledge Organization System, seeking to understand how this vocabulary is defined by the literature and how it has been applied. Methodology: exploratory analysis, documentary, and bibliographic survey, based on filling out a search protocol. Results: When defining the DQV, its descriptive purposes stand out, in explaining quality limitations, to be carried out by publishers, and in the aggregation of quality assessment results, which can be carried out by users and institutions of certification. Also noteworthy is its potential use in font selection. It has mainly been applied as an integral part of quality assessment tools, allowing the formal sharing of results from the assessment process. It also acts in the creation of new KOs, such as semantic vocabularies and data quality models. Conclusion: It is understood that the DQV is a KOS, specifically a semantic vocabulary of the data structure pattern, which enables the description of concepts and relationships derived from the quality assessment process. It is also concluded that its use can facilitate the processes of recovery, selection, and reuse of data available on the Web.

Keywords: Knowledge Organization Systems. Semantic Vocabularies. Data Quality. Data Quality Vocabulary.

Resumen

El contexto actual está marcado por la presencia masiva de datos, especialmente en la Web, actuando como insumo en varios sectores, lo que hace relevante la descripción y evaluación formal de la calidad de los datos, con el objetivo de facilitar el uso de agentes computacionales en su proceso de recuperación y selección. Los Sistemas de Organización del Conocimiento (SOCs) han actuado tradicionalmente en los procesos de representación y recuperación de recursos de información, habiéndose ampliado para actuar también en la descripción formal de conjuntos de datos. El Data Quality Vocabulary (DQV) es un vocabulario semántico creado para la descripción formal de la calidad de los datos. Objetivo: discutir el DQV como Sistema de Organización del

Conocimiento, buscando comprender cómo este vocabulario es definido por la literatura y cómo ha sido aplicado. Metodología: análisis exploratorio, levantamiento documental y bibliográfico, a partir del llenado de un protocolo de búsqueda. Resultados: Al definir el DQV, se destacan sus propósitos descriptivos, en la explicación de las limitaciones de calidad, a realizar por las editoriales, y en la agregación de los resultados de las evaluaciones de calidad, que pueden realizar los usuarios y las instituciones de certificación. También cabe destacar su uso potencial en la selección de fuentes. Se ha aplicado principalmente como parte integral de las herramientas de evaluación de la calidad, lo que permite compartir formalmente los resultados del proceso de evaluación. También actúa en la creación de nuevos SOC, como vocabularios semánticos y modelos de calidad de datos. Conclusión: Se entiende que el DQV es un SOC, específicamente un vocabulario semántico del patrón de estructura de datos, que posibilita la descripción de conceptos y relaciones derivadas del proceso de evaluación de la calidad. También se concluye que su uso puede facilitar los procesos de recuperación, selección y reutilización de los datos disponibles en la Web.

Palabras clave: Sistemas de Organización del Conocimiento. Vocabularios semánticos. Calidad de datos. Vocabulario de calidad de datos.

1 Introdução

O presente contexto tecnológico é geralmente associado a termos como *Big Data*, *Linked Data* e *Datificação*. Vive-se um processo acelerado de geração e uso de dados como insumos nos mais distintos contextos. Como apontam Martins e Valente (2019, p. 89),

Nos últimos anos, o uso e coleta de dados foram intensificados, em um processo que veio a ser denominado de “datificação”. Esse fenômeno está ancorado nas Tecnologias da Informação e Comunicação (TIC) e tem como palco central a internet [...]

São marcos do processo de datificação: o estabelecimento dos computadores portáteis, da internet, a popularização da *Web*, dos dispositivos móveis e o barateamento dos sensores, que tornaram cotidiana a geração constante e mesmo inconsciente de grandes conjuntos de dados.

Nesse contexto, a qualidade de dados tornou-se uma preocupação concreta, não apenas para os profissionais cuja atividades são tradicionalmente relacionadas com a criação e manutenção de bases de dados, como para todos os que desempenham atividades nas quais sejam empregados dados como insumos.

Embora a *Web* tenha facilitado a disponibilização desses dados, os usuários enfrentam dificuldades com os processos de busca, seleção e reuso. Os diferentes níveis de qualidade e curadoria, fazem com que, em muitos casos, os dados não sejam suficientemente descritos e representados, especialmente no que diz respeito a representação formal, ou seja, compreensível a agentes computacionais.

Sem a representação formal, o uso desses agentes na identificação e seleção dos conjuntos de dados é prejudicada. Com o volume de dados existentes e a velocidade da geração de novos dados, enfrenta-se um cenário onde a busca e seleção eficiente desses recursos apenas por usuários humanos, sem o uso de agentes computacionais, tornou-se improdutiva e até mesmo inviável (ISOTANI; BITTENCOURT, 2015).

O *Word Wide Web Consortium* (W3C), consórcio responsável pelo desenvolvimento da *Web*, estabeleceu

uma série de recomendações e instrumentos para lidar com as questões apresentadas.

Nesse processo, destaca-se o *W3C Data on the Web Best Practices (DWBP) working group*, o grupo de trabalho foi mantido entre 2013 e 2017 com o intuito de elaborar as Melhores Práticas (MP) para a publicação na web, que resultou em um guia com 35 MPs que abordam diferentes aspectos da disponibilização de dados na Web. (W3C, 2017).

Parte das MPs são relacionadas com a representação e com o fornecimento de (meta)dados a respeito dos conjuntos de dados e recursos informacionais a serem disponibilizados. Nesse aspecto, destaca-se a importância da criação e aplicação de vocabulários semânticos, dentre os quais encontra-se o *Data Quality Vocabulary* (DQV). O DQV foi elaborado para prover termos e formalizar as relações existentes entre os conceitos utilizados para a descrição formal de qualidade. (W3C, 2016)

Destacada a importância da qualidade de dados em um cenário de intensa datificação de processos e produtos, apresentadas as dificuldades de recuperação de dados na *Web* e as potenciais soluções para essa questão, a presente pesquisa propõe-se a: discutir o DQV enquanto um sistema de organização do conhecimento, buscando compreender como esse vocabulário é definido pela literatura e como tem sido aplicado. Para isso, foi realizado uma análise exploratória, um levantamento documental e bibliográfico, cujos procedimentos são apresentados na próxima seção.

2 Procedimentos metodológicos

O estudo consiste em uma pesquisa de caráter exploratório e descritivo, pautado na realização de uma análise exploratória da literatura científica, de caráter não exaustivo, e na realização de um estudo da documentação oficial do W3C sobre o vocabulário estudado.

Para o levantamento foi elaborado um protocolo de busca apresentado na tabela 1.

Tabela 1. *Protocolo de busca*

<i>Objetivos</i>	Identificar artigos, teses e dissertações cujo foco seja discutir o DQV ou suas aplicações
<i>Questão Principal</i>	Como o Data Quality Vocabulary é definido pela comunidade científica? Quais suas principais aplicações?
<i>Palavras-chave</i>	“Data Quality Vocabulary”
<i>Linguagem dos estudos</i>	Inglês, português e espanhol.
<i>Métodos de seleção</i>	Leitura do título, palavras-chave e resumo dos documentos; Leitura da introdução e conclusão; Leitura completa para Aplicação dos critérios de inclusão e de exclusão.
<i>Bases de dados</i>	BRAPCI, BDTD, SCIELO, SCOPUS, Portal de periódicos CAPES, Google scholar
<i>Critérios de inclusão (I) e exclusão (e)</i>	(I) Aplica o DQV; (I) O DQV é o foco das discussões do artigo (I) Discute o DQV (e) Não está nos idiomas estabelecidos para a pesquisa; (e) Não está no formato estabelecido para a pesquisa; (e) Não aplica ou não discute a aplicação vocabulário; (e) Apenas menciona o DQV (e) Não é possível ter acesso ao documento; (e) Não está relacionado com a Ciência da Informação.
<i>Tipos documentais</i>	Artigos em periódicos, trabalhos apresentados em eventos, teses e dissertações.
<i>Estudos iniciais</i>	Não se aplica ao estudo.
<i>Campos de extração</i>	Definição do DQV; Tipo de aplicações do DQV; Como foi feito o uso do DQV na aplicação; Contexto de criação de criação da aplicação.
<i>Sumarização dos resultados.</i>	Os dados foram agrupados em quadros-resumo de acordo com os campos de extração estabelecidos, os resultados foram abordados em categorias estabelecidas a <i>posteriori</i> .

Fonte: Autores (2023)

3 A qualidade na recuperação, seleção e reuso de dados

A qualidade de dados se caracteriza como um conceito multifacetado, cuja discussão e compreensão pode ocorrer através de diversas perspectivas.

Ela é mais comumente definida pela literatura com uma perspectiva contextual, com base na ideia de “*fitness for use*”, onde, para serem considerados de qualidade, os dados precisam ser adequados a uma aplicação. Nessa perspectiva, os mesmos dados podem ser adequados para uma tarefa e não atuarem satisfatoriamente em outro contexto (JURAN, 1988; WANG; STRONG, 1995; ZAVERI et al., 2012).

A atividade de avaliar a qualidade de um conjunto de dados para um uso pretendido é denominada avaliação de qualidade. Para realizar essa atividade, faz-se necessário identificar o que será avaliado e como os resultados serão mensurados e apresentados.

Os principais conceitos envolvidos na avaliação de qualidade são (JURAN, 1988; WANG; STRONG 1996; BIZER, 2013; NELSON, TODD; WIXOM, 2005; NOOGHABI; DASTGERDI; MELO, 2017; TALLERÁS, 2017):

- *Categoria* – Categorias que permitem o agrupamento das dimensões por semelhança em relação a sua perspectiva de avaliação. As principais perspectivas de avaliação de qualidade são: *intrínseca* (onde se observam apenas características próprias dos dados), *contextual* (onde se verifica se um conjunto de dados é adequado à determinada aplicação), *representacional* (onde busca-se avaliar se os dados apresentam as informações necessárias para sua interpretação, compreensão e aplicação) e *acessibilidade* (onde avalia-se à capacidade de acesso aos conjuntos de dados);
- *Dimensões* – são as características abrangentes a serem observadas no processo de avaliação. Ou ainda, “o que será avaliado” nesse processo;
- *Crítérios* - os atributos que descrevem os itens observados para avaliar um conjunto de dados. Podem ser compreendidos como um desmembramento de determinada dimensão, que ocorre quando essa se propõe a avaliar uma característica complexa; e
- *Métricas* - são as formas como serão mensurados e apresentados resultados, atuando como indicadores criados para qualificar e quantificar a qualidade. Algumas métricas possuem um caráter objetivo, enquanto outras necessitam de uma análise subjetiva, especialmente os que dependem da percepção do usuário. Cada dimensão e critério de qualidade podem ser associado a uma ou mais métricas.

Pode-se ainda selecionar/criar artefatos para auxiliar o processo de avaliação de qualidade. Jesus e Santarem Segundo (2022, p. 13) apontam que “Esses artefatos são plurais em muitos aspectos [...]” podendo divergir principalmente quanto ao *tipo de artefato* (frameworks; softwares; modelos, métodos, vocabulários e etc), *atividade a ser realizada* (avaliação, melhoria, avaliação e melhoria) e *forma como desempenha a atividade* (automático, semiautomático e manuais).

Para os autores, os automáticos são aqueles que desempenham o processo de avaliação sem interferência humana. Os semiautomáticos realizam parte do processo de maneira automática, mas em determinadas etapas necessitam de interferência humana. Já os manuais são guias elaborados com objetivo de orientar/facilitar a avaliação, como modelos, metodologias, ciclos de vida, etc.

Os vocabulários têm impacto direto no uso de artefatos automáticos e semiautomáticos, já que a descrição formal e explícita das características dos dados facilita o processo de uso desses agentes computacionais. Além disso, os vocabulários também podem ser empregados na comunicação dos resultados dos processos de avaliação realizados com base nesses artefatos.

Para que sejam empregados de maneira otimizada nesse processo de avaliação de qualidade, especialmente no contexto da Web, os vocabulários precisam ser elaborados visando atender recomendações do W3C, em especial as MPs para publicação de dados na Web. A próxima seção irá apresentar um breve panorama da evolução dos vocabulários, apresentando definições, seu estabelecimento enquanto SOCs, e sua aproximação com o contexto da *Web*.

4 Vocabulários enquanto Sistemas de Organização do Conhecimento e sua relação com a recuperação/reuso de dados na Web

Knowledge Organization Systems (KOS), ou em português Sistemas de Organização do Conhecimento (SOC), consiste em um termo abrangente, elaborado para agrupar estruturas terminológicas. Biscalchin (2019, p. 4-5), aponta que:

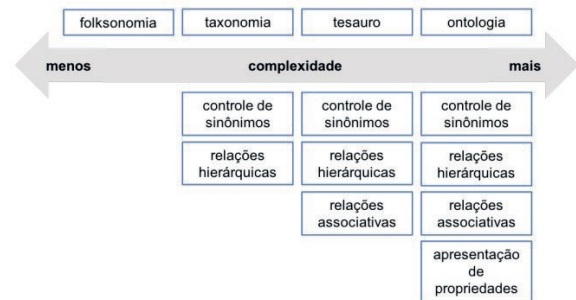
Como elemento em comum podemos afirmar que os SOC consistem em conceitos representados por termos. As variações entre as diferentes tipologias consistem na complexidade para a definição terminológica conceitual, partindo da ausência de controle terminológico (folksonomia) ao controle mais detalhado (ontologia).

Nessa definição destaca-se o objetivo dos SOC de organizar e representar recursos informacionais. Os SOCs tradicionais emergiram e se estabeleceram em um contexto de explosão da informação, primeiro para

lidar com recursos basicamente bibliográficos, e depois para contextos científicos, como periódicos e publicações seriadas. (MARCONDES, 2021).

Os SOCs são plurais em relação a sua estrutura, nomenclatura e complexidade. Moreira (2018) sistematizou os principais tipos de SOC de acordo com sua complexidade, esse modelo é apresentado na Figura 1.

Figura 1. SOCs categorizados com base em sua complexidade



Fonte: Moreira, 2018, p. 102.

Outro termo que perpassa a conceituação de SOCs é o controle de vocabulários, como aponta Fujita e Tolare (2019, p. 98):

Sistemas de organização do conhecimento (SOCs) são importantes porque realizam o controle de vocabulário em sistemas de informação que necessitam prover a recuperação da informação armazenada [...].

Barité (2015) indica que o controle de vocabulário se estabeleceu na Ciência da Informação por volta da década de 1940, sendo uma resposta a dificuldade de recuperação da informação por meio do uso de linguagem natural. O controle de vocabulário pode ser definido

[...] como um conjunto de técnicas e procedimentos aplicados a linguagem para resolver problemas de compreensão, ambiguidade, escopo e relacionamento entre os termos que expressam conceitos, e/ou entre que expressam nomes de pessoas, lugares, produtos ou instituições (BARITÉ, 2015, p. 97).

No âmbito de controle de vocabulário, entende-se, portanto, que a linguagem natural, tal como falamos e escrevemos, não é suficiente para promover a recuperação de recursos informacionais, sendo sujeita a ambiguidades, sinônimos, erros de grafia, etc. Ainda nesse contexto, tornam-se necessários esforços para a elaboração, padronização e adoção de instrumentos criados com o objetivo de estabelecer, de maneira inequívoca e livre de ambiguidades, os conceitos que representam determinado domínio, bem como suas relações com outros conceitos.

Tais instrumentos de padronização são produtos do processo de controle de vocabulário, diretamente relacionado à área de Organização do Conhecimento.

São aplicados na descrição dos recursos informacionais, de forma estruturada e padronizada, e também na orientação dos usuários no processo de busca.

Existe uma pluralidade conceitual sobre como nomear esses instrumentos, sendo utilizados termos como linguagens documentárias “[...] listas controladas, índices controlados, linguagens controladas e finalmente vocabulários controlados.” (BARITÉ, 2015, p. 101).

Nessa pesquisa, optaremos pelo termo “vocabulários controlados” para nos referir especificamente ao conceito de instrumentos criados no âmbito da Organização do Conhecimento com a finalidade de serem aplicados no processo de controle de vocabulários de recursos informacionais. Embora existam discussões a respeito dessa nomenclatura e, em diferentes abordagens, esses conceitos possam ser definidos ou não como sinônimos, optou-se por abordar SOC's como o termo mais abrangente, que engloba diversos tipos de vocabulários controlados, mas também outros sistemas criados posteriormente, inclusive como uma resposta grandes mudanças tecnológicas.

[...] pela evolução das tecnologias de informação e no contexto da internet, surgiram sistemas de organização do conhecimento mais adequados às necessidades de organização de documentos nativos digitais que desempenham um importante papel no gerenciamento e aplicações de informações digitais em geral. (FUJITA; TOLARE, 2019, p. 98).

Com a popularização da *Web* e o estabelecimento dos buscadores, a forma de busca por recursos informacionais foi alterada. Embora a recuperação da informação com uso de linguagem natural tenha se popularizado, essa se mostrou insuficiente em muitos aspectos, especialmente por problemas relacionados a representação dos recursos disponibilizados, criando silos de informação isolada.

O W3C vem, desde o início dos anos 2000, propondo tecnologias e recomendações para lidar com as questões relacionadas à recuperação da informação na *Web*. As propostas não são relacionadas apenas com a recuperação de recursos informacionais, mas buscam promover à recuperação de informações, respostas às perguntas dos usuários e ainda a geração de novas informações por meio do reuso dos dados.

Essa valorização dos dados como insumos ocorre em consonância com a intensificação dos processos de datificação, onde o funcionamento de muitos setores da sociedade passou a ser permeado pela utilização estratégica de dados, tanto no processo de tomada de decisão como em atividades operacionais (BATINI; SCANNAPIECO, 2016).

Uma das recomendações do W3C para a representação de recursos informacionais e para a disponibilização de dados na *Web* é o *Linked Data*, que consiste na publicação de dados estruturados e conectados na *Web*, respeitando os seguintes princípios: 1) Use *Uniform Resource Identifier* (URIs) como nomes para as coisas; 2) Use *Hypertext Transfer Protocol* (HTTP) URIs, para que as pessoas possam procurar esses nomes; 3) Utilize o padrão *Resource Description Framework* para fornecer informações sobre os recursos; e 4) Inclua links para outros URIs, para que eles possam descobrir mais coisas. (BERNERS-LEE, 2006).

O RDF é um modelo pautado e uma estrutura no formato de triplas, onde cada nova declaração deve seguir o modelo Recurso + Propriedade + Valor, orientando-se pela “[...] ideia de que as coisas que estão sendo descritas têm propriedades que possuem valores e que esses recursos podem ser descritos ao se fazer declarações.” (W3C, 2004, não paginado, tradução nossa, grifo nosso).

Quase tudo pode ser descrito seguindo essa estrutura, desde recursos informacionais disponibilizados na *Web*, como PDFs e informações constantes em sites, até pessoas do mundo real, sendo representadas por identificadores únicos, URIs, que irão atuar como uma representação da coisa a ser descrita, mesmo que essa coisa não esteja fisicamente na *Web*. “Hoje trata-se não só de recuperar documentos (ou suas representações) mas representações digitais de qualquer coisa [...]” (MARCONDES, 2021, p. 251). Esse tipo de vocabulário, criado com base na estrutura do *RDF*, também pode ser chamado de vocabulário semântico ou ontologia, a depender da sua complexidade. Os vocabulários semânticos são objetos constantes nas discussões do W3C, existindo orientações próprias para a sua elaboração.

Se inicialmente o foco dos SOC's eram os vocabulários controlados, criados para descrever os conteúdos dos recursos informacionais, com o estabelecimento da *Web*, e em especial dos esforços do W3C, passou-se a buscar também a padronização e formalização dos dados que representam esses recursos (metadados), bem como a criação de representações para pessoas e coisas existentes fora do ambiente *Web*.

Os vocabulários controlados continuam a ser pertinentes em muitos contextos, especialmente em domínios especializados como a área da saúde e do patrimônio cultural, existindo uma movimentação para a sua publicação como vocabulários semânticos.

Para amparar esses processos, são empregados outros vocabulários, de caráter estrutural, que apresentam as propriedades necessárias para que se converta um SOC tradicional em um vocabulário semântico.

No contexto *Web*, esses vocabulários cumprem a função de padronizar elementos de metadados, que

podem ser aplicados na descrição de determinado cenário e formalizar as relações existentes entre esses elementos. Esses vocabulários também podem ser denominados como padrões de estrutura de dados. Na ciência da informação, também são denominados como padrões de metadados. (TOMOYOSE, 2022).

Um dos aspectos a ser representado em conjuntos de dados disponibilizados na Web é a qualidade dos dados e dos metadados (W3C, 2016). Jesus e Santarem Segundo (2022) apontam que esses dados estão sujeitos a diferentes questões de qualidade, que podem ser divididas entre questões relacionadas às fontes de dados e a problemas estruturais, relacionados às declarações em RDF. Para poder fazer referência à qualidade dos dados passaram a ser elaborados vocabulários semânticos de qualidade de dados.

O DQV é a recomendação oficial do W3C para o fornecimento de informações de qualidade, sendo inclusive mencionado no documento das melhores práticas para a publicação de dados na Web. (W3C, 2016). A próxima seção apresenta e discute esse vocabulário enquanto um SOC.

5 O DQV como um SOC para dados de avaliação de qualidade

A presente seção discute os resultados do levantamento bibliográfico. Aplicando a estratégia de busca, foram recuperados 211 artigos, dos quais 33 eram duplicados, 118 foram recusados e 60 foram aceitos para compor o corpus da pesquisa. Foram aceitos artigos, trabalhos apresentados em eventos, teses e dissertações. Os documentos aceitos foram divididos em 3 critérios de inclusão: (1) Tem como foco o DQV: 1 documento (2%); (2) Discute o DQV em meio a outros vocabulários e contextos: 8 documentos (13%); (3) Propõe uma aplicação do DQV 52 documentos (85%).

A principal forma de abordagem do DQV encontrada na literatura é a aplicação. Entre esses documentos, o DQV tem maior ou menor impacto na estrutura do artefato projetado, recebendo, conseqüentemente, menor ou maior espaço nas discussões.

Entre os artigos aceitos, apenas Albertoni e Isaac (2020) tiveram como objetivo central discutir e apresentar o vocabulário, sua estrutura e limitações. Nesse artigo, é possível identificar as justificativas de criação do DQV, compreender melhor sua estrutura e potenciais aplicações.

Apresentados os resultados gerais do levantamento, a próxima subseção apresenta o DQV por meio das definições da literatura.

5.1 O Data Quality Vocabulary como um Sistema de Organização do Conhecimento: definições, características e estrutura

Essa subseção foi elaborada com base na literatura abordada na seção 2, onde se apresentam e discutem os Sistemas de Organização do Conhecimento, bem como nos resultados do levantamento bibliográfico e da análise da documentação oficial do DQV.

Para compreender melhor o vocabulário e posicioná-lo enquanto um SOC, realizou-se uma análise das suas definições apresentadas na literatura e na documentação do vocabulário. O Apêndice A apresenta a relação entre os autores e definições.

Ao observar as definições do DQV, destaca-se o termo estrutura, que aparece acompanhado de algumas ações como descrever, definir e representar.

Nesse sentido, o DQV é definido como uma estrutura geral, abrangente e interoperável. Para compreender as razões que levam o DQV a ser definido como uma estrutura geral, abrangente e em alguns casos básica, é necessário antes resgatar os conceitos relacionados ao processo de avaliação de qualidade de dados.

Como discutido anteriormente, o processo de avaliação de qualidade ocorre com base em categorias, dimensões, critérios e métricas. Os critérios e as dimensões que serão adotados, bem como o peso das métricas, irá depender das necessidades do domínio aos quais os dados pertencem e a atividade no qual eles serão empregados.

O vocabulário não comporta propriedades com dimensões e métricas específicas, que permitam a avaliação de qualidade, mas sim uma formalização dos próprios agentes envolvidos no processo de avaliação de qualidade, ou seja, entre outros, os conceitos de “classe”, “dimensão” e de “métrica”.

Enquanto um Sistema de Organização do Conhecimento, que permite refletir sintetizar os principais aspectos de um domínio, pode-se dizer então que o DQV não é um reflexo da qualidade de dados de uma maneira geral, mas sim um reflexo ou uma representação formal dos processos de avaliação de qualidade.

O DQV não busca caracterizar ou representar os conteúdos informacionais dos conjuntos de dados, mas sim permitir aos “editores de conjuntos de dados ou uma comunidade mais ampla de usuários expressar metadados de qualidade”. (XIA; XU; MAO, 2017, p. 18).

Como o DQV não fornece dimensões e métricas para a avaliação, torna-se necessário adotar ou elaborar um modelo de qualidade, uma metodologia de avaliação de qualidade, na qual serão estabelecidas as dimensões e

métricas a serem observadas e os pesos que cada métrica irá receber nessa avaliação.

Radulovic *et al.* (2017) acrescenta que o DQV pode ser empregado no processo de elaboração de um novo modelo de qualidade ou de descrição formal de um modelo já existente em uma estrutura compatível com as MPs do W3C.

Destacam-se ainda nas definições do DQV os termos ontologia, vocabulário descritivo e metadados. Esses termos fornecem pistas a caracterização do DQV enquanto um SOC. O vocabulário fornece uma estrutura ou um padrão para o registro de metadados de qualidade. Essa estrutura é baseada nos principais conceitos da avaliação de qualidade e descrita de forma a permitir sua utilização em declarações em RDF.

Um aspecto que diferencia o DQV de um vocabulário controlado tradicional é a adoção de uma estrutura pautada no RDF e no estabelecimento de URIs para identificação das classes e subclasses, bem como das declarações feitas com base no vocabulário. Com essas características, o DQV permite uma declaração formal.

Os elementos DQV podem ser aplicados não apenas para expressar metadados sobre a qualidade dos conjuntos de dados; eles também podem ser usados para expressar declarações sobre a qualidade dos próprios metadados. Isso é especialmente verdadeiro quando se trata de representar a proveniência desses metadados ou sua conformidade com os padrões de metadados estabelecidos. (W3C, 2017, não paginado).

As definições do DQV também destacam o seu objetivo de promover recuperação e facilitar a seleção de conjuntos de dados pertinentes. Por sua estrutura formal, ele facilitaria a utilização de ferramentas automática e semiautomáticas para consultar as descrições de dados fornecidas pelos publicadores (ALBERTONI; ISAAC, 2020; ARRUDA *et al.*, 2019; ADLAN, 2018; XIA; XU; MAO, 2017; XIA *et al.*, 2016).

Além de permitir que os publicadores de dados comuniquem os aspectos de qualidade de seus conjuntos de dados no momento de sua publicação, O DQV fornece uma estrutura que viabiliza que comunidade de usuários e entidades de certificação compartilhem os resultados de processos de avaliação. (W3C, 2017).

Do ponto de vista do usuário, informações de qualidade de dados são informações essenciais para apoiar a exploração, a fim de escolher o conjunto de dados certo com base em sua aplicação. (ADLAN, 2018, p.16).

O vocabulário fornece uma estrutura para que os usuários não dependam exclusivamente dos fornecedores de dados que podem, por motivos diversos, negligenciar a representação de metadados de qualidade de seus conjuntos de dados. Permite ainda que esses resultados sejam reutilizados, evitando assim

o retrabalho da realização de um novo processo de avaliação.

A descrição formal da qualidade dos conjuntos de dados tem impacto direto na sua reutilização, auxiliando consumidores de dados no processo de seleção de fontes adequadas para suas aplicações. Com essa finalidade, são elaboradas aplicações do DQV. Essas aplicações são discutidas na próxima subseção.

5.2 Aplicações do DQV

Por ser um modelo abrangente, que representa o processo de interesse à diversos domínios, bem como por sua compatibilidade com as melhores práticas para a publicação de dados na Web, o DQV tem sido empregado na elaboração de artefatos muito distintos. Para compreender melhor como o vocabulário tem sido aplicado, observou-se como e em que etapa o DQV foi utilizado nas aplicações identificadas. A tabela 2 apresenta os resultados dessa análise.

Tabela 2. Principais tipos de aplicação do DQV

<i>Categoria</i>	<i>Aplicação do DQV</i>	<i>Nº de aplicações</i>
1	Utiliza o DQV para exportar os resultados derivados do uso de um artefato de avaliação de qualidade	18
2	Amplia o DQV, criando um novo SOC	19
3	Utiliza o DQV para estruturar um artefato de avaliação de qualidade	12
4	Avaliar a qualidade de dados a serem consumidos pela ferramenta proposta e exportar esses resultados	3

Fonte: Autores (2023)

Os artefatos incluídos na categoria 1 demonstram a aplicação do DQV para o processo de geração de metadados descritivos de qualidade. Nesses casos, o DQV permite que após a realização de um processo de avaliação de qualidade, seja por uma ferramenta automática ou semiautomática, os resultados do processo sejam disponibilizados de maneira compatível com as MPs do W3C, facilitando assim o reuso dos resultados desse processo de avaliação por outros consumidores e ainda o uso desses resultados para alimentar outras ferramentas de avaliação de qualidade.

Já os artefatos incluídos na categoria 2 realizam um uso do DQV como um vocabulário para a criação de outros SOCs, utilizando o DQV de maneira semelhante à forma como seria empregado um padrão de metadados.

Esses artefatos utilizam a estrutura do DQV para elaboração de novos vocabulários semânticos e ontologias de qualidade. Geralmente, esses SOC são voltados para domínios específicos ou projetados para a utilização em algum artefato de avaliação de qualidade.

Os artefatos incluídos na categoria 3 utilizaram o DQV para estruturar e embasar a elaboração de artefatos de avaliação de qualidade, como metodologias, modelos de qualidade e ferramentas de avaliação de qualidade. Nesses artefatos, o DQV foi incluído como fundamentação para fornecer embasamento teórico na construção da ferramenta ou ainda foi integrado ao framework da ferramenta elaborada/apresentada pelos autores. Um destaque dessa categoria são os modelos de avaliação de qualidade (propostos em 5 artigos). Eles complementam o DQV na medida em que fornecem dimensões e métricas próprias à aplicação em determinado domínio.

Os artefatos incluídos na última categoria utilizam o DQV na etapa de seleção automática de dados que serão utilizados para alimentar uma ferramenta.

Apresentados os principais tipos de aplicação do DQV, a próxima seção apresenta algumas considerações finais.

6 Considerações finais

O presente estudo partiu do objetivo de discutir o DQV enquanto um sistema de organização do conhecimento, com base em suas definições e aplicações.

Em relação a sua descrição, conclui-se que o DQV pode ser definido como um vocabulário semântico, uma meta-estrutura que reflete o processo de avaliação de qualidade, como um padrão de estrutura de dados. Por ser abrangente, ele permite expressar as qualidades intrínsecas dos conjuntos de dados e sua qualidade contextual frente a determinada aplicação. Permite ainda expressar questões de qualidade representacionais e de acessibilidade por meio de avaliações dos metadados criados para descrever os conjuntos de dados avaliados. Ou seja, ele permite tanto a descrição da qualidade dos dados como dos metadados que os representam.

Em relação a sua aplicação, observou-se que, em sua maioria, os artefatos utilizam o DQV em ferramentas automáticas e semiautomáticas de avaliação de qualidade, especificamente na etapa de explicitação dos resultados. Também se destacaram aplicações onde o vocabulário é utilizado para a criação de novos SOCs como vocabulários semânticos, ontologias e modelos de qualidade de dados, e que esses produtos geralmente são criados com uma perspectiva contextual, para representar a qualidade em um domínio específico.

Conclui-se que o DQV pode facilitar os processos de recuperação, seleção e reuso dos dados disponíveis na Web, seja por meio da descrição dos conjuntos de dados pelos seus fornecedores, ou por sua aplicação na explicitação de resultados de processo de avaliação de qualidade, que pode ser realizado por usuários e agências de certificação.

Entretanto, para abranger todas essas perspectivas de qualidade, DQV depende da aplicação conjunta de um modelo de qualidade de dados, sendo necessário que esses modelos sejam criados para domínios específicos. Como estudos futuros, pretende-se avaliar o processo de elaboração desses modelos e de sua formalização com base no DQV.

Notas

Agradecemos à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento recebido para o desenvolvimento dessa pesquisa. PROCESSO FAPESP nº 2021/03349-0

Referências

- BATINI, C.; SCANNAPIECO, M. **Data quality: concepts, methodologies and techniques**. Berlin: Springer, 2006.
- ADLAN, C. A. **Linked spatial data: beyond the linked open data cloud**. 2018. 101 f. Tese (Doutorado) - Master of Science In Geo-Information Science, University Of Twente, Holanda, 2018.
- ARRUDA, N. et al. A Fuzzy approach for data quality assessment of linked datasets. INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS, [s.l.], **Proceedings**, 2019, p. 399-406.
- ALBERTONI, R.; MARTINO, M.; QUARATI, A. Documenting context-based quality assessment of controlled vocabularies. **Ieee Transactions On Emerging Topics In Computing**, [s.l.], v. 9, n. 1, p. 144-160, 1 jan. 2021.
- ALBERTONI, R.; ISAAC, A. Introducing the Data Quality Vocabulary (DQV). **Semantic Web**, [s.l.], v. 12, n. 1, p. 81-97, 19 nov. 2020. Disponível em: <https://semantic-web-journal.net/system/files/swj2079.pdf>. Acesso em: 30 mar. 2023.
- BISCALCHIN, R. Os sistemas de organização do conhecimento e os desafios frente a geração google. **Páginas A&B: Arquivos & Bibliotecas**, [s.l.] n. 11, p. 3-9, nov. 2019. Disponível em: <https://doi.org/10.21747/21836671/pag11a1>. Acesso em: 30 mar. 2023.
- BARITÉ, M. Control de vocabulario: orígenes, evolución y proyección. **Ciência da Informação**, Brasília, v. 1, n. 43, p. 95-119, 2015. Disponível em: <https://revista.ibict.br/ciinf/article/view/1421>. Acesso em: 30 mar. 2023.
- BIZER, C. **Quality driven information filtering: in the context of web based information systems**. Berlin: VDM Publishing, 2013. 196 p.

- BERNERS-LEE, T. **Linked data**, 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 14 mar. 2021.
- FUJITA, M. S. L.; TOLARE, J. B. Vocabulários controlados na representação e recuperação da informação em repositórios brasileiros. **Informação & Informação**, [s.l.], v. 24, n. 2, p. 93-125, 30 nov. 2019. Disponível em: [10.5433/1981-8920.2019v24n2p93](https://doi.org/10.5433/1981-8920.2019v24n2p93). Acesso em: 30 mar. 2023.
- ISOTANI, S.; BITTENCOURT, I. **Dados abertos conectados**. São Paulo: Novatec, 2015. 175 p. Disponível em: <http://www.icmc.usp.br/e/b0477>. Acesso em: 19 nov. 2020.
- JESUS, A. F.; SANTAREM SEGUNDO, J. E. A questão da qualidade em dados publicados como linked data: um mapeamento sistemático da literatura. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais [...]**. p. 1-16. Disponível em: <https://enancib.ancib.org/index.php/enancib/xxii/enancib/paper/view/811>. Acesso em: 30 mar. 2023.
- JURAN, J. M. **Quality Control Handbook**. New York: McGraw-Hill. 1988. 500 p.
- LANGER, A. et al. SemQuire - assessing the data quality of linked open data sources based on dqv. **Current Trends In Web Engineering**, [S.L.], p. 163-175, 2018.
- MARCONDES, C. H. Fundamentos da organização do conhecimento. **Pontodeacesso**, [s.l.], v. 15, n. 3, p. 249-282, dez. 2021. Disponível em: <https://doi.org/10.9771/rpa.v15i3.47468>. Acesso em: 30 mar. 2023.
- MARTINS, H.; VALENTE, J.. Datificação da economia e impactos nos mercados das comunicações: uma análise do google e do grupo globo. **Revista Eptic**, [s. l.], v. 3, n. 21, p. 85-100, set. 2019. Disponível em: <https://seer.ufs.br/index.php/epitic/article/view/12481/9430>. Acesso em: 30 mar. 2023.
- MIHINDUKULASOORIYA, N. **A framework for linked data quality based on data profiling and rdf shape induction**. 2020. 290 f. Tese (Doutorado) - Universidad Politécnica de Madrid, Madrid, 2020.
- MOREIRA, W. **Sistemas de organização do conhecimento: aspectos teóricos, conceituais e metodológicos**. 2018. 164 f. Livre-Docência em Sistemas de Organização do Conhecimento, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2018.
- MELO, J.O.S F. **Metodologia de avaliação de qualidade de dados no contexto do linked data**. 2017. 111 f. Dissertação (Mestrado) - Curso de Pós-Graduação em Ciência da Informação, Faculdade de Filosofia e Ciências de Marília, Universidade Estadual Paulista, Marília, 2017.
- NAYAK, A.; BOZIC, B.; LONGO, L. Linked data quality assessment: a survey. **International Conference On Web Services**, [s. l.], v. 12994, p. 63-76, jan. 2021.
- NELSON, R. R.; TODD, P. A.; WIXOM, B.H. Antecedents of information and system quality: an empirical examination within the context of data warehousing. **Journal of Management Information Systems**, [S.l.], v. 21, n. 4, p. 199-235, abr. 2005.
- NOOGHABI, M. Z.; DASTGERDI, A. F. Proposed metrics for data accessibility in the context of linked open data. **Program**, [S.L.], v. 50, n. 2, p. 184-194, abr. 2016. Disponível em: <http://dx.doi.org/10.1108/prog-01-2015-0007>. Acesso em: 26 maio 2022.
- PEREIRA JUNIOR, V. A. **Desenvolvimento de abordagem de fusão de informações utilizando técnicas semânticas para a melhoria da qualidade da informação de situações de risco**. 2017. 94 f. TCC (Graduação) - Curso de Bacharel em Ciência da Computação, Centro Universitário Eurípides de Marília, Marília, 2017.
- PEREIRA, G. M. C. **Modelo de Recuperação de Informação com Suporte Semântico e Temporal**. 2020. 92 f. Dissertação (Mestrado) - Curso de Programa de Pósgraduação em Ciência da Informação, Universidade Estadual Paulista, Marília, 2020. Disponível em: <https://repositorio.unesp.br/handle/11449/192964>. Acesso em: 30 mar. 2023.
- RADULOVIC, F. et al. A comprehensive quality model for linked data. **Semantic Web**, [s.l.], v. 9, n. 1, p. 3-24, 30 nov. 2017. Disponível em: <http://dx.doi.org/10.3233/sw-170267>. Acesso em: 03 abr. 2023.
- RÜMMLER, A.; HENZEN, C.; FIGGEMEIER, H. Modelling inhomogeneous geodata quality in a dataset's metadata. **Agile**, [s. l.], v. 59, n. 3, p. 1-6, jan. 2022.
- SILVA, J. N. et al. Desenvolvimento de ontologia ciente de qualidade de informações para o domínio de gerenciamento de emergências. **Encontros Bibli**, [s. l.], v. 23, n. 53, p. 184-200, set. 2018. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2018v23n53p182>. Acesso em: 31 mar. 2023.
- TALLERÁS, K. Quality of linked bibliographic data: the models, vocabularies, and links of data sets published by four national libraries. **Journal Of Library Metadata**, [s.l.], v. 17, n. 2, p. 126-155, abr. 2017. Disponível em: <http://dx.doi.org/10.1080/19386389.2017.1355166>. Acesso em: 26 mai. 2022.
- WANG, R. Y., STRONG, D. M. Beyond accuracy: what data quality means to data consumers. **J. Manage. Inf. Syst.** v. 12, n. 4, p 5-33, jan. 1996.
- W3C. **Data on the web best practices: Dataset Usage Vocabulary**. 2016. Disponível em: <https://www.w3.org/TR/vocab-duv/>. Acesso em: 16 ago. 2022.
- W3C. **Data on the web best practices**. 2017. Disponível em: <https://www.w3.org/TR/dwbp/#intro>. Acesso em: 16 ago. 2022.
- W3C. **Primer RDF**. 2004. Disponível em: <https://www.w3.org/TR/rdfprimer/#intro>. Acesso em: 08 set. 2020.

XIA, W.; XU, Z.; MAO, C. User-driven filtering and ranking of topical datasets based on overall data quality. **2017 14Th Web Information Systems And Applications Conference (Wisa)**, [s.l.], p. 257-262, nov. 2017. Disponível em: 10.1109/WISA.2017.24. Acesso em: 30 mar. 2023.

XIA, W. et al. DQFIRD: towards data quality-based filtering and ranking of datasets for data portals. In: WEB INFORMATION SYSTEMS AND APPLICATIONS CONFERENCE, 13., 2016, [S.L.]. **Anais [...]**. [s.l.], 2016. p. 18-23.

ZAVERI, A. et al. Quality assessment for Linked Data: a survey. **Semantic Web**, [s.l.], v. 7, n. 1, p. 63-93, 17 mar. 2015. Disponível em: <http://dx.doi.org/10.3233/sw-150175>. Acesso em: 19 ago. 2022.

Apêndice A – Definições do DQV obtidas na documentação oficial e na literatura levantada

AUTORES/ ANO	DEFINIÇÕES
Rümmler, Henzen e Figgemeier (2022)	<p>“O Vocabulário de Qualidade de Dados (DQV) é um vocabulário de dados semânticos [...] para descrever a qualidade de um conjunto de dados.” (p. 3, tradução nossa).</p> <p>“O DQV fornece um conceito genérico e extensível no que diz respeito à definição de indicadores de qualidade específicos do domínio, garantindo assim opções para definir qualquer indicador de qualidade espacial” (p. 3, grifo nosso, tradução nossa).</p>
Albertoni, Martino e Quarati (2021)	O Vocabulário de Qualidade de Dados (DQV) <i>fornece classes e propriedades para representar métricas</i> , dimensões e categorias de qualidade, fornecendo dados quantitativos ou qualitativos e informações sobre o conjunto de dados ou suas distribuições. (p. 39, grifo nosso, tradução nossa).
Nayak, Bozic e Longo (2021)	“O DQV permite [...] <i>representar</i> a avaliação da qualidade dos dados em formato de web semântica. Um usuário em potencial pode usar isso para <i>representar seus dados</i> em um relatório de avaliação de qualidade.” (p. 3-4, grifo nosso, tradução nossa)
Pereira (2020)	O DQV permite “identificar as relações das dimensões de qualidade com as instâncias existentes na <i>ontologia</i> . Este vocabulário foi desenvolvido pela W3C e objetiva as boas práticas para expressar a qualidade de dados, além de estimular, facilitar a publicação, intercâmbio e consumo de metadados de qualidade para cada etapa existente no processo informacional. (p. 62)
Albertoni e Isaac (2020)	“O DQV foi concebido como uma <i>estrutura interoperável</i> de alto nível que deve acomodar várias visões sobre a qualidade dos dados. DQV não procura determinar o que <i>significa qualidade</i> ”. (p. 1, grifo nosso, tradução nossa).
Mihindukulasoorya (2020)	<p>“Data Quality Vocabulary (DQV) é um vocabulário para <i>descrever</i> a qualidade de um conjunto de dados, seja pelo editor do conjunto de dados ou por uma comunidade mais ampla de usuários.”. (p. 11, grifo nosso, tradução nossa).</p> <p>“O DQV fornece classes e propriedades para capturar informações sobre categorias de qualidade, dimensões e métricas de um conjunto de dados, bem como sobre certificados de qualidade, padrões e proveniência relacionados a um conjunto de dados” “. (p. 11, grifo nosso, tradução nossa).</p>
Arruda <i>et al.</i> (2019)	O DQV “[...]permite que os editores de dados calculem a qualidade dos dados para que os consumidores possam <i>descobrir se os dados são adequados</i> para o uso pretendido ou não apresentam a qualidade dos dados esperado pelo consumidor.” (p. 400, grifo nosso, tradução nossa).
Adlan (2018)	“DQV não apenas fornece vocabulário para expressar a qualidade do conjunto de dados, mas também expressa declarações sobre a qualidade <i>dos metadados</i> ” (p. 16, grifo nosso, tradução nossa).
Langer <i>et al.</i> , (2018)	“O vocabulário de qualidade de dados (DQV) portanto, introduz um conjunto de propriedades para anunciar resultados de medição de qualidade”. “[...] um modelo de qualidade de dados poderia usar DQV como <i>ontologia base para representar</i> os elementos do modelo de qualidade como Linked Data. (p. 63 grifo nosso, tradução nossa).
Silva <i>et al.</i> (2018)	“O DQV é um modelo de boas práticas para expressar a qualidade de dados. Foi desenvolvido pelo Grupo do W3C (World Wide Web Consortium), visando estimular, facilitar a publicação, intercâmbio e consumo de metadados de qualidade, para cada etapa do ciclo de vida de um conjunto de dados.” (p. 190).
Pereira Junior (2017)	<p>Define o DQV como uma ontologia criada “visando estimular e facilitar a <i>publicação, intercâmbio, consumo e expressividade de metadados</i> de qualidade presentes em conjuntos de informações.” (p. 67)</p> <p>“DQV permite a criação de instâncias de categorias, dimensões e métricas para informar a qual uma determinada medida de qualidade se refere.” (p. 83).</p>
Radulovic <i>et al.</i> (2017)	<p>“[...] visa ser uma ontologia leve adequada para qualquer tipo de dados na web (por exemplo, CSV, XML, HTML, RDF, etc.), portanto, é <i>genérico</i> e não aborda as características específicas do Linked Data. (p. 2, grifo nosso, tradução nossa).</p> <p>“[...] ele apenas fornece uma <i>estrutura básica</i> para descrever métricas e medidas de qualidade, mas não define métricas de qualidade, que devem aparecer em qualidade modelos”. (p.5, grifo nosso, tradução nossa).</p> <p>“No contexto do modelo conceitual, DQV tende a fornecer os meios para capturando tanto os detalhes sobre qualidade (ou seja, características e medidas) quanto sobre valores de qualidade (resultados de avaliação)” (p. 5, grifo nosso, tradução nossa).</p>
Xia, Xu e Mao (2017)	“Fornece uma <i>estrutura [framework]</i> para descrever a qualidade de conjuntos de dados publicados em um portal de dados (<i>catálogo de dados</i>). Ele <i>define várias classes e propriedades</i> RDF usadas para implementar um modelo de qualidade personalizado para o portal de dados com base em um ou modelos de qualidade de dados mais estabelecidos (...) e permite que os editores de conjuntos de dados ou uma comunidade mais ampla de usuários <i>expressar metadados de qualidade</i> ”. (p. 258, grifo nosso, tradução nossa).
W3C (2016)	O DQV “[...] fornece uma <i>estrutura [framework]</i> na qual a qualidade de um conjunto de dados pode ser <i>descrita</i> , seja pelo editor do conjunto de dados ou por uma comunidade mais ampla de usuários. Ele <i>não fornece uma definição formal e completa de qualidade</i> , mas estabelece um meio consistente pelo qual as informações podem ser fornecidas, de modo que um usuário potencial de um conjunto de dados possa fazer seu próprio julgamento sobre sua adequação ao propósito”. (não paginado, grifo nosso, tradução nossa).

Fonte: Autores (2023)

Datificação, Inteligência Artificial e imagens: o paradigma dominante na representação do conhecimento em imagens

Datafication, Artificial Intelligence, and images: the dominant paradigm in images' knowledge representation

Pedro Cremonoz Rosa (1), Antonio Lucio Barizon Filho (2), Raquel Torrão Valentim (3),
Natália Bolfarini Tognoli (4)

(1) Universidade Estadual de Londrina, pedro.cremonoz@uel.br (2) antonio.barizon@uel.br (3) Universidade Federal Fluminense, Raquel_valentim@id.uff.br (4) nataliatognoli@id.uff.br

Resumo

Objetivo: Verificar se imagens geradas por Inteligências Artificiais generativas replicam estereótipos e preconceitos presentes no paradigma dominante. **Metodologia:** Desenvolveu-se um estudo de caso na ferramenta Leonardo.Ai, a qual gerou imagens a partir do uso de termos simples combinados, a saber: "Scientist, person"; "Cook, person"; "Doctor, person"; "CEO, person"; "Housekeeper, person"; e "Nurse, person". As imagens escolhidas foram analisadas a partir da metodologia de análise documental de imagens de Rodrigues (2007), em conjunto com as contribuições de Gemma Penn (2008), através dos critérios: sexo; faixa etária; etnia; porte físico; vestuário; e elementos circunscritos. **Resultados:** Verificou-se que as imagens geradas pela ferramenta Leonardo.Ai apresentaram uma série de características que perpetuam estereótipos e preconceitos. **Conclusões:** Visto o provável impacto que Inteligências Artificiais generativas terão na produção e compartilhamento de informações, conclui-se que é preciso discutir, além das questões éticas relacionadas à própria criação da ferramenta, modos de torná-la mais inclusiva e responsável para a representação das informações.

Palavras-chave: Inteligência Artificial. Imagem. Representação do Conhecimento em Imagens. Paradigma dominante.

Abstract

Objective: To verify whether images generated by generative Artificial Intelligence replicate stereotypes and prejudices present in the dominant paradigm. **Methodology:** We conducted a case study in the Leonardo.Ai tool, which generated images from the use of simple terms combined, namely: "Scientist, person"; "Cook, person"; "Doctor, person"; "CEO, person"; "Housekeeper, person"; and "Nurse, person". The chosen images were analyzed based on the methodology of documentary analysis of images by Rodrigues (2007), together with the contributions of Gemma Penn (2008), through the criteria: sex; age range; ethnicity; physical shape; clothing; and circumscribed elements. **Results:** The images generated by the Leonardo.Ai tool presented a series of characteristics that perpetuate stereotypes and prejudices. **Conclusions:** Given the likely impact that generative Artificial Intelligence will have on the production and sharing of information, it is concluded that, in addition to ethical issues related to the creation of the tool, ways to make it more inclusive and responsible need to be discussed to information representation.

Keywords: Artificial Intelligence. Image. Images' Knowledge Representation. Dominant paradigm.

Resumen

Objetivo: Verificar si las imágenes generadas por Inteligencia Artificial generativa replican estereotipos y prejuicios presentes en el paradigma dominante. **Metodología:** Se desarrolló un estudio de caso en la herramienta Leonardo.Ai, que generó imágenes a partir del uso de términos simples combinados, a saber: "Científico, persona"; "Cocinero, persona"; "Médico, persona"; "CEO, persona"; "ama de llaves, persona"; y "enfermera, persona". Las imágenes escogidas fueron analizadas con base en la metodología de análisis documental de imágenes de Rodrigues (2007), junto con los aportes de Gemma Penn (2008), a través de los criterios: sexo; rango de edad; etnicidad; forma física; ropa; y elementos circunscritos. **Resultados:** Se verificó que las imágenes generadas por la herramienta Leonardo.Ai presentaban una serie de características que perpetúan estereotipos y prejuicios. **Conclusiones:** Dado el probable impacto que tendrán las Inteligencias Artificiales generativas en la producción y el intercambio de información, se concluye que, además de las cuestiones éticas relacionadas con la creación de la herramienta, es necesario discutir formas de hacerla más inclusiva y responsable para la representación de la información.

Palabras clave: Inteligencia Artificial. Imagen. Representación del Conocimiento en Imágenes. Paradigma dominante.

1 Introdução

As ferramentas para geração de imagens por meio de Inteligências Artificiais generativas tornaram-se acessíveis e populares, possibilitando gerar imagens em poucos segundos a partir de qualquer comando textual. Construídas a partir da datificação de bilhões de imagens, as Inteligências Artificiais generativas replicam estilos, técnicas e características visuais de sua base de dados.

Contudo, compreendendo que a construção e o desenvolvimento do processo de aprendizagem dessas ferramentas dependem do envio prévio de dados de seus programadores, questionamos a possibilidade de as Inteligências Artificiais generativas replicarem preconceitos e estereótipos sociais presentes no paradigma dominante.

Nesse sentido, com o objetivo de verificar se imagens geradas por ferramentas desse tipo replicam tais estereótipos e preconceitos presentes nesse paradigma, desenvolvemos um estudo de caso na ferramenta Leonardo. Ai a partir dos termos combinados: "*Scientist, person*"; "*Cook, person*"; "*Doctor, person*"; "*CEO, person*"; "*Housekeeper, person*"; e "*Nurse, person*". As imagens obtidas foram analisadas segundo a metodologia de análise documentária de imagens de Rodrigues (2007) em conjunto com as contribuições de Gemma Penn (2008) através dos seguintes critérios: sexo; faixa etária; etnia; porte físico; vestuário; e elementos circunscritos.

Essa pesquisa é resultado de uma colaboração entre instituições e entre grupos de pesquisa certificados pelo CNPq: Grupo de Pesquisa Estudos Terminológicos em Ciência da Informação (Universidade Estadual de Londrina) e Grupo de Pesquisa Arquivos Bibliotecas e Organização do Conhecimento (Universidade Federal Fluminense) [1].

Como resultado verificou-se que as imagens apresentam uma série de características que perpetuam estereótipos e preconceitos, especialmente em relação à representação de gênero, idade, cor de pele, porte físico e trajas em determinadas profissões. Portanto, visto o provável impacto que as Inteligências Artificiais generativas terão na produção e compartilhamento de informações, conclui-se que, é preciso discutir, além das questões éticas relacionadas à própria criação da ferramenta, modos de torná-la mais inclusiva e responsável para a representação da informação.

2 Imagens em Representação do Conhecimento

A imagem é uma forma de comunicação que desempenha um papel fundamental na Ciência da Informação, pois permite a transmissão de informações visuais, muitas vezes de maneira mais eficaz do que o

texto escrito, por concentrar grande quantidade de informação em uma representação visual (ROBERTS, 2001). As imagens são utilizadas para registrar e documentar eventos, situações e informações relevantes, contribuindo para a construção da memória coletiva e do conhecimento científico (TORRES e MACULAN, 2019).

Além disso, a imagem também é utilizada como um recurso importante em diversas áreas da Ciência da Informação, como em museus, bibliotecas e arquivos, onde é comum encontrar coleções de imagens que são utilizadas para ilustrar, documentar e contextualizar informações em diferentes suportes, tais como livros, artigos e relatórios (MAIMONE, 2018). A análise de imagens também é uma área importante na Ciência da Informação, permitindo a extração de informações relevantes, tais como identificação de padrões, detecção de objetos e reconhecimento de caracteres, o que contribui para o desenvolvimento de novas tecnologias de busca, recuperação e análise de informação visual (MANINI, 2002).

Durante a década de 2001 a 2010, um novo conceito de uso da internet emergiu. A característica fundamental deste novo modelo foi a possibilidade de publicação e tratamento livre de informações por parte dos usuários, o que os estudiosos chamaram de Web 2.0 (PALETTA e GONZALES, 2019). Esse modelo estabeleceu-se e aprimorou-se, e as redes sociais foram uma consequência natural desse fato. A criação de ferramentas disseminou e democratizou o uso de imagens, que agora podem ser acessadas a qualquer momento e de qualquer lugar do mundo.

Essa liberdade de publicação de imagens pelos usuários na internet também trouxe a problemática sobre como tratar e recuperar essas imagens. Dada a variedade de fatores envolvidos, incluindo o tipo de imagem e o suporte em que se encontra, a literatura científica de Ciência da Informação ainda não chegou a um consenso sobre como as imagens devem ser tratadas, embora alguns modelos eficientes para indexação de imagens (MANINI, 2002; RODRIGUES, 2007) possam ser encontrados na literatura.

Manini (2002) parte das propostas de Análise Documentária de Fotografias como forma de endossar e estruturar lacunas acerca da Análise Documentária de Imagens. A autora trabalha com a transposição de métodos e técnicas da análise documentária aplicados em textos escritos para análise de fotografias. A dimensão expressiva da fotografia deve ser levada em consideração, bem como a representação escrita e a recuperação da informação imagética pelo usuário. O objetivo é identificar o conteúdo informacional da imagem fotográfica, além de entender a lacuna na análise documentária de imagens. A indexação deve ser feita com base nos atributos das imagens e pode ser aplicada a grupos de imagens.

Rodrigues (2007) propôs uma metodologia que destaca que uma imagem não apenas mostra, mas também representa algo que pode não ter uma relação direta com os objetos apresentados. Dessa forma, uma imagem pode ter dois níveis ou sentidos principais: o denotativo, que se refere àquilo que a imagem representa com certa precisão, e o conotativo, que se refere àquilo que a imagem pode interpretar em um determinado contexto, em um sentido figurado e simbólico.

A análise documentária de imagens começa com a leitura do documento fotográfico, que exige conhecimento prévio do conteúdo, mas não é uma condição para a análise. A leitura do profissional prepara a leitura do usuário, incluindo a elaboração de um resumo e a indexação. O resumo é uma etapa da análise documentária e pode ser útil para a indexação, mesmo que não contenha informações adicionais. Título e legenda podem ser indiferenciados, mas o resumo não deve ser confundido com a legenda. O resumo é construído, menos conciso e muitas vezes é abandonado, mas é importante para fornecer palavras-chave ou termos de indexação (MANINI, 2002).

A tradução do visual para o verbal é a escolha do termo de indexação, e é importante que o profissional da informação tenha conhecimento sobre o conteúdo do documento, os interesses dos usuários e a política da instituição. A análise de uma fotografia pode ser conduzida em três níveis: pré-iconográfico, iconográfico e iconológico, que remetem, respectivamente, ao referente da imagem, ao significado atribuível ao referente e aos valores simbólicos. O nível iconológico é o lugar dos valores simbólicos e remete a significados culturais, sociais, filosóficos ou ideológicos, e é onde o autor da imagem mais se afasta do leitor (MANINI, 2002).

Um ponto de extrema importância dentro do universo de Análise Documentária de imagens englobado no universo das fotografias é o uso da legenda como forma de identificação e interpretação da imagem fotográfica. As legendas descrevem o princípio da ancoragem (BARTHES, 1990), e funcionam como um limitador de significação que a imagem pode ou não pode se associar. Esta ancoragem serve como uma fixação estática das potencialidades de significado da imagem e, conseqüentemente, as limitações de sua interpretação.

Segundo Barthes (1990), a linguagem é um sistema de signos que possui uma natureza instável e ambígua, tornando-se necessária a utilização de mecanismos que fixem o sentido das palavras. Nesse contexto, a ancoragem é um processo pelo qual um elemento do discurso é fixado a um significado específico, limitando as possíveis interpretações. Assim, a ancoragem é um princípio que tem como objetivo evitar a polissemia, a ambiguidade e a incerteza na

comunicação. Para Barthes, a ancoragem é um recurso importante para tornar o discurso mais objetivo e claro, permitindo que os significados sejam compreendidos de maneira mais precisa.

A legenda é definida como palavras que acompanham a imagem e são fundamentais para a interpretação correta da fotografia. Ela pode conter o título e explicações sobre a produção da imagem fotográfica, além de refletir seu conteúdo de maneira genérica, específica ou abstrata. A legenda é condutora do olhar do espectador e pode ser utilizada para imprimir características específicas à imagem, como identificação, contexto e intenção. Apresentando certas similaridades acerca das legendas e da linguagem recém popularizada, os prompts de comando para as Inteligências Artificiais podem ser observados dentro deste procedimento de Análise Documental de Imagens.

Dentro do novo ambiente virtual, potencializado pela emergência das Inteligências Artificiais, tanto as legendas de imagens quanto os prompts de comando são formas de orientar a produção de imagens por Inteligências Artificiais. A legenda de imagem é uma descrição textual que acompanha uma imagem e visa fornecer uma interpretação do que está sendo mostrado, enquanto um prompt de comando é uma instrução que orienta um sistema de Inteligência Artificial a gerar uma imagem com determinadas características.

Ambos são importantes para a produção de imagens por Inteligências Artificiais porque ajudam a definir o contexto e as características desejadas para a imagem que se quer criar. Além disso, tanto as legendas de imagens quanto os prompts de comando podem ser ajustados para produzir diferentes resultados, permitindo a criação de imagens com variações em suas características.

No entanto, há uma diferença importante entre as duas formas de orientação. Enquanto a legenda de imagem pretende descrever o que já está na imagem, o prompt de comando procura guiar a criação de uma imagem nova. Assim, enquanto a legenda de imagem é uma interpretação do que é visto na imagem, o prompt de comando é uma instrução para gerar uma nova imagem.

3 Datificação, Inteligência Artificial e imagens

Em consonância com o desenvolvimento e o crescimento da Web 2.0, novas ferramentas e recursos digitais ganharam espaço nos hábitos sociais. Seu crescimento e novos meios de aplicações refletem-se nas ciências e nas humanidades, onde pesquisas teóricas e experimentais cada vez mais ousadas propõem a utilização de ferramentas de Inteligências Artificiais como meios para automação de atividades visando a resolução de problemas.

Em especial, na Ciência da Informação, estudos sobre Inteligências Artificiais e automação de atividades de gestão e curadoria de acervos digitais – nato-digitais ou digitalizados – passaram a ocupar espaços privilegiados nas discussões acadêmicas, sejam através de eventos do campo, produções textuais, disciplinas especializadas nos cursos de formação ou projetos inter e transdisciplinares.

No entanto, ressalta-se que apesar dessas tecnologias e seus estudos virem tornando-se ferramentas cada vez mais essenciais, o campo ainda carece de estudos aprofundados. Tal fato dá-se, sobretudo pela pouca presença do tema nos currículos dos cursos de Ciência da Informação, bem como em seus principais campos relacionados, tais como a Biblioteconomia e a Arquivologia, sobretudo, devido a pouca atuação de pesquisadores das humanidades sobre o tema (JAILLANT, 2022).

Contudo, destaca-se que os esforços para a ampliação de estudos sobre a aplicação da Inteligência Artificial na Ciência da Informação e na Organização do Conhecimento significam avanços para o tratamento e curadoria de acervos, bem como nas relações entre as instituições e os usuários e, se antes o uso do termo era evitado por suas incertezas e pelos receios de seus pesquisadores (MARTÍNEZ-ÁVILA, 2015), agora se apresenta como sinônimo de solução para diversas questões no cenário tecnológico, proporcionando a automatização e a facilitação de atividades em diversos setores.

Os primórdios dos estudos e desenvolvimentos dessa tecnologia remetem ao matemático, cientista da computação, filósofo e biólogo, Alan Turing (1912-1954). Entretanto, foi apenas em 1956, durante uma reunião de pesquisadores sobre o tema em *Dartmouth College*, em New Hampshire (EUA), que John McCarthy, cientista da computação e um dos pioneiros nos estudos sobre o tema, cunhou o termo (BARBOSA; PORTES, 2023).

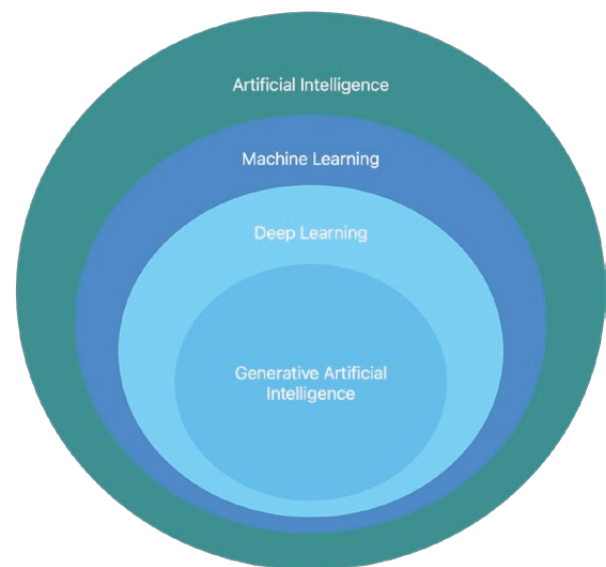
McCarthy (2007) no artigo "*What is artificial intelligence?*" apresenta perguntas e respostas básicas e essenciais sobre o tema, onde define Inteligência Artificial como "a ciência e engenharia de fazer máquinas inteligentes, especialmente sistemas de computador inteligentes" (MCCARTHY, 2007, p. 2).

Refletindo um pouco mais sobre o pensamento desse autor, podemos dizer que a Inteligência Artificial corresponde ao desenvolvimento de máquinas programadas com a capacidade de "aprendizado" a partir do uso de algoritmos, possibilitando respostas e interações a partir de dados fornecidos (DAMACENO; VASCONCELOS, 2018). Esse "aprendizado" ocorre a partir dos dados inseridos e armazenados e fórmulas do algoritmo que permitem aos softwares desencadear respostas aos problemas apresentados.

Essas ferramentas podem ser classificadas conforme suas camadas, dividindo-se em *Machine Learning* e *Deep Learning*. A primeira, como o nome sugere, promove um aprendizado contínuo da máquina, onde seus algoritmos são estruturados com equações que organizam e armazenam os dados fornecidos, promovendo respostas mais adequadas à solução de um problema. Por sua vez, *Deep Learning* configura um tipo de *Machine Learning* que executa tarefas mais complexas, tais como a identificação de imagens e o reconhecimento de falas, onde estabelece parâmetros básicos sobre os dados de entrada e os utiliza em camadas para o reconhecimento de padrões, buscando imitar o aprendizado humano e executando diversas tarefas com a "experiência" (DAMACENO; VASCONCELOS, 2018).

É nessa última categoria que encontramos as chamadas Inteligências Artificiais Generativas, isto é, uma tecnologia com base em *Deep Learning* cuja representação se assemelha mais a linguagem natural humana. A proposta desse tipo de tecnologia consiste na geração de resultados por meio da análise de dados e do treinamento armazenados no sistema, como é o caso, por exemplo, do *Generative Pre-trained Transformer*, ou *ChatGPT*, cujo uso tornou-se popular, evidenciando a pauta sobre as possibilidades de aplicações de ferramentas tecnológicas para solucionar problemas dos mais diferentes níveis, desde a busca por *streams* de filmes até respostas para o desenvolvimento de atividades de engenharia de *software*.

Figura 1. Da inteligência artificial à inteligência artificial generativa



Fonte: Elaborado pelos autores.

Mas as possibilidades desses recursos não se limitam à comunicação escrita. Além da linguagem textual,

Inteligências Generativas possibilitam a representação de conteúdos em outros meios, como através da criação de imagens. Ferramentas como a Leonardo.Ai permitem que o usuário obtenha representações visuais de assuntos a partir da interação por meio da utilização de termos simples e termos completos.

Assim, ao descrever o tipo de imagem que gostaria, com o máximo de informações para obtenção da precisão do que deseja, o software traduz as informações recebidas junto aos dados armazenados e disponibiliza o conteúdo de forma imagética.

Um componente crítico para a performance desse tipo de sistema, conforme Schuhmann et al. (2022), é o conjunto de dados, em especial em relação ao seu tamanho, utilizados para treinar a Inteligência Artificial. Enquanto os conjuntos iniciais eram constituídos de imagens capturadas da internet acompanhadas de descrições, anotações e metadados textuais inseridos por humanos, novos conjuntos se utilizam dos avanços computacionais e da própria IA para o processo de datificação, gerando conjuntos com mais de 5 bilhões de pares imagem-texto (combinação de uma imagem e sua descrição textual correspondente).

As possibilidades de uso desses recursos trazem novos horizontes para os campos das artes e designers, possibilitando novas formas de atuação. Entretanto, o crescimento do uso desse e outros tipos de Inteligências Artificiais e sua atual evidência nas pautas sociais, levam a outros questionamentos, como a violação da ética e da moral como, por exemplo, o questionamento na justiça do uso de imagens sem a autorização de seus autores nos conjuntos de dados utilizados para treinar as ferramentas.

Em 2018, o professor e pesquisador em Filosofia Marc-Antoine Dilliac da *University of Montreal*, em Quebec (Canadá), concedeu uma entrevista publicada na revista da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), onde apontou riscos éticos do uso de Inteligências Artificiais. A maior preocupação reside na possibilidade da utilização dessas ferramentas em ações discriminatórias a partir da identificação racial e da orientação sexual, bem como da análise e previsão comportamental, ainda que apresentadas de forma imprecisa e errônea, por exemplo. [2]

Tem-se, portanto, uma preocupação na garantia dos valores éticos e morais e que o paradigma dominante não prevaleça sobre as representações geradas por essas tecnologias. As discussões acerca da aplicação desses recursos devem, então, ser tratadas com cautela.

Assim, compreende-se que os debates éticos sobre o uso de tecnologias, em especial das Inteligências Artificiais, requerem mais atenção em discussões acadêmicas para a efetivação de ações imediatas, uma vez que sua popularização e uso vêm crescendo e, com

isso, novos dilemas éticos emergem com relação à representação da informação.

4 Paradigma dominante nas IA

A produção de conhecimento é um processo natural do ser humano em que as observações e transformações são estruturadas e sintetizadas em formas lógicas e metodologias para uma possível disseminação. Conforme houve a evolução desses processos e constructos, houve uma necessidade e um interesse em mapear essas concepções normativas de investigação, suas validades e viabilidades processuais. Essa vertente filosófica é denominada Epistemologia.

Para a Epistemologia, o conhecimento é uma forma de crença envolvida em processos de asserção e justificação de fenômenos observados como forma de aprisionar a realidade e, posteriormente, reproduzi-la em novas observações e fundamentações de observações semelhantes. A preocupação em compreender este percurso lógico da construção do conhecimento permite uma gama de concepções filosóficas que analisam todos os procedimentos empregados na construção do conhecimento, como seu impacto, influência, contexto histórico e evolução.

Nessa evolução histórica do conhecimento, observou-se uma virada do pensamento tido como científico em meados do século XVI. Segundo Santos (2010), a emergência da Ciência Moderna de Galileu Galilei e René Descartes e Francis Bacon, por exemplo, reconfigurou as formas e métodos de pensamento científico em um novo patamar lógico.

O empirismo, método em que o pensamento científico se baseia em um processo de experiência e observação que resultam em um método indutivo na formulação de estruturas lógicas acerca do fenômeno observado, substituiu o método dedutivo aristotélico como fundamentação do pensamento.

Nessa Revolução Científica, o conhecimento científico ainda tinha seu princípio basilar em garantir sua própria viabilidade e produção (ABBAGNANO, 2000), a partir de hipóteses, métodos de investigação, princípios de redução, objetividade, validação e princípio de falibilidade. Com essa nova configuração do pensamento científico, a estrutura lógica passava a ser endossada por um rigor absoluto e, conseqüentemente, seletivo em relação aos conhecimentos tidos como válidos e científicos.

O senso comum, neste contexto, deixou de ser visto como uma etapa importante na construção do conhecimento e, também, todo e qualquer conhecimento que não englobasse o pensamento moderno enquanto lógica de conduta científica. A ciência deixou de ser uma forma estrutural na concepção do conhecimento e passou a ser uma forma

excludente de todo e qualquer conhecimento que não se adequasse aos seus princípios de validação.

Muitos saberes deixaram de ser agregados ao conhecimento científico dado a esse caráter regulamentar do pensamento moderno. Santos (2010) descreve este pensamento como uma ordem científica hegemônica, baseado na racionalidade pautada pela Revolução Científica do século XVI. A ordem científica hegemônica, portanto, se incumbiu de delimitar a produção do conhecimento científico, só tido como válido a partir de sua construção baseada em um método rigoroso, mas também potencializou uma crise futura sobre o próprio pensamento científico pois, com o aprofundamento e especialização de conhecimentos, sinais de fragilidade na fundamentação científica do pensamento moderno foram cada vez mais vivenciadas.

O paradigma dominante se estabelece em uma forma vigente por apresentar ontologias, princípios teóricos e metodologias que sustentem sua viabilidade e continuidade. Em sua formação estrutural em que o paradigma dominante cria formas de se estruturar e reestruturar em uma configuração científica, componentes de manutenção e perpetuação deste rigor científico tornam-se ferramentas de continuidade do pensamento moderno na atualidade.

Por ser uma estrutura lógica e processual do pensamento do paradigma dominante, passa a permear inúmeros setores da sociedade para estender sua forma de domínio. O pensamento decorrente desse paradigma atinge outros aspectos da sociedade que fogem do escopo da ciência, por estar relacionado a um princípio de mentalidade resultante do pensamento moderno. Santos (2021) retrata tais aspectos que são inerentes ao pensamento científico moderno, mas que alcançam outras estruturas sociais enquanto monoculturas. Composta de cinco monoculturas, o paradigma dominante se instaura em diversos aspectos da sociedade, como: o saber (com a monocultura do rigor do saber), o tempo (com a monocultura do tempo linear), a classificação social (com a monocultura da classificação social), o princípio lógico de escala (com a monocultura da escala dominante) e a lógica produtivista (com a monocultura do produtivismo capitalista).

Cada uma destas monoculturas é resultante do pensamento moderno e do método indutivo e fazem com que a percepção da sociedade ou a forma com que a mesma se comporta seja alterada a partir desta perspectiva do paradigma dominante. As monoculturas, portanto, são constituídas a partir de princípios colonialistas, patriarcais e que desarticulam as noções que vão além do saber. O paradigma dominante dita a percepção dos indivíduos em relação à validade dos saberes, a percepção do tempo como linear, a constituição da sociedade a partir da normalização das

diferenças e das hierarquias, as escalas globais e universais como forma de visão de mundo e a lógica predatória e exploratória do capitalismo.

Nesse panorama, Santos (2010) descreve que, nesse conjunto de elementos que constituem o pensamento científico advindo do paradigma dominante, que vão além da prática científica, a Epistemologia se reconfigura, podendo ser compreendida enquanto Epistemologias do Norte.

As Epistemologias do Norte são o conjunto de práticas que perpetuam aspectos do paradigma dominante em diversos níveis na sociedade como forma de manutenção e perpetuação da mentalidade dominante. As Epistemologias do Norte tendem a valorizar o conhecimento científico e tecnológico, baseado em métodos empíricos e objetivos de observação e experimentação. Essa abordagem valoriza a objetividade e a neutralidade e tende a enfatizar a separação entre o sujeito e o objeto do conhecimento. As Epistemologias do Norte também tendem a enfatizar a importância da lógica e da razão na construção do conhecimento. Essa abordagem ao conhecimento pode ser limitada e tende a ignorar outras formas de conhecimento, como o conhecimento local e tradicional, que pode ser baseado em experiências pessoais, histórias e práticas culturais. Além disso, as epistemologias do norte tendem a assumir que o Conhecimento é Universal e pode ser aplicado a todas as sociedades, independentemente de sua cultura ou história, o que pode levar à homogeneização e à perda de diversidade cultural.

Desta forma toda a percepção da sociedade que é constituída e formalizada pelas Epistemologias do Norte sustentam e fundamentam uma visão unidimensionalizada sobre a realidade humana. As formas de assegurar a vigência do paradigma dominante são tecnologias que perpetuam as concepções das monoculturas em forma de práticas sociais. Todo e qualquer produto resultante das Epistemologias do Norte são sustentáculos do pensamento moderno e do paradigma dominante.

Ou seja, as ferramentas de Inteligências Artificiais, enquanto produtos das Epistemologias do Norte são, segundo Santos (2021), uma forma de promover e continuar o paradigma dominante. Enquanto resultado, as imagens geradas são produtos de uma tecnologia produzida em um contexto do Norte epistemológico, que por sua vez tem como base dados advindos de uma plataforma que em si também é um produto do paradigma dominante (internet).

5 Aplicação - Estudo de caso com análise da imagem

Com o objetivo de compreender de que forma as ferramentas de geração de imagens por Inteligências

Artificiais generativas contribuem para a predominância do paradigma dominante, foram geradas e analisadas imagens a partir de comandos (*prompts*) simples.

A ferramenta selecionada para o estudo de caso foi a Leonardo.Ai (www.leonardo.ai), uma aplicação web de acesso gratuito que oferece ao usuário uma quantidade de créditos renováveis diariamente para a geração de imagens. A escolha pela ferramenta teve como critério a qualidade das imagens geradas (ainda que não sejam perfeitas, que seja possível identificar claramente as características a serem analisadas), facilidade de acesso (gratuidade da ferramenta e acesso via navegador, sem necessidade de instalação, de forma que seu uso não fique restrito a um grupo de pessoas) e sua usabilidade (uma interface gráfica e comandos que não necessitem conhecimento técnico para sua utilização).

Embora em geral as ferramentas aceitem comandos em português, optou-se pela utilização do inglês pela ausência de gênero, o que permitiu uma melhor análise das imagens geradas. Neste mesmo sentido, buscou-se nos comandos a menor quantidade possível de termos, para oferecer o mínimo de informações e direcionamento.

Enquanto tema central dos termos a serem utilizados nos comandos, foram selecionadas seis profissões que possuem uma imagem estereotipada socialmente. Um primeiro teste dos termos retornou uma série de imagens com diferentes elementos e ausentes de uma figura humana, motivo pela qual foi adicionado um segundo termo (*person* - pessoa) que garantisse uma imagem em que os critérios pudessem ser analisados. O quadro 1 apresenta os comandos utilizados, seguidos de sua tradução entre parênteses.

Tabela 1. Comandos utilizados para geração de imagens

ID	Comando
01	<i>Scientist, person</i> (cientista, pessoa)
02	<i>Cook, person</i> (cozinheiro(a), pessoa)
03	<i>Doctor, person</i> (médico(a), pessoa)
04	<i>CEO, person</i> (diretor Executivo, pessoa)
05	<i>Housekeeper, person</i> (empregado (a) doméstico (a), pessoa)
06	<i>Nurse, person</i> (enfermeiro(a), pessoa)

Fonte: Elaborado pelos autores.

Após as gerações, as imagens foram analisadas a partir de um recorte da metodologia de análise e tematização de imagens de Rodrigues (2007) em conjunto com as contribuições de Gemma Penn (2008) a partir das imagens geradas pela Inteligência Artificial. A análise parte da distinção entre a linguagem denotativa – a literalidade da composição imagética –, utilizando

como critérios a verificação do sexo, faixa etária, etnia, porte físico, vestuário e elementos circunscritos, de modo a identificar se foram replicados estereótipos sociais e preconceitos, e sua linguagem conotativa como as possíveis interpretações da imagem.

As imagens, no entanto, não foram analisadas isoladamente, mas apoiadas no conceito proposto por Pierce (2005) de recursividade, que pauta a relação intrínseca que os signos possuem entre si, não existindo uma ordem sónica no processo de semiose, mas um processo de recursividade em que um signo é sustentado e configurado por outro no processo de representação.

As imagens foram geradas em abril de 2023 utilizando a configuração padrão da ferramenta, o que gera quatro imagens a partir de um único comando (Apêndice A). O quadro 2 apresenta a análise denotativa do conjunto de imagens para cada um dos termos e em seguida discorre-se sobre os aspectos conotativos.

Tabela 2. Análise denotativa das imagens geradas

ID	Análise denotativa
01	<i>Scientist, person</i> (cientista): Figuras masculinas, em idade adulta (30 anos ou mais), com maior incidência da pele caucasiana (3 das 4 figuras humanas representadas), indivíduos magros, com elementos conotativos de ideia (lâmpadas) ou invenções (maquinários), portando trajes formais com jalecos.
02	<i>Cook, person</i> (cozinheiro(a)): Maioria composta de figuras masculinas, caucasianos, trajando doma e chapéu de chef, em frente de instrumentos de cozinha ou alimentos, jovens.
03	<i>Doctor, person</i> (médico(a)): Predominância de homens, idade adulta (30 anos ou mais), com maior incidência da pele caucasiana (5 das 6 figuras humanas representadas), portando trajes formais de médico (jaleco e terno ou jaleco e conjunto hospitalar), presença de estetoscópio e objetos hospitalares.
04	<i>CEO, person</i> (diretor(a) executivo): Figuras masculinas, com maior incidência da pele caucasiana (3 das 4 figuras humanas representadas), idade adulta (30 anos ou mais), trajes formais (terno e gravata), aparentando seriedade
05	<i>Housekeeper, person</i> (empregado (a) doméstico (a)): Figuras femininas, com maior incidência da pele caucasiana (4 das 5 figuras humanas representadas), usando uniformes limpos, luvas, cabelos presos, com objetos de limpeza em mãos
06	<i>Nurse, person</i> (enfermeiro(a)): Maioria composta de figuras femininas, jovens, uniformizadas, pele caucasiana, com trajes clássicos de enfermagem

Fonte: Elaborado pelos autores.

Apoiada no conceito de recursividade, a análise se deu a partir dos comandos textuais que as geraram, de forma que a significação imagética é pautada pelo texto mas a extrapola uma vez que a linguagem visual requer para sua construção uma quantidade informativa muito superior à do texto, caracterizada por um alto nível de abstração.

É justamente essa lacuna informacional o foco da análise. Representar graficamente um termo requer especificidades em determinadas características ausentes no texto. Ao gerar a imagem do termo cientista, por exemplo, a ferramenta toma uma série de “decisões” visuais quanto a sexo, cor, postura, vestimenta, ambientação, etc., sendo estas pautadas pela ferramenta e pelos dados utilizados em seu desenvolvimento.

Dentro dos elementos levantados é possível observar a confluência de elementos visuais que resultam em aspectos conotativos da imagem, ou seja, aquilo que a imagem pode interpretar. Ficou evidente alguns padrões de repetição aplicados pela inteligência artificial para a construção de imagens representativas das profissões. Ao representarem figuras de profissionais, há a demonstração de certos padrões de gênero, idade, cor de pele, porte físico e trajes. A predominância de figuras masculinas em profissões como cientista, cozinheiro(a), médico(a) e diretor(a) executivo(a) é indicativo de uma construção estereotipada dos profissionais, reproduzido pela Inteligência Artificial, enquanto as profissões de empregado(a) doméstico(a) e enfermeiro(a) são representados por figuras femininas, também englobando aspectos estereotipados de representação, podendo ser caracterizados como manutenções do pensamento advindo do paradigma dominante.

Vale lembrar que os prompts formulados na língua inglesa não possuem especificações de gênero em sua formulação, dada a natureza da língua inglesa não possuir flexões de gênero em substantivos, desta forma, a determinação dos gêneros é exclusivamente atribuída a partir da natureza da inteligência artificial em concentrar padrões iconográficos dentro do universo imagético relacionado às temáticas requeridas.

Outros aspectos conotativos são evidentes a partir do momento em que há uma repetição de elementos estéticos nas imagens. A predominância da pele caucasiana é fruto de um mesmo padrão representativo resultante da construção por parte da inteligência artificial (das vinte imagens geradas, três apresentam indivíduo com pele não caucasiana). Os trajes referentes à cada profissão também demonstram reproduções estéticas geradas a partir da padronização de bancos de imagem associados a cada profissional. Os uniformes ou vestes são padrões conotativos indicativos de status social (no caso dos diretores executivos), sanitariedade (no caso dos empregados(as)

domésticos(as) e enfermeiros(as)), profissionalismo (médicos(as), cozinheiros(as) e cientistas).

Dentro de uma infinidade de possibilidades em que a Inteligência Artificial possui na formulação de imagens, a partir de seu banco de dados, a repetição e padronização são formas determinantes que impactam a construção final de imagens a partir dos prompts exigidos. Desta forma, o resultado oferecido é um reflexo de um discurso imagético preconcebido, decorrente de um conjunto de monoculturas que solidificam aspectos do paradigma dominante.

6 Conclusão

Com base nos resultados obtidos neste trabalho, é possível concluir que as Inteligências Artificiais reproduzem imagens que perpetuam preconceitos e estereótipos, especialmente em relação à representação de gênero, idade, cor de pele, porte físico e trajes em determinadas profissões, proporcionando desafios para a área da Organização do Conhecimento. Esses padrões de representação podem ser caracterizados como manutenções do pensamento advindo do paradigma dominante, que solidifica aspectos culturais e sociais de uma determinada época.

Essas conclusões indicam a necessidade de um olhar crítico em relação às imagens produzidas pelas Inteligências Artificiais, a fim de evitar a perpetuação de estereótipos e preconceitos. É importante serem desenvolvidos algoritmos mais conscientes e inclusivos, que possam promover a diversidade e a igualdade em suas representações imagéticas. Um olhar crítico envolve uma análise cuidadosa das implicações e consequências da utilização da Inteligência Artificial em diferentes contextos. Isso inclui a avaliação de se os dados usados para treiná-la são representativos e diversificados, se as ferramentas estão sendo usadas de forma justa e imparcial, e se os resultados gerados pela Inteligência Artificial são compreensíveis e transparentes.

No que toca à Organização do Conhecimento, ao compreender como funciona a geração de textos e também de imagens no contexto das Inteligências Artificiais, seu papel é aplicar e desenvolver estudos que garantam que as informações sejam representadas de forma justa e inclusiva pela Inteligência Artificial. Como mencionado anteriormente, a Inteligência Artificial pode perpetuar estereótipos e preconceitos, especialmente quando se trata da representação de grupos sociais marginalizados. Nesse sentido, a OC poderia ajudar a minimizar esses riscos, por meio da adoção de práticas que promovam a diversidade e a inclusão nas representações de informações.

Pontua-se que, por ser desenvolvida e treinada por humanos, a Inteligência Artificial não é neutra e reflete

as visões de mundo, valores e preconceitos dos seres humanos. Por isso, é essencial que se tenha um olhar crítico em relação ao uso da Inteligência Artificial, a fim de identificar e corrigir possíveis erros ou vieses que possam prejudicar determinados grupos sociais.

Assim, reconhecendo que a geração das imagens depende de um comando textual em que o usuário tem a possibilidade de explicitar e detalhar sua demanda, a ferramenta pode ser, também, um recurso de quebra do paradigma dominante, visto que grupos tradicionalmente ignorados ou estereotipados podem, sem necessidade de conhecimento técnico artístico, gerarem representações visuais dentro do contexto que desejarem.

Notas

- [1] Trabalho resultado de pesquisas financiadas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).
- [2] UNESCO. Inteligência artificial: entre o mito e a realidade. **O Correio da UNESCO**, n. 3, 2018. Disponível em: <https://pt.unesco.org/courier/2018-3/os-riscos-eticos-da-ia>. Acesso em: 14 abr. 2023.

Referências

- ABBAGNANO, N. **História da filosofia**. V. 5. 4 ed. Trad. Nuno Valadas e Antonio Ramos Rosa. Lisboa: Editorial Presença, 2000.
- BARBOSA, L. M.; PORTES, L. A. F. **A Inteligência Artificial**. Revista Tecnologia Educacional, n. 236, 2023.
- BARTHES, R. **O óbvio e o obtuso**: ensaios críticos III. Rio de Janeiro: Nova Fronteira, 1990.
- DAMACENO, S. S.; VASCONCELOS, R. O. **Inteligência Artificial**: uma breve abordagem sobre seu conceito real e o conhecimento popular. Caderno de Graduação - Ciências Exatas e Tecnológicas - UNIT - SERGIPE, [S. l.], v. 5, n. 1, p. 11, 2018. Disponível em: <https://periodicos.set.edu.br/cadernoexatas/article/view/5729>. Acesso em: 13 abr. 2023.
- MAIMONE, G. D. **A fotografia no contexto da organização do conhecimento**. Pesquisa Brasileira em Ciência da Informação e Biblioteconomia, v. 13, n. 2, 2018. DOI: 10.22478/ufpb.1981-0695.2018v13n2.42490 Acesso em: 14 abr. 2023.
- MANINI, M. P. **Análise documentária de fotografias**: um referencial de leitura de imagens fotográficas para fins documentários. 2002. 231 f. Tese (doutorado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo, 2002.
- MARTINEZ-AVILA, D. **Knowledge Organization in the Intersection with Information Technologies**. Knowledge Organization. Würzburg: Ergon-verlag, v. 42, n. 7, p. 486-498, 2015. Disponível em: <http://hdl.handle.net/11449/160599>. Acesso em: 14 abr. 2023.
- MCCARTHY, J. **What is Artificial Intelligence?** Stanford University. 2007. Disponível em: <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf>
- JAILLANT, L. **Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections**. Bielefeld: Bielefeld University Press, 2022.
- PALETTA, F. C.; GONZALEZ, J. A. M. **Modelos de busca, acesso, recuperação e apropriação da informação na web de dados**. Informação & Informação, [S. l.], v. 24, n. 2, p. 182–210, 2019. DOI: 10.5433/1981-8920.2019v24n2p182. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/35261>. Acesso em: 14 abr. 2023.
- PEIRCE, C. S. **Semiótica**. São Paulo: Perspectiva, 2005.
- PENN, G. **Análise Semiótica de Imagens Paradas**. In: BAUER, M. W.; GASKELL, G. Pesquisa qualitativa com texto, imagem e som: um manual prático. Tradução de Pedrinho A. Guareschi. 7. ed. Petrópolis, RJ: Vozes, 2008.
- ROBERTS, H.E. (2001), “A picture is worth a thousand words: srt indexing in electronic databases”, Journal of the American Society for Information Science and Technology, Vol. 52 No. 11, pp. 911-6.
- RODRIGUES, R. C. **Análise e tematização da imagem fotográfica**. Ciência da Informação, Brasília, v. 36, n. 3, p. 67-76, set./dez. 2007.
- SANTOS, B. **Para além do pensamento abissal**: das linhas globais a uma ecologia de saberes. In: SANTOS, B.; MENESES, M. P. Epistemologias do sul. 2010. p. 637-637.
- SCHAEFFER, J. M. **A imagem precária**: sobre o dispositivo fotográfico. Campinas: Papirus, 1996.
- SCHUHMAN, C. et al. **Laion-5b**: An open large-scale dataset for training next generation image-text models. arXiv preprint arXiv:2210.08402, 2022. Disponível em: <https://arxiv.org/pdf/2210.08402.pdf>. Acesso em 15 de abr. 2023.
- TORRES, A. A. L.; MACULAN, B. C. M. D. S. **Imagens científicas**: organização e representação de imagens para compartilhamento de conhecimento.. Múltiplos Olhares em Ciência da Informação, v. 9 No. 2, n. 2, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/137076>. Acesso em: 14 abr. 2023.

Apêndice A – Imagens geradas a partir dos comandos selecionados



A dimensão aplicada da Organização do Conhecimento no Campo da Inteligência Artificial

The applied dimension of Knowledge Organization in the field of Artificial Intelligence

Rogério Aparecido Sá Ramalho

Universidade Federal de São Carlos – PPGCI UFSCar, ramalho@ufscar.br

Resumo

Ao longo dos últimos anos a Inteligência Artificial tem se destacado como um campo fértil de pesquisas, favorecendo o desenvolvimento de novos métodos, produtos e serviços inovadores que possibilitam a automatização de tarefas inimagináveis há algumas décadas. Alguns autores definem a Inteligência Artificial como um campo universal, devido à grande variedade de aplicações e abordagens potencialmente relevantes para praticamente qualquer esfera de atividade intelectual humana, caracterizando-a como uma das principais tecnologias habilitadoras, capaz de conduzir mudanças radicais na capacidade de usuários e de toda a sociedade, favorecendo a criação de soluções disruptivas e mais eficientes. A presente pesquisa tem como objetivo identificar contribuições da área de Organização do Conhecimento no campo da Inteligência Artificial. A metodologia está pautada em uma pesquisa de cunho teórico e metodológico, de abordagem qualitativa e caráter exploratório. Os resultados indicam que a utilização de técnicas baseadas em Inteligência Artificial se apresenta como uma demanda crescente na estruturação de sistemas de representação, favorecendo a automação de tarefas de organização, classificação e identificação de padrões. Do mesmo modo, a Organização do Conhecimento pode apresentar subsídios teóricos que possibilitem representar de maneira controlada e formal aspectos semânticos fundamentais para o desenvolvimento de novas aplicações no campo da Inteligência Artificial, contribuindo a partir da disseminação de conceitos e vocabulários controlados padronizados. Conclui-se que a hipótese fundamental subjacente ao campo da Inteligência Artificial é a utilização de mecanismos de processamento e representação de informações que possibilitem a realização de inferências automáticas, estando diretamente relacionada com a temática de Sistemas de Organização do Conhecimento que apresenta princípios, teorias e métodos para a representação e organização de conceitos.

Palavras-chave: Inteligência Artificial. Organização do Conhecimento. Representação da Informação. Sistemas de Organização do Conhecimento. SOC.

Abstract

Over the past years, Artificial Intelligence has emerged as a fertile field of research, fostering the development of new methods, products, and innovative services that enable the automation of tasks previously unimaginable a few decades ago. Some authors define AI as a universal field, given its wide range of applications and potentially relevant approaches to virtually any sphere of human intellectual activity. It is regarded as one of the key enabling technologies capable of bringing about radical changes in users' capabilities and society as a whole, promoting the creation of disruptive and more efficient solutions. This research aims to identify the contributions of the Knowledge Organization field to the field of Artificial Intelligence. The methodology is based on a theoretical and methodological study with a qualitative and exploratory approach. The results indicate that the use of AI-based techniques is increasingly demanded in the structuring of representation systems, facilitating the automation of organization tasks, classification, and pattern identification. Likewise, Knowledge Organization can provide theoretical foundations to represent, in a controlled and formal manner, fundamental semantic aspects for the development of new applications in the field of AI, contributing through the dissemination of standardized controlled concepts and vocabularies. It is concluded that the fundamental hypothesis underlying the field of Artificial Intelligence is the use of information processing and representation mechanisms that enable automatic inference, directly related to the theme of Knowledge Organization Systems, which encompasses principles, theories, and methods for the representation and organization of concepts.

Keywords: Artificial Intelligence. Knowledge Organization. Information Representation. Knowledge Organization Systems. KOS.

Resumen

En los últimos años, la Inteligencia Artificial (IA) se ha destacado como un campo fértil de investigación, favoreciendo el desarrollo de nuevos métodos, productos y servicios innovadores que permiten la automatización de tareas impensables hace unas décadas. Algunos autores definen la IA como un campo universal debido a su amplia variedad de aplicaciones y enfoques potencialmente relevantes para prácticamente cualquier esfera de la actividad intelectual humana. Se la considera una de las principales tecnologías habilitadoras capaz de generar cambios radicales en las capacidades de los usuarios y de toda la sociedad, fomentando la creación de soluciones disruptivas y más eficientes. Esta investigación tiene como objetivo identificar las contribuciones del campo de la Organización del Conocimiento en el ámbito de la Inteligencia Artificial. La metodología se basa en un estudio teórico y

metodológico com um enfoque qualitativo y exploratorio. Los resultados indican que el uso de técnicas basadas en Inteligencia Artificial se presenta como una demanda creciente en la estructuración de sistemas de representación, favoreciendo la automatización de tareas de organización, clasificación e identificación de patrones. Del mismo modo, la Organización del Conocimiento puede proporcionar fundamentos teóricos para representar de manera controlada y formal aspectos semánticos fundamentales para el desarrollo de nuevas aplicaciones en el campo de la Inteligencia Artificial, contribuyendo a través de la difusión de conceptos y vocabularios controlados estandarizados. Se concluye que la hipótesis fundamental subyacente al campo de la Inteligencia Artificial es el uso de mecanismos de procesamiento y representación de información que permitan realizar inferencias automáticas, estando directamente relacionada con la temática de los Sistemas de Organización del Conocimiento, que presenta principios, teorías y métodos para la representación y organización de conceptos.

Palabras clave: Inteligencia Artificial. Organización del conocimiento. Representación de la Información. Sistemas de Organización del Conocimiento. SOC.

1 Introdução

O campo da Inteligência Artificial (IA) representa um conjunto de software, lógica, computação e disciplinas filosóficas que visa fazer com que os computadores realizem funções que, há algumas décadas, acreditava-se ser impossível realizá-las sem a interação humana. Sendo contemporânea à Ciência da Informação, teve início após a Segunda Guerra Mundial e, atualmente, abrange uma enorme variedade de subcampos, apresentando-se como um campo universal e potencialmente relevante para qualquer esfera da atividade intelectual humana (RUSSELL; NORVIG, 2004).

A partir da análise de literatura, é possível constatar o enorme potencial de transformação que a IA oferece em praticamente todos os áreas do conhecimento: pesquisas relacionadas às temáticas de gestão da cadeia de suprimentos (CHI; HUANG; GEORGE, 2020), medicina (ALI; SHRESTHA; SOAR; WAMBA, 2018), modelos de negócio (DUAN; EDWARDS; DWIVEDI, 2019), futuro do trabalho (SCHWARTZ *et al.*, 2019), Organização do Conhecimento (MARTÍNEZ-ÁVILA; SAN SEGUNDO, ZURIAN, 2014; GREENBERG; XINTONG; MONSELISE; GRABUS; BOONE, 2021) e até mesmo aprimoramento das capacidades humanas (DWIVEDI *et al.*, 2021). Na área da Ciência da Informação desde o final da década de 1980 (GORDON, 1988; KWOK, 1995; CAPUANO, 2009; FERNEDA 2009) já apontavam possibilidades de aplicações e uso de técnicas de IA para expandir as consultas e aumentar a relevância e qualidade das informações recuperadas.

Contudo, a área de Ciência da Informação também pode contribuir para uma maior disseminação da IA a partir do desenvolvimento de Sistemas de Organização do Conhecimento (KOS), os KOS mais do que nunca desempenham um papel crucial na gestão da informação digital e aplicações em geral, representando aspectos semânticos de forma controlada e documentada que favorecem uma grande variedade funções como: ferramentas para representação e indexação, suporte baseado em conhecimento para pesquisadores de informações, mapas semânticos, comunicação ferramentas que fornecem uma estrutura

conceitual e uma base conceitual para sistemas baseados em conhecimento (GNOLI, 2020).

Historicamente o campo de representação da informação sempre esteve condicionado às tecnologias vigentes, conforme destaca Lima (2015, p. 683-684). Ao longo dos últimos anos o crescimento exponencial das tecnologias digitais tem favorecido novas possibilidades de desenvolvimento de instrumentos e métodos de representação, armazenamento, organização, processamento e recuperação de informações.

Este avanço impactou primeiramente nas atividades relacionadas ao processo de indexação, com o surgimento de sistemas automatizados para criação de índices e processamento automático do texto, com os aportes da linguística, no processo de catalogação, com a criação do formato MARC, pela Library of Congress (LC), que passa a ser utilizado efetivamente em 1966, e que foi considerado pela International Standard Organization (ISO) o primeiro projeto de automação internacional a estabelecer normas de descrição bibliográfica em forma mecanicamente legível, possibilitando o intercâmbio de dados. (LIMA, 2015, p. 683-684)

Do mesmo modo, por meio da incorporação das tecnologias computacionais nos processos de classificação tornou-se possível a compatibilização de vocabulários controlados com auxílio do computador, favorecendo a integração de tesouros em Sistemas de Recuperação de Informação e consequente evolução das linguagens documentárias (ALVITE DÍEZ *et al.*, 2010). Martínez-Ávila (2015) também destaca a necessidade de uma maior interseção entre a Organização do Conhecimento (OC) e as Tecnologias da Informação (TI), fornecendo novas perspectivas e oportunidades para melhorar a recuperação e o acesso às informações.

Contudo, apesar do crescente interesse pelas inter-relações existentes entre a área de Ciência da Informação e o campo da Inteligência Artificial, verifica-se ainda um número limitado de pesquisas que abordam a dimensão aplicada da Organização do Conhecimento no campo da Inteligência Artificial, favorecendo (re)pensar os instrumentos clássicos de Organização e Representação a partir de novas perspectivas de automação de sistemas, contribuindo

para atender as demandas de uma sociedade cada vez mais caracterizada por interações digitais autônomas.

O presente trabalho apresenta resultados parciais de uma pesquisa de cunho teórico e metodológico, com abordagem qualitativa e caráter exploratório, que objetiva identificar contribuições da área de Organização do Conhecimento no campo da Inteligência Artificial. Caracteriza-se como parte de um projeto de pesquisa apoiado pelo CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil.

2 Sistemas de Organização do Conhecimento como instrumentos de apoio à Inovação

Machado, Lehmann e Araújo (2008) caracterizam inovação como: a capacidade de uma organização de articular e mobilizar recursos (humanos, financeiros, materiais etc.), para captar oportunidades e neutralizar ameaças (operacionais, mercadológicas etc.); e o processo de transformação de ideias (inovação) em soluções que atendam às necessidades existentes ou latentes; e o resultado de uma atividade criativa que resulta na forma de produtos, processos e sistemas, novos ou modificados.

A inovação, de acordo com o Manual de Oslo, é caracterizada como um processo complexo e a gama de atividades exigida para a inovação em uma organização pode variar consideravelmente. O desenvolvimento de novos produtos e serviços inovadores dentro da própria organização, contribui para o aprimoramento de processos, a partir da incorporação de tarefas automatizadas (OCDE, 2004).

Melhado (2005) destaca que a Ciência da Informação, ao estudar propriedades da informação e processos de sua construção, comunicação e uso, possibilita uma melhor compreensão de problemas sociais advindos da necessidade de informações, do surgimento da indústria da informação e do advento de novas tecnologias que estabelecem um paradigma de manipulação associado à informação.

Ao refletir sobre a definição de informação González de Gomez (2000) afirma tratar-se de um termo flutuante que, tal como o de democracia, produz diferentes efeitos de sentido em diferentes contextos. Informação, como construção e sintoma da sociedade contemporânea, designa algo que se desdobra em vários planos de realização – todos eles transcendidos por valores semânticos e quadros normativos, da ordem do ideacional, e ancorados no solo de sua inscrição ou manifestação, da ordem do visível e do temporal.

Nesta perspectiva, Schumpeter (1961) afirma que são cinco os tipos básicos de inovação: 1) desenvolvimento de um novo produto, ou de uma nova tecnologia de um bem já existente; 2) desenvolvimento de um método de

produção, ou de uma nova logística comercial; 3) desenvolvimento de um novo mercado; 4) desenvolvimento de novas fontes de suprimento das matérias-primas ou produtos semi-industrializados; 5) desenvolvimento de uma nova organização industrial, como a criação ou a fragmentação de uma posição de monopólio.

Dosi (1982) destaca que a inovação possui quatro propriedades fundamentais: (i) incerteza; (ii) crescente dependência das novas oportunidades tecnológicas no conhecimento científico; (iii) crescente formalização das atividades de pesquisa e desenvolvimento e, (iv) learn-by-doing (aprender fazendo). Pode-se analisar que nas inovações ocorrem incertezas quando existem problemas técnico-econômicos, onde as soluções são desconhecidas e deve-se atuar para diminuir estes problemas.

Assim, torna-se evidente a relação entre Ciência da Informação e os processos de Inovação, conforme já destacado nos estudos de Cervantes (2004) quanto propõe um estudo terminológico para a construção de vocabulário para a temática de Inteligência Competitiva.

No âmbito da Ciência da Informação as Liguagens Documentarias caracterizam como o instrumento mais utilizado para o controle de vocabulários, são sistemas de signos que visam a uniformização da linguagem utilizada em um Sistema de Informação, proporcionando uma representação padronizada do conteúdo informacional. Também denominadas como sistemas de organização do conhecimento, têm como função principal representar o conteúdo dos documentos favorecendo uma recuperação da informação mais pertinente (BOCCATO, 2008; SALLES 2007).

Segundo Hodge (2000) o termo Knowledge Organization Systems (KOS) foi proposto em 1998 pelo Networked Knowledge Organization Systems Working Group para englobar sistemas de classificação, cabeçalhos de assunto, arquivos de autoridade, redes semânticas e ontologias. Na taxonomia de tipos de KOS proposta por Hodge (2000) também são incluídos, entre outros, dicionários, glossários, taxonomias e tesouros.

Conforme destaca Lara (2015, p. 92) “no Brasil, não há consenso sobre a utilização de um termo que abranja o conjunto de instrumentos de organização da informação e do conhecimento”. Apesar das variações de denominação identificadas na literatura nacional para designar tais instrumentos, observa-se um crescente aumento no número de pesquisas relacionadas à esta temática.

A partir da popularização do ambiente Web, novos instrumentos de representação organização e representação têm sido propostos, provocando convergências entre aportes teóricos e metodológicos de diversos campos científicos, no intuito de atender as

demandas informacionais de uma Sociedade cada vez mais baseada em interações virtuais (RAMALHO, 2010).

Na última década foi constatado um crescente interesse de pesquisadores no estudo e desenvolvimento de ontologias, envolvendo pesquisas empenhadas principalmente na consolidação das definições de tal instrumento de representação e análise de possibilidades efetivas de uso (SANTOS; CORREA; SILVEIRA, 2013).

García-Marco (2007) ressalta que o crescente aumento do número de pesquisas relacionadas a temática de ontologias na área de Ciência da Informação constitui a “ponta do iceberg” de um processo de reconfiguração disciplinar, como resultado da integração de diversas Ciências.

No âmbito da área de Ciência da Informação o termo ontologia passou a ser utilizado no final da década de 1990, quando instrumentos de representação de informações despertaram maior interesse dos pesquisadores das mais diversas áreas do conhecimento, impulsionados pela criação do ambiente Web e desenvolvimento de novas tecnologias digitais (SOERGEL, 1999; VICKERY, 1997).

Entre as definições de ontologia mais referenciadas na literatura destaca-se a apresentada por Gruber (1993, p. 1) de que “é uma especificação explícita de uma conceitualização”. Contudo, tal definição remete o conceito de ontologia a aspectos filosóficos, favorecendo a geração de conflitos terminológicos, conforme destaca Guarino (1998, p. 5):

No sentido filosófico, podemos nos referir a uma "ontologia" como um sistema particular de categorias que versa sobre uma certa visão do mundo. Desta forma, este sistema não depende de uma linguagem particular: a ontologia de Aristóteles é sempre a mesma, independente da linguagem usada para descrevê-la. Por outro lado, em seu uso mais prevalente na IA [inteligência artificial], uma ontologia é referida como um artefato de engenharia, constituído de um vocabulário específico usado para descrever uma certa realidade e um conjunto de pressupostos explícitos relacionados com o significado pretendido para as palavras do vocabulário.

Buscando aproximar o conceito de ontologia aos pressupostos da área de Ciência da Informação, principalmente em relação aos processos de representação, tratamento, organização e recuperação de recursos informacionais, Ramalho (2010, p. 107) define ontologia como

[...] um artefato tecnológico que possibilita representar formalmente as propriedades e relacionamentos de um determinado modelo conceitual, favorecendo a utilização de inferências automáticas nos processos de organização e recuperação de recursos informacionais.

Tal definição caracteriza as ontologias no âmbito da área de Ciência da Informação como instrumentos de

nível epistemológico, concebidos a partir de linguagens computacionais com o intuito de favorecer a representação formal dos relacionamentos existentes entre os conceitos de um domínio específico.

As ontologias potencializam a representação dos relacionamentos existentes entre os conceitos que traduzem os conteúdos informacionais de modo a possibilitar representações semânticas que não poderiam ser obtidas a partir de descrições textuais, caracterizando-se como um tipo de fonte secundária, instituindo novas formas de acesso e favorecendo maior flexibilidade na recuperação de informações, por meio da realização de inferências. Quanto ao enfoque das representações semânticas, Zafalon (2013) esclarece que “é a semântica que dá conta dos processos mentais segundo os quais se produz, constitui, compreende e descreve a representação de um recurso informacional.”

Na concepção das linguagens documentárias o tesouro é o instrumento que possui maior familiaridade e relacionamento com as ontologias por serem linguagens de estruturas combinatórias, de caráter especializado, constituídos por termos providos de suas relações semânticas que possibilitam a representação temática do conteúdo de um documento, bem como sua posterior recuperação (BOCCATO; RAMALHO; FUJITA, 2008, p. 200)

Conforme afirma García Jiménez (2004), apesar das similaridades existentes entre tesouros e ontologias, ambos os instrumentos de representação possuem vinculações operativas e teóricas distintas, sendo desenvolvidos a partir de diferentes trajetórias históricas.

Sales, Campos e Gomes (2008, p. 67) destacam que outro fator determinante para a distinção das ontologias e os modelos de representação tradicionalmente utilizados no âmbito da área de Ciência da Informação é a própria natureza dos relacionamentos utilizados, ao afirmar que, “a literatura da Ciência da Informação apresenta apenas relações entre categorias – Relações Diádicas, mas as ontologias necessitam de relações mais formalizadas – Relações Triádicas”.

Tal distinção ressalta os diferentes propósitos dos modelos clássicos de representação em relação ao objetivo mais complexo das ontologias que buscam favorecer a realização de inferências automáticas, sendo necessário além de descrever formalmente, rotular e categorizar as próprias relações expressas na ontologia (RAMALHO, 2015a).

Ao longo dos últimos anos inúmeras metodologias têm sido propostas para o desenvolvimento de ontologias. Dentre elas, destacam-se a de Uschold e King (1995), a de Uschold e Gruninguer (1996), a de Fernández, Gómez-Pérez e Juristo (1997), e o método 101, proposto por Noy e McGuinness (2001). No âmbito da área de Ciência da Informação pode-se destacar o método proposto por Ramalho (2010) para o

desenvolvimento de ontologias a teoria do conceito e a teoria da classificação facetada utilizadas na elaboração de tesouros (CAMPOS; GOMES, 2006; GOMES *et al.*, 1990), bem como a norma ANSI/NISO Z39.19-2005, que apresenta diretrizes para a elaboração de vocabulários controlados.

Apesar da difusão de linguagens e ferramentas para a representação e construção de ontologias, observa-se ainda a necessidade de estudos que contemplem diretrizes mais eficientes para a identificação dos conceitos e relacionamentos existentes entre estes, favorecendo um melhor direcionamento para a construção de ontologias de qualidade (FERNANDEZ; GOMEZ-PEREZ; JURISTO, 1997; DING, 2002a; 2002b; FERNÁNDEZ-LOPES, 2001; GARCÍA-JIMÉNEZ, 2004, PASLARU-BONTAS, 2007; CAMPOS, 2010)

Construir ontologias compatíveis com as normas já existentes e hábitos da comunidade, ao mesmo tempo em que possibilite usufruir das capacidades oferecidas pelas novas tecnologias semânticas, apresenta-se como um dos principais desafios contemporâneos no campo da Ciência da Informação (DABROWSKI; SYNAK; KRUK, 2009).

Lima (2015) aponta entre as principais necessidades atuais das pesquisas na área de Ciência da Informação a necessidade de estudos sobre possíveis soluções para a interoperabilidade, visando ao compartilhamento de recursos informacionais na Web e a compatibilização entre diferentes sistemas de representação.

Assim, conforme apresenta Pastor-Sanchez, Martinez-Rodriguez e Mendez-Muñoz (2009) a exploração dos tesouros na Web é muito limitada, as diferenças na forma como recuperação de informações na Internet exigem vocabulários interoperáveis e multilíngues. Tais autores insistem que o próprio conceito de tesouro tem evoluído para adaptar-se aos novos modelos de representação da Web, abandonando o paradigma léxico em favor de um paradigma conceitual. Segundo Sanchez-Jimenez e Gil-Urdiciain (2007) esta adaptação permite superar as deficiências das relações fortemente ancoradas ao nível léxico que caracterizam os tesouros tradicionais.

Clarke e Zeng (2012) demonstram que os princípios para a construção de tesouros mudaram radicalmente da norma ISO 2788 para a ISO 25964, publicada em 2011, tais mudanças podem ser justificadas devido ao avanço das tecnologias digitais que demandam uma necessidade de interoperabilidade semântica, onde a distinção entre termo e conceito é fundamental, contribuição importante implementada na nova norma.

No ano de 2009 o World Wide Web Consortium (W3C) formalizou como Recomendação o padrão Simple Knowledge Organization System – SKOS (Sistema de Organização do Conhecimento Simples) como um

modelo de dados para a representação de Sistemas de Organização do Conhecimento no ambiente Web, buscando atender às novas demandas informacionais relacionadas ao compartilhamento e interoperabilidade de vocabulários no ambiente Web (RAMALHO, 2015b).

O modelo SKOS proporciona um modelo para a representação formal da estrutura básica e conteúdo de Sistemas de Organização: como tesouros, esquemas de classificação, listas de cabeçalho de assunto, taxonomias, entre outros tipos similares de vocabulários controlados (RAMALHO; VIDOTTI; FUJITA, 2007). Sua codificação é baseada no padrão RDF (Resource Description Framework), favorecendo a representação de informações de maneira estruturada, assim como a integração de diferentes esquemas conceituais.

Conforme relatam Pastor-Sanchez, Martínez-Mendez, e Rodriguez-Muñoz (2012) o desenvolvimento do modelo SKOS começou em 2002, com as atividades do grupo de trabalho Semantic Web Advanced Development for Europe (SWAD-E) e foi difundido publicamente a partir de novembro de 2005, com a publicação do documento SKOS Core Guide (MILES; BRICKLEY, 2005). Apenas em agosto de 2009 com a publicação do SKOS Simple Knowledge Organization System Reference (MILES; BECHHOFFER, 2009) SKOS passou ser considerado como uma recomendação oficial do W3C como modelo de dados para a disponibilização e compartilhamento de sistemas de organização do conhecimento na Web.

Tendo como principal objetivo possibilitar a representação de vocabulários no ambiente Web como base em um modelo simplificado, o SKOS não tem a função de substituir vocabulários controlados em seu contexto original de uso, e sim favorecer uma maior reutilização e interoperabilidade entre os vocabulários existentes (RAMALHO, 2015a). Segundo Isaac e Summer (2009) o padrão SKOS possibilita que conceitos possam ser identificados utilizando Uniform Resource Identifiers (URI), rotulados com sequências textuais em uma ou mais línguas naturais, documentados por meio de diferentes tipos de notas, relacionados semanticamente entre si, a partir de hierarquias informais e redes associativas; e agregados em diferentes esquemas conceituais.

O modelo SKOS foi projetado como uma alternativa de baixo custo que permita a transição de Sistemas de Organização do Conhecimento tradicionais para o ambiente Web, favorecendo a representação de aspectos semânticos dos recursos informacionais. Podendo ser utilizado de forma isolada ou em combinação com linguagens formais de representação do conhecimento, como OWL (Web Ontology Language), SKOS fornece um elo entre o formalismo lógico rigoroso das linguagens de ontologias e o

mundo caótico, informal e pouco estruturado de ferramentas de colaboração baseadas na Web.

Conforme destaca Mustafa El Hadi (2015) desde a padronização do modelo SKOS muitos vocabulários passaram a ser protegidos por direitos autorais para serem publicados como dados ligados na Web, proporcionando uma dimensão semântica praticamente inesplorada anteriormente.

Dextre Clarke e Zeng (2012, p. 23) destacam a importância do padrão SKOS

As Recomendações SKOS do W3C são projetadas para suportar publicações de vocabulários, como tesouros na Web. E na sua essência é um modelo de dados que distingue expressamente entre os conceitos e etiquetas para representar conceitos. (...) Além disso, que é veiculado pelo SKOS Core para as relações entre conceitos, a extensão fornece suporte adicional para identificar, descrever e linkar entidades lexicais.

Broughton (2010) destaca a proximidade entre SKOS e classificações facetadas, ressaltando que algumas aplicações da Web Semântica apresentam semelhanças com os tradicionais sistemas de organização facetados, de modo que a teoria da classificação facetada pode fornecer um modelo para os elementos utilizados nos sistemas de representação baseados tecnologias semânticas, bem como terminologias para habilidá-los.

Mustafa El Hadi (2015) também contextualiza que até o momento as pesquisas relacionadas ao desenvolvimento de novos sistemas de representação têm se focado em questões de interoperabilidade técnica e semântica, porém existe uma falta de um arcabouço teórico geral para considerar estes desenvolvimentos.

A pesquisa em questão se propõe a identificar e analisar teorias e métodos que possibilitem fornecer um arcabouço teórico a partir do qual seja possível representar um microtesouro na temática de Inteligência Artificial Tecnologias da Informação e Inovação em formato SKOS, contribuindo para a sistematização de conceitos e identificação das contribuições da Ciência da Informação nos processos de inovação. Tal proposição vai ao encontro da necessidade crescente de sistematização de fundamentos teóricos que possibilitem validar o desenvolvimento de novas tecnologias, favorecendo que os avanços tecnológicos possam ser devidamente sedimentados no campo teórico.

Considerando o evidente caráter interdisciplinar que delinea o corpus teórico das pesquisas relacionadas a representação e organização de informações em ambientes digitais, destaca-se a relevância de pesquisas desta natureza, que buscam contribuir para uma maior familiarização dos profissionais da informação com as novas tecnologias, para que as mesmas também possam ser desenvolvidas a partir de princípios éticos sociais e não baseadas única e exclusivamente em

conhecimentos e processos puramente técnicos, pois só assim os profissionais da informação estarão realmente assumindo e desempenhando seu verdadeiro papel como agentes sociais.

Neste sentido, concordamos com Mustafa El Hadi (2015) quando afirma que o padrão SKOS inaugura uma nova etapa de sistemas de representação, capazes de proporcionar uma dimensão semântica praticamente inesplorada em ambientes digitais. SKOS favorece uma aproximação entre instrumentos e métodos de construção de vocabulários controlados, utilizados tradicionalmente no campo da Ciência da Informação, com as novas potencialidades oferecidas pela IA.

Conforme destaca a Portaria do MCTIC Nº 1.122, de 19 de março de 2020, que define as prioridades para os projetos de pesquisa, de desenvolvimento de tecnologias e inovações para o período 2020 a 2023, a IA caracteriza-se como uma área de vanguarda com impactos produtivos, econômicos, sociais e de sustentabilidade, por isso favorecer o desenvolvimento de um KOS, fundamentado nas teorias e métodos da Ciência da Informação, seguramente é uma abordagem capaz de favorecer inúmeros impactos positivos. Do mesmo modo, o presente projeto também favorece uma maior aproximação com o setor produtivo, indo ao encontro das atualizações da legislação nacional. Em 2016, a Lei de Inovação foi alterada com a promulgação da Lei 13.243/2016 (Brasil, 2016), denominada como Marco Legal da Ciência, Tecnologia e Inovação, com o objetivo de desburocratizar as atividades de pesquisa e inovação no Brasil. Em fevereiro de 2018, foi publicado o Decreto 9.283 (BRASIL, 2018), que regulamentou as Leis n. 10.973/2004 - Inovação (BRASIL, 2004) e a n. 13.243/2016 - Marco da CT&I (BRASIL, 2016), estabelecendo medidas de incentivo à inovação e à pesquisa científica e tecnológica no ambiente produtivo, com vistas à capacitação tecnológica, ao alcance da autonomia tecnológica e ao desenvolvimento do sistema produtivo nacional e regional.

Assim, destacamos a necessidade de esforços interdisciplinares que, respeitando as especificidades de cada área do conhecimento, possam contribuir na evolução do conhecimento humano de forma integral. Deste modo, fundamentando-se nas teorias e métodos identificados na área de Organização do Conhecimento torna-se evidente a necessidade de um maior aprofundamento na dimensão aplicada da Organização do Conhecimento no campo da Inteligência Artificial

3 Considerações Finais

Corroborando com a afirmação de Qin (2020), quando realiza uma análise sobre Organização do Conhecimento e Representação sob a perspectiva da

AI, é evidente que as práticas de Knowledge Organization (KO) que fundamentam a área de Organização do Conhecimento e a perspectiva aplicada de Knowledge Representation (KR) possuem pontos de convergência, apesar de possuírem diferentes vinculações teóricas.

Assim, conclui-se que a hipótese fundamental subjacente ao campo da IA é a utilização de mecanismos de processamento e representação de informações que possibilitem a realização de inferências automáticas, estando diretamente relacionada com a temática de Sistemas de Organização do Conhecimento (SOC) que apresenta princípios, teorias e métodos para a representação e organização dos conceitos de um domínio. Assim, constatou-se que a utilização de técnicas baseadas em Inteligência Artificial se apresenta como uma demanda crescente na OC, favorecendo a automação de tarefas de organização, classificação e identificação de padrões. Deste modo, verifica-se a necessidade de estudos que possam contribuir para um melhor entendimento das técnicas de IA e suas aplicações no ambiente da OC, favorecendo o surgimento de novas interfaces de investigação eminentemente interdisciplinares, capaz de engendrar uma nova constelação epistemológica, apontando perspectivas inovadoras e novas possibilidades no fazer profissional da área.

Referencias

- ALI, O.; SHRESTHA, A.; SOAR, J.; WAMBA, S. F. Cloud computing-enabled healthcare opportunities, issues, and applications: A systematic review. **International Journal of Information Management**, 43, 146–158. 2018.
- ALVITE DÍEZ, M. L. *et al.* Propuesta de representación del tesauro EuroVoc en SKOS para su integración en sistemas de información jurídica. **Scire: Representación y Organización del Conocimiento**, n.16, v.2, p. 47-51, 2010.
- ANSI/NISO. **Z39.19-2005**: Guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005.
- BOCCATO, V. R. C. A linguagem documentária como instrumento de organização e recuperação da informação. *In*: HOFFMANN, W. A. M.; FURNIVAL, A. C. (Orgs.) **Olhar: ciência, tecnologia e sociedade**. São Paulo: Pedro e João Ed., CECH-UFSCar, p. 269-78, 2008.
- BOCCATO, V. R. C., RAMALHO, R. A. S., FUJITA, M. S. L. A contribuição dos tesauros na construção de ontologias como instrumento de organização e recuperação da informação em ambientes digitais. *In*: GARCIA MARCO, F. J. (Ed.). **Avances y perspectivas en sistemas de información y documentación – IBERSID**, 2008. Zaragoza: Universidad de Zaragoza, 2008, p. 199-209.
- BRASIL. **Lei no 10.973, de 02 de dezembro de 2004**. Dispõe sobre incentivos à inovação e à pesquisa científica e tecnológica no ambiente produtivo e dá outras providências. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/110.973.htm>. Acesso em abril de 2019.
- BRASIL. **Lei no 12.504, de 11 de outubro de 2011**. Confere ao Município de São Carlos, no Estado de São Paulo, o título de Capital Nacional da Tecnologia. Disponível em: <http://www.planalto.gov.br/CCIVIL_03/_Ato2011-2014/2011/Lei/L12504.htm> Acesso em abril de 2019.
- BRASIL. **Decreto Nº 9.283, de 7 de fevereiro 2018**. Regulamenta Novo Marco Legal da Inovação. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/decreto/d9283.htm>.
- BRASIL. **Lei nº 9.279 de 14 de maio de 1996**. Regula direitos e obrigações relativos à propriedade industrial. Disponível em: <http://www.planalto.gov.br/CCIVIL_03/Leis/L9279.htm>. Acesso em: 05 ago.
- BRASIL. Ministério da Ciência, Tecnologia e Inovação. **Portaria nº 293, de 19º de março de 2020**. Define as prioridades, no âmbito do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC). Disponível em <https://www.in.gov.br/en/web/dou/-/portaria-n-1.122-de-19-de-marco-de-2020-249437397>.
- BROUGHTON, V. Emergent vocabulary control in Web 2.0: Comparisons with conventional LIS theory and practice. **Les Cahiers du numérique**, v. 6, n. 3, p. 49-75, 2010. Disponível em: <https://www.cairn.info/revue--2010-3-page-49.htm>. Acesso em: 28 maio 2023.
- CAMPOS, M. L. A. O papel das definições na pesquisa em ontologia. **Perspectivas em Ciência da Informação**, v. 15, p. 10-20, 2010.
- CAMPOS, M. L. A.; GOMES, H. E. Metodologia de elaboração de tesauro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 3, set/dez. 2006. Disponível em: <http://www.eci.ufmg.br/pcionline/index.php/pci/article/viewFile/273>. Acesso em: 20 jun. 2014.
- CAPUANO, E. A. Inteligência competitiva e suas conexões epistemológicas com gestão da informação e do conhecimento. **Ciência da Informação**, Brasília, v. 38, n. 2, p. 19-34, 2009.
- CERVANTES, B. M. N. **Contribuição para a terminologia do processo de inteligência competitiva**: estudo teórico e metodológico. Marília, 2004. 183 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade do Estado de São Paulo, Campus de Marília, 2004.
- CHI, M., HUANG, R.; GEORGE, J. F. Collaboration in demand-driven supply chain: Based on a perspective of governance and IT-business strategic alignment. **International Journal of Information Management**, 52. 2020.
- DABROWSKI, M; SYNAK, M.; KRUK, S. R. Bibliographic Ontology. *In*: KRUK; MCDANIEL (Eds.). **Semantic Digital Libraries**. Berlin: Springer, 2009.
- DEXTRE CLARKE, S. ISO 25964: a standard in support of KOS interoperability, in: A. Gilchrist; J. Vernau (eds), *Facets*

- of Knowledge Organization, 4-5 July 2011, London. 2011. London, Emerald.
- DEXTRE CLARKE, S., ZENG, M.L. **From ISO 2788 to ISO 25964: The Evolution of Thesaurus Standards towards Interoperability and Data Modeling**, in Information Standards Quality, Special Edition: Year in Review and State of the Art of the Standards, p. 20-26, 2012.
- DING, Y. Ontology research and development. Part 2 - a review of ontology generation. **Journal of Information Science**, v.28, n.5, p. 375-388, 2002.
- DING, Y.; FOO, S. Ontology research and development. Part 1 - a review of ontology mapping and evolving. **Journal of Information Science**, v. 28, n. 2, p. 123-136, 2002.
- DOSI, G. **The nature of the innovative process**. Technical change and economic theory, London, Pinter. 1982.
- DUAN, Y.; EDWARDS, J. S.; DWIVEDI, Y. K. Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. **International Journal of Information Management**, 48, 63–71. 2019.
- DWIVEDI, Y. K.; HUGHES, L.; ISMAGILOVA, E.; AARTS, G.; COOMBS, C.; CRICK, T.; MEDAGLIA, R. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. **International Journal of Information Management**, 57. 2021.
- FÉRNANDEZ, M.; GÓMEZ-PÉREZ, A.; JURISTO, N. Methontology: From Ontological Art Towards Ontological Engineering. **Palo Alto: Association for the Advancement of Artificial Intelligence**, 1997. AAAI Technical Report. Disponível em: http://oa.upm.es/5484/1/METHONTOLOGY_.pdf. Acesso em: 20 jun. 2015.
- FERNANDEZ-LOPES, M. Overview of methodologies for building ontologies. In: WORKSHOP ON ONTOLOGIES AND PROBLEM-SOLVING METHODS: LESSONS LEARNED AND FUTURE TRENDS, 1999. Stockholm, Sweden. **Proceedings** [...]. Stockholm, Sweden: Facultad de Informática (UPM), p. 26-34, 2001.
- FERNEDA, E. Applying Genetic Algorithms in Information Retrieval. **DataGramZero**, Rio de Janeiro, v. 10, n. 1, fev. 2009. .
- GARCIA-JIMÉNEZ, A. Instrumentos de representación del conocimiento: tesauros versus ontologías. **Anales de documentación**, Mucia, n. 7, p. 79-95, 2004. Disponível em: <http://www.um.es/fccd/anales/ad07/ad0706.pdf>. Acesso em: 10 jan 2008.
- GARCÍA-MARCO, F. J. **Ontologías y organización del conocimiento**: retos y oportunidades para el profesional de la información. *El profesional de la información*, v. 16, n. 6, p. 541-550. 2007.
- GREENBERG, J., XINTONG Z., MONSELISE M., GRABUS S, BOONE. J. Knowledge Organization Systems: A Network for AI with Helping Interdisciplinary Vocabulary Engineering. **Cataloging & Classification Quarterly**, v. 59, n. 8, 2021. .
- GNOLI, C. Introduction to Knowledge Organization. Facet Publishing: London, 2020.
- GOMES, H. E. **Manual de elaboração de tesauros monolíngues**. Brasília: CNPq/PNUB, 1990.
- GONZÁLEZ DE GOMEZ, M. N. O caráter seletivo das ações de informação. **Informare**, Rio de Janeiro, v. 5, n. 2, p. 7-31, 2000.
- GORDON, M. Probabilistic and genetic algorithms for document retrieval. **Communications of the ACM**, New York, v. 31, n. 10, p. 1208-18, 1988.
- GUARINO, N. Formal Ontology and Information Systems. [Vilnius]: Vilnius University Faculty of Mathematics and Informatics, 1998. Disponível em: <http://uosis.mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKonceptinisModeliavimas/papildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf>. Acesso em: 26 jul. 2014. .
- HODGE, G. **Systems of knowledge organization for digital libraries: beyond traditional authority files**. Washington, D.C.: The Digital Library Federation Council on Library Information Resources, 2000. Disponível em: <http://old.diglib.org/pubs/dlf090/dlf090.pdf> Acesso em dez. 2013.
- ISAAC, A.; SUMMERS, E. SKOS Simple Knowledge Organization System Primer. **W3C Working Group Note 18** August 2009. Disponível em: <http://www.w3.org/TR/skos-primer/> Acesso em: fev. 2015.
- LARA, M. L. G. Propostas de tipologias de KOS: uma análise das referências de formas dominantes de organização do conhecimento. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 20, n. esp. 1, Fev. 2015. Disponível em: <http://dx.doi.org/10.5007/1518-2924.2015v20nesp1p89> Acesso em jul. 2015.
- LIMA, G. A. B. O. Organização do conhecimento: pesquisa e desenvolvimento. In: GUIMARÃES, J. A. C.; DODEBEI, V. (Org.). **Organização do conhecimento e diversidade cultural**. 1ed. Marília: ISKO-Brasil; FUNDEPE, v. 1, p. 670-687, 2015.
- MACHADO, D. Del P. N.; LEHMANN, C. Arndt; ARAUJO, B. de. Organização e cultura de inovação: elementos concretos e fatores percebidos. **Revista Alcance**, v. 15, n. 2, p. 152-168, mai/ago. 2008. .
- MARTÍNEZ-ÁVILA, D.; SAN SEGUNDO, R.; ZURIAN, F. A. Retos y oportunidades en organización del conocimiento en la intersección con las tecnologías de la informacional. **Revista Española de Documentación Científica**, v. 37, n. 3, p. e053, 2014. .
- MARTÍNEZ-ÁVILA, D. Knowledge Organization in the intersection with information technologies. *Knowledge Organization*, v. 42, n. 07, p. 486-498, 2015.
- MELHADO, J. P. **Disseminação e proteção de informações no processo de inovação tecnológica um estudo do contexto regulatório aplicado ao caso brasileiro**. 2005. Dissertação (Mestrado em Ciência da Informação) - Pontifícia Universidade Católica. Campinas, 2005.
- MILES, A.; BECHHOFER, S. **SKOS Simple Knowledge Organization System Reference**. W3C Recommendation, 2009. Disponível em: <http://www.w3.org/TR/skos-reference/>. Acesso em: 20 jun 2015.

- MILES, A.; BRICKLEY, D. **SKOS Core Guide**. W3C: 2005. Disponível em: <http://www.w3.org/TR/2005/WD-swp-skos-core-guide-20051102/>. Acesso em: jun 2015.
- MUSTAFA EL HADI, W. Cultural Interoperability and Knowledge Organization Systems. *In*: José Augusto Chaves Guimarães; Vera Dodebei. (Org.). **Organização do conhecimento e diversidade cultural**. 1.ed. Marília: ISKO-Brasil ; FUNDEPE, v. 1, p. 575-606, 2015.
- NOY, N. F.; MCGUINNESS, D. L. **Ontology development 101: a guide to creating your first ontology**. Stanford: Standford University, [2001?]. Disponível em: <http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html>. Acesso em: 20 fev. 2014.
- OCDE. **Manual de Oslo**: proposta de diretrizes para coleta e interpretação de dados sobre inovação tecnológica. Tradução da Financiadora de Estudos e Projetos. Paris: OCDE, 2004.
- PASLARU-BONTAS, E. **A contextual approach to ontology reuse: methodology, methods and tools for the semantic web**. 2007. Dissertation (Fachbereich Mathematik und Informatik) - Freien Universit"at Berlin, 2007.
- PASTOR-SANCHEZ, J. A.; MARTÍNEZ-MENDEZ, F. J.; RODRÍGUEZ-MUÑOZ, J. V. Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives. **Information Research**, v.14, n.4, 2009.
- QIN, J. Knowledge organization and representation under the AI lens. **Journal of Data and Information Science**, v.5, n.1, p. 3–17. 2020.
- RAMALHO, R. A. S. **Desenvolvimento e utilização de ontologias em bibliotecas digitais: uma proposta de aplicação**. 2010. 145f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012.
- RAMALHO, R. A. S.; FUJITA, M. S. L. Aplicabilidad de ontologías en bibliotecas digitales. **Anales de Documentación**, v. 14, n. 1, 2011. Disponível em: <http://revistas.um.es/analesdoc/article/view/120161>. Acesso em: 12 mar. 2014.
- RAMALHO, R.A.S. Análise do Modelo de Dados SKOS: Sistema de Organização do Conhecimento Simples para a Web. **Informação & Tecnologia (Itec)**, v. 2, p. 66-79, 2015a. .
- RAMALHO, R.A.S. Ontologias e Simple Knowledge Organization System (SKOS): aproximações e diferenças. *In*: José Augusto Chaves Guimarães; Vera Dodebei. (Org.). **Organização do conhecimento e diversidade cultural**. 1.ed. Marília: ISKO-Brasil; FUNDEPE, v. 1, p. 100-107, 2015b.
- RAMALHO, R. A. S.; VIDOTTI, S.A.B.G; FUJITA, M.S.L. Web semântica: uma investigação sob o olhar da Ciência da Informação. **DataGramZero - Revista de Ciência da Informação** - vol.8, n.6, dez/2007.
- RUSSEL, S.; NORVIG, Peter. **Inteligência Artificial**. 2. Ed. Rio de Janeiro: Campos, 2004.
- SALES, L.F.; CAMPOS, M. L. A.; GOMES, H. E. Ontologia de domínio: um estudo das relações conceituais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 13, n. 2, p. 62-76, maio/ago. 2008.
- SANCHEZ-JIMÉNEZ, R.; GIL-URDICIAIN, B. Lenguajes documentales y ontologías. **El Profesional de la Información**, v.16, n.6, p. 551-560, 2007. .
- SANTOS, M. T.; CORREA, R. F.; SILVEIRA, M. A. A. Estudos brasileiros sobre ontologia na Ciência da Informação. **Datagramazero**, v. 14, n. 1, 2013. Disponível em: http://www.dgz.org.br/fev13/Art_05.htm. Acesso em: 20 jun. 2014.
- SCHUMPETER, J. Capitalismo, Socialismo e Democracia. Tradução de Ruy Jungmann. Rio de Janeiro: Ed. Fundo de Cultura, 1961.
- SCHWARTZ, R.; DODGE, J.; SMITH, N.; ETZIONI, O. Green AI. **Communications of the ACM**, v. 63, n. 12, p. 54-63, 2019.
- SOERGEL, D. The rise of ontologies or the reinvention of classification. **Journal of the American Society for Information Science**, v. 50, n. 12, p. 1119-1120, Oct. 1999. Disponível em: <http://www.dsoergel.com/cv/B70.pdf>. Acesso em: 10 jan. 2014.
- USCHOLD, M.; GRUNINGER M. Ontologies: principles, methods and applications. **The Knowledge Engineering Review**, v. 11, n. 2, p. 93-136, 1996.
- USCHOLD, M.; KING, M. Towards a Methodology for Building Ontologies. *In*: IJCAI-95: Workshop on Basic Ontological Issues in Knowledge Sharing, 1995, Montreal, Canada. Proceedings. 1995. Disponível em: http://www1.cs.unicam.it/insegnamenti/reti_2008/Readings/Uschold95.pdf. Acesso em: 20 jun. 2014.
- VICKERY, B. C. Ontologies. **Journal of Information Science**, v. 23, n. 4, p. 277–286, 1997.

O domínio científico da Arquivologia na Organização do Conhecimento: uma análise a partir do acoplamento bibliográfico no periódico Knowledge Organization (KO)

The scientific domain of Archival science in the Organization of Knowledge: an analysis based on the bibliographic coupling in the journal Knowledge Organization (KO)

Amanda Marissa Soares da Silva (1), Luciana Davanzo (2), Walter Moreira (3)

(1) Universidade Estadual Paulista – Faculdade de Filosofia e Ciências, Rua Hygino Muzzi Filho, 737 – Mirante – Marília – SP, amandamarissa1@gmail.com, (2) luciana.davanzo@gmail.com, (3) walter.moreira@unesp.br

Resumo

O presente artigo tem por objetivo utilizar o método do acoplamento bibliográfico no periódico científico *Knowledge Organization (KO)* oriundo da *International Society for Knowledge Organization (ISKO)*, a partir da recuperação na base de dados Scopus. A escolha desse periódico deu-se por ser este um dos principais veículos de comunicação científica que abarcam as áreas da Ciência da Informação, Organização do Conhecimento e Arquivologia. Propõe-se identificar os principais autores/pesquisadores desse domínio específico com a finalidade de identificar a Organização do Conhecimento Arquivístico (OCA). A metodologia utilizada nesta pesquisa é a análise de domínio a partir dos estudos bibliométricos e a aplicação do método da análise relacional – acoplamento bibliográfico conforme Kessler (1963), que permite a visualização da estrutura do domínio científico analisado e a compreensão de como se dá essa relação entre os autores e suas correntes teóricas dentro desse domínio. Para fins da análise, foram identificados e recuperados artigos no universo na base de dados Scopus, e assim selecionou-se o periódico científico KO. Por oportuno, essa pesquisa busca caracterizar o estado da arte da Arquivologia na Organização do Conhecimento. Os resultados obtidos foram a identificação de nove tipos de subdomínio de autores citados, sendo que somente cinco são atuantes na literatura científica da KO, suas filiações são de origem brasileira, portuguesa, americana, canadense e outras. Por fim, a pesquisa evidencia a aproximação da Arquivologia com a OC, e isso favorece novos estudos acerca dos sistemas e dos processos de organização do conhecimento voltados ao documento de arquivo.

Palavras-chave: Organização do Conhecimento Arquivístico. Análise de Domínio. Acoplamento bibliográfico.

Abstract

This article aims to use the method of bibliographic coupling in the scientific journal *Knowledge Organization (KO)* from the *International Society for Knowledge Organization (ISKO)*, based on the recovery in the Scopus database. This journal was chosen because it is one of the main scientific communication vehicles covering the areas of Information Science, Knowledge Organization and Archivology. It is proposed to identify the main authors/researchers of this specific domain in order to identify the domain of the Organization of Archival Knowledge (OCA). The methodology used in this research is domain analysis based on bibliometric studies and the application of the relational analysis method - bibliographic coupling according to Kessler (1963), which allows the visualization of the structure of the analyzed scientific domain and the understanding of how this relationship between the authors and their theoretical currents within this domain takes place. For the purposes of the analysis, articles were identified and retrieved in the universe in the Scopus database and thus the scientific journal KO was selected. Appropriately, this research seeks to characterize the state of the art of Archivology in Knowledge Organization. The results obtained were the identification of nine types of domain networks of cited authors, only five of which are active in the scientific literature on KO, their affiliations are of Brazilian, Portuguese, American, Canadian and other origins. Finally, the research highlights the approximation of Archival Science with the KO, and this favors new studies about the systems and processes of knowledge organization focused on the archival document.

Keywords: Organization of Archival Knowledge. Domain Analysis. Bibliographic coupling.

Resumen

Este artículo tiene como objetivo utilizar el método de acoplamiento bibliográfico en la revista científica *Knowledge Organization (KO)* de la *International Society for Knowledge Organization (ISKO)*, a partir de la recuperación en la base de datos Scopus. Se eligió esta revista por ser uno de los principales vehículos de comunicación científica que abarca las áreas de las Ciencias de la Información, la Organización del Conocimiento y la Archivología. Se propone identificar a los principales autores/investigadores de este dominio específico para identificar el dominio de la Organización del Conocimiento Archivístico (OCA). La metodología utilizada

en esta investigación es el análisis de dominio con base en estudios bibliométricos y la aplicación del método de análisis relacional - acoplamiento bibliográfico según Kessler (1963), que permite visualizar la estructura del dominio científico analizado y comprender cómo funciona esta relación. entre los autores y sus corrientes teóricas dentro de este dominio tiene lugar. Para efectos del análisis se identificaron y recuperaron artículos del universo en la base de datos Scopus y así se seleccionó la revista científica KO. Oportunamente, esta investigación busca caracterizar el estado del arte de la Archivología en la Organización del Conocimiento. Los resultados obtenidos fueron la identificación de nueve tipos de redes de dominio de los autores citados, de los cuales solo cinco están activos en la literatura científica sobre KO, sus afiliaciones son de origen brasileño, portugués, estadounidense, canadiense y otros. Finalmente, la investigación destaca la aproximación de la Archivología con el KO, lo que favorece nuevos estudios sobre los sistemas y procesos de organización del conocimiento centrados en el documento archivístico.

Palabras clave: Organización del Conocimiento Archivístico. Análisis de Dominio. Acoplamiento bibliográfico.

1 Introdução

A Arquivologia vem se fazendo presente nos estudos da Organização do Conhecimento desde o século passado, incluindo-se os estudos e a criação de sistemas de organização do conhecimento (SOC) e de processos de organização do conhecimento (POC) voltados ao documento de arquivo, considerando-se os princípios que atestam a singularidade dos arquivos, que se fundamentam nos princípios da Arquivologia, tais como a proveniência e a ordem original. Segundo Barros e Sousa (2020, p. 493) os motivos dessas recentes aproximações ocorrem devido “às mudanças ocorridas nos últimos vinte anos em relação à gestão documental e à forma como as pessoas constroem conhecimento e buscam informações”.

Desse modo, o presente artigo nasce do interesse de identificar quais os autores que publicam no principal periódico científico internacional da área o *Knowledge Organization (KO)* oriundo da *International Society for Knowledge Organization (ISKO)*, tornando um importante espaço para discussões sobre as conexões entre a OC e a Arquivologia.

Para os fins dessa análise, consideramos que as citações dos artigos científicos são fontes fiéis de informação que abarcam a comunidade científica e servem como instrumento de avaliação do comportamento da produção científica de um campo/domínio.

A OC, segundo a concepção de Hjørland (2008), é um campo de pesquisa, ensino e prática [...] que aborda a descrição, representação e organização de documentos e de suas representações bem como de assuntos e conceitos ligados aos seres humanos e aos programas de computador.

Já para Mario Barité (2015) é uma área do conhecimento de formação relativamente recente, “que estuda as leis, princípios e procedimentos pelos quais o conhecimento especializado é estruturado em qualquer disciplina, visando representar tematicamente e recuperar as informações contidas em documentos de qualquer natureza, por meios eficientes que deem resposta rápida às necessidades dos usuários” (BARITÉ, 2015, p. 120).

No que tange aos procedimentos metodológicos, empregamos a Análise de Domínio (AD) (utilizando as abordagens epistemológica, histórica e bibliométrica) a partir da aplicabilidade do método da análise relacional a partir do acoplamento bibliográfico. Assim, caracteriza-se como uma pesquisa exploratória e descritiva, de natureza quali-quantitativa. A identificação, seleção e recuperação dos artigos componentes do corpus ocorreu por consulta à base de dados SCOPUS, mais especificamente contemplando o conjunto de artigos publicados no periódico *Knowledge Organization*, com cobertura do período que compreende os anos de 2006 a 2020.

Diante disso, utiliza-se nesta pesquisa o Acoplamento Bibliográfico (AB), enquanto método de pesquisa relacional que permite a visualização da estrutura científica do domínio estudado (BOYACK; KLAVANS, 2010; GRÁCIO, 2016).

Segundo Grácio (2016, p. 82), o uso dos métodos de cocitação e o do acoplamento bibliográfico “permite conhecer as relações estruturais de conectividade teórico-metodológica de um domínio, a proximidade, a vizinhança, a associação e a interlocução estabelecida entre documentos e pesquisadores”. Dessa forma, como citado anteriormente, nesta pesquisa, será abordado o AB, considerando-se sua utilidade e precisão para identificar os indicadores de similaridade ou proximidade temática.

Desse modo, objetivou-se compreender a conexão entre os autores em função da frequência de compartilhamento da referência entre eles. Poderemos, então, visualizar as redes de autores, por conseguinte identificar a existência desses subdomínios científicos da Organização do Conhecimento Archivístico (OCA) em que consideramos a definição do conhecimento arquivístico como “[...] pode ser concebido a partir do conceito de fundo (sobre o qual incidirão todos os processos de organização), caracterizado a partir da reunião e análise de documentos produzidos por uma mesma pessoa ou instituição, com base no método diplomático enquanto suporte para a metodologia da identificação arquivística” (TOGNOLI; RODRIGUES; GUIMARÃES, 2019, p. 71-72).

Por fim, saberemos os grupos atuantes, ou seja, as influências teóricas dentro desse domínio e aqui consideramos o referencial teórico advindo da OC, no entanto, está pesquisa possui como interesse o universo da OCA de modo a identificar o referencial teórico atuante e suas conexões.

2 Contextualizando a Arquivologia na Organização do Conhecimento: principais premissas sobre a Organização do Conhecimento Arquivístico

Desde o final século XX observa-se uma aproximação entre a OC e a Arquivologia, notadamente quando os arquivos passam a ser concebidos como sistemas de informação, a partir de uma perspectiva pós-custodial, compartilhando o propósito de organizar um conhecimento específico produzido e registrado pela sociedade, de maneira que permita sua portabilidade no espaço e sua permanência no tempo e, como consequência, a promoção de seu acesso. Aos arquivos é dado um novo poder, o de servir a um papel crítico em seus meios sociais, preservando e disseminando o conhecimento coletivo de suas culturas, assim como ocorre com as bibliotecas (SMIRAGLIA, 2014).

Esteban Navarro (1995) defende uma maior aproximação entre os estudos arquivísticos e a OC, considerando esse último um campo integrador, capaz de ir além da divisão tripartida entre as áreas de Biblioteconomia, Arquivologia e Documentação, pressupondo uma integração entre elas a partir do estudo das técnicas de armazenamento, tratamento e recuperação de documentos.

A Organização do Conhecimento Arquivístico (OCA) é aquela que visa o contexto de produção dos documentos (proveniência), desse modo, consideramos que esse conhecimento é único, pois conforme refletem Alencar e Cervantes (2017, p. 02), a OCA “oferece enfoques teóricos e metodológicos para que se torne possível colaborar no acesso informacional, que irá dar vida ao ciclo de conhecimento da estruturação das vivências humanas de determinado contexto específico”.

Segundo Dalhberg (1993) existem duas concepções do conhecimento, sendo um passível de ser registrado (documento) e outro inerente ao processo individual (mente), aqui trabalhamos com o conhecimento registrado, ou seja, aquele que figura em um documento de arquivo.

No âmbito dos arquivos, os processos de organização do conhecimento são a classificação e a descrição, enquanto os sistemas de organização conhecimento são os planos de classificação, quadros de arranjo e as normas de descrição arquivística.

Por oportuno, OC para a Biblioteconomia tem por objetivo trabalhar com o conteúdo extraído dos documentos, onde são atribuídos aos documentos uma represen-

tação destes, já a Arquivologia visa trabalhar com o contexto, no qual se insere o documento, respeitando, assim, os princípios arquivísticos, e principalmente a sua proveniência, a qual é a base fundamental para os eixos centrais da organização do conhecimento arquivístico - a classificação e a descrição.

Consideramos que a Arquivologia e a Biblioteconomia não comungam dos mesmos objetos e nem dos mesmos sistemas de organização do conhecimento e processos de organização do conhecimento. Desse modo, o que difere a organização do conhecimento arquivístico é seu objeto – o documento de arquivo, com vistas sempre ao contexto de produção enquanto a organização do conhecimento, propriamente dita, ligada à Biblioteconomia, é direcionada para o conteúdo do documento.

Então, o documento de arquivo é:

um artefato humano com pressupostos e características específicas. O ambiente e o conteúdo são delimitados e definidos pelo sujeito acumulador, que pode ser uma pessoa física ou jurídica. Então, quando falamos de arquivo, estamos nos referindo a um conjunto finito de documentos acumulados, que tem suas fronteiras demarcadas pela missão do criador, no caso das instituições, e pela área de atuação, no caso das pessoas físicas. Ao contrário daqueles encontrados em bibliotecas, por exemplo, os documentos arquivísticos não constituem um conjunto formado em vista de uma finalidade específica: eles representam, mais que tudo, o produto das atividades do sujeito criador (SOUSA, 2014, p. 06).

Barros e Sousa (2019) apresentam uma aproximação entre as disciplinas, defendendo uma incorporação dos estudos da Organização do Conhecimento na Arquivologia. Enquanto, os teóricos como Tognoli e Guimarães (2019) consideram a identificação dos processos arquivísticos como parte da Organização do Conhecimento.

Desta maneira, consideramos pertinente o conceito de organização do conhecimento arquivístico defendida pelos autores Tognoli, Guimarães e Tennis (2013, p.219, tradução nossa), isto é, “todo o conhecimento produzido sobre uma pessoa ou entidade específica e agrupada em um fundo”.

Não obstante, a recente aproximação, deve-se observar que a Arquivologia sempre organizou seu conhecimento a partir de princípios e conceitos específicos, a exemplo dos princípios da proveniência e da ordem original, tendo o contexto de produção como uma bússola a ser seguida nos caminhos da organização.

Dessa maneira, alguns autores compreendem que a relação entre a Arquivologia e a OC é realizada por meio da incorporação dos estudos da OC na Arquivologia (BARROS; SOUSA, 2019; ALENCAR; CERVANTES, 2017), enquanto para Tognoli e Guimarães (2019) essa relação é possibilitada a partir de uma identificação dos processos arquivísticos como parte da OC.

3 Acoplamento bibliográfico enquanto método de pesquisa

Nesse contexto, ao evidenciar a relação entre a OC e a Arquivologia, analisando a presença arquivística na literatura científica produzida no periódico KO, busca-se caracterizá-lo e compreender como interagem os membros dessa comunidade discursiva.

Segundo Mai (2005, p. 605) um domínio é caracterizado como uma “área de especialidade, um conjunto literário ou um grupo de pessoas trabalhando juntas numa organização”, enquanto uma comunidade discursiva é definida por Hjørland e Albrechtsen (1995, p. 400), como “distintos grupos sociais sincronizados em pensamento, linguagem e conhecimento, constituintes da sociedade moderna”.

Guimarães e Tognoli (2015, p. 563, tradução nossa) compreendem a análise de domínio como “uma abordagem importante para caracterizar e avaliar a ciência, pois permite identificar as condições para a construção e socialização do conhecimento científico”. A AD estuda as comunidades de discurso, por meio da compreensão da informação.

Thellefsen e Thellefsen (2004, p. 179) definem os domínios de conhecimento como “uma demarcação de determinado conhecimento, seja ele fixado num contexto profissional ou não”.

Smiraglia (2014) esclarece que o nome do teórico e professor dinamarquês Birger Hjørland é frequentemente associado aos estudos de domínio e considera que sua escrita pragmática exorta os autores de domínio a se afastarem da pragmática e assim passam a considerar outras perspectivas epistemológicas.

Para o autor, quando se trabalha com domínio não se pode tratá-los como se todos fossem fundamentalmente semelhantes, desse modo, uma abordagem teórica da CI deve considerar diferentes comunidades de discurso. Segundo Hjørland (2002) a relação entre a pesquisa e a prática na CI, ele chama de “estoque de conhecimento mais profundo e satisfatório para a pesquisa” (2002, p. 451, tradução nossa).

Considera-se que entender um domínio específico do conhecimento permite identificar elementos para a escolha de teorias e métodos singulares e específicos para o tratamento daquele domínio. Nesse sentido, têm-se subsídios para caracterizar uma comunidade de pesquisadores envolvidos na construção de uma base epistemológica, com uma estrutura discursiva específica, para a fundamentação dos estudos arquivísticos no âmbito da OC. Para os fins deste trabalho, utilizaremos somente três das sete abordagens, defendidas por Hjørland (2008), isto é, a epistemológica, histórica e bibliométrica.

A abordagem epistemológica é, segundo Hjørland (2008, p. 439), “a mais básica sem a qual qualquer outra abordagem tende a ser superficial”, uma vez que fornece o conhecimento sobre as bases do domínio e, geralmente, vem acompanhada da abordagem histórica. Já a abordagem bibliométrica permite medir, neste caso, a produção científica sobre Arquivologia na Organização do Conhecimento, reconhecendo seus atores e influências.

Dentro da bibliometria existem dois principais métodos relacionais de citação segundo Marshakova (1981), são eles: análise de cocitação e o acoplamento bibliográfico. Em suma, a “cocitação analisa os artigos citados pelo mesmo documento de forma prospectiva, enquanto o acoplamento bibliográfico consiste na união de artigos, que cita(m) o(s) mesmo(s) documento(s) desenvolvendo uma análise de cunho retrospectivo” (LUCAS; GARCIA-ZORITA, 2014, p. 31).

Justificamos a utilização desse método, nesta pesquisa, devido o interesse de fazer uma análise relacional de autores de modo a identificar as correntes teóricas da Arquivologia a partir do acoplamento bibliográfico de artigos recuperados da revista KO, no período compreendido entre os anos de 2006 a 2020, uma vez que, esse método possui um caráter retrospectivo a partir dos seus resultados.

O acoplamento bibliográfico une os artigos citando os mesmos documentos, o método de cocitação une os artigos citados pelo mesmo documento. A força da ligação de dois artigos está no número de documentos em que ambos os documentos citados aparecem simultaneamente. Nesse caso, o tempo é determinante e faz com que a força de ligação possa variar sensivelmente com o tempo. No momento da publicação, dois artigos podem não parecer muito ligados, mas suas ligações podem aparecer (e crescer com o tempo) quando esses artigos começam a serem citados na literatura científica (LUCAS, E. O.; GARCIA-ZORITA, 2014, p. 30).

Para Grácio (2016) o acoplamento bibliográfico tem por objetivo medir “a relação entre dois artigos com base no número de referências em comum citadas pelos dois artigos” (2016, p. 86) e acrescenta que a análise de citação visa medir “a relação entre dois artigos com base no número de publicações em que estes aparecem citados concomitantemente” (2016, p. 86). Desse modo, consideramos serem métodos de indicadores de similaridade, porém nos mostram que os resultados dos dois serão distintos.

Então, para Lucas e Garcia-Zorita (2014, p. 31) o AB “possibilita o estudo do desenvolvimento de cadeias ou linhas de pesquisa, permitindo identificar os núcleos de pesquisa, os pesquisadores e os artigos mais importantes de um determinado ramo da ciência”. O AB é então definido pela “quantidade de documentos citados em comum” (CASTANHA; GRÁCIO, 2020, p.05)

Esses documentos citados em comum mostram o domínio científico a partir do seu referencial teórico em comum. A partir de um estudo publicado por Grácio (2016) sobre uma revisão teórico-conceitual dos dois métodos, a autora constatou que:

o método de acoplamento bibliográfico, ao analisar as referências e autores compartilhados por dois cientistas na construção das suas pesquisas, permite identificar a sobreposição da identidade teórico e/ou metodológica desses cientistas, assim como visualizar a estrutura intelectual, social e cognitiva na perspectiva deles próprios. Evidencia, ainda, as similaridades de ambiente científico em que atuam os pesquisadores analisados. Considera-se, assim, que o acoplamento bibliográfico AB permite identificar as correntes teóricas, e seus pesquisadores componentes, em um dado momento do desenvolvimento de um domínio científico, com significativa contribuição para a visualização da estrutura intelectual e correntes que vigoram em um domínio em estudo, nos estágios iniciais do seu desenvolvimento (GRÁCIO, 2016, p. 05).

Desse modo, podemos considerar que o acoplamento bibliográfico permite visualizar e medir a conexão entre autores em função do que eles citam em comum. O AB faz a mensuração da proximidade entre dois artigos comparando suas referências e quanto maior o número de referências que compartilham, maior a similaridade entre eles, que pode ser temática, teórico, metodológica ou outra particularidade compartilhada (LUCAS et al., 2013, PEIXE; PINTO, 2022).

Por fim, essa análise relacional com as abordagens epistemológica e histórica da análise de domínio permitirá mostrar a evolução das comunidades discursivas no domínio circunscrito.

4 Análise dos dados

A busca pelos artigos iniciou-se pela consulta à base de dados SCOPUS, em seguida no campo *Start within* e na seleção do campo ISSN, foi inserido o número do ISBN (0943-7444) do periódico KO. A primeira busca resultou na recuperação de 759 documentos. Em seguida, ao atribuir o radical “arch” como estratégia de busca foram recuperados 29 documentos, desses foi atribuído também o critério de seleção de artigos que estavam relacionados a Arquivologia/arquivística/arquivo/documento de arquivo a partir da análise dos títulos, resumos e palavras-chave, assim tivemos a recuperação de 19 artigos no intervalo entre os anos de 2006 a 2020.

Dessa forma, foram identificados 30 pesquisadores oriundos de 28 filiações. Pode-se perceber que o maior número de pesquisadores é de brasileiros seguidos por estadunidenses. Desse modo, a nacionalidade ficou entre os países e o respectivo quantitativo, são eles: Brasil (5); Estados Unidos (4); Portugal (3); Canadá (2); Espanha (2); Dinamarca (1); Alemanha (1); Itália

(1); e Taiwan (1), conforme a tabela 1 - Artigos recuperados no periódico KO.

Tabela 1. *Artigos recuperados do periódico KO*

<i>Autores</i>	<i>Título</i>
Dobreski, B.; Qin, J., Resnick, M. (2020)	Depicting historical persons and identities: A faceted approach
Bäk, G.; Allard D.; Ferris, S. (2019)	Knowledge organization as knowledge creation: Surfacing community participation in archival arrangement and description
Tognoli, N.B.; Rodrigues, A.C.; Guimarães, J. A. G. (2019)	Archival knowledge: Conceptual frameworks for recent terminology in the KO domain
Chen, S.-J. (2019)	Semantic enrichment of linked archival materialst
Barros, T.H. B.; Sousa R.T.B. (2019)	Archival science and knowledge organization: Mapping methodological relationships
Souza, R. M. (2019)	The representation of archival information in controlled vocabularies: The context of the archival institutions in Rio de Janeiro
Simões; Graça. M.; Freitas M.C.V.; Rodríguez-Bravo B. (2016)	Theory of classification and classification in libraries and archives: Convergences and divergences
Guimarães, J.A.C.; Tognoli N.B. (2015)	Provenance as a Domain Analysis Approach in Archival Knowledge Organization
Henttonen, P. (2015)	Dimensions of Contextual Records Management Classifications
Silva A.P.; Guimarães J.A.C.; Tognoli N.B. (2015)	Ethical values in archival arrangement and description: An analysis of professional codes of ethics
McQueen, K. (2015)	Ethical issues of knowledge organization in designing a metadata schema for the Leo Kottke Archives
Ribeiro; F. (2014)	The use of classification in archives as a means of organization, representation and retrieval of information
Rosati, L.; Schena A.; Masacesi R. (2013)	Childhood and adolescence between past and present. Using knowledge organization to bridge the different channels of a cultural institution: The case of the Istituto degli Innocenti, Firenze
Gilliland, A.T. (2012)	Contemplating co-creator rights in archival description
Zhang, J. (2012)	Archival context, digital content, and the ethics of digital archival representation

<i>Autores</i>	<i>Título</i>
Silva, A.M.D.; Ribeiro F. (2012)	Documentation/information and their paradigms: Characterization and importance in research, Education, and professional practice
Giménez-Chornet, V.; Escrig-Giménez M. (2011)	Designing a thesaurus to give visibility to the historical archives in the archivo del reino in Valencia
Hjørland, B. (2008)	What is Knowledge Organization (KO)?
Foscarini, F. (2006)	Records classification and functions: An archival perspective

Fonte: Autores.

Os primeiros artigos sobre a temática em análise foram publicados no periódico entre os anos de 2006 e 2008. Diante disso podemos considerar o início dos anos 2000 como o período em que começaram a se propagar na OC a necessidade de a Arquivologia se fazer presente nas discussões da área, e principalmente na ISKO. Desse modo, a absorção de problemas característicos do documento de arquivo foram cruciais para as pesquisas futuras, como, por exemplo, no desenvolvimento de SOC e POC voltados para a arquivística.

O primeiro artigo publicado na KO direcionado a Arquivologia foi da autora Fiorella Foscarini, do Canadá, sob o título *Records classification and functions: An archival perspective*, no ano de 2006. A essa publicação seguiu-se a do autor Birger Hjørland, da Dinamarca, sob o título *What is Knowledge Organization (KO)*, no ano de 2008. Considerando que esses autores são referências na área da OC, então nos mostram que a Arquivologia necessitaria se fazer presente no universo da KO.

Identificamos a presença de autores brasileiros na emergência de uma comunidade discursiva sobre a OCA e assim sinaliza a solidez desses pesquisadores no âmbito internacional. Desse modo, concretizam-se também nos estudos de Silva; Tognoli (2022) em que essas autoras evidenciam uma elite de pesquisa dessa comunidade discursiva no âmbito da OCA, e especificamente através do universo analisado da ISKO, sendo seus capítulos nacionais e regionais e seus eventos internacionais.

Acredita-se essa contribuição genuinamente brasileira à atuação dos pesquisadores da Universidade Estadual Paulista em um primeiro momento, notadamente a partir da ligação dos autores com a Isko (60% dos autores são ou foram filiados à ISKO Brasil em algum momento) e com as pesquisas desenvolvidas no âmbito do Programa de Pós-Graduação em Ciência da Informação da instituição, de onde advém a maioria dos autores (Barros, Tognoli, Guimarães, Moraes, Troitinho) que compõem a elite de pesquisa em OCA (SILVA; TOGNOLI, 2022, p. 15).

No apêndice A - apresentamos uma figura produzida com o recurso da ferramenta VOSviewer, na qual permite visualizar o acoplamento bibliográfico identificado a partir da coleta dos dados, por meio da exportação em arquivo CSV dos resultados obtidos na base Scopus e assim caracterizar a análise dessas redes de autores.

A rede para o acoplamento bibliográfico foi construída com destaque na conexão entre os artigos que compartilham no mínimo uma referência em comum. Então, o número mínimo de documentos por autor foi de trinta, na produção de no mínimo um artigo citado. Contudo, observamos que dos trinta autores da rede, somente vinte e um autores estão conectados entre si.

Podemos identificar que existem nove categorias de conexões entre esses autores, em que se evidenciam os tipos de correntes de autores na OCA, ou seja, suas linhas de pesquisas, as ligações entre os autores que são acoplados pelos autores citados (vejamos no apêndice A).

Consideramos, que existem nove subdomínios presentes nesta rede de acoplamento bibliográfico. Desses, cinco subdomínios estão conectados entre si, por pelo menos, um autor em comum. Para compreender a análise, foram organizados os seguintes grupos consoante aos tipos investigação teórica seguidas pelos autores, são eles: a) 1º grupo: Guimarães, J. A. G. Tognoli, N.B.; Rodrigues, A.C.; Silva A.P, este grupo é composto por pesquisadores brasileiros com a proposta de discutir os estudos epistemológicos e histórico da OCA a partir da junção dos estudos da Arquivologia com a Diplomática; b) 2º grupo: Silva, A.M.D.; Ribeiro F., este grupo é composto por pesquisadores portugueses que também estudam os aspectos epistemológico e histórico da arquivística; c) 3º grupo: Foscarini, F.; Barros, T.H. B.; Sousa R.T.B.; Souza, R. M. e; Henttonen, neste grupo observamos a composição de quatro brasileiros, um finlandês e uma autora canadense que se propõe a estudar os aspectos epistemológicos e aplicados da arquivística na OC; d) 4º grupo: Chen, S.-J.; Simões; Graça. M.; Freitas M.C.V.; Rodríguez-Bravo B.; McQueen, K. e; Zhang, J., este grupo se compõe de dois portugueses, dois estadunidenses, uma espanhola e um taiwanês, cujos estudos são direcionados aos estudos aplicados da teoria arquivística; e) 5º grupo: Bäk, G.; Allard D.; Ferris, S.; Gilliland, A.T., este grupo é composto por três canadenses e um estadunidense que se propõem a estudar os aspectos aplicados da arquivística através da descrição de documentos.

Hjørland (2013) diz que o acoplamento bibliográfico entre autores permite identificar o grau de sobreposição da identidade de citação desses autores, ou seja, quanto mais autores são citados em comum, mais poderemos identificar os domínios de acordo com a identidade. Consideramos que a composição dos grupos 1º, 2º e 3º são atuantes na literatura científica da OCA e, sobretudo

do, possuem redes conectadas por mais de uma referência em comum, na qual os autores citam os demais participantes destes. Esses grupos trabalham com autores de estudos de cunho epistemológico e aplicado da Arquivologia na OC.

Inferimos que quatro desses grupos não estão conectados entre si, pois apresentam linhas de pesquisas diferentes ou não são pesquisadores da própria Arquivologia porque não possuem nenhum autor citado em comum com os demais. Então, existem poucas conexões como nos demonstram na rede a saber: Dobreski B.; Qin J., Resnick M. (2020); Bäk, G.; Allard D.; Ferris, S. (2019); Giménez-Chornet, V.; Escrig-Giménez M. (2011) e; Hjørland, B. (2008), desse modo não nos aprofundaremos sobre eles.

3 Conclusão

Este artigo teve o objetivo de caracterizar o estado da arte da Arquivologia na Organização do Conhecimento, especificamente no periódico científico internacional *Knowledge Organization (KO)*. Nota-se, assim, a aproximação entre as áreas nas últimas décadas, quando vários autores passam a defender essa relação, emergindo uma comunidade discursiva comprometida com uma relação que se demonstra ser profícua para ambos os lados.

Para efetivar a análise, utilizou-se do acoplamento bibliográfico de autores, sendo esse considerado um dos métodos mais importante para a visualização de um domínio. Desse modo, pode-se visualizar por meio de nove tipos de subdomínios (nomeamos de grupos) na KO. Contudo, somente cinco grupos são atuantes na área a fim de estudar e solidificar as discussões em torno da presença da Arquivologia na construção de SOC e POC voltados ao documento de arquivo.

Ressaltamos que este artigo teve por objetivo estudar a aplicabilidade do método relacional – acoplamento bibliográfico de modo que permitisse utilizá-lo a partir do uso da ferramenta VOSviewer para a criação e visualização das redes do domínio científico da OCA na KO. Contudo, relatamos algumas dificuldades encontradas para a análise dos dados, mas consideramos que este estudo abriu o leque para o interesse de aprofundá-lo futuramente.

As redes colocam em evidência os autores brasileiros atuantes na KO. Destarte, quando os pesquisadores e professores Tognoli, Rodrigues e Guimarães (2019) definiram o conceito de OCA, esses autores fundamentam debates subjacentes à Arquivologia nos estudos da OC. Assim, evidenciou-se nesta pesquisa a visualização das redes de citação entre os membros desse domínio científico composto por pesquisadores que na sua maioria são brasileiros atuantes na KO, como: Natália

B. Tognoli, José Augusto C. Guimaraes, Ana Célia Rodrigues e Thiago H. B. Barros.

Por fim, temos a presença de autores portugueses e estadunidenses também atuantes na OCA, na qual evidenciam o interesse pelos estudos epistemológicos e aplicados da Arquivologia na OC. Contudo, podemos perceber o interesse de autores atuantes na OC, mas apresentam a sua preocupação em relação aos estudos dos arquivos/documentos de arquivo.

Notas

[1] Ferramenta utilizada: Site: www.scopus.com

[2] Ferramenta utilizada: Site: www.vosviewer.com

Referências

- ALENCAR, M. F.; CERVANTES, B. M. N. A importância da organização do conhecimento arquivístico no acesso à informação: um olhar para tesouros funcionais. *In: XVIII ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - ENANCIB*, 2017, Marília. *Anais [...]*. Marília: UNESP, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/105535>. Acesso em: 10 mar. 2023.
- BARROS, T. H. B.; SOUSA, R. T. B. Archival science and knowledge organization: mapping methodological relationships. *KO KNOWLEDGE ORGANIZATION*, v. 46, n. 7, p. 493-501, 2020.
- CASTANHA, R. G.; GRÁCIO, M. C. C. Indicadores de acoplamento bibliográfico para a avaliação da proximidade teórico-metodológica em redes de genealogia acadêmica: um estudo aplicado aos descendentes bolsistas PQ de Aldo Barreto. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, v. 18, p. e020039-e020039, 2020.
- ESTABAN NAVARRO, M. A. La representación y la organización del conocimiento em los archivos. *In: GARCIA MARCO, F. J. (coor). Organización del conocimiento en sistemas de información y documentación: actas del I Encuentro de ISKO-España*, Madrid, 4 y 5 de noviembre de 1993, págs. 65-90, 1995.
- GRÁCIO, M. C. C. Acoplamento bibliográfico e análise de cocitação: revisão teórico-conceitual. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, v. 21, n. 47, p. 82-99, 2016.
- GUIMARÃES, J. A. C.; TOGNOLI, N. B. Provenance as a Domain Analysis Approach in Archival Knowledge Organization. *Knowledge Organization*, v. 42, n. 8, p. 562- 69, 2015.
- GUIMARÃES, J. A. C. Análise de domínio como perspectiva metodológica em organização da informação. *Ciência da Informação*, v. 43, n. 1, 2014.
- GUIMARÃES, J. A. C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo

científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação**, v. 1, n. 1, p. 77-99, 2008.

HJØRLAND, B. **Domain analysis in Information Science: Eleven approaches traditional as well as innovative**. JASIS, 58 (4): 422-62, 2002.

HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. **Journal of the American Society for Information Science**, v. 46, n. 6, p. 400-425, 1995.

HJØRLAND, B. What is knowledge organization (KO)? **Knowledge Organization**, v. 35, n. 2-3, p. 86-101, 2008.

LUCAS, E. O.; GARCIA-ZORITA, J. C. Produção Científica sobre Capital Social: estudo por acoplamento bibliográfico. **Em Questão**, v. 20, n. 3, ed. esp., p. 27-42, 2014.

LUCAS, E. O.; GARCIA-ZORITA, J. C.; SANZ-CASADO, E. Evolução histórica de investigação em informetria: ponto de vista espanhol. **Liinc em Revista**, v. 9, n. 1, p. 255-270, maio, 2013.

MARSHAKOVA, I. V. Citation networks in information science. **Scientometrics**, Budapest, v.31, n.1, p.13-16, 1981.

MAI, J.-E. Analysis in indexing: document and domain centered approaches. **Information Processing and Management**, v.41, n.3, p.599-661, 2005.

SMIRAGLIA, R. P. The elements of knowledge organization. **Springer**, 2014.

SILVA, A. M. S. **A Arquivologia na organização do conhecimento: uma análise de domínio**. 2021. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal Fluminense, Rio de Janeiro, 2021.

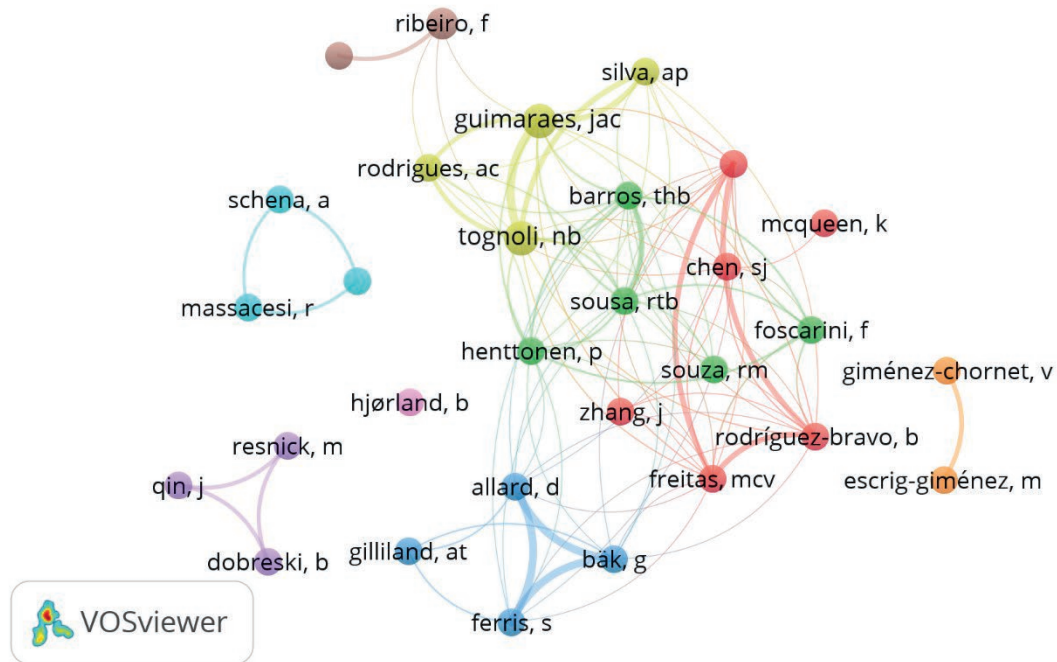
SILVA, A. M. S.; TOGNOLI, N. B. A organização do conhecimento arquivístico: a emergência de uma comunidade discursiva brasileira. **Acervo**, Rio de Janeiro, v. 35, n. 2, p. 1-17, maio/ago. 2022.

HELLEFSEN, T. L.; THELLEFSEN, M. M. Pragmatic semiotics and knowledge organization. **Knowledge Organization**, v. 31, n. 3, p.177-187, 2004.

TOGNOLI, N. B.; GUIMARÃES, J. A. C. Provenance as a knowledge organization principle. **Knowledge Organization**, 46, no. 7: 558-68, 2019.

TOGNOLI, N. B.; RODRIGUES, A. C.; GUIMARÃES, J. A. C. Definindo o conhecimento arquivístico: estruturas conceituais. **Informação & Informação**, [S.l.], v. 24, n. 2, p. 58-75, nov. 2019.

Apêndice A – Rede de acoplamento bibliográfico de autores da Arquivologia no periódico científico *Knowledge Organization (KO)*



Fonte: Elaborado pelos autores a partir da ferramenta VOSviewer, 2023.

Epistemologia do Mapa Conceitual em Sistema de Organização do Conhecimento por meio da historicidade dos conceitos

Epistemology of the Concept Map in a Knowledge Organization System through the historicity of concepts

Maria Rosemary Rodrigues, Brígida Maria Nogueira Cervantes

Universidade Estadual de Londrina (UEL), Londrina, rosemmaryrodrigues42@gmail.com

Resumo

O Mapa Conceitual tem se demonstrado como um Sistema de Organização do Conhecimento porque sua epistemologia se revela por meio da historicidade dos conceitos como mapas de conhecimento em outros períodos. O objetivo: O objetivo geral deste estudo é apresentar a epistemologia do Mapa Conceitual por meio da historicidade dos conceitos para a representação do Mapa Conceitual no contexto do Sistema de Organização do Conhecimento. Metodologia: A metodologia utilizada é a pesquisa descritiva, de abordagem qualitativa por meio da Teoria Fundamentada e Análise de Domínio. Resultados: A teoria-metodológica estabeleceu a sua estrutura lógica no âmbito do Sistema de Organização do Conhecimento pelo estudo histórico de Heidegger, percebendo que a sociedade científica representa o conhecimento de algum conteúdo temático em algum tempo na história. O estudo epistemológico por Aristóteles e Kant trouxe uma visão teórico-metodológica por meio de sistemas conceituais. O estudo terminológico, fundamentado em Peirce, desenhou a representação dos termos para comunicá-los, pela linguagem, à temática de algum documento. Conclui-se que a fundamentação na epistemologia do conhecimento desenhado demonstrou-se o significado da estrutura lógica do Mapa Conceitual por meio da historicidade dos conceitos de forma ampla, por ser estrutura e função dos sistemas inseridos no Sistema de Organização do Conhecimento, e de forma restrita, por ser enunciado lógico de materialização do conceito.

Palavras-chave: Historicidade; Conceitos; Mapa Conceitual; Sistemas de Organização do Conhecimento; Organização e Representação do Conhecimento.

Abstract

The Conceptual Map has been shown to be a Knowledge Organization System because its epistemology is revealed through the historicity of concepts as maps of knowledge in other periods. The objective: The general objective of this study is to present the epistemology of the Conceptual Map through the historicity of the concepts for the representation of the Conceptual Map in the context of the Knowledge Organization System. Methodology: The methodology used is descriptive research, with a qualitative approach through Grounded Theory and Domain Analysis. Results: The theory-methodology established its logical structure within the scope of the Knowledge Organization System through Heidegger's historical study, realizing that the scientific society represents the knowledge of some thematic content at some time in history. The epistemological study by Aristotle and Kant brought a theoretical-methodological vision through conceptual systems. The terminological study, based on Peirce, designed the representation of terms to communicate them, through language, to the theme of a document. It is concluded that the foundation in the epistemology of designed knowledge demonstrated the meaning of the logical structure of the Conceptual Map through the historicity of the concepts in a broad way, as it is the structure and function of the systems inserted in the Knowledge Organization System, and in a restricted way, as it is a logical statement of materialization of the concept.

Keywords: Historicity; Concepts; Conceptual map; Knowledge Organization Systems; Organization and Representation of Knowledge.

Resumen

El Mapa Conceptual se ha mostrado como un Sistema de Organización del Conocimiento porque su epistemología se revela a través de la historicidad de los conceptos como mapas del conocimiento en otras épocas. El objetivo: El objetivo general de este estudio es presentar la epistemología del Mapa Conceptual a través de la historicidad de los conceptos para la representación del Mapa Conceptual en el contexto del Sistema de Organización del Conocimiento. Metodología: La metodología utilizada es la investigación descriptiva, con enfoque cualitativo a través de la Teoría Fundamentada y el Análisis de Dominio. Resultados: La teoría-metodología estableció su estructura lógica en el ámbito del Sistema de Organización del Conocimiento a través del estudio histórico de Heidegger, percibiendo que la sociedad científica representa el conocimiento de algún contenido temático en algún momento de la historia. El estudio epistemológico de Aristóteles y Kant trajo una visión teórico-metodológica a través de sistemas conceptuales. El estudio terminológico, basado en Peirce, diseñó la representación de términos para comunicarlos, a través del lenguaje, al tema de un documento. Se concluye que la fundamentación en la epistemología del conocimiento diseñado demostró el significado de la estructura lógica del Mapa Conceptual a través de la historicidad de los conceptos de manera amplia, como es la estructura y función de los sistemas insertos en el Sistema de Organización del Conocimiento, y de forma restringida, ya que es un enunciado lógico de materialización del concepto.

Palabras clave: Historicidad; Conceptos; Mapa conceptual; Sistemas de Organización del Conocimiento; Organización y Representación del Conocimiento.

1 Introdução

Este estudo¹ buscou investigar a epistemologia do Mapa Conceitual, no domínio da Organização do Conhecimento (OC) e no subdomínio dos Sistemas de Organização do Conhecimento (SOC), portanto partiu da seguinte questão: Quais as bases epistemológicas no domínio da OC que podem resultar na fundamentação da Epistemologia do Mapa Conceitual por meio da historicidade dos conceitos, com ênfase no processo de construção e sua aplicabilidade?

Neste sentido procurou comprovar que a historicidade do Mapa Conceitual, por meio de conceitos, pode gerar a sua fundamentação no contexto do SOC utilizando-se do processo e seus diferentes usos como estrutura lógica de representação conceitual. Isso porque, o Mapa Conceitual é pertinente à SOC porque sua epistemologia se revela por meio da historicidade dos conceitos de diferentes maneiras e em diferentes períodos, contextualizando, através do processo, seus sentidos e aplicações. O objetivo geral deste estudo apresentou a epistemologia do Mapa Conceitual por meio da historicidade dos conceitos para a representação do Mapa Conceitual no contexto do SOC. A metodologia utilizada é a pesquisa descritiva, de abordagem qualitativa por meio da Teoria Fundamentada e Análise de Domínio.

2 Mapa Conceitual

O que é Mapa Conceitual? Para Novak (2000), Novak e Cañas (2008) o Mapa Conceitual é ferramenta gráfica para organizar e representar o conhecimento como possibilidade de mostrar os significados dos conceitos e as relações que existem entre esses conceitos de algum domínio do conhecimento.

Neste estudo, o Mapa Conceitual é conceituado em conformidade a ênfase dada ao seu uso no decorrer do seu processo de construção no domínio da Organização do Conhecimento, tais como: ferramenta, instrumento, modelo conceitual, método/técnica e âmbito pedagógico, porque “[...] os mapas construídos hierarquicamente ajudam-nos a abraçar significados conceptuais mais específicos nos domínios dos conceitos mais amplos, mais gerais.” (NOVAK; GOWIN, 1984, p. 47).

No formato arborescente, os conceitos mais gerais ficam próximos à raiz e os mais específicos se estendem até suas folhas. (PINOTTE; CURY; ZOUAQ, 2015). A definição descritiva do Mapa Conceitual em analogia com um mapa de árvores, desenha e visualiza o conhecimento hierárquico, em que suas partes estão unidas por linhas que significam as vias de comunicação por conceitos dentro de círculos ou caixas composto pelos elementos conceito, proposição e palavras de ligação.

Conceição e Correia (2020) argumentam que as proposições em Mapa Conceitual podem ser descritivas ou explicativas. A proposição em Mapa Conceitual Descritivo auxilia na definição ou entendimento de algum conceito. A questão para a proposição descritiva é: “o que é determinada coisa?” A proposição em Mapa Conceitual Explicativo acontece por meio de processos cognitivos complexos e elaborados, bem como: compreender, aplicar, analisar e avaliar. A questão para a proposição explicativa é: “como ou porque” algo acontece?

A construção de um Mapa Conceitual evolui conforme as tipologias de estrutura para Mapa Conceitual. Já sua apresentação compreende três formas: paisagem, mandala e multidimensional ou 3d.

Segundo Novak (2000), a construção de um Mapa Conceitual evolui conforme sua proposta mais simplificada, por meio do processo para elaboração constituído de seis fases: 1- Identificar o tema ou a pergunta de enfoque que se vai representar; 2- Verificar os conceitos; 3- Ordenar os conceitos por meio de lista; 4- Agrupar e arranjar os conceitos; 5- Estabelecer ligações ou proposições; 6- Rever a estrutura do mapa conceitual.

Neste contexto do Mapa Conceitual, a epistemologia revela-se essencial para a OC. Fundamentada em Zins (2002), a explicação é porque, possibilita, aprofundar a construção de conhecimento e desenvolve mapas de conhecimento por meio da atividade de inteirar-se com elementos, devido a intenção de reconhecê-los, segundo o conhecimento já formado na mente, por meio de processos cognitivos a fim de representá-los.

O Mapa Conceitual, desenvolvido por Joseph Donald Novak, inicia-se a partir de diálogos com Ausubel, denominada de Aprendizagem de Conceitos, no ano de 1965. Os diálogos a respeito da Aprendizagem de Conceitos auxiliaram Novak nos *insights* da sua teoria e aplicação na educação para existir uma aprendizagem significativa. Para isso, foi necessário entender como ocorre o processo de transformação do conhecimento, bem como a importância que os processos mentais assumem nesse desenvolvimento. Mais tarde, Novak acrescenta que as aprendizagem mecânica e aprendizagem significativa deveriam ser consideradas um *continuum*, porque a primeira pode ser o início para o aprendizado mais duradouro. Além disso, passou a ver a criatividade como um nível mais alto da aprendizagem significativa.

Seguido por Marco Antônio Moreira, precursor a respeito de Mapa Conceitual no Brasil, menciona a visão clássica de David Ausubel, a visão humanista de Joseph Novak e a sua visão crítica. A visão clássica de Ausubel, é descrita como um processo de aprendizagem para assimilar a formação de significados, bem como, a organização do conhecimento e o entendimento individual). A visão humanista de Novak integra

Rodrigues, Maria Rosemary; Cervantes, Brígida Maria Nogueira. Epistemologia do Mapa Conceitual em Sistemas de Organização do Conhecimento por meio da historicidade dos conceitos. *Organização e Representação do Conhecimento em diferentes contextos: desafios e perspectivas na era da datificação*, 2023.

pensamentos, sentimentos e ações, uma vez que isso remete ao aprendiz uma sensação boa e que manifesta mais interesse para novas aprendizagens. Já a visão crítica do próprio Moreira envolve conhecimento prévio, relacionamento social, questionamento, linguagem, diversificação de estratégias de ensino. Finalizando, Albert Cañas, criador do *software Cmap Tools*, o denomina como “Mapas de Conhecimento”, pois refere-se a um conjunto de mapas e compartilhamento de conhecimento, isto é, uma coleção de Mapas Conceituais a respeito de um determinado domínio. Tem como função apoiar o desenvolvimento de Mapas de Conhecimento de todos os tamanhos, sem limitações de onde os recursos e mapas residem fisicamente.

Na epistemologia o conhecimento e a produção de conhecimento são estruturas em evolução de conceitos e proposições pelas permutas de relações conceito-conceito, proporcionado pelo discurso, compartilhamento das estruturas conceituais individuais e, assim, a mudança de significados pode ser alcançada. A organização do processo cognitivo de estruturas de conceitos dá-se pelo processo metacognitivo para entender os conceitos e os novos conceitos a fim de organizá-los e estabelecer suas relações. Assim, Novak utilizou dessa fundamentação para o desenvolvimento da ferramenta em 1972, o que, agora, chamamos de Mapas Conceituais.

2.1 Epistemologia do Conhecimento Desenhado

Por isso, mapear epistemologicamente, segundo Francelin (2018, p. 91), é “[...] tentar saber como o termo ‘epistemologia’ é usado [...]” porque o termo, quando possui alicerce conceitual, pode ir além da palavra oferecendo um entendimento melhor do seu significado. Assim, investigar a historicidade das ciências e interpretá-la fundamenta objetos e conhecimentos. Francelin (2021, p. 2) afirma que “Epistemologicamente, os domínios do conhecimento podem ser domínios conceituais.”

Lima (2014) menciona que, o mapa arbóreo é adotado ao longo dos séculos para mapear os diversos domínios de conhecimento. Também relata que é resultado da cartografia, uma vez que, sua evolução mostra-se mais enraizada e inerente do que parece. Assim, a figura da árvore vem a ser uma das metáforas mais vitalícia para mapear a informação, porque o mapa como a maneira mais antiga, tanto de representação quanto de comunicação, antecede portanto, a linguagem escrita e numérica. (OKADA, 2008, 2006). Na opinião de Lima (2014), desenhar essas pontes só é possível por meio da epistemologia, pois “Aprender o raciocínio equivale, então, a ter compreendido a ‘história’”. (OKADA, 2008, p. 43). Portanto, “conceitos, ideias, pensamento, argumentação, diálogo, imagens visuais e suas múltiplas conexões.” (OKADA, 2008, p. 43) são os elementos categóricos cognitivos utilizados para construir a

representação de diferentes associações, terminologias e mapas.

Para Aristóteles, o conhecimento científico promove o primeiro saber, o qual se refere à cognição prévia, uma vez que é um princípio do conhecimento científico. (PEREIRA, 2014). O conhecimento é gradual, assim, para vir a entender a causa é necessário que as informações estejam organizadas sistematicamente, representadas pela estrutura lógica demonstradas pelo silogismo. O processo de conhecer perpassa operações mentais tendo como produto lógico do pensamento e da expressão, as 10 categorias de Aristóteles, as quais fundamentam o conhecimento com propriedade ao desenhar o sistema categorial conceitual hierárquico.

Segundo Kant, o conhecimento científico é sintético *a priori*, necessário e universal, na medida em que é fenômeno e é percebido pelas estruturas *a priori* da cognição humana. Para uma melhor explicação, o fenômeno nada mais é que em um âmbito restrito, dentro de um âmbito vasto, a coisa tal como aparece à cognição humana.

Gorner (2017, p. 43) aponta que a hermenêutica de Heidegger “[...] interpreta a historicidade do *Dasein*, [...]” porque é a possibilidade histórica. Para a análise do *Dasein*, Heidegger utiliza o método fenomenológico hermenêutico, que, para representar, é necessário estabelecer antes uma relação. Heidegger explica que a representação é um processo mental e a relação é a prática de lidar com a realidade. Assim, pela hermenêutica, é preciso compreender o *ente* ora como função do *ser*, ora como contexto de relações para pressupor a representação. Nesse sentido, a representação pode ser quando se estabelece relação de utensílio, uso próprio e rede de significados. Assim, a historicidade revela o ‘como acontece’ e ‘como se dá’, uma vez que a ‘verdade’ experimenta a passagem do tempo no *Dasein*, ela é, de fato, seu passado.

A Semiótica de Peirce é fundamental na formação de hábitos, porque ela explica um conceito em que se começa pela emoção de forma ampla, passa pelas associações mentais e chega-se ao conceito de algum signo, o qual procede em um conceito de algum domínio de conhecimento. Santaella (1995, p. 46) menciona que “[...] o objeto é aquilo que nele promove sempre volta a insistir porque resiste na sua diversidade.” e a sua historicidade é determinada pelo interpretante (processo de composição) do objeto do signo. Desse modo, a semiótica deve ser contextualizada pela compreensão dos signos de todos os tipos, pois, ao serem organizados sistematicamente por meio de mapa, a significação acontece quando objeto e signo se relacionam. A semiótica na Organização do Conhecimento se refere a um raciocínio lógico e, conforme o objeto vai se aprofundando, vem a necessidade de organizar, porque é um procedimento de mapear, auxiliado pela tarefa de desenhar um tipo de visualidade.

2.2 Organização do Conhecimento

A Organização do Conhecimento constitui-se pelos conceitos de OC e de RC, pois a junção desses conceitos compreende o objeto (conhecimento) e a ação (atividade de organizar e representar, gerar instrumentos, processos e produtos para o uso em ambientes institucionais). Cervantes (2009, p. 19; 21) explica que a união desses conceitos envolve “‘conhecimento’ no sentido de ‘conhecer’ e ‘organização’ no sentido de ordenação de objetos, [porque] tanto sob ponto de vista de processamento quanto de representação”, o conhecimento fundamenta-se no entendimento, no sistema e na representação (CERVANTES, 2009).

Si concebimos a una noción o concepto como un iceberg, de modo tal que lo que queda por debajo de la línea de agua sea la conceptualización misma (la delimitación nocional de algo), y lo que se exhiba por encima de esa línea sea la palabra o el símbolo que expresa ese concepto [...] (BARITÉ, 2000, p. 33).

De acordo com Barité (2000), a definição de conceitos mais comum ocorre nos estudos científicos pela compreensão ou intenção. Eles são utilizados para padronizar enunciados que se complementam do geral para o específico, pois a compreensão das características é relevante para definir algo, porque são ordenados por proposições que vão se agregando conforme características da proposição anterior. Logo, o conceito é a síntese de características de algum objeto escolhido por um termo. Assim, quando a mente cria ou entende um conceito universal ajustando-o de forma igual a todas as coisas em sua extensão, pode ser expressos por meio de termos idênticos e se referir a conceitos diferentes, em razão de representar as relações entre objetos da realidade, pois o conceito é o significado da palavra relacionado no contexto que determina seu significado. A sistematização por meio de conceitos ordena um domínio desde a “[...] sua representação até a sua recuperação, estudam-se, primeiramente, os conceitos que compõem esse campo do conhecimento e as relações que se estabelecem entre eles” (CERVANTES, 2006, p. 26). Assim, a operação conceitual de compreensão se refere ao processo de estudar o conceito para representá-lo.

A expressão composta RC empenha-se em materializar o pensamento humano que possibilita a descrição e explicação de fenômenos a partir de ferramentas utilizadas e suas funções no ato de construir modelo de construção de domínios de algum conhecimento. A representação é uma função cognitiva, um constructo sociocultural proporcionando mudanças do conhecimento (GONZALEZ DE GOMEZ, 1993). A RC compõe-se de conceitos e de estruturas para representar símbolos, estabelecendo vínculos estruturais, sistemáticos, de associação e de distinção (SAN SEGUNDO MANUEL, 2003). Logo, os conceitos podem referenciar a teoria da microestrutura porque é o elemento central para estruturar o conhecimento e as

relações com todos os outros conceitos existentes dentro de um domínio (BARITÉ, 2001). A representação da estrutura lógica coaduna com a teoria da macroestrutura, pois refere-se às disciplinas de domínio temático mais geral, as quais explicam a estruturação dos sistemas de conceitos. O autor discorre ainda que métodos para representar o conhecimento são os sistemas de conceitos fundamentados pelas teorias e suas funções.

Ao definir sistemas conceituais, denominados SOC, apresenta-se a compreensão da realidade, promovendo a recuperação e acesso aos conteúdos de objetos informacionais numa constante produção de processos, instrumentos e produtos de formas de se ver o mundo e seus significados (SOUZA; SALES; SALDANHA, 2022). Os autores Zeng (2008), Mazzocchi (2017), Moreira (2018) e Hjørland (2021) referem-se ao SOC como ferramentas semânticas que representam determinado domínio por meio da sistematização dos conceitos e das relações semânticas que se estabelecem entre eles.

Neste cenário, diferentes SOC percorrem configurações variadas do saber-fazer do gesto bibliotecário com fundamentação teórico-aplicada passando pela filosofia, classificação, linguagens, terminologia, semântica e estruturas. O SOC atua ora em seu sentido amplo como organização sociocultural, ora em seu sentido restrito como organização intelectual por meio de princípios e teorias, sendo que, no sentido restrito corresponde ao significado de ferramentas semânticas, uma vez que são compostas por conceitos/termos/relações semânticas. (RODRIGUES, CERVANTES, 2017); (BARITÉ, 2022). Sua definição ampla denota diferentes estruturas de sistemas que vão de plano a multidimensional (ZENG, 2008), com diferentes funções e no contexto de diferentes comunidades com a intenção de apoiar a OC. Neste sentido, Zeng (2008) apoiou-se nos estudos de Hodge (2000) e agrupou os SOC mediante suas estruturas e suas funções: - Estruturas: Listas de Termos; Classificação e Categorização; Modelos de Relações. - Funções: Eliminar a ambiguidade; Controle de sinônimos; Estabelecem relacionamentos hierárquicos; Estabelecem relacionamentos associativos; Apresentando as propriedades.

Para esclarecer a respeito dos elementos dos SOC, utiliza-se da Terminologia (Linguagem de Especialidade) para analisar os termos relacionados à temática dos documentos. Krieger e Finatto (2020) explicam a terminologia com ‘t’ minúsculo, significa termo técnico-científico, porque expressa um conjunto de unidades lexicais de alguma área, técnica, campos de estudos. Terminologia com ‘T’ maiúsculo manifesta a organização formal das terminologias por meio de instrumentos. Assim, terminologia refere-se a um conjunto de termos e Terminologia compreende campo de estudo ou disciplina. Barité (2022) complementa ser um conjunto de práticas e métodos usados para a coleta e descrição e apresentação de termos. Assim, nas línguas

de especialidade, a Terminologia/terminologia pode ser verificada segundo: signficante e o significado; relações com outros termos; valor sociolinguístico pelo uso, porque, a partir da terminologia, define-se o conceito, por meio do processo cognitivo de relacionar as categorias para, dessa maneira, representar a realidade.

Nesse sentido, Barité (2011) destaca que os SOC se apresentam como estrutura e representação, pois trata-se de uma reprodução do conhecimento acumulado de forma epistemológica, filosófica ou particular, isso porque desmembra conhecimento de alguma especialidade.

3 Metodologia

Para apresentar a epistemologia do Mapa Conceitual por meio da historicidade dos conceitos para a representação do Mapa Conceitual no contexto do SOC abordou-se a natureza de pesquisa básica denominada de pesquisa teórica. As tipologias foram descritiva e histórica. A descritiva para apresentar uma nova percepção do que já se tem e a histórica para perceber as alterações na evolução de fenômenos influenciadas por fatos e pela sociedade. Caracterizou-se como pesquisa bibliográfica, pois proporcionam suporte teórico e reflexivo para conduzir a pesquisa. Sua abordagem foi qualitativa, as quais aconteceram pelos procedimentos metodológicos: coleta de dados pelos elementos da própria tese; estratégia de dados utilizou da Teoria Fundamentada.

Dessa forma, seguiu-se as instruções das etapas da Teoria Fundamentada de Strauss e Corbin (1990): 1 - Da amostragem teórica surgiram os *insights*, em um constante ir e vir, já proporcionados pela própria tese e, assim, explicou-se a historicidade dos conceitos dos Mapas Conceituais para estabelecer a sua estrutura lógica no âmbito do SOC por meio dos principais conceitos determinados pela escolha dos filósofos para este estudo: Aristóteles, Kant, Heidegger e Peirce; 2 - A Codificação Aberta: buscou-se identificar códigos substantivos, suas propriedades e dimensões a respeito das bases epistemológicas interdisciplinares para a representação do Mapa Conceitual; 3 - A Codificação Axial refletiu-se em um processo indutivo e dedutivo com a intenção de relacioná-las, a fim de identificar e explicar a epistemologia do Mapa Conceitual por meio da historicidade dos conceitos para o estabelecimento de sua estrutura lógica no âmbito do SOC; 4 - A Codificação Seletiva desenhou a representação do pensamento analítico. A Redação da Teoria descreveu uma história conceitual, especificando as relações de conceitualização para especificar as bases e relações epistemológicas interdisciplinares das estruturas lógicas, a fim de demonstrar no Mapa Conceitual, por meio da historicidade dos conceitos.

Já, a análise de dados ocorreu pela Análise de Domínio de Hjørland, entretanto, as escolhidas foram a sexta abordagem – estudos históricos; a oitava abordagem – estudos epistemológicos e críticos; e a nona abordagem – estudos terminológicos, linguagem para fim especial (LSP), semântica de base de dados e estudos do discurso. Para complementar é importante reportar-se a Tennis (2003) que se preocupa com os procedimentos de operacionalização de definição de domínios propostos por dois dispositivos analíticos chamados eixos. Para esses eixos, Tennis (2012) denomina-os de: Áreas de Modulação - (Nomeação (Definição do Domínio) Foco (Propósito: análise para compreender o domínio) e Graus de Especialização - Extensão (Escopo e Alcance: nº de produções) e Finalidade (Responde à pergunta).

4 Análises e Resultados: Epistemologia do Mapa Conceitual

A "Epistemologia do Mapa Conceitual" surgiu para comprovar que a historicidade dos conceitos pode gerar a sua fundamentação no contexto do SOC, utilizando-se do processo e seus diferentes usos como estrutura de representação conceitual.

4.1 Análises e Resultados: Epistemologia do Mapa Conceitual pela Teoria Fundamentada

Mediante a Teoria Fundamentada observou-se pelas: estrutura arbórea; sistema ou produto categorial; processo mental ou cognitivo; probabilidade relacional histórica e representação do significado que, conforme o objetivo, memorandos e diagramas podem ter diferentes formas. Para este estudo, escolheu-se os Diagramas lógicos, porque a “[...] representação visual do pensamento analítico mostra a evolução das relações lógicas entre categorias e subcategorias em termos de características de paradigma.”. Entretanto, deixa-se aberta “[...] à imaginação.” a intenção de descobrir novas propriedades e dimensões por meio dos significados dos conceitos. (STRAUSS; CORBIN, 1990, p. 197). Em razão disso apresenta-se o que são:

Diagrama Lógico Arbóreo? são técnicas para organizar o saber; Diagrama Lógico Categorial? é apresentar “o que é” e “o que não é” ao saber a causa de algo; Diagrama Lógico Mental? percebeu a possibilidade de compreensão de um conceito por meio da cognição humana pelas categorias mentais; Diagrama Lógico Histórico? é a compreensão do modo de ser histórico para explicar os fenômenos; Diagrama Lógico Representacional? é usado para tornar nossas ideias claras diante de significados que afetam nossa conduta de hábito, pela interpretação e a representação, ao usar o método lógico de análise de conceitos.

Neste contexto, ao utilizar os Diagramas Lógicos verificaram-se nos diferentes desenhos das epistemologias filosóficas escolhidas: processos,

categorias e conceito/conhecimento. Processos referem-se às etapas das operações mentais no intelecto humano pelos aspectos básicos do sentido fenomenológicos, ontológicos, ônticos e pela semiose para a construção do conhecimento. Categorias são as representações gráficas espaciais de objetos e o sistema de compreensão humano de modelos visuais no tempo através da cognição humana. Conhecimento por meio de mapas é quando o sujeito realiza a abstração no intelecto, usando conceitos adquiridos pela historicidade por meio da constituição, da combinação, da interpretação e da representação, os quais procedem em um conceito.

Já, a combinação do Mapa Conceitual com os mapas arbóreos desenham representam e promovem a visualização de Mapa Conceitual como desenho gráfico e sistemático de conceitos passa pelo tempo e pelo espaço em que se desenvolveram habilidades sociais e intelectuais. Fundamentado na cartografia, o Mapa Conceitual é um esquema de conhecimento.

A combinação de Sistema ou Produto Categorial com o Mapa Conceitual acontece ao utilizar as 10 categorias em que a Substância é o Mapa Conceitual, o conceito mais geral e as outras categorias representam os seus predicados por meio de proposições. Qualidade: O Mapa Conceitual é uma ferramenta para representação lógica de conceito; Quantidade: No decorrer do seu processo de construção, é uma técnica, a qual pode cumprir diferentes objetivos em áreas, disciplinas e assuntos, conforme uso: ferramenta, instrumento, modelo conceitual, método/técnica, âmbito pedagógico; Relação: Mapa Conceitual registra as relações entre os conceitos, determinados por verbos ou pequenas frases de ligação, os quais estão indicados nas linhas, formando uma proposição sendo sua característica particular; Ação: O Mapa Conceitual age como diagrama lógico para aprendizagem significativa, desenha e visualiza o conhecimento hierárquico e as suas partes estão unidas por linhas que significam as vias de comunicação; Sofrer: O Mapa Conceitual vai se modificando, conforme os entendimentos dos relacionamentos entre os conceitos forem se alterando; Posição: O Mapa Conceitual, como diagrama, representa o *continuum* da aprendizagem, dos sentimentos e da criatividade, pois captura a essência do conceito profundamente importante da aprendizagem significativa; Estado: O Mapa Conceitual é uma maneira de organizar e representar o conhecimento por meio de conceitos; Lugar: O Mapa Conceitual, relacionado ao conceito como dimensão interna, encontra-se na organização do conhecimento e, relacionado ao conceito como dimensão externa, na representação do conhecimento; Tempo: O Mapa Conceitual foi desenvolvido por Joseph Novak, na década de 1970.

A combinação de Sistema Mental com o Mapa Conceitual busca a possibilidade do conhecimento humano pela característica essencial do Mapa

Conceitual, as quais correspondem à proposição. O processo de construção do Mapa Conceitual é produzido pela interpretação do conceito e pela cognição humana para expressar o pensamento discursivo. O Mapa Conceitual apresenta: no sentido interno – o tempo – como processo mental para conceber o significado do conceito. No sentido externo – o espaço – ao ordenar os conceitos em uma estrutura de conceitos. O Mapa inserido nas Categorias de Kant: Quantidade: dentro de um SOC, como estrutura lógica de conhecimento, o Mapa Conceitual se torna um conceito geral. Qualidade: como conceito geral, o Mapa Conceitual se constrói a partir de seus atributos. Relação: o Mapa Conceitual pressupõe a presença de seus elementos. Modalidade: o Mapa Conceitual, especialmente em suas relações, envolve todo o conhecimento verdadeiro de um dado conhecimento particular em relação com o pensamento em geral.

A combinação de Probabilidade Relacional Histórica com o Mapa Conceitual sustenta sua compreensão no tempo, o que tange sua historicidade. Pelas categorias, compreendeu-se que a sua historicidade inicia-se pela possibilidade relacional de ser um Mapa Conceitual no contexto de Sistemas Conceituais, com valores socioculturais dentro do SOC que o antecede. O método fenomenológico hermenêutico responde: Como acontece? Por meio de predicções de termos dados pelas proposições. Como se dá? Pela essência do significado. O que é a verdade a respeito do Mapa Conceitual? As suas relações de significados, proporcionando e causando de forma mais adequada a compreensão e melhor visualização dos conceitos, correlacionando-se aos SOC.

A combinação Representacional com O Mapa Conceitual ocorre por meio das 10 classes, as quais, Peirce dividiu as em três tricotomias de relações: sintáticas, semânticas e pragmáticas dos signos. Neste sentido as quatro primeiras tricotomias referem-se ao desejo do Mapa conceitual vir a ser uma possibilidade de estrutura lógica. É um gráfico de visualidade por meio da historicidade dos conceitos no contexto do SOC. As tricotomias cinco, seis e sete, revelam-se como experiência de ser uma realidade concreta de estrutura lógica, uma visualidade gráfica de conhecimento. As últimas tricotomias, oito nove e dez, conceitua o Mapa Conceitual como uma ferramenta gráfica que possui estrutura lógica de organização e representação do conhecimento firmada por meio da historicidade dos conceitos no contexto do SOC. Assim, evidencia-se por meio do argumento, a razão final: O Mapa Conceitual é uma estrutura lógica firmada por meio da historicidade dos conceitos que organiza e representa o conhecimento porque proporciona visualidade geral do conhecimento e, em razão das diferentes semelhanças existentes, pode ser considerado um SOC.

4.2 Análise e Resultados pela Análise de Domínio de Hjørland (2002)

A Análise de Domínio foi usada para reconhecer características elementares, práticas e semânticas ou distinções da epistemologia do Mapa Conceitual, por meio da historicidade dos conceitos, para o estabelecimento de sua estrutura lógica no âmbito do SOC.

Hjørland (2002, p. 1) aponta que o “[...] critério de sucesso de sistemas de informação são aqueles que identificam e comunicam o conhecimento necessário [...]”. Essa filosofia implícita beneficia diferentes abordagens de diversos domínios, considerando a visão analítica de domínio e o enunciado lógico da materialização de vários conceitos. Nesse sentido, este estudo utilizou a sexta abordagem – estudos históricos; a oitava abordagem – estudos epistemológicos e críticos; e a nona abordagem – estudos terminológicos, linguagem para fim especial (LSP), semântica de base de dados e estudos do discurso.

- 6ª abordagem - estudos históricos: O estudo histórico de Heidegger contribuiu para a possibilidade relacional histórica e para um melhor entendimento a respeito do Mapa Conceitual. Essa contribuição se deu ao (ORC) organizarem e representarem o conhecimento por intermédio da sociedade científica (OC) pelo ato de classificação (conhecimento: conceitos) mediante conceitos já construídos para conceber novos conceitos (RC) por *representatio*, ao substituir um conceito para materializar o pensamento humano (SOC) por meio de processo conceitual, uma operação mental ao classificar conforme interpretação e entendimento de algum conteúdo temático em algum tempo.

- 8ª abordagem - estudos críticos e epistemológicos: A epistemologia, a partir de Aristóteles e Kant, fundamentou e sustentou as relações entre a teoria e a prática para a compreensão do processo de conhecimento elaborado por meio do Mapa Conceitual. Isso ocorreu porque: a (ORC) organização do conhecimento ordena o conceito envolvendo processo, entendimento e sistema para contribuir com a representação, por meio de uma (OC) visão teórico-metodológica para conhecer o (conhecimento: conceito), com a interferência do processo mental, para (RC) representar conceitos, estabelecendo vínculos estruturais, sistemáticos, de associação e de distinção, fazendo uso de (SOC), da figura da árvore do conhecimento, que, no século XVII, dá lugar ao uso do termo Organização do Conhecimento, o qual significa sistemas.

- 9ª abordagem - estudos terminológicos: O filósofo Peirce assegurou a representação pragmática do significado e da comunicação conceitual por meio Terminologia/terminologia pelo Mapa Conceitual, porque: pela (ORC), os conceitos relacionam-se aos termos que desenham o mundo por meio de valores

culturais de todas as épocas (OC) para a representação pela síntese, pela (conhecimento: conceito) comunicação de termos, por (RC) intermédio da linguagem (SOC) do termo que representa a temática do documento pela Linguagem Documentária pré-coordenada e pós-coordenada. Ao usar a Terminologia (Linguagem de Especialidade), como Terminologia (T), é campo de estudo ou disciplina por meio de instrumentos e denomina-se Macroestrutural = dedica-se a obter o mapeamento de um domínio. Como terminologia (t), é usada como fundamentos aplicados e refere-se a um conjunto de termos. A terminologia pontual é Microestrutural, é o estudo de um termo.

Uma abordagem ampara a outra, proporcionando uma análise distinta no contexto histórico e social, incluindo processos categoriais e cognitivos para enunciados lógicos de materialização na construção e representação de conhecimento.

4.3 Análise e Resultados pelos Eixos de Tennis (2012)

Tennis (2003) que se preocupa com os procedimentos de operacionalização de definição de domínios. Em razão disso, apresentam-se os elementos da epistemologia do Mapa Conceitual, por meio da historicidade dos conceitos para o estabelecimento de sua estrutura lógica no âmbito do SOC, para comporem a sua delimitação e a sua finalidade. Assim, para ajudar, foram propostos dois dispositivos analíticos chamados eixos.

- EIXO 1 - Áreas de Modulação (nomeação e extensão) tem como definição de domínio: O Mapa Conceitual como desenho de uma estrutura lógica gráfica que relaciona processos, categorias e conhecimento para organizar e representar a linguagem especializada de enunciado lógico de materialização de conceito. - Extensão ou Escopo e Alcance: Extensão: compõem-se pelo *corpus* da tese. Exclusão: não houve. Rótulo: refere-se ao Mapa Conceitual pela historicidade dos conceitos.

- EIXO 2 - Graus de Especialização (propósito e finalidade) tem como foco ou propósito: analisar a base epistemológica no domínio da OC, pelos filósofos Aristóteles, Kant, Heidegger e Peirce, e pelo subdomínio de SOC, por meio de processos, categorias e conceito – conhecimento. - Finalidade: responde à questão: “Quais as bases epistemológicas no domínio da OC que podem resultar na fundamentação da Epistemologia do Mapa Conceitual por meio da historicidade dos conceitos, com ênfase no processo de construção e sua aplicabilidade?”.

A epistemologia do Mapa Conceitual por meio da historicidade dos conceitos fundamenta-se tanto pela filosofia quanto pela ciência, a partir da elaboração de árvores do conhecimento. No domínio da OC, a epistemologia é um *continuum* para a construção de conceitos teóricos e aplicados para representação dos

significados. Os conceitos teóricos decorrem da intenção dos conceitos ordenados, ao longo do tempo, pela ciência. Já os conceitos aplicados surgem pela extensão dos conceitos em que se aprende a ordem do conhecimento semântico social na dimensão epistemológica pela história.

Observa-se que as posições epistemológicas influenciam na OC, por serem essenciais para desenhar o SOC. Os SOC são desenhados de forma teórico-metodológica para demonstrar a verdade, utilizando a linguagem, classificação e Sistema de Representação do Conhecimento ao analisar, aplicar, reaplicar e nomear, decompondo no tempo e ordenando no espaço.

Os SOC apresentam características metacientíficas, tais como: categorial, mental, histórica e representacional para analisar a forma de saber a verdade, por meio de mapas, que desenharam a complexidade do cérebro, compreensão do processo categórico de conhecimento e a representação e comunicação do conhecimento.

Assim, a epistemologia do conhecimento desenhado vem alicerçar conceitualmente a historicidade do Mapa Conceitual.

5 Considerações Finais

Apresentar a epistemologia do Mapa Conceitual por meio da historicidade dos conceitos para o estabelecimento de sua estrutura lógica no âmbito do SOC evidenciou que, os conceitos se deslocam, peregrinando de um local a outro sem limites, e assim, são ressignificados (FRANCELIN, 2010). O mapa do conhecimento especializado se move, buscando inserir novos conceitos e novos nomes em estruturas nocionais preexistentes. Em escala social, são produzidas terminologias alternativas. Inúmeras expressões também são criadas ou recriadas para dar nome a situações da realidade. (BARITÉ, 2022, *slides*).

O Apêndice A - Diagrama Lógico Conceitual da apresentação da Epistemologia do Mapa Conceitual por meio da historicidade dos conceitos representa o desenho da epistemologia do Mapa Conceitual, em que se exploraram as bases epistemológicas interdisciplinares para compreender e representar o Mapa Conceitual com o auxílio de estrutura arbórea, sistema ou produto categorial, processo cognitivo, probabilidade relacional histórica e representação do significado indicando um novo *status* para o Mapa Conceitual, ao ser amparado, por meio de sua estrutura lógica, no âmbito do SOC mediante a definição: Desenho de uma estrutura gráfica lógica que relaciona processos, categorias e conhecimento para organizar e representar uma linguagem especializada de enunciado lógico de materialização de conceito.

Assim, representado no Mapa Conceitual conforme o Apêndice B - Epistemologia do Mapa Conceitual por

meio da Historicidade dos Conceitos para o estabelecimento de sua Estrutura Lógica no âmbito do Sistema de Organização do Conhecimento com fundamentação na epistemologia do conhecimento desenhado, demonstrou-se o significado da estrutura lógica do Mapa Conceitual por meio da historicidade dos conceitos usando ações intelectuais, as quais permitiram compreendê-lo: de forma ampla, por ser estrutura e função dos SOC, e de forma restrita, por ser enunciado lógico de materialização do conceito revelando sua epistemologia por meio da historicidade dos conceitos.

Verifica-se que no tempo e no espaço é necessário saber o que é, como foi gerado e realizado o enunciado lógico de materialização de conceito. Logo, a epistemologia, proporcionou o ato de conhecer e apresentar as relações entre o saber construído e acumulado pelo homem organizado pelo mapa do saber fazendo uso de influências dos filósofos: Aristóteles, Kant, Heidegger e Peirce, além do intermédio da historicidade conceitual.

Este estudo busca-se contribuir com a Linha de Pesquisa Organização e Representação da Informação e do Conhecimento, no domínio da OC, no subdomínio do SOC, em razão de considerar o Mapa Conceitual como um Sistema de Organização do Conhecimento pela sua historicidade, sua estrutura lógica e suas aplicabilidades. Recomenda-se, com base nas epistemologias interdisciplinares e com o desenvolvimento teórico-metodológico, motivar, propiciar e ampliar o domínio estudado. O estímulo para estudos posteriores volta-se, em especial, para o Mapa Conceitual como SOC, intensificando, assim, seus diversos usos e funções de forma ampla ou de forma restrita.

Notas

[1] Este estudo é relativo à tese de doutorado concluída pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (PPGCI), originalmente apresentado tendo como referência: RODRIGUES, M. R. **Epistemologia do mapa conceitual em sistema de organização do conhecimento por meio da historicidade dos conceitos**. 2023. 237 f. Tese (Doutorado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual de Londrina, Londrina, 2023.

Referências

BARITÉ, Mario. Los conceptos y su representación: una perspectiva terminológica para el tratamiento temático de la información. *Scire*, Uruguay, v. 6, n. 1, p. 31-53, jun. 2000. Disponível em: <https://ibersid.eu/ojs/index.php/scire/article/view/1123/1105>. Acesso em: 01 mar. 2021.

BARITÉ, Mario. Organización del conocimiento: un nuevo marco teórico-conceptual en bibliotecología y

documentacion. In: CARRARA, Kester (Org.). **Educação, universidade e pesquisa**: textos completos do III simpósio em filosofia e ciência: paradigmas do conhecimento no final do milênio. Marília: Unesp-Marília-Publicacoes; São Paulo: FAPESP, 2001. p. 35-60.

BARITÉ, Mario. La garantía cultural como justificación en sistemas de organización del conocimiento: aproximación crítica. **Palabra Clave**, La Plata, Buenos Aires, v. 1, n. 1, p.2-11, 2011. Disponível em: <https://www.palabraclave.fahce.unlp.edu.ar/article/view/PCv1n1a02/1868>. Acesso em: 05 fev. 2023.

BARITÉ, Mario. **Las lecciones aprendidas por la terminología en la pandemia**. 2022. (slides).

CERVANTES, Brígida Maria Nogueira. **Terminología do processo de inteligência competitiva**: estudo teórico e metodológico. Londrina: EDUEL, 2006.

CERVANTES, Brígida Maria Nogueira. **A construção de tesouros com a integração de procedimentos terminográficos**. 2009. 198 f. Tese (Doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, 2009.

CONCEIÇÃO, Adriano Nardi, CORREIA, Paulo Rogério Miranda. Explicar é muito mais do que descrever: estratégias para estimular a criação de proposições dinâmicas. **Caminhos da Educação Matemática em Revista**: Ed. IFS: Aracaju, v.1, 2020. Disponível em: <https://mapasconceituais.com.br/pesquisa/artigos/explicar-e-muito-mais-do-que-descrever-estrategias-para-estimular-a-criacao-de-proposicoes-dinamicas/>. Acesso em: 03 ago. 2022.

FRANCELIN, Marivalde Moacir. Epistemologia da Ciência da Informação: evolução da pesquisa e suas bases referenciais. **Perspectivas em Ciência da Informação**, v. 23, n. 3, p.89-103, jul./out. 2018. Disponível em: http://old.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362018000300089. Acesso em: 09 jun. 2022.

FRANCELIN, Marivalde Moacir. **Ordem dos conceitos na organização da informação e do conhecimento**. 220f. Tese (Doutorado em Cultura e Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2010. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27151/tde-29102010-125600/publico/5504324.pdf>. Acesso em: 21 jun. 2021.

FRANCELIN, Marivalde Moacir. Ciência da informação e organização do conhecimento: tensões conceituais e dinâmicas epistemológicas para o século XXI. In: Encontro Nacional de Pesquisa em Ciência da Informação, 21. **Anais...** Rio de Janeiro: Rio de Janeiro: ANCIB. 2021.

GONZÁLEZ DE GOMEZ, Maria Nélide. A representação do conhecimento e o conhecimento da representação: algumas questões epistemológicas. **Ciência da Informação**, [S. l.], v. 22, n. 3, 1993. Disponível em: <http://revista.ibict.br/ciinf/article/view/479>. Acesso em: 06 out. 2020.

GORNER, Paul. **Ser e tempo**: uma chave de leitura. Petrópolis: Vozes, 2017. (Chaves de Leitura).

HJORLAND, Birger. Domain analysis in information Science: eleven approaches, tradicional as well as innovative.

Journal of Documentation: Dinamarca, n. 58, v. 4, 2002, p. 422-462. Disponível em: https://www.researchgate.net/publication/249366184_Domain_analysis_in_information_science_Eleven_approaches_-_Traditional_as_well_as_innovative. Acesso em: 06 out. 2020.

HJORLAND, Birger. Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science. **Information**, n. 12, v. 3, 2021. Disponível em: <https://doi.org/10.3390/info12030135>. Acesso em: 10 jun. 2021.

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny. **Introdução à Terminologia**: teoria & prática. 2. Ed. São Paulo: Contexto, 2020.

LIMA, Manuel. **The book of trees**: visualizing branches of knowledge. New York: Princeton Architectural Press, 2014.

MAZZOCCHI, Fulvio. **Knowledge organization system (KOS)**. ISKO Encyclopedia of Knowledge Organization (EIKO), 2017. Disponível em: <https://www.isko.org/cyclo/kos>. Acesso em: 10 jun. 2021.

MOREIRA, Walter. **Sistemas de organização do conhecimento**: aspectos teóricos, conceituais e metodológicos. 2018. 164f. Tese (livre-docência) - Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, 2018. Disponível em: <https://repositorio.unesp.br/bitstream/handle/11449/190878/0/00908213.pdf?sequence=1&isAllowed=y>. 22 nov. 2020.

NOVAK, Joseph Donald. **Aprender criar e utilizar o conhecimento**: mapas conceituais como ferramentas de facilitação nas escolas e empresas. Lisboa: Plátano Editora, 2000.

NOVAK, Joseph Donald; GOWIN, Dixie Bob. **Aprender a aprender**. Lisboa: Plátano Editora, 1984.

NOVAK, Joseph Donald; CAÑAS, Albert J. **The Theory Underlying Concept Maps and How to Construct and Use Them**. Florida: Institute for Human and Machine Cognition, 2008. Disponível em: <http://cmap.ihmc.us/Publications/>. Acesso em: 20 jul. 2013.

OKADA, Alexandra Lilaváti Pereira. **Cartografia investigativa: interfaces epistemológicas comunicacionais para mapear conhecimento em projetos de pesquisa**. 315f. Tese (Doutorado em Educação) - Pontifícia Universidade Católica de São Paulo, São Paulo, 2006. Disponível em: <https://repositorio.pucsp.br/jspui/handle/handle/9889>. Acesso em: 12 abr. 2022.

OKADA, Alexandra. O que é cartografia cognitiva e por que mapear redes de conhecimento? In: OKADA, Alexandra. **Cartografia Cognitiva**: mapas do conhecimento para pesquisa, aprendizagem e formação docente. Porto Cuiabá, Mato Grosso: KCM, 2008. p. 37-65. Disponível em: https://dl1wqtxts1xzle7.cloudfront.net/44475875/Cartografia_Cognitiva-libre.pdf?1459948258=&response-content-disposition=inline%3B+filename%3DCAPITULO_1_O_que_e_cartografi_a_cognitiv.pdf&Expires=1677854835&Signature=MhWZ1aDZUreQkr1VYyvm1RMa1~xQEHI1RkfPiIKDSy5Oo~EQesWk30HDsk3r73m5Lp1E0dCg2k3JDiDXzVh23Px9wknmaMxOmq1GRhWiXB7nUdJhrGMAQYOXDGRRD-F38P46lCpSecT-vJ1-V-uEsIbXSTLHpkMcQQ2GFuJ9inm1DyZTft5IITFJSnYnbY9p

y4pu7B8qyqoKkLLInrnAawHnvdLn2Mb9u5VgcZdOlkzLF7hPuo1LK0NSYJXtwQ-SXK6XKKBQ~xwI1zTZQ2vPeGMYhtTt-voaBcVL1Tawvl-YFtlGiye8pQcHshBPJO~rNWgQiUlyl3gjOZdXkSQ6KUG_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA. Acesso: 12 abr. 2022.

PEREIRA, Clevisson Junior. Conhecimento, racionalidade e simbólico: Aristóteles e Cassirer, distintas interpretações sobre o conhecer humano e suas possíveis reflexões na epistemologia geográfica. **GeoTextos**, v. 10, n. 2, dez. p. 161-189, 2014. Disponível em: <https://periodicos.ufba.br/index.php/geotextos/article/view/11273>. Acesso em: 18 jun. 2020.

PINOTTE, Guilherme Nicchio; CURY, Davidson; ZOUAQ, Amal. OntoMap: de mapas conceituais a ontologias OWL. In: **Nuevas Ideas en Informática Educativa**, 20. Anais... Santiago, Chile: TISE. 2015. Disponível em: <http://www.tise.cl/volumen11/TISE2015/TISE%202015.pdf>. Acesso em: 27 abril 2022.

RODRIGUES, Maria Rosemary; CERVANTES, Brígida Maria Nogueira. Os mapas conceituais e as múltiplas aplicações para a organização e representação do conhecimento. **Informatio**, Uruguay, v. 22, n. 2, p. 101-121. 2017.

SAN SEGUNDO MANUEL, R. Nueva concepción del conocimiento. In: RODRIGUES, G.M.; LOPES, I.L. (Org.). **Organização e representação do conhecimento na**

perspectiva da Ciência da Informação. Brasília: Thesaurus, 2003. (Estudos Avançados em Ciência da Informação, v.2).

SANTAELLA, Lúcia. **A teoria geral dos signos: semiose e autogeração**. São Paulo: Ática, 1995.

SOUZA, Rosali Fernandez de; SALES, Luana Farias; SALDANHA, Gustavo Silva (Org.). **Sistemas de Organização do Conhecimento: territórios**. Rio de Janeiro: IBICT, 2022. (Coleção PPGCI 50 anos).

STRAUSS, Anselm L; CORBIN, Juliet. **Basics of qualitative research: grounded theory, procedures and techniques**. Newbury: SAGE, 1990.

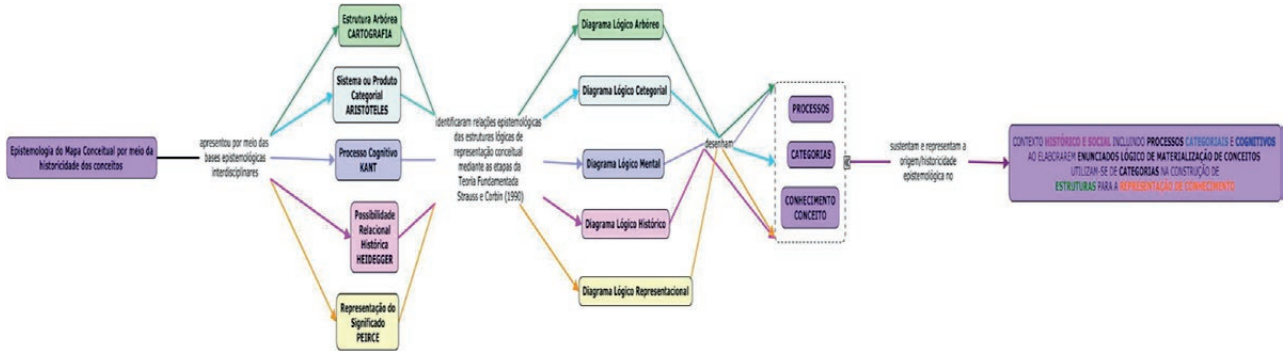
TENNIS, Joseph T. Two Axes of Domains for Domain Analysis. **Knowledge Organization**, v. 30, n. 3, n. 4, 2003. p. 191-195. Disponível em: https://www.researchgate.net/publication/244476999_Two_Axes_of_Domains_for_Domain_Analysis. Acesso em: 09 jul. 2020.

TENNIS, Joseph T. Com o que uma análise de domínio se parece no tocante a sua forma, função e gênero?. **Brazilian Journal of Information Science**, n. 1, v. 6 No 1, 2012.

ZENG, Marcia Lei. Knowledge organization systems (KOS). **Knowl. Org.** v. 35, n. 2, n. 3, p. 160-182. 2008.

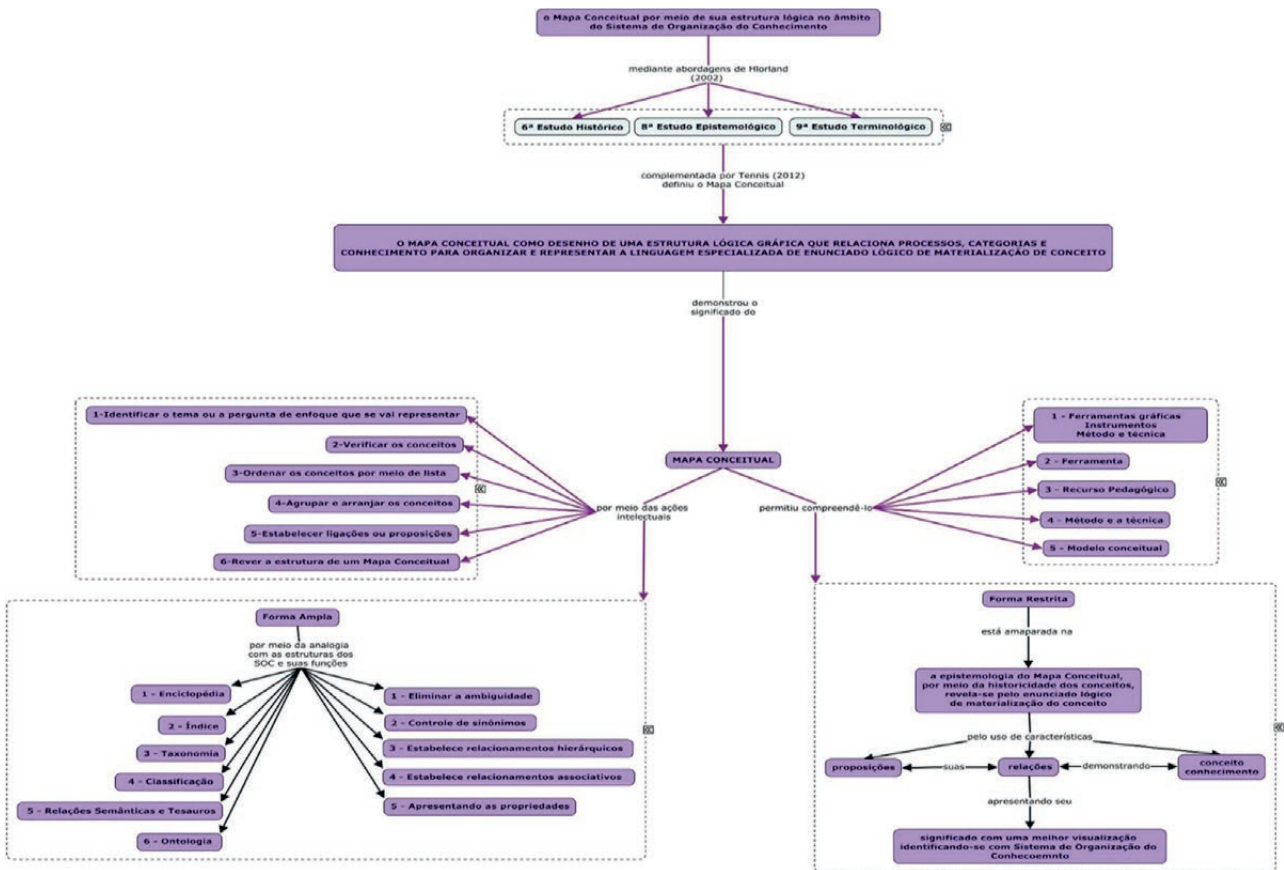
ZINS, Chaim. **Organização do Conhecimento: uma perspectiva epistemológica**. 2002.

Apêndice A – Diagrama Lógico Conceitual da apresentação da Epistemologia do Mapa Conceitual por meio da historicidade dos conceitos



Fonte: Elaborado pelas autoras.

Apêndice B – A Epistemologia do Mapa Conceitual por meio da historicidade dos conceitos para o estabelecimento de sua estrutura lógica no âmbito do Sistema de Organização do Conhecimento



Fonte: Elaborado pelas autoras.

Estratégias de pesquisa e análise de dados no domínio da Organização do Conhecimento: análise de abordagens metodológicas utilizadas nos trabalhos do eixo dimensão aplicada da ISKO-Brasil

Research strategies and data analysis in the field of Knowledge Organization: analysis of methodological approaches used in ISKO-Brasil's applied dimension work.

Bruna Lessa (1), Eneida Santana (2), Gercina Ângela de Oliveira Lima (3), Elisângela C. Aganette (4), Celsiane Araújo (5)

(1) Universidade Federal da Bahia (UFBA), Programa de Pós-graduação em Ciência da Informação, R. Cláudio Fontenelle, 1946 - Canela, Salvador - BA, 40110-070, brunalesa@ufba.br. (2) Programa de Pós-graduação em Difusão do Conhecimento, Avenida Reitor Miguel Calmon, Vale do Canela, s/n, Salvador, Bahia. 40110-100, eneidasantana@gmail.com. (3) Universidade Federal de Minas Gerais, Programa de Pós-Graduação em Gestão e Organização do Conhecimento, Av. Pres. Antônio Carlos, 6627 – Pampulha, Belo Horizonte – MG, 31270-901, glimaacci.ufmg.br. (4) elis@eci.ufmg.br, (5) celsianevaraujo@gmail.com

Resumo

Introdução: A *International Society for Knowledge Organization* (ISKO), por meio de ações para comunicação e divulgação científica, vem ressignificado de modo transversal os estudos que envolvem o domínio da Organização do Conhecimento, levando-se em consideração a construção teórica a partir das dimensões epistemológica, aplicada e cultural das pesquisas. **Objetivo:** Pretende-se com esse estudo, identificar as abordagens metodológicas utilizadas nos trabalhos apresentados no eixo Dimensão Aplicada, no capítulo brasileiro do Congresso da ISKO, entre os anos 2017 e 2021. **Metodologia:** O estudo é caracterizado como pesquisa exploratória, com abordagem quali-quantitativa. O procedimento para coleta e análise de dados foi realizado em três etapas, utilizando-se da técnica de análise de conteúdo, para categorização metodológica, e do *software Gephi* para visualização e análise das relações criadas entre as estratégias de pesquisa, e análises de dados, e as temáticas dos trabalhos coletados, representada em uma rede semântica de três modos. **Resultados:** Com o *corpus* de 61 trabalhos analisados, a *Pesquisa Exploratória* e a *Pesquisa bibliográfica e documental*, são as estratégias de pesquisa mais adotadas entre os estudos. As técnicas para análises de dados que possuem maior conexão na rede estão associadas a base teórica da Organização Conhecimento, evidenciando práticas também adotadas no âmbito da Arquivologia. **Conclusão:** A caracterização tridimensional da rede evidenciou a inexistência de um padrão de uso de estratégias de pesquisa e análise de dados. Os temas mais conectados, distribuídos por estratégias de pesquisa, apontam para tendências de pesquisa, e relações interdisciplinares na dimensão aplicada dos trabalhos apresentados nas últimas edições da ISKO-Brasil.

Palavras-chave: ISKO-Brasil. Dimensão Aplicada. Abordagem metodológica. Organização do conhecimento. Rede de três modos.

Abstract

Introduction: The *International Society for Knowledge Organization* (ISKO), through actions for communication and scientific dissemination, has re-signified in a transversal way, the studies that involve the domain of Knowledge Organization, taking into account the theoretical construction from the epistemological, applied and cultural dimensions of research. **Objective:** This study aims to identify the methodological approaches used in the works presented in the Applied Dimension axis, in the Brazilian chapter of the ISKO Congress, between the years 2017 and 2021. **Methodology:** The study is characterized as exploratory research, with a quali-quantitative approach. The procedure for data collection and analysis was carried out in three stages, using the content analysis technique, for methodological categorization, and the Gephi software for visualization and analysis of the relationships created between research strategies, data analysis, and the thematic of the collected works, represented in three-mode semantic network. **Results:** With the corpus of 61 studies analyzed, Exploratory Research, Literature and Documentary Research are the most adopted research strategies among the studies. The techniques for data analysis that have a greater connection in the network are associated with the theoretical basis of the Knowledge Organization, evidencing practices also adopted in the scope of Archivology. **Conclusion:** The three-dimensional characterization of the network evidenced the non-existence of a pattern of use of research strategies and data analysis. The most connected themes, distributed by research strategies, point to research trends, and interdisciplinary relations in the applied dimension of the works presented in the last editions of ISKO-Brazil.

Keywords: ISKO-Brasil. Dimension Applied. Methodological approach. Organization of knowledge. Three-mode network.

Resumen

Introducción: La Sociedad Internacional para la Organización del Conocimiento (ISKO), a través de acciones de comunicación y divulgación científica, viene replanteando transversalmente los estudios que involucran el dominio de la Organización del Conocimiento, teniendo en cuenta la construcción teórica a partir de las dimensiones aspectos epistemológicos, aplicados y culturales de la investigación. **Meta:** La meta de este estudio es identificar los enfoques metodológicos utilizados en los trabajos presentados en el eje Dimensión Aplicada, en el capítulo brasileño del Congreso ISKO, entre los años 2017 y 2021. **Metodología:** El estudio se caracteriza como investigación exploratoria, con un enfoque cualitativo-cuantitativo. El procedimiento de recolección y análisis de datos se realizó en tres etapas, utilizando la técnica de análisis de contenido para la categorización metodológica, y el software Gephi para la visualización y análisis de las relaciones creadas entre las estrategias de investigación, el análisis de datos y las temáticas de los trabajos recolectados, representados en una red semántica de tres maneras. **Resultados:** Con un corpus de 61 trabajos analizados, la Investigación Exploratoria y la Investigación Bibliográfica y Documental son las estrategias de investigación más adoptadas entre los estudios. Las técnicas de análisis de datos que tienen mayor conexión en la red están asociadas a la base teórica de la Organización del Conocimiento, evidenciando prácticas también adoptadas en el ámbito de la Archivología. **Conclusión:** La caracterización tridimensional de la red mostró la falta de un uso estándar de estrategias de investigación y análisis de datos. Los temas más conectados, distribuidos por estrategias de investigación, apuntan tendencias de investigación y relaciones interdisciplinarias en la dimensión aplicada de los trabajos presentados en las últimas ediciones de ISKO-Brasil.

Palabras clave: ISKO-Brasil. Dimensión aplicada. Enfoque metodológico. Organización del conocimiento. Red de tres modos.

1 Introdução

A Organização do Conhecimento (OC) é uma área que investiga a organização e representação da informação com a proposta de melhorar a recuperação da informação em sistemas de informação. Para isso, conta com teorias e métodos, a fim de avançar nas práticas que permeiam as pesquisas que tratam sobre como e por que organizar um domínio de conhecimento. Na perspectiva de Hjørland (2003), a organização do conhecimento abordada pela Biblioteconomia e pela Ciência da Informação, foca nos estudos referentes a processos e às construções instrumentais da organização do conhecimento, tais como descrição de documentos, classificação e indexação.

Nesse contexto, a *International Society for Knowledge Organization* - ISKO, fundada no final da década de 1980, tendo como idealizadora e precursora a pesquisadora Ingetraut Dahlberg, vem se consolidando como referência nas ações da comunicação científica do domínio da Organização e Representação do Conhecimento, a qual envolve, além da Ciência da Informação e a Biblioteconomia, pesquisadores de distintas áreas de estudo, tais como Filosofia, Sociologia, Ciência da Computação, Informática, Linguística, História, Matemática, Comunicação, entre outros (ARBOIT, 2014).

Desde sua criação, a ISKO, especificamente a ISKO-Brasil, precursora na proposição de dimensões epistemológica, aplicada e cultural, além da discussão em nível internacional, por meio dos capítulos nacionais, contribui para a construção e revisão permanente de questões conceituais, epistemológicas e técnicas da organização do conhecimento. Nesse percurso, a organização de temas interdisciplinares a partir de dimensões de estudo e abordagem - epistemológica, aplicada e sociocultural, traduz-se na constante resignificação e, conseqüente, validação da área como um campo de estudo. De fato, o crescimento

da produção científica possibilitou ampliar os estudos de maneira transversal, gerando, no entanto, complexidade na totalidade das abordagens metodológicas, modelos, produtos e instrumentos utilizados, com ênfase especial para a dimensão aplicada das pesquisas no contexto da Organização do Conhecimento (OC).

Dessa maneira, vários métodos têm sido aplicados no âmbito da OC, desde a análise qualitativa, quanto métodos bibliométricos e processamento de linguagem natural, que tem orientado a direção das pesquisas realizadas pelos pesquisadores da área. Desperta-se, portanto, o interesse e motivação de pesquisa que contempla a problemática por análises mais detalhadas sobre as características e tendências utilizadas nas abordagens metodológicas nos estudos em OC. Para iniciar tal discussão, e ampliar esse espaço investigativo, esta pesquisa buscou identificar as abordagens metodológicas utilizadas nas pesquisas aplicadas à Organização do Conhecimento, a partir dos estudos apresentados no capítulo brasileiro do Congresso da ISKO, dos três últimos eventos realizados nos anos 2017-2021.

Neste estudo, de modo específico, caracterizado como pesquisa exploratória, com abordagem qualitativa, procura-se analisar as relações entre as estratégias de pesquisa e as análises de dados utilizadas no *corpus* desta pesquisa, ordenando os estudos analisados conforme categoria metodológica, a fim de possibilitar a visualização da interconexão existente entre as abordagens utilizadas e suas aplicações práticas, além das possíveis relações interdisciplinares na dimensão prática dos trabalhos da ISKO-Brasil.

2 Bases teórico-metodológicas em Organização do Conhecimento

Sabe-se que os estudos sobre a organização do conhecimento, oriunda dos princípios da classificação,

remontam-se a Platão e Aristóteles, porém, o termo organização do conhecimento foi referenciado, primeiramente, por volta de 1900, com os autores Cutter (1879), Richardson (1901), Sayers (1915), e Bliss (1929) como o uso das expressões “classificação do conhecimento”, “ordem das coisas”, “ordem das ideias” no âmbito dos estudos e construções de sistemas de classificação. Com seus estudos, esses autores foram os precursores da abordagem teórica na área, pois consideraram que toda classificação deveria ser tanto sistemática quanto conceitual.

Nos anos de 1970, novos estudos começaram a surgir principalmente com Dahlberg (1993,1995,2006) e Soergel (1971), nos quais ambos autores, também, comungavam com a ideia que a teoria da classificação deve-se basear em estruturas conceituais e formação de conceitos.

O termo “Organização do Conhecimento”, utilizado na *International Society of Knowledge Organization* (ISKO), criada em 1989 por Ingetraut Dahlberg, está diretamente relacionado ao seu uso já nos anos 1929 e 1933 por Henry Evelyn Bliss (DAHLBERG, 2006), e que passa a ter caráter de disciplina científica, representando um domínio do conhecimento, que pode colaborar em várias esferas sociais, conforme Dahlberg (1993), tais como na política e na indústria. Na ocasião do texto publicado na década 1990, a autora elenca 10 (dez) grupos principais de estudos em OC, indicando não somente suas relações epistemológicas, mas suas aplicações que vão desde o ensino e serviços bibliográficos, ao desenvolvimento de produtos e sistemas em OC. Nesta perspectiva, os estudos na área deveriam se preocupar em como as unidades do conhecimento estão relacionadas e organizadas dentro de um determinado contexto, compreendendo suas características e aplicações.

Em 2006, Dahlberg define a Organização do Conhecimento (OC) como uma ciência que ordena a estruturação e sistematização dos conceitos, de acordo com suas características, que podem ser definidas como elementos de herança do objeto, e a aplicação dos conceitos e classes dos conceitos ordenados pela indicação de valores, dos referentes conteúdos dos objetos ou assuntos (DAHLBERG, 2006).

Nesse contexto, Hagar Gomes (2009), em seu estudo “Tendências da pesquisa em organização do conhecimento”, aponta que:

A OC é uma área de múltiplas aplicações. As pesquisas especificamente na área apontam para, no mínimo, dois aspectos importantes: de um lado, a abordagem teórica-filosófica voltada para questões epistemológicas ao analisar um domínio juntamente com as aplicações na organização de sistemas/tabelas de classificação e, de outro lado, os métodos empregados na elaboração propriamente dita dos sistemas, atividades que

alguns autores consideram complementar, mas que tem suas bases teóricas já sedimentadas (GOMES, 2009, p.63).

Em um contexto mais social, Hjørland (2013) afirma que

[...] o conhecimento é criado por humanos para alguns propósitos específicos e serve a alguns interesses melhor do que a outros. Conceitos e relações semânticas não são a priori ou neutros, mas devem ser examinados em relação às suas implicações para os usuários que eles devem servir (HJORLAND, 2013, p.174).

Seguindo as discussões em torno da construção teórica e prática no âmbito das pesquisas sobre OC, sobretudo porque dão base para a estruturação da comunicação científica da ISKO-Brasil, Guimarães (2017) apresenta a evolução das abordagens de investigação, caracterizadas em três dimensões: epistemológica, aplicada e cultural.

Em relação à dimensão aplicada, especialmente destaca-se neste estudo, que para Guimarães (2017) há

[...] perspectivas promissoras como a promoção cada vez mais intensa da interoperabilidade de sistemas e a promoção de uma recuperação da informação rápida, eficaz e **culturalmente significativa** (GUIMARÃES, 2017, p.92, destaque nosso).

A autoria ainda indica, como possível desafio para a aplicação das técnicas e teorias da OC, a gestão dos processos de organização e de representação do conhecimento, que envolve práticas mais minuciosas de identificação de volumes informacionais e, acrescenta-se aqui, sobretudo no contexto virtual, diretamente relacionado à criação, desenvolvimento e uso de instrumentos que venham a viabilizar este gerenciamento.

Acredita-se que, tal concepção sobre o que se considera [e o que se espera] da dimensão aplicada na OC, em destaque no contexto brasileiro, predominantemente alinhada ao campo da Ciência da Informação, coaduna-se a perspectiva de Hjørland (2003), de que as abordagens metodológicas na OC, em suas multiplicidades orientadas nas funções dentro de um sistema e/ou comunidade, devem seguir a construção epistemológica da área, considerando a evolução dos significados da linguagem e produtos socioculturais.

Sob uma análise mais detalhada das abordagens identificadas nas pesquisas em OC, Barros e Vital (2019), apontam para os enfoques metodológicos adotados nessas pesquisas, a partir de trabalhos apresentados no Grupo de Trabalho 2 do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB) entre os anos 2013 e 2017. Para as autoras, o resultado deste estudo revela que a maioria dos métodos identificados não estão diretamente associados à construção científica da OC.

Na próxima seção, apresenta-se o percurso metodológico deste estudo, seguindo-se a análise e discussão dos dados.

3 Procedimentos metodológicos

O *corpus* deste estudo reúne os trabalhos apresentados no eixo Dimensão Aplicada da Organização do Conhecimento, do capítulo brasileiro do Congresso da *International Society for Knowledge Organization* (ISKO), nos anos 2017, 2019 e 2021. A escolha deste eixo, justifica-se por meio do objetivo proposto neste estudo, que conforme já mencionado, consiste em identificar as abordagens metodológicas utilizadas nas pesquisas aplicadas na Organização do Conhecimento e suas relações com os objetos de estudo.

A partir da coleta nos Anais dos Congressos realizados em 2017 e 2019, e nos resumos das apresentações realizadas em 2021 (quando o evento foi remoto em função da pandemia da Covid-19), cedidas pela Diretoria ISKO-Brasil, formou-se o conjunto de dados com 61 trabalhos no eixo analisado. O conjunto de dados inclui informações sobre autorias, instituição, título, abordagens metodológicas, e palavras-chave.

Fez-se a leitura integral dos textos, pois nem sempre os resumos indicavam com detalhes as abordagens metodológicas utilizadas, sendo necessário, em alguns casos, a tradução dos termos apresentados no percurso metodológico adotado nestas pesquisas, com base na literatura sobre metodologia científica (CERVO, 2007; MARCONI; LAKATOS, 2021; PINTO; VIDOTTI; CAVALCANTE, 2015).

Optou-se pela modelagem dos dados coletados, sobre as abordagens metodológicas, com base em dois aspectos:

- a) a maneira que a investigação é conduzida para obtenção e análise dos dados - estratégia de pesquisa (GIL, 2022; LAKATOS; MARCONI, 2021; MARTINS; THEÓPHILO, 2007);
- b) As técnicas, práticas, processos e instrumentos utilizados para apresentação do fenômeno investigado - análise de dados.

Nesse contexto, identificar a abordagem utilizada para visualização e interpretação desses dados, este estudo contemplou 3 (três) etapas:

1. Análise de conteúdo (BARDIN, 2011): categorização dos elementos textuais que indicavam as abordagens metodológicas, após a leitura integral dos manuscritos, criaram-se categorias de abordagens metodológicas, quanto a estratégia de pesquisa e análise de dados;
2. Tratamento textual: tradução do texto para extração de termos semânticos dos títulos, com

base em vocabulários controlados – organização dos termos em um arquivo de texto, de formato *.txt*;

3. Tratamento computacional: para criação da rede e visualização das relações, criou-se uma planilha (em formato *.csv*), para organizar os relacionamentos entre os termos, aplicando-se algoritmos para análise estrutural a partir de suas relações e similaridades por meio da abordagem de análise de redes sociais (ARS).

Ressalta-se que os trabalhos analisados em idioma estrangeiro, traduziu-se para língua portuguesa a fim de padronizar os termos na visualização e análise das relações criadas na rede.

Para a etapa do tratamento computacional, considerou a Teoria de Redes (WASSERMAN; FAUST, 1994; NEWMAN, 2003; BARABÁSI, 2016) para análise dos índices e visualização da rede criada, utilizando-se o *software Gephi*.

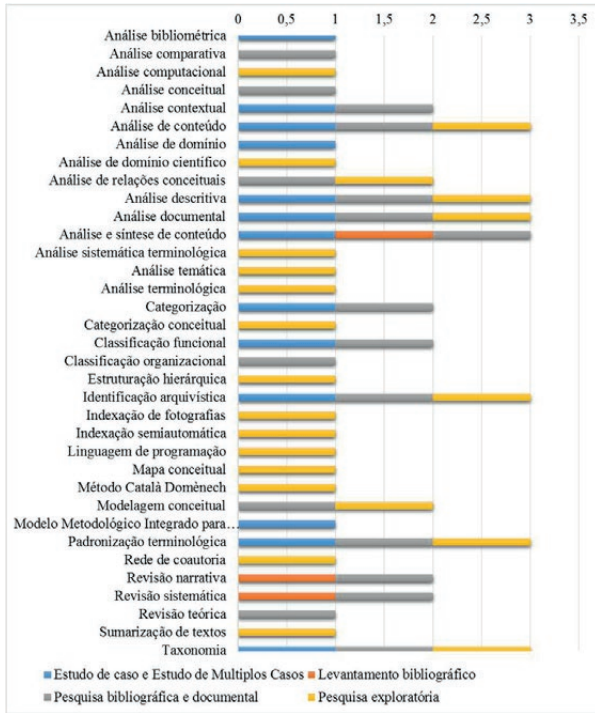
4 Apresentação e discussão dos dados

Após estruturação do conjunto de dados, considerando o aspecto quantitativo das abordagens metodológicas identificadas nos trabalhos, verificou-se que 33 estudos, em relação às estratégias de pesquisa adotadas, apresentam-se como *Pesquisa exploratória*, seguindo-se *Pesquisa bibliográfica e documental* (21), *Estudo de caso* (18), e *Levantamento bibliográfico* (4). Em alguns dos estudos analisados, o uso de tais estratégias são apresentadas de maneira múltipla, justificando com isso o somatório maior que o número de trabalhos coletados.

Para melhor compreensão das conexões entre as estratégias de pesquisa e análise de dados utilizados nos trabalhos, e possibilitar a análise sobre as categorias identificadas, utilizou para representação a matriz de adjacência, conforme o modelo apresentado na Figura 1.

Na extração apresentada na Figura 1, percebe-se, por exemplo, que *Análise de conteúdo*, *Análise descritiva*, *Análise documental*, *Análise e síntese de conteúdo*, *Identificação arquivística*, *Padronização terminológica* e *Taxonomia* são as técnicas para análises de dados mais utilizadas, ou ainda, que possuem maior conexão, com as estratégias de pesquisa identificadas nos trabalhos. Do mesmo modo, a estratégia *Pesquisa Exploratória* é a que mais se relaciona, de maneira única, com as técnicas, práticas, processos e instrumentos utilizados para apresentação dos dados investigados.

Figura 1. Gráfico da Matriz de adjacência das relações entre as abordagens metodológicas



Fonte: Autores.

A partir disso, a representação destas conexões e suas propriedades estruturais, ao ser analisada em rede, pode ser visualizada por meio da rede semântica criada dos termos categorizados no âmbito das abordagens metodológicas identificadas nos trabalhos e as temáticas dos estudos.

Na Figura 2 (Apêndice A), a rede de três modos (*three-mode*) entre abordagens metodológicas e trabalhos científicos da ISKO-Brasil (2017-2021), compreende as conexões (arestas) entre vértices (termos) de diferentes categorias (estratégia de pesquisa, análise dos dados e temáticas dos estudos) do sistema (rede), caracterizada como complexa devido às multiplicidades presentes no contexto dessas interações.

A rede é composta por 243 vértices (termos presentes na rede) e 2411 arestas (relação entre os termos). As características de tridimensionalidade da rede são refletidas na medida estatística do Grau $\langle k \rangle$ em seu maior $24,5 \langle k \rangle$. O grau de uma rede tridimensional refere-se ao número de conexões que cada modo (ou dimensão) da rede possui. Portanto, se uma rede *three-mode* possui grau $24,5 \langle k \rangle$, significa que cada modo da rede possui 24,5 conexões, demonstrando a robustez da leitura da rede por modos quando conectados entre si, visto que redes com grau mais elevado tendem a ser mais densas e conectadas, enquanto redes com grau mais baixo podem ter maior variação na sua conectividade e densidade.

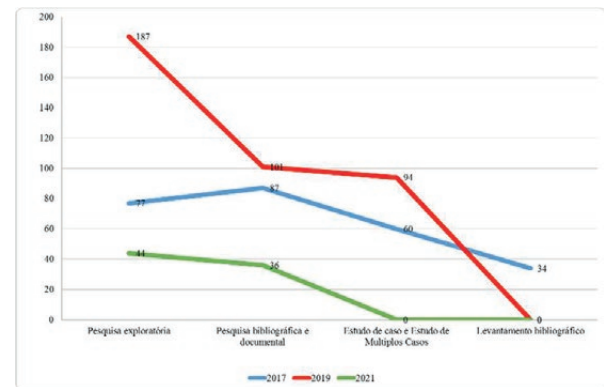
As conexões entre os modos da rede possuem direcionamento onde cada vértice pode ter um ou mais vértices que apontam para ele. Neste estudo, os vértices predecessores são categorizados por meio das abordagens metodológicas, enquanto estratégias de pesquisa e análise de dados, seguidos dos temas abordados que compõem os vértices sucessores da rede. Essa direção nas arestas pode ser utilizada para modelar sistemas que possuem fluxos unidirecionais de informação, tais como apresentada nesta rede.

Após análise dos termos individualizados da Figura 2, observou-se que a medida algorítmica do *Clustering Coefficient* $\langle c \rangle$ dos vértices possui variação entre $0,076 \langle c \rangle$ - $0,128 \langle c \rangle$, apontando que apenas cerca de 12,8% dos pares de vizinhos de um nó estão conectados entre si. Isso preconiza que a rede tem uma conectividade relativamente baixa entre os vértices adjacentes, o que pode ser indicativo de uma rede com estrutura mais dispersa ou heterogênea, ou seja, conclui-se que não há um padrão relacional entre as estratégias de pesquisa, o uso de técnicas para análise de dados, associados a temas de pesquisa.

A classificação da rede como heterogênea também é identificada na Distribuição de Grau $\langle k \rangle$ apresentada nas Figuras 3 e 4. Ao associar as medidas $\langle k \rangle$ e $\langle c \rangle$ identifica-se a estrutura topológica de livre escala o que garante a tendência de redes mais robustas.

A distribuição de grau $\langle k \rangle$ apresentada na Figura 3, evidencia significativa variação nas adesões às estratégias de pesquisa pelos participantes da ISKO-Brasil.

Figura 3. Distribuição de grau - Estratégias de pesquisa (tendências)



Fonte: Autores.

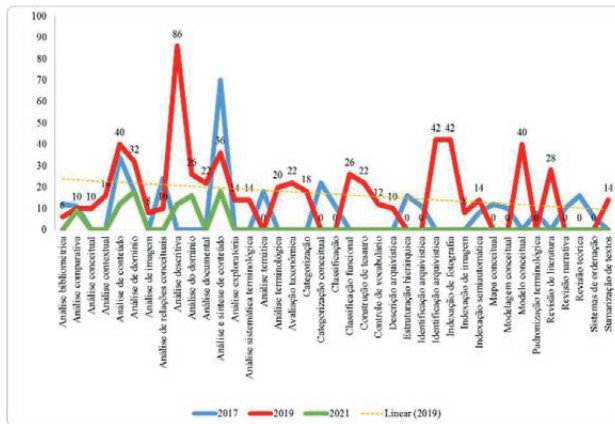
Verifica-se que na categoria estratégias de pesquisa, por exemplo, *Levantamento bibliográfico*, estabeleceu $187 \langle k \rangle$ conexões com outras temáticas no ano de 2019, tendo uma redução significativa de conexões com outros termos das pesquisas no ano de 2021. Já *Pesquisa bibliográfica e documental* $\langle 101 \rangle$, apresentou um representativo aumento de conexões, despontando como

uma tendência significativa para os trabalhos apresentados na dimensão aplicada da ISKO-Brasil.

Outra tendência caracterizada pela distribuição de grau <k>, está relacionada ao quantitativo de conexões das abordagens para análise dos dados de pesquisa. Levando-se em consideração a limitação da pesquisa, buscou-se identificar tal tendência por meio do corte de linearidade marcada entre conexões de 10<k> e 16<k>.

E ainda, na Figura 4, vê-se que *Análise e síntese de conteúdo*, *Análise exploratória*, *Análise sistemática terminológica*, *Análise temática*, *Análise terminológica*, *Avaliação taxonômica*, *Categorização*, *Categorização conceitual*, *Controle de vocabulário*, *Descrição arquivística*, *Mapa conceitual* e *Sistemas de ordenação*, são abordagens significativas para a categoria análise dos dados, dos estudos aplicados, apresentados no ISKO-Brasil (2017 a 2021).

Figura 4. Distribuição de grau– *Análise de dados* (tendências)



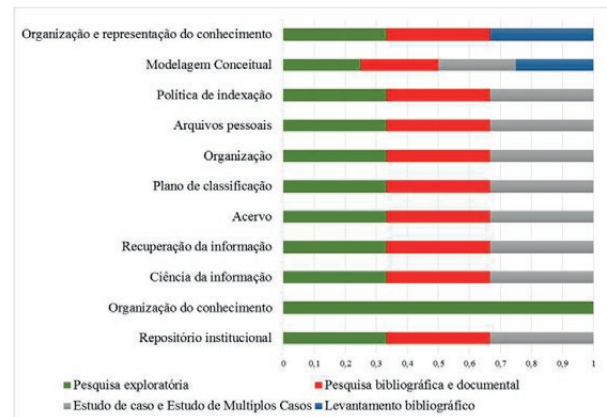
Fonte: Autores.

Já na Figura 5, apresenta-se os 10 temas mais conectados na rede, ou seja, aqueles que possuem um maior número de conexões com outros vértices na rede, e relação direta com as estratégias de pesquisa e análise dos dados.

Os temas com maiores conexões (relacionados) são propagadores de informações da rede, visto que a alta conexão entre eles e os outros vértices, pode afetar a velocidade com que se tenha acesso a outro tema, atuando como "pontes" entre partes distintas da rede.

Este estudo, traz ainda como resultado, a análise de rede ego, apresentada na Figura 6 (Apêndice B), a rede ego representa a relação da temática Modelagem Conceitual, 40<K>. Uma rede ego oriunda de uma rede *three-mode* é segundo Brandes e Erlebach (2005), uma rede de múltiplos modos, é uma sub-rede que representa as conexões entre um conjunto específico de vértices de um dos modos, e os vértices do outro modo. Essa rede permite analisar a estrutura e a importância das conexões entre nós.

Figura 5. Temas mais conectados distribuídos por Estratégias de pesquisa - Distribuição de grau



Fonte: Autores.

Para ilustrar, na Figura 6, tem-se a rede ego em primeira camada com propagação dos 20 principais vértices. Entre eles, pode-se observar a relação com as estratégias da pesquisa bibliográfica e da pesquisa exploratória.

Destaca-se aqui as conexões do vértice (termo) *Modelagem conceitual*, pois além de figurar como um dos temas mais conectados na rede criada, é um processo que utiliza de bases teórico-práticas interdisciplinares. Uma das técnicas voltadas para análise cognitiva utilizadas na Organização do Conhecimento, tal como empregada neste estudo, por meio da organização das abordagens metodológicas em categorias para representação e visualização de suas relações.

3 Conclusão

Este estudo buscou identificar a interconexão entre as estratégias de pesquisa e as análises de dados utilizadas nos estudos da dimensão aplicada, apresentados no capítulo brasileiro da ISKO-Brasil (2017-2021). Com o *corpus* de 61 trabalhos analisados, foi possível perceber o uso da *Pesquisa Exploratória* e *Pesquisa bibliográfica e documental*, como estratégias de pesquisa mais adotadas entre os trabalhos analisados.

Como tendência no âmbito das técnicas e processos voltados para análise dos dados, as abordagens estão diretamente relacionadas às bases teóricas da OC, com destaque aqui para *Identificação arquivística*, *Padronização terminológica* e *Taxonomia*, indicando a construção teórica para a Organização do Conhecimento Arquivístico, já que tais termos estão relacionados no *corpus* analisado a trabalhos no contexto da Arquivologia.

Com uso da ARS, realizou-se a modelagem dos dados por meio da rede semântica, possibilitando visualizar e explorar relações entre as abordagens metodológicas e as temáticas dos trabalhos. A caracterização topológica

da rede indicou a heterogeneidade nas relações entre os termos, evidenciando a inexistência de um padrão de uso de estratégias de pesquisa e análise de dados, relacionados às temáticas de pesquisas em OC, sobretudo na dimensão aplicada.

Os temas mais conectados, distribuídos por estratégias de pesquisa, apontam para tendências de assuntos nas últimas edições da ISKO-Brasil com maior aplicação prática.

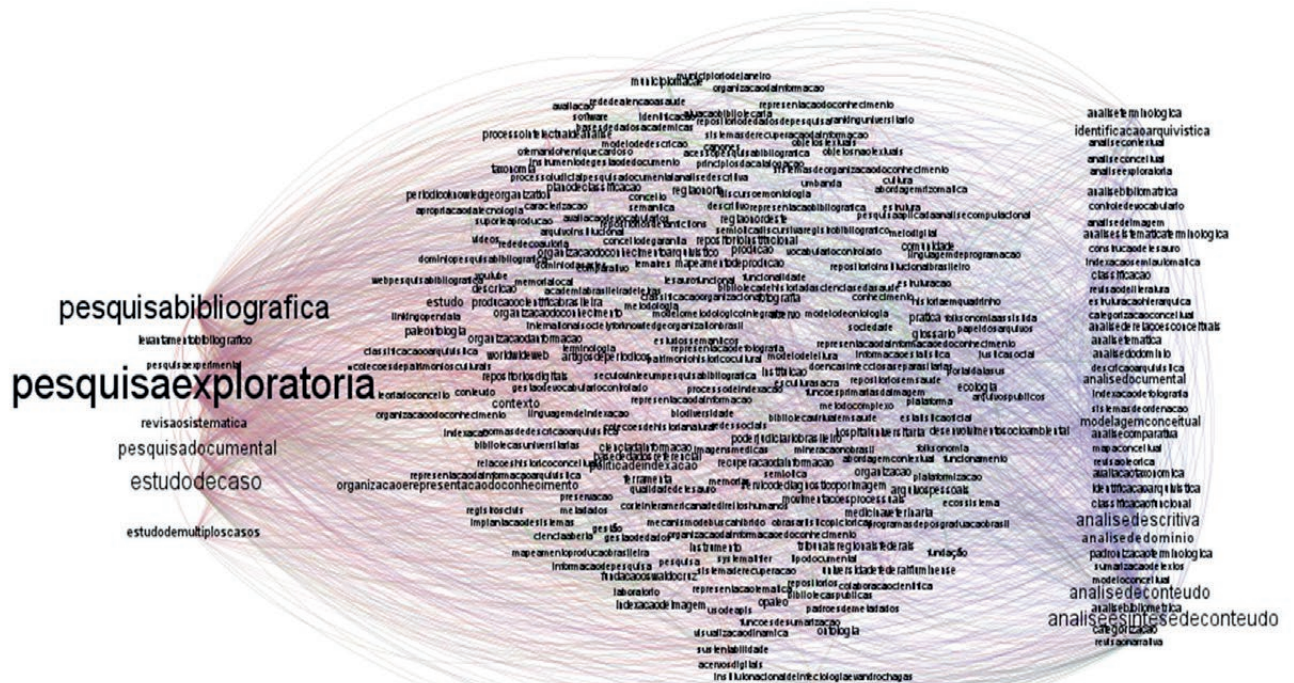
Limita-se a esta pesquisa outras análises sobre o delineamento das técnicas utilizadas e suas bases teóricas com a OC, além das tendências de uso a partir do vínculo dos pesquisadores e suas instituições/grupos de pesquisa, que se espera incluir em estudos futuros.

Referências

- BARABÁSI, A.-L. **Network Science**. Cambridge: Cambridge University Press, 2016. Disponível em: <http://networksciencebook.com>. Acesso em: 10 abr. 2021.
- BARDIN, Laurence. **Análise de conteúdo**. São Paulo: Edições 70, 2011.
- BARROS, C. M.; VITAL, L. P. Abordagens metodológicas das pesquisas em organização e representação do conhecimento no contexto brasileiro. **Revista Digital de Biblioteconomia & Ciência da Informação**, Campinas, SP, v. 17, p. e019018, 2019. DOI: 10.20396/rdbci.v17i0.8653734. Acesso em: 23 mar. 2023.
- BRANDES, U.; ERLEBACH, T. **Network Analysis: Methodological Foundations**. Springer Science & Business Media, 2005.
- CERVO, A. L. **Metodologia Científica**. 6. ed. São Paulo: Pearson Prentice Hall, 2007.
- CUTTER, C. A. Classification on the Shelves: with some account of the new scheme prepared for the Boston **Athenaeum**. **Library Journal**, 4, p. 234-243, 1879.
- BLISS, H. E. **The organization of knowledge and the system of sciences**. New York: Henry Holt & Co., 1929.
- DAHLBERG, I. Knowledge organization: its scopes and possibilities. **Knowledge Organization**, Frankfurt, v. 20, n. 4, p. 211-222, 1993.
- DAHLBERG, I. Current trends in knowledge organization. In: GARCIA MARCO, Francisco Javier (org.). **Organización del conocimiento en sistemas de información y documentación**. Zaragoza: Universidad de Zaragoza, v.1, p. 7-26, 2005.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, Frankfurt, v. 33, n. 1, p. 11-19, 2006.
- GUIMARÃES, J. A. C. Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. **Informação & Informação**, v. 22, n. 2, p. 84-98, 2017. DOI: 10.5433/1981-8920.2017v22n2p84. Acesso em: 23 mar. 2023.
- GOMES, H. E. Tendências da pesquisa em organização do conhecimento. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 2, n. 1, p. 60-88, 2009. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/169>. Acesso em: 23 mar. 2023.
- HJØRLAND, B. Classification. **Knowledge Organization**, Würzburg, v.30, n.2, p.87-111, 2003.
- HJØRLAND, B. Theories of knowledge organization: theories of knowledge. **Knowledge Organization**, v. 40, n. 3, p. 169-181, 2013.
- MARCONI, M. A.; LAKATOS, E. M. **Técnicas de pesquisa**. 9. ed. São Paulo: Atlas, 2021
- RICHARDSON, E. C. **Classification: theoretical and practical**. New York: Charles Scribner's Sons, 1901.
- SAYERS, W. E. B. **Canons of classification applied to "the subject", "the expansive", "the decimal", and "the Library of Congress" classifications: a study of bibliographical classification method**. London: Grafton & Co., 1915.
- WASSERMAN, S.; FAUST, K. **Social Network analysis: methods and applications**. Cambridge University Press, Cambridge, 1994.

Apêndice A

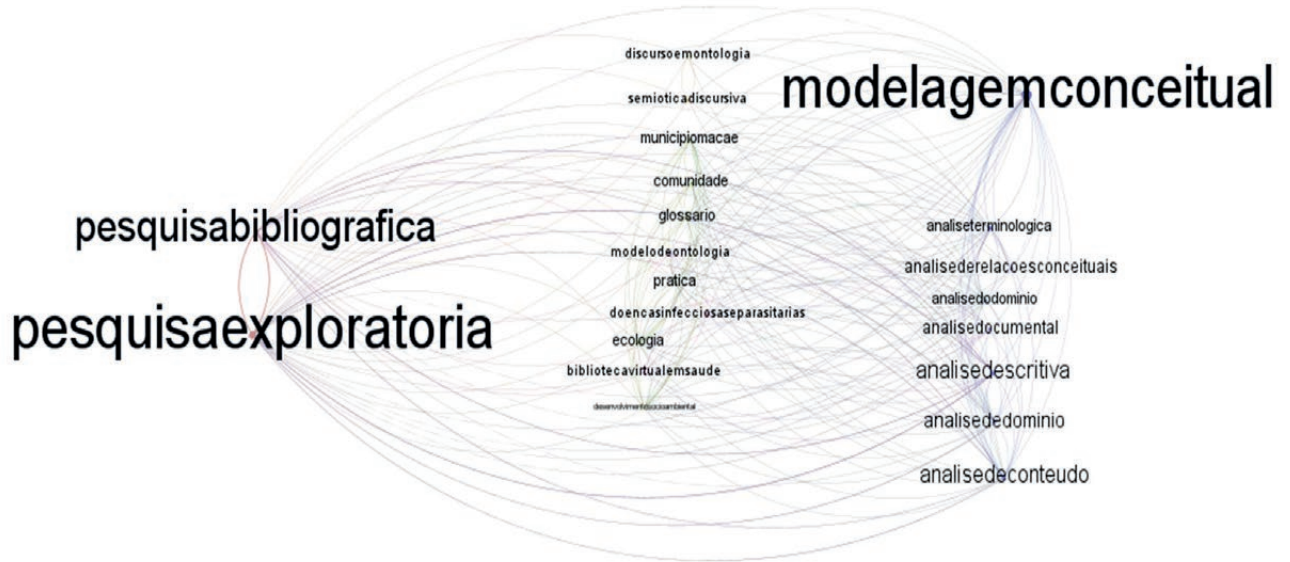
Figura 2. Rede three-mode entre abordagens metodológicas e trabalhos científicos da ISKO-Brasil (2017-2021)¹



¹ Link para navegação na rede encr.pw/redeISKOBrasil

Apêndice B

Figura 6. Sub-rede three-mode para o termo “Modelagem conceitual”



Ética da informação nos compiladores de tendências das plataformas de mídia digital

Information ethics in the trend compilers of digital media platforms

Sergio Mari Junior (1), Rosane Suely Alvares Lunardelli (2), Francisco Carlos Paletta (3)

(1) Universidade Estadual de Londrina (UEL), Londrina/PR, prof.sergiomari@gmail.com, (2) lunardelli@uel.br, (3) Universidade de São Paulo (USP), São Paulo/SP, fcpaletta@usp.br

Resumo

Introdução: Plataforma é um modelo de negócio amplamente adotado por empresas de mídia digital. Nesse modelo a empresa atua de modo indireto. Primeiro se conquista a atenção de grande número de pessoas por meio de seus próprios produtos. Depois essa atenção capturada é usada para ofertar produtos e serviços de outras empresas. Alguns produtos componentes das plataformas de mídia digital utilizam processos de organização e representação do conhecimento para compilar e indicar assuntos que são tendências entre seus usuários. **Objetivo:** Essa pesquisa teve o objetivo de compreender como esses compiladores de tendências das plataformas de mídia digital podem ser abordados à luz da ética da informação, em suas dimensões macro e micro. **Metodologia:** Para isso foi realizada uma revisão narrativa de literatura e, a partir dos pressupostos colhidos nessa revisão, o apontamento de algumas questões e problemas éticos suscitados pela oferta de produtos desse tipo. **Resultados:** Compreendeu-se que a ética da informação integra reflexões em dois níveis: o da microética e o da macroética, e constatou-se que a ação das plataformas por meio dos compiladores de tendências levanta questionamentos nos dois níveis. **Conclusão:** Na dimensão microética evidencia-se problemas nos processos de representação da informação adotados por esses produtos. Na dimensão macroética os questionamentos decorrem do empobrecimento que eles podem provocar nos processos informacionais que deveriam aprimorar, por exemplo, ao favorecerem a desinformação.

Palavras-chave: Ética da informação. Ética na organização e representação do conhecimento. Mídias digitais.

Abstract

Introducion: Platform is a widely adopted business model for digital media companies. In this model the company operates indirectly. First, the attention of a large number of people is captured through its own products. Then, this captured attention is used to offer products and services from other companies. Some products that are part of digital media platforms use information organization and representation processes to compile and indicate subjects that are trends among their users. **Objectives:** This research aimed to understand how this trends compilers on digital media platforms can be approached from information ethics, in its macro and micro dimensions. **Methodology:** For this, a narrative review of the literature was carried out and, based on the assumptions collected in this review, the appointment of some ethical questions raised by this kind of products. **Results:** It was understood that information ethics integrates reflections on two levels: microethics and macroethics, and it was found that the action of the platforms through trends compilers raises questions at both levels. **Conclusion:** In the microethical dimension, problems are evident in the information representation processes adopted by these products. In the macroethical dimension, the questions arise from the impoverishment that they can cause in the informational processes that they should improve, for example, by favoring disinformation.

Keywords: Information ethics. Ethics in knowledge organization and representation. Digital media.

Resumen

Introducción: Plataforma es un modelo de negocio ampliamente adoptado por las empresas de medios digitales. En este modelo, la empresa opera indirectamente. En primer lugar, se capta la atención de una gran cantidad de personas a través de sus propios productos. Luego, esta atención capturada se utiliza para ofrecer productos y servicios de otras empresas. Algunos productos que forman parte de las plataformas de medios digitales utilizan procesos de organización y representación de información para compilar e indicar temas que son tendencia entre sus usuarios. **Objetivos:** Esta investigación tuvo como objetivo comprender cómo se pueden abordar los compiladores de tendencias en plataformas de medios digitales desde la ética de la información, en sus dimensiones macro y micro. **Metodología:** Para ello, se realizó una revisión narrativa de la literatura y, en función de los supuestos recopilados en esta revisión, se plantearon algunas cuestiones éticas generadas por este tipo de productos. **Resultados:** La ética de la información integra reflexiones en dos niveles: microética y macroética, y se encontró que la acción de las plataformas a través de los compiladores de tendencias plantea cuestiones en ambos niveles. **Conclusión:** En la dimensión microética, son evidentes los problemas en los procesos de representación de la información adoptados por estos productos. En la dimensión macroética, surgen preguntas sobre el empobrecimiento que pueden causar en los procesos informativos que deberían mejorar, por ejemplo, favoreciendo la desinformación.

Palabras clave: Ética de la información. Ética en la representación del conocimiento. Medios digitales. Desinformación.

1 Introdução

Plataformas de mídia digital são importantes elementos da vida contemporânea conectada. São componentes constituintes do ambiente informacional digital utilizados por bilhões de pessoas, todos os dias, ao redor do mundo, para desempenharem cada vez mais atividades da vida moderna. Por isso, o olhar da ética para essas plataformas, suas dinâmicas e seus efeitos, é essencial.

O termo plataforma pode ser compreendido de vários modos diferentes. Entre outros usos, é empregado para nomear um modelo de negócios segundo o qual uma empresa passa a atuar de modo indireto, proporcionando a conexão de possíveis consumidores com os produtos e serviços de outras empresas.

O modelo de plataforma foi favorecido pelas novas tecnologias da informação e da comunicação e tem sido aplicado nos mais diversos mercados e setores da economia, sendo aclamado como uma abordagem revolucionária de gestão estratégica para as empresas que enfrentam os desafios do mundo contemporâneo (PARKTER; ALSTYNE; CHOUDARY, 2019). Contudo, trata-se de um modelo controverso que tem suscitado uma série de questionamentos éticos em muitos dos contextos em que é empregado.

Schwab (2016) afirma que as plataformas já não são uma exclusividade do mundo digital, uma vez que o modelo já se estendeu para mercados do mundo físico. O autor chama de efeito plataforma as consequências negativas da aplicação deste modelo nos diversos setores da economia. Entre essas consequências está a acentuação da desigualdade e da injustiça social em escala mundial, uma vez que o modelo promove a concentração de vantagens econômicas nas mãos de um pequeno número de empresas que dominam seus mercados. Para o autor essa é uma importante questão ética pois um mundo desigual e injusto cria as condições para a agitação social e, eventualmente, para o extremismo violento (SCHWAB, 2016).

Algumas corporações aplicam o modelo de plataformas para intermediar não somente as relações entre empresas e seus consumidores, mas também entre elas e seus trabalhadores. A chamada platformização do trabalho têm acentuado a precarização das relações trabalhistas uma vez que, nesse modelo, o vínculo entre empresas e trabalhadores se afrouxa pois se torna indireto, dificultando o acesso a muitos direitos que existem em vínculos formais de trabalho Srnicek (2016).

No campo da cultura, as empresas controladoras das plataformas podem se valer de sua posição de intermediadoras para direcionar a atenção de seus usuários não para os produtos culturais que correspondam a seus gostos, mas para os que representem maior vantagem econômica para o negócio. Por exemplo, as plataformas podem permitir que artistas

paguem para ter mais visibilidade, como ocorreu na já extinta Last.fm (SANTINI, 2020) e na atual Spotify (AGUIAR; WALDFOGEL, 2018). Essa prática, que pode ser entendida como um tipo de suborno, é conhecida no Brasil como jabaculê e nos EUA como *payola*.

Entre as diversas aplicações das plataformas, essa pesquisa destaca aquela que se configura no mercado das mídias digitais, compreendendo sistemas de busca na Web, mídias sociais, portais de notícias e aplicativos para as mais diversas finalidades.

Uma plataforma de mídia digital consiste em um amplo conjunto de produtos e serviços, ofertados por uma empresa controladora, projetados para atuarem de maneira integrada, para atender a certas demandas dos usuários dos meios digitais a partir das suas próprias funcionalidades, e também, do intermédio de produtos e serviços ofertados por outras empresas.

Plataformas desse tipo desenvolvem produtos capazes de reter a atenção de grande número de usuários por uma expressiva quantidade de tempo. Essa atenção capturada é, então, utilizada para apresentar aos usuários os produtos e serviços de outras empresas. A atenção do usuário é capturada por meio da oferta de recursos úteis, como o acesso a notícias, compartilhamento de mensagens, busca por páginas na Web, interação social com outros usuários e entretenimento.

Cada plataforma é controlada por uma corporação e as companhias controladoras de plataformas tecnológicas de mídia digital são conhecidas como *big techs* (MOROZOV, 2018). Schwab (2016) lembra que, nas últimas décadas, essas corporações substituíram as petrolíferas, montadoras de automóveis e bancos no *ranking* das empresas mais valiosas do mundo.

A aplicação do modelo de plataforma no mercado das mídias digitais suscita questões éticas que são específicas deste setor, mas que também revelam aspectos controversos do próprio modelo. O fato de as *big techs* figurarem entre as empresas mais valiosas do mundo atualmente já é um indício dessas questões. Por que a posição de intermediária entre consumidores e outras empresas é tão lucrativa?

A nebulosidade e até a perversidade embutida nessa relação é discutida por Lanier (2018). O argumento do autor é que, além de atenderem as demandas de seus usuários, essas companhias precisam também atender aos interesses de seus acionistas e das empresas que intermedeiam. Nesse conflito de interesses, os usuários são sistematicamente preteridos, pois são o polo mais fraco, sem poder de negociação e com menos proteção de regulamentações governamentais (LANIER, 2018).

As plataformas digitais frequentemente adotam práticas para evitar que seus usuários desistam de utilizá-las, oferecendo a eles mecanismos para prevenir desconfortos e inseguranças.

Um desses mecanismos são os algoritmos de filtragem de conteúdo (SANTINI, 2020). Sob o pretexto de oferecer informações classificadas por algum critério de relevância, são criados algoritmos que selecionam e apresentam apenas informações alinhadas com os interesses e gostos pessoais de cada usuário. Esses algoritmos ficaram conhecidos como “filtros bolha”, uma vez que acabam por manter cada usuário circunscrito a um ambiente informacional pobre em termos de diversidade de visões de mundo (PARISER, 2012).

Outro mecanismo utilizado para oferecer conforto e segurança aos usuários das plataformas de mídia digital são os compiladores de tendências. É esse mecanismo em específico que é abordado nessa pesquisa. Trata-se, portanto, da proposição de reflexões éticas a respeito das práticas de organização e representação do conhecimento adotadas no contexto desses sistemas.

Os compiladores de tendências são produtos componentes das plataformas de mídia digital que organizam a informação, misturando processos automatizados por meio de algoritmos e a curadoria humana, com o propósito de indicar tendências informacionais nos ambientes digitais. Entre os mais conhecidos estão o Trends [1] da plataforma Google, o Trending Topics [2] da plataforma Twitter, o Browse [3] da plataforma TikTok, e o Em Alta [4] da plataforma YouTube.

Muitos questionamentos relacionados a aspectos éticos podem ser feitos acerca da oferta desse tipo de produto, seus impactos nos ambientes informacionais e as práticas que adotam para organizar e representar a informação, muitas vezes por meio de algoritmos com critérios opacos para seus usuários. Essas práticas diferem daquelas que tradicionalmente têm merecido o olhar da Ciência da Informação, como a representação do conhecimento a partir da construção de índices de assunto com base em vocabulários controlados. Porém a reflexão ética sobre elas à luz da ética da informação se justifica por se tratar da aplicação de processos informacionais que ganham cada vez mais relevância social.

Essa pesquisa de caráter exploratório tem o objetivo de compreender como os compiladores de tendências das plataformas de mídia digital podem ser abordados à luz da ética da informação, em suas dimensões macro e micro. Para isso, realizou-se uma revisão narrativa de literatura e, a partir dos pressupostos colhidos nessa revisão, foram apontadas algumas questões e problemas éticos suscitados pela oferta desses produtos.

2 Desenvolvimento

2.1 Ética da informação

A proposta dos compiladores de tendências das mídias digitais é a oferta de informação organizada. São produtos de informação ao mesmo tempo em que afetam

o ambiente informacional. Por isso, cabe a reflexão sobre eles à luz da ética da informação.

Gómez (2017) inicia sua reflexão a respeito da ética da informação fazendo referência à origem etimológica similar das palavras ética e moral, cujos significados remetem a usos e costumes. A Ética, pontua a autora, é considerada uma indagação filosófica, crítica e reflexiva acerca da moral, que congrega questões sobre as normas e valores que motivam e orientam as ações dos homens, em direção à justiça e à boa convivência.

Capurro explica que ética é a proposição de reflexões críticas sobre as regras e costumes vigentes em uma sociedade ou “[...] uma problematização da moral vigente [...]” (CAPURRO, 2017, p. 49). Essa ação de problematização da moral (ética) pode ser aplicada em diferentes contextos, inclusive na ética da informação que é entendida como “[...] uma reflexão sobre as normas e valores vigentes nas sociedades da informação e conhecimento” (CAPURRO, 2017, p. 49).

A ética da informação diz respeito às escolhas que levam a um agir responsável nos processos informacionais. Na sociedade da informação, agir responsabilmente no tratamento da informação é uma importante questão ética, afinal, “[...] ser ético é agir de acordo com os valores morais de uma determinada sociedade” (PALETTA; SILVA, 2017, n.p.).

Capurro aponta também que essas problematizações relacionadas à ética da informação devem ser ajustadas às tecnologias e processos comunicacionais em questão, uma vez que cada nova tecnologia muda as possibilidades de como um indivíduo pode se fazer presente no mundo e, com isso, provoca “[...] uma transformação mais ou menos fundamental de tradições morais e culturais vigentes” (CAPURRO, 2017, p. 49).

A reflexão ética acerca dos compiladores de tendência das mídias digitais e suas práticas informacionais deve, portanto, levar em conta esses dois aspectos: o agir responsável no tratamento da informação por parte das plataformas e o ajuste à realidade tecnológica em que eles se inserem como elementos constituintes.

Floridi (2010) sistematizou as reflexões éticas acerca das práticas informacionais propondo que elas se dariam em três direções (que o autor chama de setas). A primeira é a que trata da informação como um recurso, ou seja, a necessidade que um determinado agente tem de obter informação, considerando que “[...] quanto mais bem informado o agente estiver, mais provável é que faça a coisa certa” (FLORIDI, 2010, p.105, tradução nossa). Aqui se incluem questões como o excesso de informação, a falta de acesso e a exclusão digital.

O olhar ético nessa direção legitima a existência dos compiladores de tendência nas mídias digitais. Seus usuários precisam estar bem-informados para guiarem de maneira adequada suas ações. Os compiladores

oferecem essa ajuda ao indicarem a respeito de quais assuntos os demais usuários estão conversando.

A segunda direção compreende a informação como produto, ou a informação produzida por um agente para a consecução de seus objetivos. Aqui se incluem questões como a responsabilidade, o plágio e a desinformação (FLORIDI, 2010). Essa perspectiva aponta o olhar ético para as corporações que controlam as plataformas e oferecem os compiladores de tendências, e levanta questionamentos sobre sua responsabilidade ao fazê-lo.

A terceira direção é mais generalista e aborda a informação como alvo, ou seja, o modo como as ações de um agente afetam o ambiente informacional. Inclui questões como privacidade, segurança da informação, liberdade de expressão, censura, filtragem e controle de conteúdo (FLORIDI, 2010). Aqui a abordagem ética abrange os efeitos que a oferta dos compiladores e os usos que se fazem dele provocam no ambiente informacional. A gestão de um recurso como esse implica na adoção de procedimentos que se situam na conturbada fronteira entre a moderação de conteúdo, a liberdade de expressão e a censura.

Ambas as três direções de Floridi, portanto, oferecem caminhos para o apontamento de questões e problemas éticos que emanam dos compiladores de tendências.

2.2 Microética e macroética da informação

Gómez (2017) explica que Floridi classifica essas direções, quando olhadas individualmente, como microéticas da informação, e “[...] abre o debate sobre a necessidade de outras perspectivas, com o escopo e a abrangência das macroéticas da informação [...]” (GÓMEZ, 2017, p. 39). Para fazer essa transposição, Floridi parte da concepção de uma Ecologia da Informação, pois:

Uma ética da informação ecológica não só pode obter uma visão global de todo o ciclo vital da informação, superando assim os limites de outras abordagens microéticas, mas também pode reivindicar um papel como uma macroética, isto é, como uma ética que diz respeito a todo o reino da realidade. (FLORIDI, 2010, p. 111)

O autor coloca a ética da informação em paralelo com a já estabelecida ética ambiental. Esta última considera que a vida tem valor intrínseco e avalia toda ação moral de modo orientado ao paciente, ou seja, àquele que sofre as consequências das ações. É moral toda ação que permite ou aumenta o bem-estar de um ser vivente. Não são morais as ações que infringem sofrimento a qualquer forma de vida (FLORIDI, 2010).

Essa perspectiva é estendida para a compreensão da ética da informação, de modo que a existência de uma entidade informacional tem valor intrínseco e essa entidade “[...] tem o direito de persistir em seu estado atual e o direito de florescer, isto é, de aumentar e enriquecer sua existência e essência” (FLORIDI, 2010,

p. 112, tradução nossa). Assim, é moral toda ação que promove o favorecimento da existência e o florescimento de uma realidade informacional e não é moral tudo o que provoca a entropia, aqui compreendida como “[...] todo tipo de destruição, corrupção, poluição, e esgotamento de objetos informacionais” (FLORIDI, 2010, p. 112, tradução nossa).

A classificação das questões e problemas éticos em micro e macro não é uma forma de valoração que visa separar questões maiores de questões menores, mas sim um caminho para sua melhor compreensão. Por exemplo, questões relacionadas a práticas de organização e representação do conhecimento poder ser compreendidas como microéticas se tomadas a partir da segunda direção proposta por Floridi (informação como produto), mas também como macroéticas, uma vez que impactam o ambiente informacional de modo amplo provocando entropia quando mal realizadas.

As duas dimensões se aplicam ao olhar ético a respeito dos compiladores de tendências. Tanto a reflexão microética, quanto às suas práticas e seus usos, como a macroética, quanto a seus efeitos no ambiente informacional tecnológico em que se inserem, permitem trazer à tona as questões e problemas que lhes são próprios.

2.3 Problemas éticos na organização e representação do conhecimento

Para nortear a reflexão ética quanto aos compiladores de tendências, a revisão de literatura aqui proposta buscou identificar questões e problemas já conhecidos nas práticas informacionais. Não se trata de uma tentativa de listar exaustivamente todas as demandas éticas que possam surgir em ambientes que adotam práticas de tratamento da informação, mas sim de evidenciar aquelas já conhecidas para examinar as possibilidades de sua manifestação também nesse ambiente.

A partir da investigação a respeito do modo como questões éticas relacionadas à organização e representação do conhecimento (ORC) foram tratadas em artigos publicados em uma série de periódicos entre os anos de 1995 e 2004, Guimarães, Milani e Pinho (2008) propuseram a categorização dos valores e dos problemas éticos que se apresentam em decorrência da negação desses valores. Em outro trabalho, Guimarães, Milani e Evangelista (2015) atualizaram categorização.

Esses trabalhos tratam a organização e representação do conhecimento como uma atividade que deve ser orientada por alguns valores, sendo o principal deles a garantia do acesso. Decorrem desse valor primordial outros valores relacionados à atividade informacional, à conduta do profissional da informação e às práticas de organização e representação do conhecimento. O desrespeito aos valores supracitados implica em problemas éticos que os autores dividem em três categorias: quanto à linguagem, quanto à postura

profissional e quanto à unidade ou sistema informacional.

O quadro 1 apresenta os problemas relacionados à linguagem utilizada na representação do conhecimento. Os itens desta categoria decorrem principalmente dos conflitos que surgem na adoção de sistemas padronizados de classificação e de vocabulários controlados frente ao dinamismo dos processos de produção de conhecimento e das transformações sociais.

Quadro 1. *Problemas éticos quanto à linguagem na ORC*

1. Categorias dicotômicas
2. Categorias remanescentes, gerando empecilhos ao fluxo informacional
3. Falta de garantia cultural
4. Moralidade classicista determinando as estruturas principais das classificações
5. Terminologia preconceituosa
6. Traduções inadequadas
7. Vocabulários controlados desatualizados, com lacunas, discriminatórios e anglo cêntricos

Fonte: Elaborado pelos autores a partir de Guimarães, Milani e Evangelista (2015).

A segunda categoria de problemas éticos diz respeito à postura do profissional da informação. Esses problemas dizem respeito à ineficiência, falta de comprometimento ou crenças equivocadas por parte do profissional e seus efeitos na ORC. Os problemas inclusos nessa categoria são apresentados no Quadro 2.

Quadro 2. *Problemas éticos quanto à postura do profissional na ORC*

1. *Biases* na representação de assunto
2. Censura
3. Desconhecimento das necessidades informacionais do usuário
4. Despreocupação com questões de gênero
5. Direcionamento informacional (com o objetivo de produzir silêncio, dificultando a recuperação)
6. Distorção da informação
7. Falta de espírito profissional coletivo
8. Marginalização na representação
9. Negligência
10. Parcialidade não declarada
11. Utilização de hierarquias reforçando preconceitos
12. Crença em linguagens estáticas
13. Crença na necessidade de linguagens universais
14. Crença na neutralidade dos processos, produtos e instrumentos de representação
15. Crença na objetividade como sinal de imparcialidade na representação de assunto
16. Crença na tradição cartesiana como única lógica de classificação
17. Crença no bem-fazer do indexador como algo intrínseco ao profissional
18. Manutenção da visão ocidental em contextos orientais
19. Presunções teóricas (tomar postulados como verdade sem reflexão)

Fonte: elaborado pelos autores a partir de Guimarães, Milani e Evangelista (2015).

Por fim, a terceira categoria de problemas éticos elencados nesse esforço diz respeito às práticas adotadas pela unidade ou sistema informacional. São problemas que emanam das políticas e das práticas estabelecidas para o funcionamento de bibliotecas e arquivos, mas que também podem se manifestar nas práticas das corporações que mantêm as plataformas de mídia digital. Em essência, tratam de sua responsabilidade perante a oferta da informação do modo mais eficiente e perante o bem-estar de seus usuários. O Quadro 3 enumera os problemas apontados nessa categoria.

Quadro 3. *Problemas éticos quanto à unidade ou sistema informacional na ORC*

1. Falta de organização informacional em sites de busca, gerando inacessibilidade
2. Sistema de browsing genéricos
3. Difamação
4. Veiculação inadvertida de plágios
5. Violação da privacidade (incluindo vigilância, monitoramento informacional e censura)
6. Violência (quando a unidade de informação viola os direitos do usuário, a ponto de causar-lhe dano)

Fonte: elaborado pelos autores a partir de Guimarães, Milani e Pinho (2008) e Guimarães, Milani e Evangelista (2015).

Nem todos esses problemas, enumerados a partir das práticas de ORC em bibliotecas e arquivos, se manifestam nas plataformas digitais, uma vez que essas possuem finalidades e dinâmicas diferentes. Do mesmo modo, uma listagem como essa não é capaz de esgotar todas as possibilidades nesse sentido, uma vez que o dinamismo dos ambientes tecnológicos faz surgir novas questões o tempo todo. Porém, eles servem como ponto de partida para a reflexão ética a respeito dos compiladores de tendências se eles forem compreendidos como produtos de informação.

2.4 Problemas éticos próprios das TICs

Atendendo à necessidade apontada por Capurro (2017), de que o olhar ético precisa se ajustar ao ambiente tecnológico em que as práticas a serem analisadas se inserem, é preciso complementar esse inventário de problemas éticos com questões próprias das tecnologias.

Fernández-Molina *et al.* (2005) avaliaram o modo como as questões éticas relacionadas ao uso das tecnologias da informação e da comunicação (TICs) foram tratados em artigos publicados no periódico *Journal Ethics and Information Technology* entre os anos de 1999 e 2004. Os autores compilaram duas listas de categorias: uma para valores éticos e outra para problemas éticos. Essas duas listas foram atualizadas e ampliadas posteriormente por Chardulo (2022), que repetiu o procedimento realizado no trabalho original para os anos de 2012 a 2020. O Quadro 4 apresenta em ordem alfabética os problemas éticos enumerados nesses trabalhos.

Quadro 4. *Valores éticos relacionados às TICs*

1. Abuso e abandono de menores
2. Alienação
3. Aprisionamento (“bolha”)
4. Armadilha e hostilidade <i>online</i> , discurso de ódio e <i>trollagem online</i>
5. Assassinato virtual (“cancelamento”)
6. Bloqueio de anúncios
7. Burocracia
8. Caixas pretas e opacidade
9. Censura, filtragem, regulação de conteúdo e controle de acesso
10. Comportamento de veículos autônomos
11. Consentimento informado
12. Controle de armas, robôs assassinos, autômatos letais e robôs militares
13. Crime, punição e crimes sexuais
14. <i>Cyber-bullying</i>
15. Decepção e traição
16. <i>Deep fake</i> e falsificação
17. Desemprego tecnológico
18. Desencantamento humano
19. Desigualdade e desequilíbrio
20. Difamação
21. Dilema do bonde
22. Divisão/exclusão digital
23. <i>Doxing</i> e assédio
24. Encriptação
25. Exclusão de dados de pessoas mortas
26. <i>Gatekeeping</i>
27. Guerra
28. Guerra de informação
29. Injustiça
30. Julgamento da capacidade (capacitismo) de tomada de decisão
31. Misoginia <i>online</i>
32. Personalização para segmentação e direcionamento excessivo.
33. Persuasão, mudança de comportamento e manipulação
34. Pirataria e plágio
35. Policiamento preditivo
36. Pornografia e obscenidade
37. Preconceito e discriminação
38. Racismo
39. Regulação paternalista
40. Representação em realidade virtual
41. Robô cuidador de idosos
42. Simulação de cérebro de personalidades, roubo de cérebro e neurocrime
43. <i>Spamming</i>
44. Terrorismo
45. Vandalismo
46. Vigilância, controle, interceptação de comunicação e mineração de dados
47. Violência

Fonte: elaborado pelos autores com base em Fernández-Molina et al. (2005) e Chardulo (2022).

É interessante notar que algumas questões enumeradas nessas listagens estão presentes também naquelas elaboradas por Guimarães, Milani e Evangelista (2015), o que evidencia a interação das práticas informacionais

com as características da tecnologia na produção dos problemas éticos nos ambientes digitais.

2.5 *Microética da informação nos compiladores de tendências*

Os trabalhos citados elencaram um total de 79 problemas éticos conhecidos em processos de informação e em plataformas tecnológicas. Não é o propósito dessa pesquisa verificar a ocorrência ou não de cada um deles nos compiladores de tendências das plataformas de mídia digital, mas sim apontar que pelo menos alguns deles se manifestam também nesses ambientes.

Retomando a sistematização proposta por Floridi (2010), na dimensão microética é possível verificar quais dos problemas elencados são passíveis de ocorrer com a oferta dos compiladores de tendências em cada uma das três direções da informação. Na dimensão macroética pode-se verificar os efeitos entrópicos que tais problemas podem provocar no ambiente informacional das plataformas de mídia digital.

A primeira direção é aquela que analisa a informação como fonte. Essa costuma ser a justificativa utilizada pelas plataformas para ofertar produtos como os compiladores de tendências. Por exemplo, ao defender a oferta do compilador Google Trends na plataforma do sistema de buscas na Web Google, Rogers (2016) afirma que o sistema “[...] está ajudando as pessoas ao redor do mundo a explorar a reação global aos grandes eventos [...]” e que a ferramenta “[...] fornece uma perspectiva única sobre o que as pessoas estão interessadas e curiosas” (ROGERS, 2016, tradução nossa). O recurso é tido como um serviço ao usuário, pois

Os dados de tendências podem fornecer uma lente poderosa sobre o que os usuários do Google estão curiosos e como as pessoas ao redor do mundo reagem a eventos importantes. (ROGERS, 2016, tradução nossa)

Visto dessa forma, os compiladores podem ser considerados uma ação moral das plataformas, pois representam um esforço de melhor informar seus usuários para que, assim, possam melhor agir. Mas essa postura pode também evidenciar algumas vulnerabilidades éticas, como o desconhecimento das necessidades informacionais do usuário (quadro 1, item 3) e a crença no bem-fazer do indexador como algo intrínseco ao profissional (ou à plataforma, neste caso) (quadro 1, item 17).

Ainda a respeito do Google Trends, Rogers (2016, tradução nossa), afirma que “os dados de tendências são uma amostra imparcial de nossos dados de pesquisa do Google”, o que aponta para outras questões éticas levantadas anteriormente, como a crença na neutralidade dos processos, produtos e instrumentos de representação (quadro 1, item 14); a crença na objetividade como sinal de imparcialidade na representação de assunto (quadro 1, item 15).

No aspecto tecnológico, a oferta de um compilador de tendências para guiar os usuários não resolve e talvez até mesmo acentue problemas estruturais do acesso à informação. Usuários que acessam esses recursos passam a ter melhores condições de uso das plataformas do que os que não as acessam com frequência, aumentando a divisão e exclusão digital (quadro 4, item 22) ou o desemprego tecnológico (quadro 4, item 17).

A segunda direção proposta por Floridi (2010) trata a informação como produto e talvez seja a que mais possa nortear as reflexões a respeito dos compiladores de tendência se eles forem compreendidos como produtos informacionais. Essa linha aponta para a qualidade dos produtos e para a responsabilidade de quem os oferta.

No que diz respeito à qualidade, a reflexão ética se direciona para o modo como é feita a seleção e representação dos assuntos apontados como tendências. Quanto à responsabilidade pela oferta, as reflexões apontam para as corporações que controlam esses produtos, que não podem ser vistas como neutras nesse processo informacional.

A qualidade da representação é uma das principais questões nesse aspecto. Os compiladores se utilizam de algoritmos para fazer a seleção dos assuntos e sua classificação como tendências. Por exemplo, a respeito de seu compilador Trending Topics, o Twitter afirma que “as tendências são determinadas por um algoritmo [...]. Esse algoritmo identifica tópicos que são populares [...]” (TWITTER, s/d, tradução nossa). A seleção automática dos assuntos não utiliza vocabulários controlados, podendo incorrer em redundância, duplicidades e no uso de categorias dicotômicas (quadro 1, item 1).

No caso do Twitter Trending Topics, os assuntos são selecionados com base na classificação feita pelos próprios usuários, numa aplicação da folksonomia, por meio das *hashtags*, mas também com base na análise do texto completo, com o próprio algoritmo identificando quais são os assuntos abordados em cada publicação. Mari Junior e Paletta (2022) identificaram que entre 2018 e 2022 o produto passou a privilegiar a seleção de assuntos a partir do texto completo, em detrimento das *hashtags*, o que torna ainda mais opaco o modo como os assuntos são selecionados.

Os algoritmos que criam os índices de tendências fazem a filtragem do conteúdo disponível nos ambientes informacionais das plataformas buscando aquilo que é relevante coletivamente (uma tendência), mas também o que é relevante para cada usuário individualmente. A respeito de seu compilador Trending Topics, o Twitter afirma que “as tendências [...] por padrão, são personalizadas para você com base em quem você segue, seus interesses e sua localização” e, caso não queira essa personalização, cada usuário pode “[...] optar por ver tendências [...] de um local específico” (TWITTER, s/d, tradução nossa). Desse tipo de prática decorrem

problemas como a personalização para segmentação e direcionamento excessivo (quadro 4, item 32), que colabora para a formação do efeito bolha (PARISER, 2012) (quadro 4, itens 2 e 3).

A busca pela personalização implica em outro problema conhecido: o excesso de dados coletados. Para que possam filtrar a informação que seja mais relevante é necessário, antes, coletar um grande volume de dados acerca das decisões e hábitos dos usuários ao longo de um grande período de tempo (SANTINI, 2020). Essa prática força o limite entre o que é público e o que é privado. A questão ética colocada aqui é que os dados que permitem a personalização se tornaram muito valiosos, produzindo o que Zuboff (2018) chama de capitalismo de dados. Diante de tanto valor as plataformas podem se sentir tentadas a desrespeitar a privacidade do usuário para obter os dados à sua revelia.

Schwab (2016) acrescenta que a prática de coleta massiva de dados, pode tomar a forma de vigilância e “[...] dar origem a autoridades públicas com excesso de poder em suas mãos” (SCHWAB, 2016, p. 73). Prática da qual decorre problemas como o policiamento preditivo (quadro 4, item 35) e a vigilância, controle, interceptação de comunicação e mineração de dados (quadro 4, item 46).

Um dos maiores desafios individuais colocados pela internet [...] diz respeito à privacidade. [...] A internet pode ser uma ferramenta sem precedentes de libertação e de democratização e, ao mesmo tempo, um facilitador da fiscalização maciça, indiscriminada. (SCHWAB, 2016, p. 104-105)

Outro problema de qualidade comum dos algoritmos de filtragem de conteúdo são sua opacidade, que faz com que os usuários não consigam saber quais critérios estão sendo utilizados para a determinação da relevância das informações. Esse problema é uma manifestação da parcialidade não declarada (quadro 2, item 10) e o direcionamento informacional com o objetivo de produzir silêncio, dificultando a recuperação (quadro 2, item 5).

No aspecto tecnológico a opacidade é citada nominalmente no item 8 do quadro 4 e indiretamente sobre a forma de censura, filtragem, regulação de conteúdo e controle de acesso (quadro 4, item 9); *gatekeeping* (quadro 4, item 26), que consiste exatamente no poder que um agente informacional adquire de determinar o que é e o que não é importante; o que pode ser feito com o objetivo de provocar persuasão, mudança de comportamento e manipulação (quadro 4, item 33).

O processo de filtragem de conteúdo de acordo com os critérios de relevância estabelecidos pelos algoritmos está sujeito a problemas éticos já conhecidos. O’Neil (2020) denuncia o modo como os vieses reproduzidos por esses algoritmos fortalecem mecanismos de segregação social, prejudicam o acesso à educação de qualidade e ameaçam a democracia. Noble (2021) e

Silva (2022) discutem como o racismo se manifesta nesses mecanismos e como as companhias que os controlam lucram ele. Essas questões são reflexos de alguns problemas citados no levantamento anterior, como o uso de terminologia preconceituosa (quadro 1, item 5); a ocorrência de *biases* (vieses) na representação de assunto (quadro 2, item 1); a despreocupação com questões de gênero (quadro 2, item 4) e a marginalização na representação (quadro 2, item 8).

Em termos tecnológicos, compiladores podem amplificar o efeito de conteúdos discriminatórios ao apontá-los como tendência. Por isso esses problemas figuram também no quadro 4, como por exemplo a difamação (quadro 4, item 20); o *doxing* e assédio (quadro 4, item 23); o capacitismo (quadro 4, item 30); a misoginia (quadro 4, item 31); o preconceito e discriminação (quadro 4, item 37) e o racismo (quadro 4, item 38).

Ao permitir que esse tipo de problema se manifeste em seus produtos, as plataformas estão incorrendo nos problemas éticos apontados no Quadro 3, como a difamação (quadro 3, item 3) e a violência (quadro 3, item 6). Esse tipo de difamação e violência se caracterizam como ações propositais por parte da unidade de informação. Seria o caso de as plataformas estarem conscientemente agindo assim. A questão ética que se coloca aqui é: se uma plataforma permite que um conteúdo difamatório ou violento apareça em seu produto, há por parte dela alguma intenção ou responsabilidade?

Além da qualidade da RI, nesta abordagem cabe também reflexões quanto à qualidade do próprio produto informacional. O Quadro 3 apresenta alguns problemas, como a falta de organização informacional em sites de busca, gerando inacessibilidade (item 1) e a adoção de sistema de *browsing* genéricos (item 2). Esses dois temas colocam a falta de acessibilidade, de usabilidade e de encontrabilidade da informação como problemas éticos, uma vez que a ausência dessas qualidades provoca prejuízo ao sujeito informacional.

Como exemplo desse problema, é possível citar uma limitação do compilador Trending Topics da plataforma Twitter. Nele o usuário vê somente os tópicos classificados como tendência no instante do seu acesso, não sendo possível ver o que era considerado tendência em qualquer momento no passado (TWITTER, s/d). Não é possível saber quando um assunto se tornou tendência ou por quanto tempo ele está nessa condição, ofertando-se, portanto, uma informação parcial e sem contexto que pode induzir os usuários ao erro.

A terceira e última direção de Floridi (2010) toma a informação, os produtos e os ambientes informacionais como alvos ou como destino das ações de um agente. Aqui se inserem as questões que surgem em decorrência das ações que têm o objetivo ou a potência de afetar o ambiente informacional em si.

Essa direção engloba atos que promovam a censura (quadro 2, item 2) e a violação da privacidade (quadro 3, item 5), principalmente quando esta se dá a partir de ações relacionadas à segurança da informação, como o *hacking* ou o vazamento de dados que deveriam ser privados e a prática conhecida como *doxing* (quadro 4, item 23). Esse tipo de questão pode envolver também terrorismo tecnológico (quadro 4, item 44).

Ainda nessa direção, enquadram-se as práticas de vandalismo (quadro 4, item 45) aplicadas sobre os ambientes informacionais, que nos meios digitais frequentemente ganham a forma de *spamming* (quadro 4, item 43). Nessa prática um agente busca inserir ou aumentar o volume de informações artificialmente no ambiente para favorecer seus interesses.

Na plataforma Twitter, por exemplo, é conhecida a ação de robôs (ou *bots*) empregados para esse fim. Roth e Pickles (2020), falando em nome da plataforma, explicam que “[...] em suma, um *bot* é uma conta automatizada [...]”, que pode vir a ser usada para a manipulação de plataforma “[...] por pessoas em posições de poder político para manchar as opiniões de pessoas que possam discordar delas ou da opinião pública *online* [...]” (ROTH; PICKLES, 2020, tradução nossa).

2.6 Macroética da informação nos compiladores de tendências

A terceira direção, por tratar de questões de efeito mais amplo, direciona as discussões para o que Floridi (2010) denominou macroética da informação. Essa abordagem macroética é especialmente útil para abarcar questões éticas nos novos ambientes digitais e tecnológicos, incluindo os compiladores de tendências da mídia digital, e suas questões peculiares, como a autonomia e a privacidade do usuário, bem como a transparência e a responsabilidade das plataformas.

A perspectiva macroética da informação busca julgar as ações de um agente como moral ou não a partir de um procedimento similar ao aplicado na ética ambiental. Nesta última, não é moral tudo aquilo que causa qualquer tipo de sofrimento a um ser vivente. Para a ética da informação não é moral, ou seja, é um problema ético, tudo aquilo que provoca entropia em um objeto ou ambiente informacional.

A desinformação é um exemplo dessas questões amplas. Floridi afirma que “[...] desinformação (i.e., a produção e distribuição deliberada de conteúdo falso ou errôneo) é um problema ético que diz respeito às três setas informacionais” (2010, p. 108, tradução nossa).

Desinformação não está relacionada à falta de informação, mas à oferta de informação que não cumpre seu papel moral de bem-informar para o bem-agir. Em termos macroéticos, ela se manifesta não apenas como ação consciente e declarada de um agente mal-intencionado, o que caracterizaria um tema para a

terceira direção de Floridi, mas também como efeito holístico do modo como o ambiente informacional é arranjado e utilizado.

Por exemplo, os algoritmos de filtragem de informação, atuam “[...] como uma sofisticada solução sociotécnica e econômica – não somente para o excesso de informação, mas para o gerenciamento dos mercados de bens de consumo” (SANTINI, 2020, p.102), ou seja, são desenvolvidos para atender as necessidades de acesso à informação por parte do usuário, mas também para servirem como canais para a oferta de produtos.

Esse duplo papel dos algoritmos pode ser classificado como entrópico ou como um tipo de desinformação, uma vez que nem sempre o usuário está ciente de que os critérios de relevância adotados por eles também privilegiam os interesses comerciais das plataformas (LANIER, 2018). Assim ele pode ser induzido a tomar como relevante para si uma informação que é relevante apenas para os interesses comerciais das plataformas.

Outra questão ampla que requer o olhar no nível da macroética da informação diz respeito à fronteira entre a liberdade de expressão e a moderação com adoção de regras de conduta por parte da plataforma para seus usuários. Não se pretende aqui discutir o aspecto jurídico do problema, mas sim suas implicações éticas. Nesse sentido, alguns questionamentos que surgem são: as plataformas devem ou podem criar regras de conduta para seus usuários? Elas podem moderar o conteúdo publicado, decidindo elas mesmas o que pode e o que não pode ser compartilhado? Ou mesmo, elas podem, por si, decidirem quem tem ou não tem direito de utilizar seus ambientes?

Essa questão é relevante pois, se é permitido ou até desejável, que as plataformas façam esse tipo de moderação por suas próprias regras, se estaria conferindo a elas um poder que legalmente elas não possuem. Esse tema é bem discutido por Peres-Neto e Pereira:

Ao banirem determinado tipo de conteúdo e outros não, os próprios sites têm o poder de determinar quem tem espaço e quem não tem, quais assuntos são ‘legítimos’ ou não. (PERES-NETO, PEREIRA, 2019)

Dick (2022) lembra que, diante dessa questão, as plataformas têm se valido de um tópico da legislação dos EUA, a seção 230 da Lei de Telecomunicações de 1996, que as garante imunidade em relação ao conteúdo publicado por terceiros, mesmo que ela aplique processos de filtragem e curadoria sobre esses conteúdos.

Isso oferece a elas um poder ainda maior, uma vez que estariam duplamente imunes: caso optem por não moderar o conteúdo dos usuários, não podem ser responsabilizadas e, caso optem pelo contrário, estão livres para fazer a moderação de acordo com suas próprias regras e interesses.

Em termos macroéticos, ambas as situações têm potencial para prejudicar o ambiente informacional como um todo, desinformando o usuário e aumentando a entropia. Schwab (2016), Lanier (2018) e Noble (2021) estão entre os que aponta algum nível de regulamentação em nível governamental para amenizar essa questão, mas não a consenso sobre esse tema.

3 Conclusão

Essa pesquisa teve o objetivo de refletir acerca de como os compiladores de tendências das plataformas de mídia digital podem ser abordados à luz da ética da informação, em suas dimensões macro e micro conforme proposto por Floridi (2010). Para isso foi feito um levantamento exploratório de problemas éticos conhecidos em processos de RI, bem como em ambientes informacionais tecnológicos.

Nenhum levantamento é capaz de esgotar essa questão, pois os meios digitais oferecem ambientes extremamente dinâmicos, que forçam os limites da moral e da ética o tempo todo. De qualquer modo, foram listados 79 problemas conhecidos. Em seguida esses problemas foram confrontados com cada uma das três direções da ética da informação proposta por Floridi, caracterizando sua abordagem microética. Por fim, tratou de como essas questões, somadas a temas mais amplos como a desinformação, a liberdade de expressão e a responsabilidade das plataformas afeta esses ambientes.

Constatou-se que muitos dos problemas éticos conhecidos na ORC e no uso das tecnologias podem também se manifestam nos compiladores de tendências das plataformas de mídia social, seja na promessa de informação organizada (primeira direção), na qualidade dos produtos informacionais e na responsabilidade de quem os oferta (segunda direção), ou no modo como esses ambientes podem ser afetados por ações mal-intencionadas por parte dos usuários ou da própria plataforma (terceira direção).

Nesse contexto, essa pesquisa deve prosseguir no sentido de demonstrar com dados empíricos a manifestação de fato desses problemas nos compiladores de tendências das plataformas de mídia digital.

Notas

- [1] O Google Trends pode ser acessado no endereço <https://trends.google.com>.
- [2] O Twitter Trending Topics pode ser acesso no endereço <https://twitter.com/explore/tabs/trending>.
- [3] O TikTok Browser pode ser acessado no endereço: <https://www.tiktok.com/browse>.
- [4] O YouTube Em Alta pode ser acessado no endereço: <https://www.youtube.com/feed/trending>.

Referências

- AGUIAR, Luis; WALDFOGEL, Joel. Platforms, promotion, and product discovery: evidence from Spotify playlists. **NBER Working Paper**, n. 24713, jun. 2018. Disponível em: <https://www.nber.org/papers/w24713>. Acesso em 20 nov. 2022.
- CAPURRO, Rafael. A liberdade na era digital. In: GÓMEZ, Maria Nélide Gonzáles de; CIANCONI, Regina de Barros (orgs.). **Ética da informação: perspectivas e desafios**. Niterói: PPGCI/UFF, 2017.
- CHARDULO, Suzana Mayumi Iha. **Implicações éticas no uso dos algoritmos: um estudo exploratório na infosfera**. 2022. Dissertação (Mestrado em Cultura e Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2022.
- DIJCK, José van. Ver a floresta por suas árvores: visualizando plataformização e sua governança. **MATRIZES**, v. 16, n. 2, p. 21-44, 2022. Disponível em: <https://www.revistas.usp.br/matrizes/article/view/201591>. Acesso em: 30 jan. 2023.
- FERNÁNDEZ-MOLINA, J. Carlos *et al.* Aspectos éticos de las nuevas tecnologías de la información y la comunicación y su reflejo en la organización y representación del conocimiento. In: CONGRESO ISKO-ESPAÑA, 7, 2005, **Anais [...]**. Barcelona: ISKO, 2005.
- FLORIDI, Luciano. Information: a very short introduction. Nova Iorque: Oxford University Press, 2010.
- GÓMEZ, Maria Nélide Gonzáles de. Reflexões sobre ética da informação: panorama contemporâneo. In: GÓMEZ, Maria Nélide Gonzáles de; CIANCONI, Regina de Barros (orgs.). **Ética da informação: perspectivas e desafios**. Niterói: PPGCI/UFF, 2017.
- GUIMARÃES, José Augusto Chaves; MILANI, Suellen Oliveira; EVANGELISTA, Isadora Victorino. Valores éticos na organização e representação do conhecimento. **Enc. Bibli.** Florianópolis, n. 20, n. esp. 1, p. 19-32, fev. 2015.
- GUIMARÃES, José Augusto Chaves; MILANI, Suellen Oliveira; PINHO, Fabio Assis. Aspectos éticos em organização e representação do conhecimento (ORC): uma análise preliminar de valores e problemas a partir da literatura internacional da área. **Enc. Bibli.** Florianópolis, v. 13, n. 25, p. 124-135, 2008.
- LANIER, Jaron. **Dez argumentos para você deletar agora suas redes sociais**. Rio de Janeiro: Intrínseca, 2018.
- MARI JUNIOR, Sergio; PALETTA, Francisco Carlos. Informação como proposta de valor em produtos de plataformas de mídia digital: estudo sobre o Twitter Trending Topics. In: COLÓQUIO EM ORGANIZAÇÃO, ACESSO E APROPRIAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO (COAIC), VI, 2022, **Anais [...]**. Londrina: UEL, 2022.
- MOROZOV, Evgeny. **Big Tech: a ascensão dos dados e a morte da política**. São Paulo: Ubu, 2018.
- NOBLE, Safiya Umoja. **Algoritmos da opressão: como o Google fomenta e lucra com o racismo**. Santo André: Rua do Sabão, 2021.
- O'NEIL, Cathy. **Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia**. Santo André: Rua do Sabão, 2020.
- PALETTA, Francisco Carlos; SILVA, Armando Malheiro da. A complexidade da era digital desafia a ética. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB, XVIII, 2017. **Anais [...]**. Marília, 2017.
- PARISER, Eli. **O filtro invisível: o que a internet está escondendo de você**. Rio de Janeiro: Zahar, 2012.
- PARKTER, Geoffrey G.; ALSTYNE, Marchall W. van; CHOUDARY, Sangeet Paul. **Plataforma: a revolução da estratégia**. Rio de Janeiro: Alta Books, 2019.
- PERES-NETO, Luiz; PEREIRA, Fabriela Agostinho. Ética, liberdade de expressão e discurso de ódio de gênero em sites de redes sociais. **E-compós**, v.22, jan-dez 2019, p. 1-25.
- ROGERS, Simon. **What is Google Trends data - and what does it mean?**. Google News Lab, jul. 2016. Disponível em: <https://medium.com/google-news-lab/what-is-google-trends-data-and-what-does-it-mean-b48f07342ee8>. Acesso em: 30 jan. 2023.
- ROTH, Yoel; PICKLES, Nick. **Bot or not? The facts about platform manipulation on Twitter**. Twitter Blog, 18 mai. 2020. Disponível em: https://blog.twitter.com/en_us/topics/company/2020/bot-or-not. Acesso em: 30 jan. 2023.
- SANTINI, Rose Marie. **O algoritmo do gosto: os sistemas de recomendação online e seus impactos no mercado cultural**. Curitiba: Appris, 2020.
- SCHWAB, Klaus. **A quarta revolução industrial**. São Paulo: Edipro, 2016.
- SILVA, Tarcizio. **Racismo algorítmico: inteligência artificial e discriminação das redes digitais**. São Paulo: Sesc São Paulo, 2022 (e-book).
- SRNICEK, Nick. **Platform capitalism**. Cambridge: Polity Press, 2016.
- TWITTER. Twitter Trends FAQs. S/d. Disponível em: <https://help.twitter.com/pt/using-twitter/twitter-trending-faqs>. Acesso em: 30 jan. 2023.
- ZUBOFF, Shoshana. Big Brother: capitalismo de vigilância e perspectivas para uma civilização da informação. In: BRUNO, Fernanda; CARDOSO, Bruno; KANASHIRO, Marta; GUILHON, Luciana; MELGAÇO, Lucas (orgs.). **Tecnopolíticas da vigilância: perspectivas da margem**. São Paulo: Boitempo, 2018.

Fake News e a organização do conhecimento: aproximações possíveis

Fake News and the knowledge organization: possible approaches

Leilah Santiago Bufrem (1), Carla Conforto de Oliveira (2), Ely Francina Tannuri de Oliveira (3)

(1) Universidade Federal de Pernambuco, Av. Prof. Moraes Rego, 1235, Recife, santiagobufrem@gmail.com. (2) Universidade Estadual Paulista, Av. Hygino Muzzi Filho, 737, Marília, carla.conforto@unesp.br, (3) etannuri@gmail.com

Resumo

Considera as dimensões pelas quais é possível analisar o fenômeno denominado *fake News*, no atual contexto teórico metodológico das ciências sociais, questionando a possibilidade de construir uma análise multidimensional, para compreender os aspectos característicos dessa realidade na literatura científica. Objetiva caracterizar a produção científica internacional sobre *fake News*, de 2017 a 2022, para verificar como a Organização do Conhecimento pode absorver a problemática. Especificamente, analisa a produção científica sobre *fake News*, do ponto de vista da predominância temática e das subáreas das ciências sociais às quais se vincula, dos países mais prolíficos no tema e dos artigos mais citados dessa produção. Como procedimento de pesquisa, organiza os dados retirados da base Scopus, a partir dos 841 artigos levantados com o termo *fake News*, na produção da área de Ciências Sociais. A estratégia de busca inclui a exclusividade das línguas inglesa, espanhola e portuguesa e dos artigos em acesso aberto, resultando um corpus de 315 artigos. Considera os 18 artigos com maior quantidade de citações para analisar seu conteúdo a partir das categorias levantadas. Como resultados apresenta a palavra-chave mais frequente *fake News*, por ter sido utilizada na busca, seguida dos termos *journalism news*, *disinformation*, *credibility* e as diferentes mídias sociais. Constata a relação dessas palavras com as subáreas de Comunicação, Computação, Artes e Humanidades, Engenharia, entre outras. Apresenta os países mais prolíficos sobre o tema, com destaque para Espanha, Estados Unidos, Reino Unido e Austrália e como idioma predominante o inglês.

Palavras-chave: *Fake News*. Organização do conhecimento. Produção científica. Comunicação.

Abstract

The study considers the dimensions through which is possible to analyze the phenomenon of fake News in the current theoretical and methodological context of Social Sciences, questioning the possibility of constructing a multidimensional analysis to understand the characteristic aspects of this reality in the scientific literature. It aims to characterize the international scientific production on fake News, from 2017 to 2022, to verify how Knowledge Organization can absorb the issue. Specifically, it analyzes the scientific production on fake News, from the point of view of thematic predominance and the subareas of Social Sciences to which it is linked, the most prolific countries in the subject and the most cited articles in this production. As a research procedure, the investigation organizes the data taken from the Scopus base from 841 articles surveyed with the term fake News in the production of Social Sciences area. The search strategy includes, exclusively, English, Spanish and Portuguese languages and open access articles, resulting in a corpus of 315 articles. It considers the 18 articles with the highest number of citations to analyze their content from the surveyed categories. As a result, it presents the most frequent keyword fake News, for having been used in the search, followed by the terms journalism news, disinformation, credibility, and the different social media. It verifies the relationship of these words with the subareas of Communication, Computer, Arts and Humanities, Engineering, among others. It presents the most prolific countries on the subject, with emphasis on Spain, United States, United Kingdom and Australia, with English as the predominant language.

Keywords: Fake News. Knowledge Organization. Scientific production. Communication.

Resumen

Considera las dimensiones por las cuales es posible analizar el fenómeno denominado *fake News*, en el actual contexto teórico-metodológico de las ciencias sociales, cuestionando la posibilidad de construir un análisis multidimensional, para comprender los aspectos característicos de esta realidad en la literatura científica. Tiene como objetivo caracterizar la producción científica internacional sobre *fake News*, de 2017 a 2022, para verificar cómo la Organización del Conocimiento puede absorber la problemática. Específicamente, analiza la producción científica sobre *fake News*, desde el punto de vista del predominio temático y de las subáreas de las ciencias sociales a las que se vincula, de los países más prolíficos en el tema y los artículos más citados de esa producción. Como procedimiento de investigación, organiza los datos tomados de la base Scopus de los 841 artículos planteados con el término *fake News*, en la producción del área de Ciencias Sociales. La estrategia de búsqueda incluye la exclusividad de los idiomas inglés, español y portugués y artículos de acceso abierto, resultando en un corpus de 315 artículos. Considera los 18 artículos con mayor número de citas para analizar su contenido a partir de las categorías planteadas. Como resultado presenta la palabra clave más frecuente *fake News*, por haber sido utilizada en la búsqueda, seguida de los términos *journalism news*, *disinformation*, *credibility* y los

diferentes meios sociais. Consta a relação de estas palavras com as subáreas de Comunicação, Computação, Artes y Humanidades, Ingeniería, entre otras. Presenta los países más prolíficos en el tema, destacando España, Estados Unidos, Reino Unido y Australia, y el inglés como idioma predominante.

Palabras clave: *Fake News*. Organización del conocimiento. Producción científica. Comunicación.

1 Introdução

O fenômeno denominado *fake News* atinge a sociedade em múltiplas dimensões. Ele ocorre hoje em todas as civilizações e, como objeto de estudo atualmente vem ganhando proeminência entre os profissionais da informação. Considerando-se as dimensões pelas quais é possível analisar o fenômeno *fake News*, no atual contexto teórico metodológico, questiona-se a possibilidade de construir uma análise multidimensional, para compreender os aspectos característicos dessa nova realidade. Tal como o estudo de qualquer fenômeno, este implica em observar e analisar a realidade concreta da qual participa, sem o abstrair da realidade.

Ao se tentar contribuir para a compreensão das características do fenômeno *fake News*, parte-se da produção científica sobre o tema, na área de ciências sociais, procurando transcender a visão específica de uma área, diante da possibilidade de identificar as relações interdominiais mais evidentes nos artigos identificados na janela temporal analisada.

Por meio deste estudo, procura-se, também, aplicar uma proposta de estrutura para a análise de pesquisas na área, identificando vertentes cuja proeminência é visível na literatura científica, mais amplamente nas ciências sociais. Busca-se, também, compreender, a partir da busca na literatura sobre o tema *fake News*, as suas principais características na produção científica, analisando-o por seu caráter informativo e na dimensão aplicada da Organização do Conhecimento, por meio de um estudo bibliométrico. Destaca-se que, entre as 11 abordagens pelas quais se analisa um domínio, uma delas são os estudos bibliométricos (HJØRLAND, 2002).

Para tanto, realizou-se uma pesquisa na base de dados *Scopus*, resultando a identificação inicial de 841 artigos científicos.

Apresenta-se a seguinte questão de pesquisa- é possível utilizar uma proposta de organização do conhecimento produzido sobre o tema, de modo a identificar suas vertentes, dimensões e categorias temáticas, a partir dos aspectos levantados em recorte da recente literatura analisada?

Definem-se como objetivos gerais, analisar a produção científica sobre *fake News* em âmbito internacional, na janela temporal de 2017 até 2022, para verificar como a OC pode absorver estas questões. De forma específica propõe-se caracterizar a produção científica sobre *fake News* e as temáticas mais citadas, a quais subáreas das ciências sociais se vinculam, quais os países mais proli-

ficos no tema, bem como os artigos que mais impactaram por meio das citações.

Justifica-se esta pesquisa, especialmente neste momento em que se intensifica a impregnação de questões como as relações entre a realidade das *fake News* e as TIs, as mídias, a literatura, os meios televisivos e o reconhecimento do modo como elas nos afetam. As situações políticas e questões ideológicas, causadas pelas *fake News*, vem repercutindo nos processos de organização do conhecimento, devido às pressões provocadas pelo volume e rapidez exacerbados pelo ritmo da indústria que as produz. Considera-se que o aprofundamento da crítica às práticas de organização do conhecimento sobre a temática, pode revelar suas relações de produção, assim como questões relativas às suas condições de divulgação.

2 Fundamentos teóricos

2.1 Organização do conhecimento: dimensão aplicada

Da percepção da Ciência da Informação (CI) como uma ciência social aplicada, capaz de integrar domínios convergentes, decorre uma postura investigativa compatível com sua natureza social, cultural, política e econômica, ao se discutir e compreender fenômenos relacionados à organização da informação. Ao se eleger a concepção da história apresentada por Marx e Engels (1989) como um dos pressupostos mais importantes para o estudo das *fake News* na literatura das ciências sociais, tem-se como objeto de estudo o processo efetivo da produção científica sobre o tema em sua relação com as questões dele decorrentes.

Nesse movimento, é possível apresentar a força da presença do tema na produção científica e a materialidade de sua existência imediata, percebida no campo das ciências sociais. Aspectos relativos ao tema nuclear, elementos teóricos e epistemológicos, entre outras possibilidades analíticas, ensejam a compreensão do processo de surgimento do fenômeno *fake News* no atual contexto político-econômico, social e cultural. A partir dessa compreensão, é possível apresentar questões relativas às *fake News* em suas dimensões ética, política e ideológica. Leituras assistemáticas fornecem um retrato de múltiplas perspectivas e tratamentos dados ao binômio “ética e informação”, sugerindo a ampliação das discussões sobre a ética na produção científica. Pode-se apontar para o fenômeno das *fake News*, um novo produto de consumo no mercado das ideias, como uma das causas para essa expressiva preocupação. No debate sobre as *fake News* são considerados os mecanismos de

poder econômico e político, seu potencial para esvaziar as formas tradicionais de resistência e, sobretudo, os meios de manipulação simbólica das consciências, tendendo a reduzi-las à sua dimensão positiva, legitimadora do *status quo*. Essa formação histórico-social do campo influencia diretamente o modo de consideração do objeto de estudo, ou seja, a eleição do procedimento analítico de pesquisa, a partir de um corpus de literatura científica representativo do fenômeno *fake News*, composto de artigos relativos aos impactos e à complexidade do fenômeno no atual contexto. Procura-se relacioná-lo, mais amplamente às dimensões sociopolítica, ética e ideológica, sob um enfoque aplicado da Organização do Conhecimento, aqui considerada um domínio voltado à organização do conhecimento e aos processos de ordenação, estruturação e sistematização dos conceitos, de acordo com suas características, passíveis de definição como elementos de herança do objeto e a aplicação dos conceitos e classes dos conceitos ordenados pela indicação de valores.

No caso deste estudo, com a intenção de compreender o fenômeno das *fake News* e suas características na produção das ciências sociais, as dimensões pertinentes relacionam-se aos procedimentos de Organização do Conhecimento, considerada por Dahlberg (2006), como a ciência ordenadora da estruturação e sistematização dos conceitos, de acordo com suas características. Essa caracterização é possível, segundo a autora, a partir da definição dos elementos de herança do objeto e da aplicação dos conceitos e classes ordenados pela indicação de valores e princípios reconhecidos (DAHLBERG, 2006). Com essa compreensão, justificam-se as pesquisas relacionadas às questões sociais, cognitivas, epistemológicas e metodológicas, incluindo-se entre elas os estudos voltados à ética e à política, no contexto da produção do conhecimento levantada. Isso contribui para compreender e ampliar os cenários de pesquisa sobre o tema na Organização do Conhecimento.

O necessário retorno à ética, na concepção de Badiou (1995, p. 16), designando hoje um princípio de relação com “o que se passa”, alerta para a tendência filosófica do momento, algo como um niilismo. Já no segundo momento, nega-se a vincular a ética a categorias abstratas, defendendo sua referência a “situações” e reconstruindo o que ele considera um “conceito admissível de ética, que subordina sua máxima ao devir das verdades” (BADIOU, 1995, p. 98).

Mais ou menos otimistas em relação ao futuro, os autores referentes para este estudo buscam o equilíbrio necessário entre os avanços tecnológicos e a postura ética. Targino e Cavalcanti referem-se ao projeto ideológico subjacente ao avanço das *fake News*, enquanto uma ameaça ao significado da ética da informação, junto ao imaginário social. Para tanto, profissionais de informação, demais especialistas que lidam com notícias, Governo e governantes devem fortalecer delinea-

mentos mais sólidos em direção à ética da informação (TARGINO; CAVALCANTE, 2023).

2.2 Conceito de “fake News”

No final da década de 1990 e início dos anos 2000, a *internet* passou por um processo de popularização em âmbito mundial. Isto se deu a partir da introdução das primeiras redes sociais e fez com que discursos migrassem cada vez mais para o meio digital. Consequentemente, proporcionou o avanço da globalização, novos meios de comunicação, interações *online* e inúmeras fontes de informações. O debate sobre as fronteiras da liberdade de expressão intensificou-se, devido à rapidez da produção, divulgação e disseminação de informações em *websites* e mídias sociais.

Houve, portanto, a migração do uso da imprensa tradicional como principal meio informacional para a mídia digital. D’Ancona (2018, p. 20) alega que “Os sites conspirativos e a mídia social tratam com desdém os jornais impressos ou a grande mídia [...] considerando-os a voz desacreditada de uma ordem ‘globalista’; uma ‘elite liberal’, cujo tempo já passou”, isto corrobora com esse processo de mudança do físico para o digital. Além disso, este fato favoreceu a “explosão informacional” (BUSH, 2011), em palavras gerais, houve um intenso crescimento na produção de documentos que compõem uma gama imensurável de informações às quais a população tem acesso, colaborando com o aumento de pessoas conectadas em rede (WORD, 2016), isto dificulta cada vez mais a conferência da veracidade da informação acessada. Associado a isso, a sociedade se depara com a infodemia, que consiste no intenso fluxo de informações de diferentes fontes e credibilidade que circulam nos meios digitais. A infodemia altera negativamente a compreensão e avaliação dos indivíduos em relação ao que são as notícias verdadeiras, o que contribui para o surgimento de rumores, falácias e manipulação da informação (ALMEIDA, 2020).

Durante esse processo de intensificação do uso de tecnologias, em 2016, nas eleições estadunidenses, o corpo social deparou-se com o fenômeno atualmente conhecido como “*fake News*”. Já no contexto nacional, ele tornou-se popular a partir de 2018, durante a corrida eleitoral. Apesar de informações falsas circularem na sociedade há anos, foi somente a partir desses acontecimentos que a “*fake News*” teve intensificada sua popularidade, tornando-se um fenômeno cada vez mais presente no dia a dia da população mundial. Para compreender do que se trata a *fake News*, cuja expressão é proveniente da língua inglesa e utilizada no Brasil tanto em sua forma original quanto traduzida para “notícias falsas”, Santos (2018, n.p), relata que “*fake News* é a arte de manipular as multidões em virtude de sua linguagem fácil e destinada a um público que já tenha uma opinião desfavorável em relação aos personagens envolvidos na mentira criada”. Em complemento, Iretton e Posetti (2018, p. 15) afirmam que “*Fake News* é

hoje muito mais do que um rótulo para informações falsas e enganosas, disfarçadas e divulgadas como notícias. Tornou-se um termo emocional, armado para debilitar e depreciar o jornalismo”. Nota-se que os produtores de *fake News* têm como principal objetivo, enganar e ludibriar a população, fazendo com que ela acredite em informações descontextualizadas ou até mesmo falsas e, conseqüentemente, compartilhem com outras pessoas, criando assim, uma rede de compartilhamento.

Como visto, criar e disseminar informações falsas e/ou mentiras, não é algo novo, essa prática existe há anos, o que há de novo em relação a esse fenômeno, é a forma como as notícias são criadas, compartilhadas e interpretadas. Quando a mídia tradicional era uma das principais fontes de informação da população, as matérias divulgadas tinham certa credibilidade, visto passarem por um conselho editorial e serem regidas por critérios, regulações e padrões editoriais aos quais são submetidos antes da publicação. Já com o processo e a intensificação da digitalização, percebe-se que “a internet e as redes sociais instauram uma lógica inédita imensamente facilitadora para a publicação e o compartilhamento” (SANTAELLA, 2018, p. 30), ou seja, informações de procedência duvidosa que são constantemente difundidas nas mídias sociais visando confundir o leitor.

A produção de *fake News* tem o potencial e objetivo de manipular as massas a partir da mimetização do estilo jornalístico. Diante disso, presencia-se a produção de *fake News* em diversas esferas e destacam-se duas motivações: ideológica e financeira (TANDOC JR; LIM; LING, 2018). A primeira está intrínseca no âmbito político, em que há um processo de decadência da verdade, isto ocorre, pois, a imparcialidade, seriedade, integridade e a transparência não prevalecem em trocas políticas (D’ANCONA, 2018). Já a segunda, a financeira, utiliza a técnica “caça-cliques”, que ocorre quando notícias viralizam nas mídias sociais e geram receita para os produtores quando os indivíduos clicam para ler a matéria na íntegra. A produção de lucro faz com que a motivação financeira seja o maior incentivo à criação de *fake News* (TRUZZI, 2019).

2.3 Indicadores de produção e citação temáticas

A análise da produção científica, em determinada janela temporal, envolve um amplo conjunto de indicadores bibliométricos, que se agrupam em indicadores de produção, indicadores de citação e indicadores de ligação (OKUBO, 1997; SPINAK, 1998; NARIN *et al.*, 1994; CALLON *et al.*, 1993). Entre os diversos aspectos analisados, os indicadores evidenciam os pesquisadores, as instituições, as temáticas, as áreas do conhecimento, os países mais férteis ou mais produtivos, assim como a frente de pesquisa de um campo de conhecimento, as redes de colaboração entre cientistas, grupos, instituições ou países e as redes de citação ou cocitação, bem como acoplamento bibliográfico.

Nesta pesquisa, são analisados, em âmbito internacional, os artigos encontrados na base de dados Scopus, na temática “*fake News*”. Para tanto, utilizam-se os indicadores bibliométricos de produção e citação (GREGOLIN, 2005), a partir dos quais se sinaliza o que é mais importante ou significativo dentro de um campo ou contexto científico. Por meio da análise de citações e análises das subtemáticas em questão, avaliam-se as diferentes direções e o comportamento apresentado pelo tema na produção científica. Deste modo, evidenciam-se o referencial teórico-epistemológico dominante e as relações existentes, constituindo-se um dos instrumentos metodológicos que contribuem para a visualização do comportamento da ciência.

Os indicadores básicos de produção constituídos pela contagem do número de publicações do pesquisador, grupo de pesquisadores, instituição ou país, objetivam refletir seu impacto junto à comunidade científica à qual pertencem, dando visibilidade àqueles mais produtivos e às temáticas mais destacadas na área do conhecimento analisada.

As temáticas apresentadas, por meio dos artigos recortados da base *Scopus*, refletem os temas e subtemas relevantes investigados, bem como explicitam os interdomínios observáveis. Sob a luz de Análise de Domínio, o conjunto das temáticas da literatura no recorte temporal de seis anos, proporciona o estudo da comunidade discursiva e do papel que ela desempenha na ciência, bem como reflete o universo conceitual quanto à linguagem e pensamento de uma dada comunidade discursiva, mostrando suas tendências epistemológicas e conceituais.

Quanto aos indicadores de citações, o processo de comunicação científica, tem na citação um de seus pilares fundamentais, embora seja um tema polêmico. Todo o processo da pesquisa científica, independente do assunto em apreço, baseia-se na literatura anterior, já construída sobre o tema. Assim, a pesquisa é um processo de construção coletiva determinado e tem como fundamento a literatura mais expressiva do referencial teórico anterior. A partir desta proposição, torna-se quase impossível escrever ou publicar sem citar os mais relevantes trabalhos de outros pesquisadores que o precederam. Considera-se assim que a construção do conhecimento científico novo acontece pela apropriação do conhecimento anteriormente socializado e construído.

Assim, um texto científico não se sustenta sozinho. Citar um trabalho anterior significa dar-lhe crédito, como também possibilita preservá-lo para a posteridade (ZIMAN, 1979). Investigar as citações que os pesquisadores recebem de suas publicações (SILVEIRA; BAZI, 2008), significa avaliar o impacto causado pela publicação em questão: um trabalho muito citado considera-se mais relevante que um trabalho pouco citado, embora se saiba que são inúmeras as razões para que se

cite um documento, que vão desde a qualidade do texto até questões menos relevantes, como simpatias e aproximações pessoais (MACIAS-CHAPULA, 1998).

Uma citação é tomada como indicador objetivo e claro da comunicação científica. Ela permite a identificação de grupos de cientistas e suas publicações, com a finalidade de evidenciar os pesquisadores de maior impacto de uma área, apontando seus paradigmas, procedimentos metodológicos pertinentes e os pesquisadores de “vanguarda” que constroem o novo conhecimento na área. Desse modo, a análise de citação contribui para o entendimento de uma comunidade científica, ao identificar os pesquisadores com maior impacto na área e dar visibilidade às referências teóricas que a sustentam, seus conceitos, objetos e métodos (GRÁCIO; OLIVEIRA; MATOS, 2009).

Em síntese, a “análise de citações mapeia a comunicação científica” (VANZ; CAREGNATO, 2003, p. 248), dá indicadores de como está ocorrendo a comunicação de uma área do conhecimento, contribuindo para a construção da rede de relações e explicita a comunicação e o relacionamento entre os pesquisadores.

A análise de citações se presta a avaliar as produções mais referenciadas (mencionadas) pelos demais pesquisadores de modo a prever seu impacto e estimar a absorção dos seus resultados. Assim, os indicadores de impacto das pesquisas mais citadas, podem acontecer pelos seus pesquisadores, periódicos, instituições ou países dominantes. Isto posto, quanto mais citado é o artigo, compreende-se que ele é mais aceito pela comunidade científica, portanto ele causa maior impacto.

3 Procedimentos metodológicos

Como procedimento de pesquisa, utilizou-se o estudo bibliométrico, na Organização do Conhecimento, em sua dimensão aplicada, considerando-se a possibilidade de conhecer as interrelações entre os artigos, entre os pesquisadores e os temas relevantes, assim como perceber os interdomínios.

Para tanto, realizou-se uma pesquisa na base de dados *Scopus*, em 17 de fevereiro de 2023, resultando a identificação inicial de 841 artigos científicos. Utilizaram-se três filtros, a saber: o termo *fake News*; a exclusividade dos idiomas inglês, espanhol e português, na janela temporal de 2017 a 2022 e apenas artigos em acesso aberto, na área de ciências sociais.

Foi utilizada a seguinte sintaxe: *TITLE-ABS-KEY (“Fake news”) AND PUBYEAR >2016 AND PUBYEAR <2023 AND (LIMIT-TO (OA , “all”)) AND (LIMIT-TO (DOCTYPE , “ar”)) AND (LIMIT-TO (LANGUAGE , “English”;) OR LIMIT- TO (LANGUAGE , “Spanish”;) OR LIMIT- TO (LANGUAGE , “Portuguese”;)) AND (LIMIT-TO (SUBJAREA , “SOCI”))*.

Inicialmente o recorte dos artigos foi exportado da *Scopus* em *Plain text*, e transformados em *Word*. A seguir os dados foram refinados e organizados em ordem decrescente de citação, sendo que uma certa quantidade deles nunca foi citada, provavelmente por serem artigos de produção recente. Foram retirados os artigos com zero citações e aqueles com menos de seis citações. Justifica-se este corte tomando-se como subsídio o conceito de Frente de Pesquisa de Braga (1973, p. 12), que afirma como condição suficiente e necessária para pertencer à Frente de pesquisa, o autor ter recebido pelo menos uma citação por ano. Estendeu-se o conceito à limitação do corpus dos artigos, resultando um conjunto de 315 artigos.

Para a análise temática, feita por meio das palavras-chave do autor, encontraram-se 19 artigos sem palavras-chave, construindo-se a nuvem de palavras com 296 artigos, retirados os 6,03% correspondentes aos 19 artigos, e representados pela nuvem de palavras, por meio do site *WordClouds* (<https://www.wordclouds.com>), desenvolvido pela *Zygomatic*, um site *online* e gratuito, que conta com diversas ferramentas.

Para a análise da produção e temática, considerou-se o corpus dos 315 artigos, agrupados por temáticas mais frequentes, quer sejam: 1. motivações: ideológica e/ou, política e/ou financeira; 2. subáreas a qual pertencem; 3. países mais prolíficos no tema em estudos, observando que as categorias de 2 e de 3 foram tiradas da *Scopus*, apresentados por meio de um quadro e análise dos 18 artigos mais citados. Seguiram-se os procedimentos inerentes ao site, o documento foi transformado em PDF e importado para o site. Foram feitas as respectivas análises dentro de cada grupo. Por fim, apresentam -se, do corpus de 315 artigos, os 18 mais citados e que impactaram mais, em ordem decrescente de citação, com a devida análise. Justifica-se este corte como extensão do conceito de Elite de pesquisa, relativa a pesquisadores, considerando-se um conjunto de tamanho N, cuja "elite" é igual à raiz quadrada de N. Assim, do conjunto de 315 artigos, a raiz quadrada seria aproximadamente 18 artigos, os quais representam a "elite" de pesquisa.

4 Apresentação e análise dos dados

Para análise de dados seguiu-se a seguinte ordem: apresentação e análise da nuvem de palavras, apresentação do quadro relativo às categorias e síntese dos 18 artigos mais citados, com destaque para as seguintes características: título do artigo, ano de publicação, número de citações, objetivos do artigo e análises pertinentes. Para finalizar, realizou-se uma síntese final geral dos 18 artigos.

4.1 Nuvem de palavras

Entre os países mais prolíficos, o destaque é para Espanha (69), que também aparece na Figura 1, nas palavras-chave aparece 11 vezes, seguem os EUA (66), Reino Unido (52) e Austrália (20). O Brasil aparece com 9 artigos, refletindo assim a pesquisa da comunidade científica sobre o tema.

Em relação ao idioma, o inglês aparece com 263 artigos, o espanhol com 51 e o português com 3, refletindo a predominância da língua a inglesa.

4.3 Síntese dos 18 artigos mais citados

Dentre o corpus utilizado para esta pesquisa, destaca-se o artigo intitulado “*Fake news and the economy of emotions: problems, causes, solutions*”, (2018). Recebeu 345 citações. A pesquisa visa analisar a campanha eleitoral estadunidense de 2016 com o intuito de identificar quais fatores intervenientes e possibilidades de solucionar questões relacionadas a *fake News*. Possui nitidamente a motivação política já comentada no referente teórico. Recomenda que seja dada maior atenção ao papel da publicidade digital em causar e combater tanto o fenômeno contemporâneo de notícias falsas, quanto a variante de notícias falsas automatizadas. É possível que o fato de ter sido publicado em 2018, tenha colaborado por ser o mais citado.

O segundo e o terceiro artigos mais citados, ambos publicados em 2020, contam com 221 citações cada e tratam de questões sobre a Covid-19. O segundo, denominado “*Impact of covid-19 on the media system. Communicative and democratic consequences of news consumption during the outbreak*”, descreve o consumo de informação da Covid-19 e a credibilidade dos meios de comunicação. O objetivo da investigação é analisar a influência do novo coronavírus no consumo de notícias, a credibilidade dada pelos cidadãos aos meios de comunicação, bem como a sua capacidade de detecção de notícias falsas. Os resultados da pesquisa confirmam o impacto da Covid-19 no sistema de mídia. Avaliou o papel da mídia, em especial da televisão. As desigualdades existentes no consumo de notícias entre os cidadãos foram reduzidas, em parte. Isso gera benefícios potenciais para a democracia em termos de igualdade e acessibilidade aos assuntos públicos. O terceiro artigo mais citado “*Whats drives unverified information sharing and cyberchondria during the COVID-19 pandemic?*”, é motivado pelo compartilhamento da população em informações falsas acerca do novo coronavírus.

O quarto, quinto e sexto textos (2019) retratam a problemática da disseminação de *fake News* nas mídias sociais. O quarto com 217 citações tem enfoque principalmente no *Facebook* e *Twitter*, denominado “*Trends in the diffusion of misinformation on social media*” (2019) e o quinto intitulado “*Fake news game confers psychological resistance against online misinformation*”, com 191 citações, cita um jogo desenvolvido

para tentar solucionar essa adversidade: a disseminação da desinformação. O próximo artigo, o sexto, de 2019, citado 175 vezes e denominado “*The Brexit Botnet and User-Generated Hyperpartisan*”, também aborda as mídias sociais, em especial o *Twitter*, porém seu foco é analisar *bots* (robôs) políticos concernentes ao referendo de adesão do Reino Unido à União Europeia.

O sétimo artigo (2018), com 162 citações denominado “*Dark participation*”, o oitavo (2021) com 141 citações intitulado “*Fake news detection: A hybrid CNN- RNN based deep learning approach*” e o nono artigo (2020), com 139 citações, denominado “*The Covid-19 ‘infodemic’: a new front for information professionals*” respectivamente, trabalham questões relativas à alguns aspectos essenciais da participação *dark*, entende-se aqui por participação obscura, tais como discursos de ódio, ataques a indivíduos ou grupos, disseminação estratégica de informação falsa e teorias da conspiração (DEYERL, 2022), bem como os benefícios de olhar para o lado perverso das coisas. Novas técnicas de detecção automática foram pesquisadas com base em inteligência artificial e aprendizado de máquina. A proposta é de um novo modelo híbrido de aprendizagem que combina redes neurais convolucionais e recorrentes para classificação de notícias falsas.

Outro artigo, o décimo (2020), com 134 citações, denominado “*Desinformación en tiempos de pandemia: Tipología de los bulos sobre la Covid- 19*” combina a Covid-19 e mídias sociais com foco no *WhatsApp*. O 11º artigo (2018), com 131 citações, denominado “*Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood*”, argumenta que o conceito de “notícias falsas” se tornou um componente importante nas lutas políticas contemporâneas, para desacreditar, atacar e deslegitimar opositores políticos. Portanto, é um artigo com conotação nitidamente política.

O 12º artigo (2021), com 124 citações, denominado “*An exploratory study of COVID-19 misinformation on Twitter*”, apresenta um estudo exploratório no *Twitter* sobre a proliferação de conteúdos relacionados ao novo coronavírus. A seguir, o 13º artigo (2018), com 123 citações, intitulado “*Do tabloids poison the well of social media? Explaining democratically dysfunctional news sharing*” combina dados de pesquisa sobre o comportamento de compartilhamento de notícias em redes sociais. O 14º artigo (2020), citado 98 vezes e denominado “*Misinformation and Polarization in a High-Choice Media Environment: How Effective Are Political Fact-Checkers?*”, visa investigar os efeitos da desinformação política integrando três abordagens teóricas: (1) desinformação, (2) polarização e (3) exposição seletiva. O 15º artigo (2020), com 87 citações, denominado “*We Are the People and You Are Fake News: A Social Identity Approach to Populist Citizens’ False Consensus and Hostile Media Perceptions*” obje-

tiva investigar as relações entre as atitudes populistas dos cidadãos, as percepções da opinião pública e as percepções dos principais meios de comunicação, baseando-se na teoria da identidade social.

O 16º artigo (2018), também com 87 citações, denominado “*Discursive Deflection: Accusation of “Fake News” and the Spread of Mis- and Disinformation in the Tweets of President Trump*” analisa os *Tweets* realizados por Donald Trump e como ele opera enquanto um disseminador de desinformação.

Por último, apresentam-se o 17º e 18º artigos, ambos com 82 citações cada. O 17º artigo (2021), denominado “*Partisan Polarization Is the Primary Psychological Motivation behind Political Fake News Sharing on Twitter*” analisa as motivações dos indivíduos para disseminar notícias falsas e conclui que são alimentadas pelas mesmas motivações psicológicas que impulsionam outras formas de comportamento partidário, incluindo as notícias partidárias. O 18º artigo (2020), intitulado “*Fake News Early Detection: A Theory-driven Model*” analisa a disseminação massiva de notícias falsas com a finalidade de corroer a democracia e seu potencial, fomentando a demanda por detecção por notícias falsas. O artigo mostra que há uma forte necessidade de desenvolver procedimentos e abordagens que possam detectar notícias falsas. O método investiga o conteúdo das notícias, em nível: léxico, sintaxe, semântico e do discurso, de modo a permitir a detecção precoce de notícias falsas, quando há informações de conteúdo limitadas.

A partir da síntese dos 18 artigos mais citados, destaca-se que a grande maioria trata de assuntos que relacionam a pandemia da Covid-19 e às mídias sociais, com objetivo predominantemente político, visto que o mundo vivenciou nos três últimos anos o isolamento social, o que auxiliou na intensificação do uso das mídias sociais, interações *online* e geração de notícias tendenciosas.

Uma das principais características da produção intelectual analisada é a vinculação das reflexões teóricas - a crítica midiática e política - e a análise contextual, especialmente sobre Covid-19 e eleições, bem como o papel dos profissionais da informação e da comunicação e das instituições na promoção da democracia. Visões alternativas foram questionadas para enfrentar as *fake News* como fenômeno social. Para tanto, profissionais de informação, demais especialistas que lidam com notícias, governo e governantes devem fortalecer delineamentos mais sólidos em direção à ética e à política da informação.

5 Conclusões

Diante das dimensões pelas quais foi possível analisar o fenômeno denominado *fake News*, no atual contexto

teórico metodológico das ciências sociais, a primeira questão proposta nesta pesquisa foi sobre a possibilidade de construir uma análise multidimensional, com apoio da organização do conhecimento, no sentido de compreender os aspectos característicos dessa realidade na literatura científica. Assim, confirma-se essa possibilidade de caracterizar a produção científica internacional sobre *fake News*, dos anos de 2017 a 2022, para verificar como a Organização do Conhecimento pode absorver a problemática, no contexto de uma sociedade em crise, com destaque para as dimensões ética e política, especialmente quando da interpretação de seus contornos e movimentos.

A análise dessa produção concreta foi mapeada em corpus, do ponto de vista da predominância temática e das subáreas das ciências sociais, dos países mais prolíficos no tema e dos artigos mais citados dessa produção. Observou-se grande dispersão em relação às autorias dos 315 artigos, significando assim, que há uma pulverização autoral, pois somente um autor, J. Roozenbeek participou de quatro artigos, M. Hamellers e Van der Linden assinaram três e R. Zafarani participou com dois. Os demais autores contribuem com apenas um artigo. Isto sugere que não há um corpo consignado de autores na temática, considerando o recorte de 315 artigos.

Quanto à predominância temática, foi possível perceber a evidência dos termos relacionados aos domínios distintos das Ciências Sociais, configurando-se interdomínios como o da comunicação midiática e a política, o da Ciência da Computação com Artes e Humanidades e com a Engenharia. A ética impõe-se como presença inegociável e permeia interdomínios nos quais são acolhidos conceitos representados pelos termos: verdade, confiança, informação falsa, teoria da conspiração, credibilidade, manipulação e atuação profissional. Constata a relação dessas palavras com as subáreas de Computação, Artes e Humanidades, Engenharia, Ciências da Decisão, Gestão, Psicologia, Ciência Ambiental e Economia, apresentando os países mais prolíficos sobre o tema, com destaque para Espanha, Estados Unidos, Reino Unido e Austrália e como idioma predominante o inglês. Particularmente nos domínios da informação e da comunicação, dada a natureza de seu objeto de estudo e trabalho, a ética permeia os estudos, envolvendo questões como a veracidade da informação como valor para sobreviver na sociedade contemporânea, direito de todos, condição de cidadania – seja no texto acadêmico, seja no tempo/espço da atuação laboral.

Referências

ALLCOTT, H; GENTZKOW, M. Social media and fake news in the 2016 election. **Journal of Economic Perspectives**. v. 31, n. 2, p. 211-236, 2017.

- ALMEIDA, A. S. As bibliotecas universitárias no combate à infodemia. **Revista Informação & Universidade**, v. 2, p. 1-19, 2020.
- BADIOU, A. **Ética: um ensaio sobre a consciência do mal**. Rio de Janeiro: Relume Dumará, 1995.
- BRAGA, G. M. Relações bibliométricas entre a frente de pesquisa (research front) e revisões da literatura: estudo aplicado a ciência da informação. **Ciência da Informação**, v.2, n.1, p.9-26, 1973.
- BUFREM, L. S.; FREITAS, J. L. Interdomínios na literatura periódica científica da ciência da informação. **DataGrama-Zero**, v. 16, n. 4, 2015. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/8295>. Acesso em: 26 mar. 2023.
- BUSH, V. Como podemos pensar. **Rev. latinoam. psicopatol. fundam.**, São Paulo, v. 14, n. 1, p. 14-32, mar. 2011. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415. Acesso em fev. de 2023.
- CALLON, M.; COURTIAL, J.-P.; PENAN, H. **Cientometria: el estudio cuantitativo de la actividad científica: de la bibliometria a la vigilancia tecnológica**. Gijón: Ediciones Trea, 1993.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, n.33, v.1, p.11-19. 2006.
- D'ANCONA. **Pós-Verdade: a nova guerra contra os fatos em tempos de fake news**. Barueri: Faro Editorial, 2018.
- DEYERL, Kilian. **Lecture on Dark Participation by Dr. Lena Frischlich**. 2022. Disponível em: <https://www.explore-interactions.de/cohesion/blog/2022/04/20/lecture-dark-participation-lena-frischlich/>. Acesso em: 18 abr. 2023.
- IRETON, C.; POSETTI, J. **Jornalismo, Fake News & Desinformação: manual para educação e treinamento em jornalismo**. UNESCO, 2019.
- GREGOLIN, J. A. R. Análise da produção científica a partir de indicadores bibliométricos. In: LANDI, F. R. **Indicadores de ciência, tecnologia e inovação em São Paulo**. São Paulo: Fundação de Amparo à Pesquisa do Estado de São Paulo, 2005. p. 5-43.
- GRÁCIO, M. C. C.; OLIVEIRA, E. F. T.; MATOS, G. I. Visibilidad de los investigadores en el tema Estudios Métricos: análisis de citación y cocitación en los periódicos del SciELO. IBERSID: **Revista de Sistemas de Información y Documentación**, Zaragoza, v. 3, p. 81-86, 2009.
- HJØRLAND, B. Domain analysis information Science: eleven approaches-tradicional as well as innovative. **Journal of Documentation**. London, v. 58, n. 4, p. 422-462, 2002.
- MACIAS-CHAPULA, C. O papel da informetria e da cienciometria sua perspectiva nacional e internacional. **Ciência da Informação**, Brasília, DF, v. 27, n. 2, p. 134-140, 1998.
- MARX, K.; ENGELS, F. A ideologia alemã. In: FERNANDES, Florestan (Org.). **Marx e Engels: história**. São Paulo: Ática, 1989. p. 182-214.
- NARIN, F.; OLIVASTRO, D.; STEVENS, K. S. Bibliometric theory, practice and problem. **Evaluation Review, Beverly Hills**, v. 18, n. 1, p. 65-76, 1994.
- OKUBO, Y. **Bibliometric indicators and analysis of research systems: methods and examples**. Paris: OECD, 1997.
- SANTAELLA, L. **A pós-verdade é verdadeira ou falsa?** Editora estação das letras e cores, 2018.
- SANTOS, J. C. S.; SOUZA, V. M. R.; LAVIGNE, F. C. Desinformação, pós-verdade e comportamento humano: discussões plausíveis. **BIBLOS**, v. 34, n. 2, 2020.
- SANTOS, R. R. O. **Fake news como produto da pós-verdade**. 2018. Disponível em: <http://observatoriodaimprensa.com.br/comunicacao-social/fake-news-como-produto-da-pos-verdade/>. Acesso em: 13 fev.2023.
- SILVEIRA, M. A. A.; BAZI, R E. R. A Ciência da Informação no Brasil e sua frente de pesquisa: estudo cienciométrico sob a ótica da institucionalização da pesquisa científica (1995-2005). **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 13, n. 26, p. 1-16, 2008. Disponível em: <https://www.redalyc.org/articulo.oa?id=14712794002>. Acesso em: 21 abr. 2020.
- SPINAK, E. Dicionário enciclopédico de bibliometria, cienciometria e informetria. Caracas: UNESCO, CII/II, 1996.
- TANDOC JR, E. C.; LIM, Z. W.; LING, R. Defining “fake news” A typology of scholarly definitions. **Digital Journalism**, v. 6, n. 2, p. 137-153, 2018.
- TARGINO, M. D. G.; CAVALCANTE, A. V. B. Admirável mundo novo da ética da informação 2.0 em tempos de fake news. **Informação em Pauta**, v. 5 No 1, n. 1, p. 33-53, 2020.
- TRUZZI, G. **Fake news: quais os impactos das notícias falsas no seu negócio?** 2019. Disponível em: <https://www.istoedinheiro.com.br/fake-news-quais-os-impactos-das-noticias-falsas-no-seu-negocio-como-esse-fenomeno-pode-acabar-com-a-reputacao-da-sua-empresa/>. Acesso em: 10 fev. 2023.
- VANZ, S. A. de S.; CAREGNATO, S. E. Estudos de citação: uma ferramenta para entender a comunicação científica. **Em Questão**, Porto Alegre, v. 9, n. 2, p. 295-307, jul./dez. 2003.
- ZIMAN, J. M. **Conhecimento público**. Belo Horizonte: São Paulo, 1979.

A importância da organização e representação de informação para a *Search Engine Optimization*

The importance of information organization and representation for Search Engine Optimization

Gustavo Camossi (1), Heytor Diniz Teixeira (2), Cecilio Merlotti Rodas (3)

(1) Universidade Estadual Paulista (UNESP), Av. Hygino Muzzi Filho, 737 – Mirante – Marília/SP – CEP 17.525-900, gustavo.camossi@unesp.br, (2) hd.teixeira@unesp.br, (3) cecilio.rodas@unesp.br

Resumo

Este estudo tem como objetivo discutir a relação entre Organização e Representação do Conhecimento (ORC) e as técnicas de *Search Engine Optimization* (SEO), a fim de demonstrar a importância de uma boa estruturação de informações para a visibilidade e posicionamento de *sites* nos resultados de pesquisa. A metodologia utilizada consistiu em uma revisão bibliográfica, em que foram selecionados artigos científicos relacionados à ORC e SEO, avaliados quanto à relevância e analisados em sua totalidade, visando à identificação das informações relevantes para a construção de um panorama abrangente sobre a relação entre esses dois conceitos. A análise dos artigos selecionados permitiu identificar a importância da organização e representação adequada de informações para a otimização de mecanismos de busca. Foram observados que a implementação de estratégias de ORC resultou em melhorias significativas nos resultados de pesquisa. Também foi destacada a importância da utilização de metadados, palavras-chave e informações estruturadas para a identificação e indexação de conteúdo relevante pelos mecanismos de busca. Conclui-se que a ORC é um fator crucial para a otimização de mecanismos de busca, pois permite que os mecanismos de pesquisa encontrem e indexem conteúdo relevante e apresentem aos usuários resultados mais precisos e relevantes. A aplicação de técnicas de ORC, como a utilização de palavras-chave, metadados e informações estruturadas, pode resultar em melhorias significativas nos resultados de pesquisa e, conseqüentemente, no posicionamento e visibilidade de *sites* na *web*.

Palavras-chave: Organização e Representação do Conhecimento. *Search Engine Optimization*. Posicionamento nos resultados de busca.

Abstract

This study aims to discuss the relationship between Knowledge Organization and Representation (KRO) and Search Engine Optimization (SEO) techniques, in order to demonstrate the importance of good information structuring for the visibility and positioning of websites in search results. The methodology used consisted of a literature review, in which scientific articles related to ORC and SEO were selected, evaluated for relevance, and analyzed in their entirety in order to identify relevant information for the construction of a comprehensive overview of the relationship between these two concepts. The analysis of the selected articles allowed us to identify the importance of the adequate organization and representation of information for search engine optimization. It was observed that the implementation of ORC strategies resulted in significant improvements in search results. The importance of using metadata, keywords, and structured information for the identification and indexing of relevant content by search engines was also highlighted. It is concluded that ORC is a crucial factor for search engine optimization, as it enables search engines to find and index relevant content and present users with more accurate and relevant results. The application of ORC techniques, such as the use of keywords, metadata and structured information, can result in significant improvements in search results and, consequently, in the positioning and visibility of websites.

Keywords: Knowledge Organization and Representation. Search Engine Optimization. Positioning in search results.

Resumen

Este estudio tiene como objetivo discutir la relación entre la Organización y Representación del Conocimiento (ORC) y las técnicas de Search Engine Optimization (SEO), con el fin de demostrar la importancia de una buena estructuración de la información para la visibilidad y posicionamiento de los sitios en los resultados de búsqueda. La metodología utilizada consistió en una revisión bibliográfica, en la que se seleccionaron artículos científicos relacionados con ORC y SEO, se evaluó su relevancia y se analizaron en su totalidad con el fin de identificar información relevante para construir una visión global de la relación entre estos dos conceptos. El análisis de los artículos seleccionados permitió identificar la importancia de la adecuada organización y representación de la información para el posicionamiento en buscadores. Se observó que la aplicación de estrategias de ORC producía mejoras significativas en los resultados de búsqueda. También se destacó la importancia del uso de metadatos, palabras clave e información estructurada para la identificación e indexación de contenidos relevantes por parte de los motores de búsqueda. Se concluye que el ORC es un factor crucial para la optimización de los motores de búsqueda, ya que les permite encontrar e indexar contenidos relevantes y presentar a los usuarios resultados más precisos y pertinentes. La aplicación de técnicas de ORC, como el uso de palabras clave, metadatos e información

estructurada, puede dar lugar a mejoras significativas en los resultados de las búsquedas y, en consecuencia, en el posicionamiento y la visibilidad de los sitios web.

Palabras clave: Organización y representación del conocimiento. Optimización de motores de búsqueda. Posicionamiento en los resultados de búsqueda.

1 Introdução

A internet se tornou uma fonte inesgotável de informação, e a capacidade de encontrar e acessar informações relevantes é essencial para o sucesso de muitas empresas e organizações. Nesse contexto, a Organização e Representação do Conhecimento (ORC) e o *Search Engine Optimization* (SEO) se destacam como campos fundamentais para a gestão de informações na *web*.

A ORC é um campo que se preocupa em criar sistemas para organizar, representar e recuperar informações, enquanto o SEO é uma estratégia de *marketing* digital que visa melhorar o posicionamento de um *site* nos resultados de busca. Embora pareçam áreas distintas, a relação entre a ORC e o SEO é mais estreita do que se imagina.

Um dos propósitos da Organização e Representação do Conhecimento é eficientemente organizar e tornar a informação acessível (RODRIGUES, 2016). Para isso, é necessário que a informação seja estruturada de maneira clara e coerente, utilizando padrões e normas que facilitem a recuperação da informação pelos usuários. A partir daí, é possível aplicar técnicas de SEO para otimizar o posicionamento do *site* nos resultados de busca.

A seleção cuidadosa de palavras-chave pertinentes é uma das principais técnicas de SEO empregadas para aprimorar o posicionamento de um *site* nos resultados de pesquisa (MARTINS, 2018). No entanto, essa técnica também tem tudo a ver com a organização e representação de informação. Quando as palavras-chave são definidas com base nas necessidades e expectativas do usuário, a organização e representação da informação torna-se mais eficiente, contribuindo para a otimização do SEO.

Uma outra técnica importante de SEO é a otimização da arquitetura do *site* (BEAL, 2020). Isso envolve a criação de uma estrutura clara e organizada do *site*, facilitando a navegação, a recuperação e a encontrabilidade de informações pelos usuários. A aplicação de técnicas de ORC pode contribuir significativamente para a otimização da arquitetura do *site*, tornando-o mais fácil de navegar e mais amigável para os usuários.

Além disso, a aplicação de metadados é outra técnica de ORC que pode ser utilizada para otimizar o SEO. Os metadados são informações descritivas que auxiliam na identificação e descrição do conteúdo do *site* (RODRIGUES, 2016). A utilização de metadados precisos na descrição pode melhorar a visibilidade e a posição do *site* nos resultados de busca.

A usabilidade do *site* é um fator importante a ser considerado para o SEO (BEAL, 2020). A usabilidade se refere à facilidade de uso do *site* pelos usuários, o que inclui a clareza da linguagem, a organização dos conteúdos e a facilidade de navegação. Além disso, a otimização para mecanismos de busca também depende de uma boa estruturação e organização da informação, com a utilização de palavras-chave e metadados precisos.

Nesse contexto, a ORC pode ser vista como uma área fundamental para o sucesso de uma estratégia de SEO. A ORC busca organizar e representar informações de forma coerente e estruturada, possibilitando a recuperação eficiente da informação. Dessa forma, a aplicação de técnicas de ORC pode ajudar a melhorar a visibilidade dos conteúdos em mecanismos de busca, tornando-os mais facilmente encontrados pelos usuários.

Por outro lado, a SEO pode auxiliar na otimização da apresentação dos conteúdos na *web*, por meio da escolha adequada de palavras-chave, da estruturação das informações e da utilização de técnicas para aumentar a relevância dos conteúdos para os mecanismos de busca. A junção dessas áreas pode contribuir para um melhor desempenho na busca por informações na *web*, permitindo que os usuários encontrem de forma mais rápida e eficiente os conteúdos desejados.

Nesse sentido, este trabalho tem como objetivo analisar a relação entre a ORC e a SEO, demonstrando como a aplicação de técnicas de ORC pode contribuir para uma melhor performance de SEO, e como a utilização de técnicas de SEO pode auxiliar na melhoria da organização e apresentação dos conteúdos na *web*. Para isso, será realizada uma revisão bibliográfica que aborda os conceitos e práticas relacionadas a ORC e SEO, buscando identificar pontos em comum e possíveis integrações entre as duas áreas.

2 Organização e Representação do Conhecimento e a Gestão da Informação

A Organização e Representação do Conhecimento (ORC) é uma área de estudo que se dedica a analisar e desenvolver estratégias para a organização, armazenamento, recuperação e disseminação da informação em ambientes digitais. De acordo com Souza e Tarapanoff (2014), a ORC se fundamenta na teoria da representação da informação, que estuda como a informação é estruturada e organizada para fins de armazenamento, recuperação e disseminação.

A importância da ORC para a gestão da informação se deve ao seu papel fundamental em garantir a qualidade da informação e sua acessibilidade, possibilitando que os usuários encontrem e utilizem a informação de forma eficiente e eficaz. Hjørland (2017) destaca que a ORC é crucial para a organização e acesso à informação em ambientes digitais.

A organização da informação compreende a análise e estruturação dos dados e informações para facilitar sua recuperação e utilização pelos usuários, enquanto a representação da informação envolve a criação de descrições e metadados que permitam a identificação e recuperação dos recursos informacionais em ambientes digitais (CHAVES; SALES, 2018).

Na era digital, a ORC tornou-se especialmente relevante diante da vasta e complexa quantidade de informações disponíveis na rede. Tarapanoff (2015) destaca a importância da ORC para a garantia da acessibilidade e qualidade da informação disponível na *web*, permitindo que as informações sejam encontradas e utilizadas de forma eficiente pelos usuários.

Além disso, a ORC é essencial para a interoperabilidade entre sistemas de informação, possibilitando que diferentes sistemas se comuniquem e compartilhem informações de forma padronizada. Zeng e Qin (2008) apontam que a interoperabilidade é fundamental para a integração de sistemas de informação em diferentes áreas, tais como a biblioteca digital, a gestão de documentos e a gestão do conhecimento.

Portanto, a ORC desempenha um papel fundamental na organização, armazenamento, recuperação e disseminação da informação em ambientes digitais, tendo impacto direto na qualidade e acessibilidade da informação, na interoperabilidade entre sistemas e na sua utilização pelos usuários.

2.1 As técnicas de organização e representação de informação

A organização e representação da informação são fundamentais para a gestão da informação e para a sua recuperação posterior. Entre as técnicas utilizadas para tal, destacam-se a indexação, a catalogação e a classificação.

A indexação é uma técnica que consiste em atribuir termos representativos ao conteúdo da informação, com o objetivo de torná-la mais fácil de ser recuperada. Segundo Lancaster (1993), indexação é "[...] a atividade que consiste em identificar e descrever o conteúdo da informação, por meio de um vocabulário controlado, de modo a permitir sua recuperação posterior". A indexação pode ser realizada de forma manual ou automática, sendo a segunda opção mais comum atualmente, com o uso de algoritmos e técnicas de processamento de linguagem natural.

A catalogação, por sua vez, é uma técnica que consiste em descrever e identificar um documento de acordo com um padrão de descrição. A catalogação tem como objetivo fornecer informações suficientes para identificar e localizar um documento em um sistema de recuperação de informação (SLATER, 2014). A catalogação é realizada segundo normas e padrões internacionais, como a RDA (*Resource Description and Access*), que oferece instruções a respeito de como os elementos necessários para a descrição bibliográfica devem ser registrados.

Já a classificação é uma técnica que tem como objetivo agrupar documentos semelhantes em classes, tornando a recuperação da informação mais eficiente. A classificação é realizada de acordo com um sistema que atribui um número ou código a cada documento, de acordo com suas características (DAHLBERG, 2010). Dentre os sistemas de classificação mais utilizados destacam-se a Classificação Decimal de Dewey e a Classificação Decimal Universal.

Além das técnicas citadas acima, existem outras técnicas de organização e representação da informação que podem ser utilizadas, como a ontologia, que consiste em uma representação formal e explícita de conceitos e suas relações em um determinado domínio. Segundo Guarino e Welty (2002), uma ontologia "[...] é uma especificação explícita de uma conceitualização compartilhada".

Outra técnica importante é a taxonomia, que consiste em uma estrutura hierárquica de categorias que representam a organização do conhecimento em um determinado domínio. De acordo com Vakkari (2001), "[...] uma taxonomia é um esquema classificatório com uma estrutura hierárquica que tem por finalidade agrupar objetos e conceitos em categorias, de acordo com suas semelhanças e diferenças".

A organização e representação da informação também podem ser realizadas através do uso de metadados, que são informações descritivas sobre um objeto informacional, como título, autor, data de criação, entre outras. Segundo NISO (2004), "[...] os metadados são informações estruturadas que descrevem os recursos informacionais, tornando-os mais fáceis de localizar, recuperar, gerenciar e utilizar".

Por fim, destaca-se a importância do uso de vocabulários controlados para a organização e representação da informação. Segundo Zeng e Qin (2008), um vocabulário controlado é um conjunto de termos padronizados que são utilizados para indexar e representar conceitos em uma determinada área do conhecimento.

Todas essas técnicas têm como objetivo principal tornar a informação mais acessível e fácil de encontrar para o usuário, seja por meio de sistemas de busca ou de navegação em ambientes virtuais. Além disso, essas técnicas são fundamentais para a gestão da informação, permitindo que as instituições organizem seus recursos de

informação de forma eficiente e eficaz, facilitando a tomada de decisões e o compartilhamento de conhecimento (LIMA-MARQUES, 2017).

Além das técnicas tradicionais de organização e representação de informações, é importante destacar que o uso de SEO também pode contribuir para a melhoria da visibilidade e acessibilidade de conteúdo na *web*. A utilização de palavras-chave, *meta tags* e URLs amigáveis, por exemplo, podem auxiliar no processo de indexação e ranqueamento de páginas nos mecanismos de busca. Dessa forma, é fundamental que profissionais da informação e da área de *marketing* digital estejam atualizados e capacitados para aplicar essas técnicas em seus trabalhos, visando aprimorar a disseminação do conhecimento e a visibilidade de suas organizações e/ou projetos na *web*.

3 Search Engine Optimization

As técnicas de *Search Engine Optimization* (SEO) são essenciais para melhorar a visibilidade de *sites* em mecanismos de busca e, conseqüentemente, aumentar o tráfego e a relevância do conteúdo. Dentre as principais técnicas utilizadas, a escolha de palavras-chave é um fator crucial para que o conteúdo seja facilmente encontrado pelos usuários. Segundo Zhang *et al.*, (2019), é importante escolher palavras-chave relevantes e que correspondam ao conteúdo, para que os mecanismos de busca possam classificar o *site* com mais precisão.

Outra técnica importante é a otimização do conteúdo, que envolve a organização e estruturação de forma a facilitar a compreensão pelos usuários e pelos mecanismos de busca. De acordo com Farooq *et al.*, (2020), a otimização do conteúdo inclui a utilização de títulos e subtítulos claros, além da inserção de imagens e vídeos relevantes. Além disso, a utilização de meta descrições é importante para fornecer um resumo do conteúdo da página e atrair a atenção dos usuários nos resultados de pesquisa.

A construção de *links* também é uma técnica relevante em SEO, que consiste em obter *links* de outros *sites* que apontem para o seu conteúdo. Segundo Xie *et al.*, (2019), essa técnica pode melhorar a autoridade do domínio do *site*, além de aumentar a relevância e a credibilidade do conteúdo para os mecanismos de busca. No entanto, é importante destacar que a qualidade dos *links* é mais importante do que a quantidade, e que a utilização de *links* de baixa qualidade pode prejudicar o ranqueamento do *site*.

A utilização de *meta tags* também é uma técnica importante em SEO, que envolve a inserção de informações adicionais sobre o conteúdo da página, como o título, descrição e palavras-chave. De acordo com Moz (2021), a utilização adequada de *meta tags* pode ajudar os mecanismos de busca a entender melhor o conteúdo da

página e classificá-la de forma mais precisa nos resultados de pesquisa.

Outra técnica relevante é a utilização de URL amigáveis, que são URLs que incluem palavras-chave e que são facilmente compreendidas pelos usuários e pelos mecanismos de busca. De acordo com Chen *et al.*, (2019), a utilização de URLs amigáveis pode melhorar a experiência do usuário e facilitar a indexação do conteúdo pelos mecanismos de busca.

A otimização de imagens, para que elas carreguem mais rapidamente e tenham um tamanho adequado para o uso na página, também é uma técnica importante em SEO. Segundo Feng *et al.*, (2020), a otimização de imagens pode melhorar a experiência do usuário e reduzir o tempo de carregamento da página, o que é um fator importante para a classificação nos resultados de pesquisa.

A utilização de redes sociais também é uma técnica relevante em SEO, que envolve a criação e compartilhamento de conteúdo para atrair tráfego para o *site*. Segundo Olmedilla *et al.*, (2019), a utilização adequada das redes sociais pode aumentar a visibilidade do conteúdo e atrair novos usuários para o *site*, o que pode resultar em um aumento no tráfego orgânico e, conseqüentemente, na melhoria do ranqueamento nos mecanismos de busca. Para isso, é importante que o conteúdo seja otimizado para as redes sociais, levando em consideração as particularidades de cada plataforma.

Outra técnica importante de SEO é a construção de *links*, que consiste em obter *links* externos apontando para o *site*, o que é visto pelos mecanismos de busca como um sinal de relevância e autoridade. Segundo Liu e Lu (2018), a construção de *links* é uma das estratégias mais eficazes para melhorar o ranqueamento em mecanismos de busca, porém é importante que os *links* sejam obtidos de forma natural e relevante, evitando práticas consideradas *spam*.

Por fim, é importante destacar que as técnicas de SEO devem ser utilizadas de forma ética e transparente, evitando práticas consideradas *spam* ou que possam prejudicar a experiência do usuário. Segundo Cho e Cheon (2019), a ética é um aspecto fundamental do SEO, pois a credibilidade e a reputação do *site* também são levadas em consideração pelos mecanismos de busca.

Portanto, as técnicas de SEO são importantes para aumentar a visibilidade e a relevância de um *site* nos mecanismos de busca. A escolha de palavras-chave adequadas, a otimização do conteúdo, a construção de *links* e a utilização adequada das redes sociais são algumas das técnicas mais eficazes para melhorar o ranqueamento nos mecanismos de busca e atrair mais tráfego orgânico. No entanto, é importante que as técnicas sejam utilizadas de forma ética e transparente, levando em consideração a qualidade do conteúdo e a experiência do usuário.

4 Metodologia

Para realizar este estudo, utilizamos uma metodologia baseada em revisão bibliográfica. Segundo Lopes *et al.*, (2020), a revisão bibliográfica é uma técnica utilizada para sistematizar e integrar as informações já disponíveis sobre um tema específico, permitindo identificar as lacunas de conhecimento existentes na literatura e propor novas abordagens e perspectivas para o assunto em questão.

Para selecionar as fontes bibliográficas utilizadas neste estudo, foi realizada uma pesquisa em bases de dados online, tais como *Scopus* e *Web of Science*, utilizando os seguintes termos de busca: "organização do conhecimento", "representação do conhecimento", "*Search Engine Optimization*", as buscas foram realizadas entre os dias 01 de março de 2023 e 01 de abril de 2023. A partir dessa pesquisa inicial, foram selecionados os artigos que apresentavam maior relevância para o tema em questão e que continham informações e conceitos pertinentes para a análise proposta.

Quadro 1. *Números de artigos recuperados nas bases de dados*

<i>Base de dados</i>	<i>Strings de busca</i>	<i>Recuperados</i>
<i>Scopus e Web of Science</i>	"organização do conhecimento"	65
<i>Scopus e Web of Science</i>	"representação do conhecimento"	18
<i>Scopus e Web of Science</i>	" <i>Search Engine Optimization</i> "	717

Fonte: Os autores (2023).

Segundo Cooper (2017), a revisão bibliográfica pode ser dividida em três etapas: seleção, análise e síntese das informações. Na etapa de seleção, foram avaliados os títulos e resumos dos artigos encontrados na pesquisa, verificando sua relevância para o tema proposto. Na etapa de análise, foram lidos e avaliados os artigos selecionados, buscando identificar as informações e conceitos relevantes para a análise proposta. Na etapa de síntese, foram integradas as informações e conceitos identificados nas etapas anteriores, visando a construção de um panorama abrangente sobre a relação entre ORC e SEO.

Além disso, foram utilizados critérios de inclusão e exclusão para a seleção dos artigos, visando garantir a qualidade e a relevância dos trabalhos utilizados. Foram incluídos artigos publicados em periódicos científicos de reconhecimento internacional e que apresentassem relevância para o tema proposto. Foram excluídos trabalhos que apresentassem baixa qualidade metodológica ou que não estivessem diretamente relacionados com o tema em questão.

Como resultado, 15 documentos foram selecionados para compor a análise, uma vez que os assuntos abordados possuem relações e semelhanças com a temática abordada pelo estudo.

Com base nessa metodologia, foi possível realizar uma análise crítica e aprofundada sobre a relação entre ORC e SEO, visando identificar possíveis pontos de convergência e integração entre as duas áreas.

5 Discussão dos resultados

A organização e a representação adequadas de informações são fundamentais para a otimização de mecanismos de busca (SEO), pois permitem que encontrem e indexem conteúdo relevante e apresentem aos usuários resultados mais precisos e relevantes.

Nos documentos recuperados, pode-se observar que a organização e a representação de informações foram cruciais para a otimização de SEO. No estudo realizado por Rizwan *et al.*, (2020), os autores analisaram a importância do SEO para a visibilidade e rentabilidade de lojas virtuais. Eles coletaram dados de 150 lojas virtuais de diversos setores e mediram o desempenho delas em termos de visibilidade nos mecanismos de busca e rentabilidade financeira.

Os resultados mostraram que as lojas que implementaram estratégias de SEO tiveram um desempenho melhor tanto em termos de visibilidade quanto de rentabilidade. Além disso, a pesquisa identificou que as palavras-chave têm um papel crucial na otimização do SEO, e que a otimização de imagens e a melhoria da experiência do usuário também são fatores importantes.

Um outro estudo de SEO realizado por Ma *et al.*, (2019), os autores realizaram uma revisão sistemática da literatura sobre estratégias de SEO aplicadas ao setor de turismo. A pesquisa teve como objetivo identificar as melhores práticas para a otimização de *sites* de turismo nos mecanismos de busca, a fim de melhorar a visibilidade e aumentar a quantidade de tráfego orgânico. Entre os resultados encontrados pelos autores, destacam-se a importância de palavras-chave relevantes e conteúdo de qualidade, bem como a relevância das *metas tags*, URLs amigáveis e *links* externos de qualidade. Além disso, os autores apontaram que a otimização para dispositivos móveis e a presença nas mídias sociais são cada vez mais importantes para a melhoria do SEO no setor de turismo.

Já no estudo realizado por Luh *et al.*, (2016) e Erdmann *et al.*, (2022), o estudo se baseou na análise de uma ampla variedade de fatores de SEO, como palavras-chave, *links* externos e títulos de página, entre outros. Os resultados mostraram que a otimização dos mecanismos de busca pode ter um impacto significativo no posicionamento de um *site* nos resultados de pesquisa do Google, mas que o algoritmo de classificação do Google é muito

complexo e leva em consideração uma ampla variedade de fatores além dos relacionados ao SEO.

Os autores Berman e Katona (2013), apresentam uma análise sobre o papel do SEO, com o objetivo de entender como a utilização de técnicas de otimização de busca pode afetar a efetividade das estratégias de *marketing* digital. A pesquisa utiliza dados reais de uma empresa de comércio eletrônico para avaliar o impacto da implementação de estratégias de SEO na visibilidade dos anúncios da empresa nos resultados de busca do Google. Os resultados mostram que a implementação de técnicas de SEO aumenta significativamente a visibilidade dos anúncios, bem como a taxa de cliques e as vendas da empresa.

Para Killoran (2013), aborda que ao utilizar as técnicas de otimização de busca, há uma melhora da velocidade de carregamento do *site*. O autor também destaca a importância de uma estratégia de *marketing* de conteúdo bem planejada, que inclui a criação de conteúdo relevante e original, a promoção em redes sociais e a obtenção de *backlinks* de qualidade.

No trabalho desenvolvido por Drivas *et al.*, (2020), os autores discutem a importância do uso de *big data analytics* na área de SEO, e argumentam que as estratégias de SEO podem se beneficiar do uso de análises de *big data*, que podem fornecer informações sobre o comportamento do usuário, o uso de palavras-chave e outros fatores que afetam as classificações dos mecanismos de pesquisa. O artigo discute várias ferramentas e técnicas para coletar e analisar dados, bem como como essas informações podem ser usadas para melhorar o desempenho do SEO.

No artigo desenvolvido por Onaifo e Rasmussen (2013), Shevchenko (2018) e Lee *et al.*, (2016) os autores discutem a aplicação de técnicas de SEO para melhorar a descoberta e a visibilidade do conteúdo da biblioteca na *web*. Os autores argumentam que as bibliotecas precisam implementar práticas de SEO para se manterem competitivas na era digital e oferecem recomendações para o desenvolvimento de uma estratégia de SEO. O artigo inclui um estudo de caso da implementação de SEO em uma biblioteca universitária, o que resultou em aumento do tráfego para o *site* da biblioteca e melhor classificação nos mecanismos de busca. O estudo conclui que o SEO é uma ferramenta essencial para que as bibliotecas se conectem com os usuários e potencializem o impacto de seus recursos.

Suchanek (2010), enfoca os fundamentos da criação de um *site* de comércio eletrônico de sucesso com ênfase na otimização de mecanismo de busca (SEO). A metodologia de pesquisa utilizada foi um estudo de caso de uma loja online de sucesso, analisando suas funcionalidades, interface com o usuário e técnicas de SEO. Os resultados mostraram que um *site* bem estruturado, com navegação clara e conteúdo útil, combinado com

palavras-chave e metadados adequados, contribui para uma classificação mais alta nas páginas de resultados dos mecanismos de pesquisa (SERPs) e maior visibilidade para clientes em potencial. O estudo fornece informações sobre as melhores práticas de SEO para *sites* de comércio eletrônico e a correlação entre SEO, vendas e sucesso nos negócios online. As principais técnicas de SEO incluem a escolha cuidadosa de palavras-chave relevantes, a criação de títulos e descrições atraentes para as páginas de produtos, a otimização de imagens e a criação de conteúdo relevante e de alta qualidade. A análise também revelou a importância do uso de análises de dados para monitorar o desempenho do SEO e ajustar as estratégias de acordo com as mudanças nos algoritmos de busca. Em resumo, o estudo destaca a necessidade de implementar uma estratégia eficaz de SEO para *sites* de comércio eletrônico a fim de aumentar a visibilidade, as vendas e a competitividade no mercado online.

An e Jung (2021), discutem uma abordagem heurística para recomendar metadados para melhorar a otimização de mecanismo de busca. Os autores propõem um modelo de recomendação baseado em árvore de decisão para selecionar metadados relevantes e importantes para cada página da *web*, levando em consideração fatores como palavras-chave, título da página, descrição, entre outros. Os resultados mostraram que a abordagem proposta é eficaz na recomendação de metadados para melhorar a visibilidade das páginas da *web* nos resultados de busca.

Mesa; Gomez e Valladares (2021), abordam as estratégias de SEO e como elas podem ajudar no posicionamento dos *sites* nos resultados de busca. O estudo analisa a importância da otimização de palavras-chave, títulos e descrições, além de apresentar técnicas para melhorar o *ranking* do *site* nos mecanismos de busca. O artigo conclui que, embora não haja garantia de posicionamento, a aplicação correta de técnicas de SEO pode aumentar significativamente a visibilidade do *site* nos resultados de busca.

Katumba e Coetzee (2017), exploram como as técnicas de SEO podem ser usadas para melhorar a descoberta de recursos geoespaciais na *web*. Os autores discutem a importância do uso de metadados geoespaciais e técnicas de SEO para melhorar a visibilidade e a classificação de recursos geoespaciais em mecanismos de busca, como o Google. Eles também destacam a necessidade de melhorar a qualidade dos metadados geoespaciais para garantir que as informações relevantes sejam incluídas nos resultados da pesquisa. Além disso, os autores fornecem exemplos práticos de como as técnicas de SEO podem ser aplicadas a recursos geoespaciais específicos, como mapas e imagens de satélite, para aumentar a visibilidade na *web*.

Hnat e Hnatova (2021), estudo examina os fatores que afetam a visibilidade de *sites* de museus na Eslováquia nos resultados de busca orgânica dos mecanismos de busca, com foco em palavras-chave, conteúdo e

estrutura do *site*. Foram realizadas análises quantitativas e qualitativas de 20 *sites* de museus, incluindo a análise de palavras-chave, análise de conteúdo, análise de estrutura do *site* e análise de *backlinks*. Os resultados mostraram que o conteúdo relevante e atualizado é o fator mais importante para a visibilidade do *site*, seguido pela otimização de palavras-chave e estrutura do *site*. A análise de *backlinks* mostrou que a maioria dos *sites* tinha poucos *backlinks*, o que afetou negativamente a visibilidade nos resultados de busca.

Nos artigos selecionados, percebe-se que a organização e a representação adequadas de informações foram essenciais para a otimização de SEO. Isso envolve a utilização de palavras-chave relevantes, descrições precisas, URLs amigáveis e informações estruturadas para ajudar os mecanismos de busca a encontrar e indexar o conteúdo relevante. Além disso, a criação de conteúdo estruturado e bem organizado permite que os mecanismos de busca identifiquem a relevância e a autoridade do conteúdo, o que pode aumentar sua classificação nos resultados de pesquisa.

Portanto, a organização e a representação adequadas de informações são fundamentais para a otimização de SEO, pois permitem que os mecanismos de busca encontrem e indexem conteúdo relevante e apresentem aos usuários resultados mais precisos e relevantes. Isso pode levar a um aumento significativo no tráfego e na visibilidade do *site*, resultando em maior alcance e sucesso.

6 Conclusão

O artigo discutiu a importância da organização e representação do conhecimento para a otimização de mecanismos de busca (SEO). Foram apresentados casos que ilustram como a implementação de técnicas de organização e representação adequadas de informações podem melhorar a classificação nos resultados de pesquisa e, conseqüentemente, aumentar a visibilidade de *sites* de empresas.

Destaca-se a relevância da relação entre organização e representação do conhecimento e SEO, pois, por meio dessas técnicas, é possível tornar o conteúdo mais acessível e legível para os mecanismos de busca, facilitando a indexação e melhorando a relevância e precisão dos resultados apresentados aos usuários.

Como possíveis aplicações práticas dos resultados obtidos, sugere-se a adoção de técnicas de organização e representação de informações em *sites* e empresas, como a utilização de palavras-chave relevantes, títulos e descrições precisas, URLs amigáveis, além da adoção de uma abordagem estruturada para a criação de conteúdo. Ademais, é importante destacar a relevância da gestão da informação nas organizações para aprimorar a eficiência e eficácia na tomada de decisões e no compartilhamento de conhecimento.

Em resumo, a adoção de técnicas adequadas de organização e representação de informações é fundamental para a otimização do mecanismo de busca e, conseqüentemente, para a visibilidade e sucesso de *sites* e empresas na internet.

Referências

- AN, S.; JUNG, J. J. A heuristic approach on metadata recommendation for search engine optimization. **Concurrency And Computation-Practice & Experience**, [s. l.], v. 33, n. 3, 2021.
- BEAL, V. Search engine optimization (SEO). In: LI, X.; HAO, F. **Encyclopedia of database systems**. New York: Springer, p. 88-95, 2020. Disponível em: https://doi.org/10.1007/978-1-4614-8265-9_902-1. Acesso em: 12 abr. 2023.
- BERMAN, R.; KATONA, Z. The Role of Search Engine Optimization in Search Marketing. **Marketing Science**, [s. l.], v. 32, n. 4, p. 644–651, 2013.
- CHAVES, G. C.; SALES, L. A. Organização e representação da informação digital. **Encontros Bibli**, Florianópolis, v. 23, n. 51, p. 97-111, maio/ago. 2018.
- CHEN, H. *et al.* Research on optimization of SEO technology for e-commerce website. **Journal of Physics: Conference Series**, v. 1210, n. 1, p. 012081, 2019.
- CHO, Y. J.; CHEON, H. J. Ethical Issues of Search Engine Optimization (SEO) in Online Marketing. **Journal of Distribution Science**, v. 17, n.2, p. 51-60, 2019.
- COOPER, H. **Synthesizing Research: A Guide for Literature Reviews**. 5th ed. Sage Publications, 2017.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, v. 37, n. 2, p. 89-96, 2010.
- DRIVAS, I. C. *et al.* Big Data Analytics for Search Engine Optimization. **Big data and cognitive computing**, [s. l.], v. 4, n. 2, 2020.
- ERDMANN, A.; ARILLA, R.; PONZOA, J. M. Search engine optimization: The long-term strategy of keyword choice. **Journal of Business Research**, [s. l.], v. 144, p. 650–662, 2022.
- FAROOQ, M. *et al.* A comprehensive review of search engine optimization: Past, present, and future trends. **IEEE**, v. 8, p. 99743-99764, 2020.
- FENG, C. *et al.* Research on image optimization strategy for improving web page loading speed. **International Journal of Online Engineering**, v. 16, n. 2, p. 11-19, 2020. Disponível em: <https://doi.org/10.3991/ijoe.v16i2.11509>. Acesso em: 14 abr. 2023.
- GUARINO, N.; WELTY, C. A. An overview of OntoClean. In: STAAB, S.; STUDER, R. (Ed.). **Handbook on Ontologies in Information Systems**. Berlin: Springer, p. 151-172, 2002.
- HNAT, A.; HNATOVA, J. Aspects Influencing The Visiblitiy Of Museum Websites In The Organic Search In Slovakia. **International Journal On Information Technologies And Security**, [s. l.], v. 13, n. 2, p. 75–86, 2021.

- HJØRLAND, B. The foundation of the concept of relevance. **Journal of the Association for Information Science and Technology**, v. 68, n. 4, p. 999-1013, Apr. 2017.
- LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos/Livros, 2004.
- LIMA-MARQUES, M. R. Competência em gestão da informação na perspectiva de especialistas. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 22, n. 2, p. 32-50, maio/ago. 2017. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362017000200032&lng=en&nrm=iso. Acesso em: 19 abr. 2023.
- KATUMBA, S; COETZEE, S. Employing Search Engine Optimization (SEO) Techniques for Improving the Discovery of Geospatial Resources on the Web. **Isprs International Journal Of Geo-Information**, [s. l.], v. 6, n. 9, 2017.
- KILLORAN, J. B. How to Use Search Engine Optimization Techniques to Increase Website Visibility. **IEEE Transactions On Professional Communication**, [s. l.], v. 56, n. 1, p. 50-66, 2013.
- LEE, S. *et al.* Search engine optimization A case study using the bibliographies of LG Science Land in Korea. **Library Hi Tech**, [s. l.], v. 34, n. 2, p. 197-206, 2016.
- LIU, Y.; LU, W. Building backlinks for website search engine optimization: A review. **Journal of Business Research**, 88, p. 101-109, 2018.
- LOPES, L. F. D. *et al.* Aplicação de revisão sistemática da literatura em uma pesquisa na área de gestão. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 25, n. 1, p. 59-79, jan./mar. 2020.
- LUH, C. J.; YANG, S. A.; HUANG, T. L. D. Estimating Google's search engine ranking function from a search engine optimization perspective. **Online Information Review**, [s. l.], v. 40, n. 2, p. 239-255, 2016.
- MA, X. *et al.* A review of search engine optimization in tourism. **Current Issues in Tourism**, v. 22, n. 3, p. 297-313, 2019.
- MARTINS, F. A. B. Estratégias de Search Engine Optimization (SEO) para otimização de sites. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 16, n. 2, p. 52-68, 2018. Disponível em: <https://rdg.revista.ibict.br/index.php/rdg/article/view/2353>. Acesso em: 12 abr. 2023.
- MESA, Y. R.; GOMEZ, A. M. M.; VALLADARES, L. A. Search engine optimization (SEO) and search engine positioning guarantee. **Medisur-Revista De Ciencias Medicas De Cienfuegos**, [s. l.], v. 19, n. 1, p. 188-192, 2021.
- MOZ. **Utilização de meta tags na otimização de motores de busca**. Disponível em: <https://moz.com/learn/seo/meta-description>. Acesso em: 14 abr. 2023.
- NISO. **Understanding Metadata**. 2004. Disponível em: <https://www.niso.org/publications/press/understanding-metadata>. Acesso em: 14 abr. 2023.
- OLMEDILLA, D., GARCÍA-SÁNCHEZ, Á.; Martínez-Torres, M. R. Social media optimization and its influence on search engine optimization. **Journal of Business Research**, v. 98, p. 424-437, 2019.
- ONAIFO, D; RASMUSSEN, D. Increasing libraries' content findability on the web with search engine optimization. **Library hi tech**, [s. l.], v. 31, n. 1, p. 87-108, 2013.
- RIZWAN, M.; AHMAD, S.; MOHAMED, S. S. Sustainability assessment of search engine optimization practices in e-commerce. **Sustainability**, v. 12, n. 6, p. 1-20, 2020.
- RODRIGUES, L. F. B. A organização do conhecimento e a recuperação da informação na web. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 21, n. 2, p. 21-38, 2016. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2642/1862>. Acesso em: 12 abr. 2023.
- SHEVCHENKO, L. Search engine optimization: Experience of State Public Library for Science and Technology of the Russian Academy of Sciences Siberian Branch. **Nauchnye I Tekhnicheskie Biblioteki-Scientific And Technical Libraries**, [s. l.], n. 3, p. 25-33, 2018.
- SOUZA, R. R.; TARAPANOFF, K. Organização e representação do conhecimento na sociedade da informação. **Ciência da Informação**, Brasília, v. 43, n. 1, p. 67-77, jan./abr. 2014.
- SUCHANEK, P. The Fundamentals Of A Prosperous E-Shop In Connection To Search Engine Optimization. **E & M Economie A Management**, [s. l.], v. 13, n. 2, p. 92-103, 2010.
- TARAPANOFF, K. O papel da organização e representação da informação para a web semântica. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 20, n. 2, p. 38-56, maio/ago., 2015.
- VAKKARI, P. Taxonomies for information retrieval. In: WILLIAMS, M. E.; BURNETT, K. (Eds.). **Introduction to modern information retrieval**. 3. ed. London: Facet Publishing, p. 233-252, 2010.
- XIE, H.; WANG, X. The Research of SEO Technology Based on Search Engine Algorithms. *In: 2019 IEEE 3RD INTERNATIONAL CONFERENCE ON COMPUTER AND COMMUNICATIONS (ICCC)*, 2019, Índia. **Anais [...]**. Erode, Índia, 2019. p. 1351-1355.
- ZENG, M. L.; QIN, J. Exploring and integrating diverse digital resources. **Journal of the American Society for Information Science and Technology**, v. 59, n. 9, p. 1517-1534, 2008.
- ZENG, M. L.; QUIN, D. The nature of concept and its representation in a controlled vocabulary. **Journal of the American Society for Information Science and Technology**, v. 59, n. 8, p. 1213-1224, 2008.
- Zhang, Q.; WANG, Y.; LIU, Q.; CAI, Y. A study on the optimization of web page keywords based on search engine. **Journal of Physics: Conference Series**, v. 1212, n. 2, p. 022070, 2019.

Indexação Social como estratégia para visibilidade de conteúdos científicos no serviço de rede social *online* Tiktok

Social indexing as a strategy for scientific content visibility in online social network service Tiktok

Larissa Lima da Silva (1), Amanda Garcia Gomes (2), Débora Matni Fonteles (3), Fernando de Assis Rodrigues (4)

(1) Universidade Federal do Pará, R. Augusto Corrêa, 01- Guamá, Belém – PA, larissasilva@ufpa.br. (2) Universidade Estadual Paulista, V. Hygino Muzzi Filho, 737 - Marília – SP, garcia.gomes@unesp.br. (3) Universidade Federal do Pará, R. Augusto Corrêa, 01- Guamá, Belém – PA, dmatnif@gmail.com, (4) Universidade Federal do Pará, R. Augusto Corrêa, 01- Guamá, Belém – PA, deassis@ufpa.br

Resumo

Objetivo: Analisar a utilização de *hashtags* no contexto da divulgação da ciência brasileira em Serviços de Redes Sociais *Online*, a fim de identificar as características e as estratégias de visibilidade baseada em indexação social. **Metodologia:** Estudo descritivo e exploratório, de caráter qualitativo e quantitativo, sobre a utilização de *hashtags* por divulgadores científicos. Como universo de pesquisa, foi delimitado a divulgação científica de perfis brasileiros no TikTok. Foram coletadas 1.586 *hashtags*, com 8.106 ocorrências, a partir de 17 perfis de divulgadores científicos brasileiros selecionados de uma amostra de 172 perfis do TikTok. Os critérios de seleção dos perfis foram: a) presença da indexação social; b) conteúdo produzido por especialistas ou instituição científica e cultural, e; c) perfis que se enquadrem no contexto da divulgação científica. Após classificação quantitativa, as 12 *hashtags* mais populares foram descritas quanto às suas características, termos de indexação e popularidade. **Resultados:** Observou-se que existem ocorrências das *hashtags* nas nove áreas do conhecimento categorizadas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Existem seis tipos de *hashtags* envolvidas nas estratégias de visibilidade do conteúdo: a) internas ao contexto acadêmico; b) genéricas e populares; c) vinculadas a uma área do conhecimento; d) elaboradas para promover o conteúdo educacional; e) que identificam ou personalizam a busca por informações veiculadas a um perfil, e; f) de cunho político. **Conclusão:** As *hashtags* são parte de um sistema de classificação de conteúdo do TikTok com potencial de representar o conteúdo informacional dos vídeos, possibilitando reunir publicações que possuem assuntos relacionados

Palavras-chave: Indexação Social. Visibilidade de Conteúdo. Divulgação Científica. Serviço de Rede Social *Online*

Abstract

Objective: To analyze the use of hashtags in Brazilian science dissemination in *Online* Social Network Services, to identify the characteristics and strategies of visibility based on social indexing. **Methodology:** Descriptive and exploratory study, with qualitative and quantitative nature, on hashtags used by science popularizers. The research universe is the scientific dissemination of Brazilian profiles on TikTok. It collected 1,586 hashtags, with 8,106 occurrences, from 17 profiles of Brazilian science popularizers selected from a sample of 172 profiles on TikTok. The profile selection criteria were the presence of social indexing, the content produced by specialists or scientific and cultural institutions, and the profiles that fit into the context of scientific dissemination. After quantitative classification, it described the 12 most popular hashtags according to their characteristics, indexing terms, and popularity. **Results:** It observed hashtags occurrences in the nine areas of knowledge categorized by the Coordination for the Improvement of Higher Education Personnel. Content visibility strategies apply six types of hashtags: internal to the academic context, generic and popular, linked to an area of knowledge, designed to promote educational content, that identifies or customize the search for information conveyed to a profile, and of a political nature. **Conclusion:** Hashtags are part of TikTok's content classification system and have the potential to represent the informational content of videos, making it possible to gather publications that have related subjects.

Keywords: Social Indexing. Content Visibility. Scientific diffusion. *Online* Social Network Service

Resumen

Objetivo: Analizar el uso de hashtags en el contexto de la divulgación de la ciencia brasileña en los Servicios de Redes Sociales en Línea, con el fin de identificar las características y estrategias de visibilidad a partir de la indexación social. **Metodología:** Estudio descriptivo y exploratorio, de carácter cualitativo y cuantitativo, sobre el uso de hashtags por parte de divulgadores científicos. Como universo de investigación, se delimitó la divulgación científica de los perfiles brasileños en TikTok. Se recolectaron 1.586 hashtags, con 8.106 ocurrencias, de 17 perfiles de divulgadores científicos de la ciencia seleccionados de una muestra de 172 perfiles en TikTok. Los criterios de selección del perfil fueron: a) presencia de indexación social; b) contenido producido por especialistas o institución científica y cultural, y; c) perfiles que encajen en el contexto de la divulgación científica. Luego de una clasificación cuan-

titativa, se descreveram os 12 hashtags mais populares segundo suas características, termos de indexação e popularidade. **Resultados:** Se observou que há ocorrências de hashtags nas nove áreas de conhecimento categorizadas por a Coordenação de Perfeccionamento do Pessoal de Educação Superior. Há seis tipos de hashtags envolvidos nas estratégias de visibilidade de conteúdo: a) internos ao contexto acadêmico; b) genérico e popular; c) vinculado a um área de conhecimento; d) destinados a promover conteúdos educativos; e) que identifiquem ou personalizem a busca de informação transmitida a um perfil, y; f) de caráter político. **Conclusión:** os hashtags são parte do sistema de classificação de conteúdo de TikTok e têm o potencial de representar o conteúdo informativo dos vídeos, o que permite recopilar publicações que têm temas relacionados.

Palabras clave: Indexação Social. Visibilidade do Conteúdo. Difusão científica. Serviço de Redes Sociais em Linha

1 Introdução

As Tecnologias de Informação e Comunicação (TIC), ao se desenvolverem e se popularizarem, ampliaram o acesso às ferramentas de comunicação, especialmente por meio da infraestrutura da Internet e serviços *Web*. A apropriação destes serviços alterou a forma e a velocidade da produção e do compartilhamento de informações.

A possibilidade de interação mediada pelas TIC é um dos caminhos de investigação no qual a Ciência da Informação (CI) tem direcionado seu olhar. Esta interação, para a produção e para o compartilhamento de informação, proporcionou o papel de protagonista dos usuários de TIC nas formas de representar e organizar conteúdo na *Web*, utilizando marcadores de assunto para a recuperação, também denominados como etiquetas, *tags* e *hashtags*.

Hashtags são formadas por palavras ou frases escritas sem o uso de espaço característico da escrita formal e obrigatoriamente precedidas pelo símbolo # (e.g. #ciência e #aprendanotiktok). Desempenham a função de agregadores de informação e são comumente observadas no contexto dos Serviços de Redes Sociais *Online* (SRSO) [1] para representar e categorizar informação na *Web*.

Esses marcadores são similares aos descritores e as palavras-chave do processo de indexação tradicional, praticado por bibliotecas e centros de documentação. Todavia, o processo de indexação por *hashtags* é livre e trata-se de uma prática da qual não há uma visão completa de seus benefícios e contribuições (VIANA; DAL'EVEDOVE; TARTAROTTI, 2019).

Este processo recebe diferentes denominações na literatura científica, tais como: indexação social, folksonomia, classificação social, classificação distribuída, etnoclassificação, etiquetagem, *social bookmarking*, ontologias sociais e taxonomia dinâmica (MATHEUS, 2004; GUEDES; DIAS, 2010; CATARINO; BAPTISTA, 2009; HOLSTROM, 2018; GOMES *et al.*, 2021). Mesmo que não sejam estabelecidas as diferenças entre esses termos, é um denominador comum a natureza das suas representações, cuja base é a linguagem natural empregada em conteúdos pelo usuário da *Web*.

Nesta pesquisa, o termo indexação social é adotado para representar o processo de indexação de conteúdo de divulgação científica realizado no SRSO TikTok.

O TikTok é um SRSO de compartilhamento de vídeos curtos, lançado em 2017 pela empresa chinesa de tecnologia *ByteDance*, como uma versão internacional do *Douyin*, que ganhou popularidade global ao adquirir o *Musical.ly* em 2018 (CECI, 2022).

O TikTok é um SRSO que facilita a busca e a produção de vídeos. Tem auxiliado estudantes e educadores nas atividades de estudo e de ensino, contribuindo para democratização do acesso a conteúdo educativo na *Web*, apesar de ser conhecido por conteúdos voltados ao entretenimento (TIKTOK, 2021). Além disso, o serviço permite que usuários atribuam marcadores de assunto aos vídeos, por meio de *hashtags*, para que seus conteúdos possam ser recuperados.

É possível que as características do TikTok, no contexto educativo e da divulgação científica, sejam úteis para ações de difusão do conhecimento na *Web*. A divulgação científica corresponde ao processo de compartilhamento de conteúdo científico com o público não especializado, ou seja, prioriza a comunicação com o público fora da esfera acadêmica (CANDOTTI, 2002; MULLER, 2002; PINHEIRO, 2003; GONÇALVES, 2012).

Diante desse contexto, a indexação social, por ser um processo de marcação livre e elaborado pelo produtor do conteúdo, pode ofuscar a identificação na interface do serviço sobre quais são as estratégias adotadas para dar visibilidade aos conteúdos ligados à ciência – problema desta pesquisa.

O objetivo da pesquisa é analisar a utilização de *hashtags* no contexto da divulgação da ciência em SRSO, a fim de identificar as características e estratégias de visibilidade baseada em indexação social.

Como universo de pesquisa, foi delimitado a divulgação científica no SRSO TikTok – um dos SRSO mais utilizados no Brasil – e perfis de divulgadores científicos brasileiros, detalhados na seção de Procedimentos metodológicos.

2 Indexação social no contexto do TikTok

A indexação pode ser compreendida como um processo de escolha dos termos representativos mais apropriados para um determinado documento (LIMA, 2020). Integra um processo amplo, que envolve a Organização e Representação do Conhecimento e da Informação que no contexto da cibercultura ganham novos contornos e aplicabilidades extrapolando ambientes tradicionais de informação (e.g. bibliotecas e arquivos), sejam eles físicos ou digitais.

A indexação tem duas fases: a análise conceitual e a tradução de conceitos. A análise conceitual é o processo de identificação em que se define qual é e sobre o que trata um recurso. Já a tradução de conceitos consiste na conversão dos termos extraídos em linguagem natural do documento para a linguagem documentária presente nos índices de assuntos e segue os moldes tradicionais aplicados em instituições (LANCASTER, 2004).

No processo de Representação da Informação, a indexação social serve aos propósitos da recuperação da informação. É uma forma de indexar conteúdos por meio da atribuição de *hashtags*. Também se enquadra na Representação do Conhecimento, ao ser examinada sob o ponto de vista do uso e construção de terminologias provenientes da análise de assuntos e do contexto em que foi utilizada.

Tanto os termos estruturados de um vocabulário controlado, quanto os elencados em uma lista de tópicos populares em um SRSO podem ser considerados instrumentos para tradução de conceitos. Um exemplo da aplicação de um processo de tradução em SRSO é a procura para representar um conteúdo por *hashtags* existentes, por *hashtags* populares ou por *hashtags* em alta no momento do envio do conteúdo. Todos são critérios para a escolha de uma representação, além do usuário também ser livre para criar sua própria representação.

Um dos fatores que possibilitam a recuperação da informação é a indexação, que necessita de uma estrutura conceitual prévia que oriente a escolha dos termos ou palavras-chave. Estas estruturas podem ser hierárquicas e controladas, tal como ocorre nos tesouros ou podem ser estruturas conceituais mentais que se expressam na atribuição de *hashtags* em SRSO.

A atribuição de *hashtags* por meio de linguagem natural em ambientes colaborativos (e.g. SRSO e *wikis*) visa a recuperação da informação e dispensa a utilização de vocabulários controlados, sendo, portanto, uma forma de representação da indexação social (SANTOS; CORRÊA, 2017). Possibilita a criação de etiquetas que representam os conteúdos produzidos e disseminados na *Web*, de modo que esses possam ser recuperados por seus produtores ou por outros usuários do sistema com interesses em comum, sendo uma característica da inde-

xação social (SANTARÉM SEGUNDO; VIDOTTI, 2011).

A sua principal característica diz respeito à elaboração a partir de termos que estão fora de uma hierarquia (relacionamento pai e filho ou irmão diretamente especificado) entre esses termos. O que existem são *hashtags* que se relacionam, geradas de forma automática e que agrupam *hashtags* com base em usos comuns; diferente do que ocorre com taxonomias formais, em que existem vários relacionamentos explícitos entre os termos gerais, os amplos e os específicos (MATHEUS, 2004). As estratégias de visibilidade elaboradas por meio de *hashtags* seriam o resultado de uma indexação orientada pela demanda.

A visibilidade é algo desejável no universo do SRSO, sobretudo em relação à divulgação científica. A visibilidade em SRSO corresponde a ser conhecido e o conteúdo ser considerado relevante. No contexto da comunicação científica, a visibilidade pode ser mensurada pelo impacto que causa na comunidade científica como, por exemplo, por meio de citações.

Em termos gerais, um artigo citado indica que foi lido e considerado pertinente para embasar novas pesquisas. Na divulgação científica que ocorre nos SRSO, esta visibilidade pode ser mensurada por meio do número de visualizações que o vídeo possui e pelo engajamento que possui com seu público.

O engajamento em SRSO está relacionado às interações apresentadas em forma de curtidas, comentários, compartilhamentos (PARISER, 2012; PAULINO; VENTURA, 2021). No contexto do TikTok, essas interações auxiliam no melhoramento do algoritmo (popularmente conhecido como treinamento) para recomendar conteúdo.

A indexação social e a visibilidade de conteúdo possuem uma relação intrínseca no TikTok. As recomendações para obter visibilidade incluem o aproveitamento de legendas, sons e *hashtags*. As *hashtags* no TikTok estão relacionadas com a ideia de se conectar com o público certo.

A divulgação científica pode ser compreendida como um conjunto de recursos, técnicas, processos e produtos destinados à veiculação de informações científicas ao público não especializado. Dentre os recursos tradicionais estão: textos, vídeos, feiras, palestras e exposições em museus. Na Sociedade em Rede (CASTELLS, 2021) surgem novas possibilidades de disseminação de informação, que possibilitam alcançar o público não especializado (CANDOTTI, 2002; MUELLER, 2002; TORRESI; PARDINI; FERREIRA, 2012; BUENO, 2010; GONÇALVES, 2012).

Sendo assim, é possível que as características do TikTok, no contexto educativo e da divulgação científica sejam úteis para ações de difusão do conhecimento

na *Web*. Neste processo, a indexação social pode ser considerada uma estratégia de visibilidade em SRSO, pautada na encontrabilidade de conteúdo por meio de representação da informação.

3 Procedimentos metodológicos

A pesquisa se caracteriza como um estudo descritivo e exploratório, de caráter qualitativo e quantitativo sobre a utilização de *hashtags* por divulgadores científicos brasileiros.

Os dados utilizados nesta pesquisa fazem parte do resultado de uma coleta mais extensa, que envolveu buscas por perfis de divulgação científica no TikTok, por meio de estratégias de pesquisa simples e combinadas, realizadas entre os meses de outubro e novembro de 2022, aplicadas ao motor de busca *Google Search* e ao TikTok.

Os perfis recuperados neste processo foram retirados de listas compiladas de *blogs*, sites jornalísticos e artigos publicados por instituições científicas que citam divulgadores científicos que utilizam o TikTok, bem como pela aplicação dos termos na barra de pesquisa do próprio SRSO.

Os resultados da etapa de descoberta de perfis foram reunidos e organizados em planilha eletrônica para um processo de análise e classificação, quanto a sua pertinência a pesquisa (Tabela 1), e conforme as áreas do conhecimento da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES): Ciências Agrárias, Ciências Biológicas, Ciências da Saúde, Ciências Exatas e da Terra, Ciências Humanas, Ciências Sociais Aplicadas, Engenharias, Linguística, Letras e Artes, e Multidisciplinar.

Dos 172 perfis identificados, 70 perfis foram descartados por não ser possível identificar e verificar as informações sobre a formação acadêmica dos produtores de conteúdo, sendo este um dos critérios de elegibilidade da amostra.

Com o intuito de diminuir a quantidade de perfis a serem analisados, 30 perfis foram descartados, por serem menos interessantes a pesquisa, tomando por base a utilização de indexação social em seus vídeos; 22 perfis foram descartados por não fazerem divulgação científica; 9 perfis foram descartados por apresentarem pouca ou nenhuma indexação social em seus vídeos; 8 perfis não possuíam como objetivo principal a produção de conteúdo destinado à divulgação científica. Neste caso, trata-se de produtores de conteúdo que utilizam seu perfil do TikTok para assuntos variados, entre eles conteúdos ligados à ciência; 8 perfis foram descartados por serem estrangeiros, tendo em vista a delimitação da pesquisa que está restrita a perfis brasileiros; 5 perfis foram descartados por estarem desatualizados; 3 perfis foram descartados por não pertencerem a especialistas

ou instituições científicas da área do conhecimento abordado nas postagens, totalizando 155 perfis descartados dos 172 perfis recuperados.

Tabela 1. *Categorias utilizadas para a seleção dos perfis*

<i>Categorias</i>	<i>Justificativas</i>	<i>Quantidade de perfis</i>
	Não foi possível identificar e verificar as informações acadêmicas	70
	Optou-se por outro perfil na mesma grande área do conhecimento	30
	O perfil não faz divulgação científica	22
	Pouca Indexação Social	9
Descarte	O perfil não tem como objetivo principal a produção de conteúdo destinado à divulgação científica.	8
	Perfil estrangeiro	8
	Perfil desatualizado	5
	O perfil não pertence a um(a) especialista ou instituição científica da área do conhecimento abordado nos vídeos	3
<i>Total de descartados</i>		<i>155</i>
Seleção	O perfil atende aos critérios de elegibilidade	17
<i>Total Geral</i>		<i>172</i>

Fonte: Autores (2023).

Assim, foram identificados 172 perfis do TikTok no total, dos quais 17 foram selecionados para análise por atenderem os critérios de seleção estabelecidos: @biafumelli, @cesarfavacho, @ciencia.livre, @eliasjabbourtv, @gabriel.minera, @itamarragro, @joaojustopires, @jufurno, @mulheresnaciencia, @museucasadeportinari, @museudeartesacramt, @odirfontoura, @oficialfiocruz, @olivia.sensata, @opedroloos, @professor.branco e @uaifisica.

Os critérios de seleção dos 17 perfis foram: a) presença da indexação social; b) conteúdo produzido por especialistas ou instituição científica e cultural; c) perfis que se enquadrem no contexto da divulgação científica. Os demais critérios, ligados à popularidade e a abordagem de temas diversos foram utilizados para o desempate, quando necessário. Ao final, foram escolhidos dois perfis para cada grande área do conhecimento da CAPES, sendo que Engenharia foi a única área do conhecimento com apenas um perfil que atendeu aos critérios de seleção.

Para análise, as *hashtags* foram extraídas de vídeos publicados por esses perfis no ano de 2022. Observou-se o comportamento das *hashtags* por meio dos *insights* de público [2] do SRSO. Foram coletadas 1.538 *hashtags* diferentes dos 17 perfis selecionados, que somam 8.106 ocorrências no total. Para esta análise, foi considerado um corte de utilização de *hashtags* quando a ocorrência de uso era igual ou superior a 100 ($n \geq 100$), analisadas sob duas perspectivas: a) as propriedades observadas quanto a indexação social, e b) seus aspectos linguísticos.

Três propriedades da indexação social que as *hashtags* desempenham nas publicações foram eleitas para análise: i) *hashtags* relacionadas ao contexto acadêmico; ii) *hashtags* que representam assuntos gerais e, iii) *hashtags* que representam assuntos específicos.

Quanto aos aspectos linguísticos, foram considerados dois elementos: i) *hashtags* que representam as especificidades do perfil ou áreas do conhecimento; ii) *hashtags* que possuem erros ortográficos e/ou abreviações.

4 Resultados e Discussão

A primeira dimensão de análise está relacionada a frequência de uso das *hashtags*. A Tabela 2 (em apêndice) apresenta as 12 *hashtags* analisadas e foram ranqueadas de acordo com sua frequência das nove áreas do conhecimento. As *hashtags* foram classificadas pela ocorrência, em ordem decrescente.

A segunda dimensão apresenta as características das *hashtags*, a partir da indexação social, seguindo a ordem decrescente de ocorrências (Quadro 1, em apêndice). As características são: propriedades da indexação social e aspectos linguísticos. O que foi intitulado de ‘Demais *hashtags*’ representa as *hashtags* com ocorrências menores que o critério adotado para esta pesquisa. Assim, foram desconsideradas 5.845 *hashtags* para esta análise.

A *#ciencia* foi a mais utilizada por divulgadores científicos, mencionada 492 vezes pelas áreas do conhecimento: Ciências Exatas e da Terra (252), Multidisciplinar (107), Ciências da Saúde (68), Ciências Biológicas (47) e Ciências Humanas (18). Não utilizam a *hashtag* as áreas do conhecimento das Ciências Agrárias, Ciências Sociais Aplicadas, Engenharias e Linguística, Letras e Artes.

É possível observar que a *#ciencia* foi uma das *hashtags* mais populares na categoria educação, especialmente entre pessoas de 18 a 24 anos. Entre os meses de fevereiro de 2022 e fevereiro de 2023, a *hashtag* obteve mais de 2 bilhões de visualizações sendo identificada ao menos em 29 mil publicações (TIKTOK, 2023).

Sob a perspectiva da indexação social, pode ser compreendida como um termo geral, sendo possível encontrar outros termos que poderiam representar especificações do conceito de ciência. Na amostra, houve ocorrência em diferentes áreas do conhecimento, sem a especificação de uma temática.

É uma *hashtag* interna ao contexto acadêmico, que contribui para identificar o conteúdo informacional do vídeo. Assim, seu uso aumenta as chances de alcance e visualização para o público-alvo, nesse caso, o emprego da *#ciencia* pode ser considerada uma estratégia de visibilidade baseada em indexação social, ainda que não seja possível afirmar a intenção do produtor do vídeo.

A *#ciencia* apresenta outra propriedade da indexação social no contexto digital: modificações linguísticas de escrita, (denominado nesta pesquisa como linguagem da *Web*). Por exemplo, a flexibilidade ou ser opcional a escrita formal e as regras ortográficas (e.g. a utilização de acentos) como foi observado para *#ciencia*.

A segunda *hashtag* com maior ocorrência é *#aprendanotiktok*. Aparece na amostra de pesquisa 334 vezes utilizada em vídeos de áreas do conhecimento: Ciências Exatas e da Terra (176), Ciências Sociais Aplicadas (65), Ciências Biológicas (47), Ciências Humanas (45) e Ciências da Saúde (1). Não utilizam a *#aprendanotiktok* as áreas do conhecimento: Ciências Agrárias, Multidisciplinar, Engenharias e Linguística, Letras e Artes.

A *#aprendanotiktok* surgiu visando promover a criação de conteúdos com caráter educativo de diferentes tipos no TikTok. A *hashtag* é parte do programa de aceleração elaborado pelo TikTok para apoiar a produção de conteúdo educativo no Brasil, iniciada em setembro de 2021 (TIKTOK, 2021). Os produtores de conteúdo, a convite e por meio de seleção, participaram de *workshop* voltado ao aprimoramento e capacitação para produzir conteúdo.

Embora não seja utilizada por todas as áreas do conhecimento, é possível inferir que a *#aprendanotiktok* foi absorvida pelos divulgadores científicos e encarada como parte da estratégia de visibilidade, a partir da posição que a *hashtag* ocupa na amostra de pesquisa. Prova disso é a permanência de sua utilização em 2023, pois entre os meses de fevereiro de 2022 e fevereiro de 2023 a *#aprendanotiktok* teve mais de 8 bilhões de visualizações e foi identificada em 188 mil postagens no Brasil (TIKTOK, 2023). Foi possível compreender que sua utilização relaciona os vídeos a uma produção de cunho educacional ou científica [3].

No que se refere a *insights* de público, foi possível observar que o maior alcance da *#aprendanotiktok* foram pessoas entre 18 e 24 anos (60% do público), seguida de pessoas entre 25 e 34 anos (21% do público) e pessoas com mais de 35 anos (19% do público) (TIKTOK,

2023). Vale destacar que o TikTok não incluiu nestas métricas usuárias menores de 18 anos. Portanto, é possível que menores de 18 anos sejam alcançados pelos conteúdos indexados por esta *hashtag*.

Quanto aos aspectos da indexação social, a *#aprenda-notiktok* pode ser considerada um termo composto característico da linguagem da *Web* (a falta de espaços entre palavras). A existência de espaços entre as palavras que compõem uma *hashtag* refletirá na perda de sua funcionalidade como um recurso de pesquisa e categorização de assuntos.

Pode ser considerada uma *hashtag* interna ao contexto acadêmico, embora sua aplicação não esteja presente apenas em publicações relacionadas à ciência. Não está restrita a um nicho científico, sendo utilizada tanto em vídeos que realizam divulgação científica quanto em vídeos destinados ao ensino de temas diversos (e.g. maquiagem) [4]. A *hashtag* não tem capacidade de especificar sozinho um assunto do vídeo, mas serve como uma marcação que relaciona o conteúdo à ideia de aprendizagem. De certa forma, especifica a intenção do produtor de conteúdo em ensinar algo.

A *hashtag* *#fyp* é a terceira mais frequente entre os divulgadores científicos que compõem a amostra, mencionada 276 vezes. Teve ocorrência nas áreas do conhecimento como: Ciências Exatas e da Terra (153), Ciências Humanas (79), Multidisciplinar (23), Linguística, Letras e Artes (16), Ciências da Saúde (4), Ciências Agrárias (1). Não utilizam a *#fyp* áreas do conhecimento: Ciências Biológicas, Ciências Sociais Aplicadas e Engenharias.

A *#fyp* corresponde à abreviação da frase em inglês *for your page* (para sua página em português). Refere-se a um dos elementos que compõem a interface de usuário do TikTok. Na interface de usuário *For You* (Para Você em português), o algoritmo do TikTok sugere conteúdo a partir de uma seleção (classificação) baseada nas preferências do usuário. A partir do histórico de interação com o aplicativo, o algoritmo realiza recomendações de perfis que o usuário não segue ou desconhece.

Os *insights* de público mostraram que essa *hashtag* alcançou mais pessoas entre 18 e 24 anos (44% do público), seguida de pessoas entre 25 e 34 anos (29% do público) e, por último, pessoas com mais de 35 anos (27% do público). Sendo visualizada mais de 618 bilhões de vezes sendo identificada em 35 milhões de publicações no Brasil (TIKTOK, 2023).

A *#fyp* tem uma característica especial no TikTok, pois é tida como uma *hashtag* geral, comumente utilizada pelos divulgadores científicos e perfis que produzem outros tipos de conteúdo para o TikTok. Essa característica é um dos fatores ligados à grande quantidade de visualizações. A utilização da *#fyp* em conteúdos ligados à educação, conforme os resultados do TikTok.Ads

[5], foi acompanhada por outras *hashtags* que complementam a *#fyp*.

A *#fy* foi a quarta *hashtag* em ocorrência entre os divulgadores científicos e corresponde à abreviação da frase em inglês *for you* (para você em português). É uma variação da *#fyp* e foi mencionada 267 vezes no total. Teve ocorrência nas áreas do conhecimento: Ciências Exatas e da Terra (127); Ciências Humanas (74); Multidisciplinar (30); Linguística, Letras e Artes (18); Ciências Agrárias (15). Não utilizam a *#fy* áreas do conhecimento como: Ciências Biológicas, Ciências da Saúde, Ciências Sociais Aplicadas e Engenharias.

Quanto aos aspectos da indexação social, as *#fyp* e *#fy* são termos compostos, embora apareçam como abreviações. Possuem características comuns: são externas ao contexto acadêmico, incapazes de especificar ou situar os conteúdos relacionados à ciência e educação, e são linguagens da *Web*, frequentemente usadas no TikTok. Podem ser consideradas uma estratégia de visibilidade baseada em indexação social, na perspectiva do produtor de conteúdo.

A partir do entendimento que estas duas *hashtags* são variações da mesma estratégia de visibilidade e utilizadas por serem populares no TikTok, foi realizada uma comparação entre as suas ocorrências na amostra desta pesquisa:

- i. Ciências Agrárias utilizaram com maior frequência *#fy* do que *#fyp*, sendo esta última mais popular no contexto geral da amostra;
- ii. Ciências da Saúde utilizaram apenas *#fyp*;
- iii. Ciências Exatas e da Terra e Ciências Humanas tiveram, respectivamente, 26 e 5 ocorrências a mais na utilização da *#fyp* em relação a *#fy*, valores pouco representativos;
- iv. Linguística, Letras e Artes e Multidisciplinar utilizaram mais *#fy* em relação a *#fyp*, também com valores pouco representativos;
- v. Ciências Biológicas, Ciências Sociais Aplicadas e Engenharias não utilizaram a *#fy* e *#fyp*.

Observou-se que o aparecimento dessas duas *hashtags* entre as cinco mais utilizadas na amostra apresenta indícios da sua utilização como estratégia de visibilidade. Isso sugere que os divulgadores científicos utilizaram essas *hashtags* como uma estratégia que potencialmente pode impulsionar seus conteúdos para a *For You* dos usuários do TikTok, como sugestão do algoritmo.

No entanto, a atribuição destas *hashtags* não influencia no impulsionamento de publicações (TIKTOK, 2020). A sugestão feita pelo algoritmo do TikTok para impulsionar as publicações têm centralidade na relevância e na adequação das *hashtags* ao conteúdo informacional do vídeo, na participação de *hashtags* em alta (destaca-

das nos tópicos de tendências) ou nas *hashtags* mais populares em determinados períodos. Em relação à visibilidade, a utilização de *hashtags* mais alinhadas a representação do conteúdo dos vídeos ou populares no momento tem maiores chances de serem sugeridas para o usuário e relevante para a visibilidade do conteúdo.

A *#curiosidades* foi a quinta *hashtag* mais popular, mencionada 126 vezes. Teve ocorrências nas áreas do conhecimento: Multidisciplinar (61); Ciências da Saúde (41); Ciências Biológicas (21); Linguística, Letras e Artes (2) e Ciências Sociais Aplicadas (1). Não utilizaram *#curiosidades*: Ciências da Terra, Ciências Humanas e Engenharias.

Os *insights* de público mostraram que o maior alcance dessa *hashtag* foram pessoas entre 18 e 24 anos (58% do público), seguida de pessoas entre 25 e 34 anos (24% do público) e pessoas com mais de 35 anos (18% do público) (TIKTOK, 2023). Quanto a popularidade, teve 17 bilhões de visualizações e esteve presente em 191 mil postagens no Brasil, entre os meses de fevereiro de 2022 e fevereiro de 2023 (TIKTOK, 2023).

Sob os aspectos da indexação social, *#curiosidades* pode ser considerada um termo geral no contexto do TikTok. É possível atribuir a *hashtag* junto a outro termo que torne a representação mais específica, como, por exemplo, *#curiosidadesdeterminadinca*, que é uma *hashtag* interna ao contexto acadêmico, pois remete a curiosidade um atributo do pesquisador ou de quem deseja aprender, ainda que o emprego da *hashtag* não represente algo exclusivo à ciência.

Além disso, está *hashtag* pode ser utilizada pelo divulgador científico para vincular seus vídeos a outros com temas externos ao contexto acadêmico, contribuindo para que o conteúdo seja disseminado para outros públicos que tenham interesse por informação científica. Por isso, a utilização pode ser considerada uma estratégia de visibilidade baseada em indexação social.

A sexta *hashtag* mais utilizada foi a *#agro*, com ocorrência registrada apenas na área do conhecimento de Ciências Agrárias (122). Apesar de ter como característica o pertencimento a uma área do conhecimento específica, possuiu ocorrências suficientes para estar classificada entre as 12 *hashtags* mais populares.

No que se refere a *insights* de público, é possível observar que o maior alcance dessa *hashtag* foram de pessoas entre 18 e 24 anos (47% do público), seguida de pessoas entre 25 e 34 anos (27% do público) e pessoas com mais de 35 anos (26% do público) (TIKTOK, 2023). Em relação à popularidade da *hashtag*, teve mais de 6 bilhões de visualizações e esteve presente em 419 mil publicações no Brasil, entre os meses de fevereiro de 2022 e fevereiro de 2023 (TIKTOK, 2023).

Diante disso, percebeu-se que a *hashtag* é um termo alinhado a área do conhecimento das Ciências Agrárias

(*e.g.* agronomia), porém foi compreendida como uma representação que não remete ao contexto acadêmico, tal qual as *#ciencia*, *#curiosidades* ou *#aprendanotiktok*; embora funcione adequadamente para sua área do conhecimento como um termo genérico, pois se considera que esta *hashtag* não consegue especificar um conteúdo como ligado à ciência.

Quanto aos aspectos linguísticos, não se trata de uma abreviação, mas da escrita incompleta de uma palavra. Ou seja, apenas prefixo da palavra agronomia compõe a *hashtag*. É possível inferir que a *#agro* seja uma estratégia de indexação social, que utiliza um termo genérico para representar uma área do conhecimento específica. A utilização da *hashtag* favorece a busca por temas relacionados às ciências agrárias, tais como: agrotóxicos, agropecuária e agronegócios.

A sétima *hashtag* mais utilizada na amostra de pesquisa foi *#economia*, mencionada 112 vezes, com ocorrência apenas em duas áreas do conhecimento: Ciências Sociais Aplicadas (98) e Ciências Humanas (14). Nas demais áreas do conhecimento não houve ocorrências dessa *hashtag*.

Os *insights* de público mostraram que o maior alcance dessa *hashtag* foram pessoas entre 18 e 24 anos (42% do público), 25 e 34 anos (26% do público) e com mais de 35 anos (32% do público). No que diz respeito a popularidade da *hashtag*, teve mais de 1 bilhão de visualizações e esteve presente em 49 mil postagens no Brasil, entre os meses de fevereiro de 2022 e fevereiro de 2023 (TIKTOK, 2023).

Na área do conhecimento das Ciências Humanas, *#economia* aparece devido à temática dos perfis que tratam de história e geografia. Observou-se que a sua frequência mais expressiva foi nas Ciências Sociais Aplicadas e se deve ao perfil *@olivia.sensata*.

Quanto aos aspectos de indexação social, a *#economia* pode ser considerada um termo geral no contexto do TikTok, porque é possível pensar em outros termos que poderiam contribuir para especificar o assunto (*e.g.* *#economiademercado*). A *#economia* pode contribuir para a viabilidade do conteúdo justamente por seu caráter generalista, pois ao expressar o assunto geral do conteúdo pode ligá-lo a outros vídeos que falam sobre o assunto e, conseqüentemente, aumentar suas chances de recuperação e sugestão para usuários que tenham interesse em temas ligados à economia. Assim, é possível compreender que *hashtags* gerais devem existir e precisam de outras mais específicas para contribuir na representação e recuperação de vídeos.

Quanto aos aspectos linguísticos, não possui erros ortográficos e nem apresenta nenhuma característica que destaca desvios da norma-padrão, observados na linguagem da *Web*.

A oitava *hashtag* mais utilizada foi a #biologia, mencionada 110 vezes. Teve ocorrências nas áreas do conhecimento: Ciências Biológicas (97); Multidisciplinar (12) e Ciências da Saúde (1). Não houve ocorrência nas áreas do conhecimento Ciências Agrárias, Ciências Exatas e da Terra, Ciências Humanas, Ciências Sociais Aplicadas, Engenharias, Linguística, Letras e Artes.

Os *insights* de público demonstraram que o maior alcance dessa *hashtag* foram pessoas entre 18 e 24 anos (77% do público), seguida de pessoas entre 25 e 34 anos (18% do público) e pessoas com mais de 35 anos (5% do público) (TIKTOK, 2023). Sobre a popularidade da #biologia, teve mais de 2 bilhões de visualizações e foi presente em 23 mil postagens no Brasil, entre os meses de fevereiro de 2022 e fevereiro de 2023 (TIKTOK, 2023).

Quanto aos aspectos de indexação social, a #biologia pode ser considerada um termo geral no contexto do TikTok, pois é possível pensar em outros termos que poderiam representar especificações para o assunto (e.g. #biologiamarinha). É uma *hashtag* interna ao contexto acadêmico, que ajuda a identificar a área de conhecimento no qual o conteúdo do vídeo está inserido, possibilitando segmentar, representar e demarcar o universo a que pertence àquela informação. A combinação com outros termos mais específicos contribui para a representação e a recuperação dos vídeos.

Quanto aos aspectos linguísticos, não possui erros ortográficos e nem apresenta nenhuma característica que destaca desvios da norma-padrão, observados na linguagem da *Web*.

A nona *hashtag* mais utilizada na amostra de pesquisa foi a #cienciatododia, teve 109 ocorrências apenas na área do conhecimento Multidisciplinar. Segundo o TikTok (2023), é uma das *hashtags* populares no Brasil e em Portugal, sendo o criador de conteúdo mais expressivo que a utiliza é @predoloos, um dos perfis selecionados nesta pesquisa para a área do conhecimento Multidisciplinar. Observa-se que esta *hashtag* é um termo de representação próprio de um perfil de divulgação científica. Nesta pesquisa, isso permite compreender a *hashtag* como uma representação interna ao contexto acadêmico, sendo visualizada mais de 17,6 milhões de vezes em publicações, entre os meses de fevereiro de 2022 e fevereiro de 2023.

Os *insights* de público mostraram que o maior alcance dessa *hashtag* foram pessoas entre 18 e 24 anos (65% do público), entre 25 e 34 anos (20% do público) e pessoas com mais de 35 anos (15% do público) (TIKTOK, 2023).

Quanto aos aspectos de indexação social observados, a #cienciatododia pode ser considerada um termo composto porque une palavras em torno de um sentido específico e apresenta características da linguagem da *Web*. É possível inferir que esta *hashtag* seja uma es-

tratégia de indexação social, pois houve a criação de uma identidade, ou seja, foi criado um termo que diferencia o seu conteúdo, ou seja, o específica.

A #cienciatododia pode ser considerada uma *hashtag* personalizada. Esta característica pertence a termos de marcação que não oferecem suporte à pesquisa de maneira eficaz (GUY; TONKIN, 2006). Contudo, é possível que essa personalização seja útil ao divulgador científico, pois a forma como o TikTok autocompleta buscas as *hashtags* pode ser um marcador que relaciona o conteúdo ao assunto ciência. Destaca-se que a relação entre termos excessivamente personalizados e as particularidades do TikTok não é objeto desta pesquisa.

A décima *hashtag* mais utilizada da amostra foi a #leftiktok (TikTok à esquerda em português), com 108 ocorrências, exclusivamente em publicações da área das Ciências Humanas. Trata-se de uma *hashtag* que representa a temática dos vídeos em relação ao contexto sociopolítico. Está relacionada à polarização política, observada em 2022 (ano de eleições no Brasil).

Os *insights* de público do TikTok mostraram que o maior alcance da #leftiktok foram pessoas entre 18 e 24 anos (67% do público), pessoas entre 25 e 34 anos (17% do público) e pessoas com mais de 35 anos (16% do público) (TIKTOK, 2023).

Quanto aos aspectos de indexação social, a #leftiktok, pode ser considerada um termo composto que reflete uma característica da linguagem da *Web*. Não pode ser considerada uma *hashtag* interna ao contexto acadêmico, pois sua função é demarcar conteúdo político.

A #leftiktok pode ser compreendida como um termo pertinente para representação de conteúdos, isso contribui para encontrar publicações relacionadas a assuntos desse contexto sociopolítico. No entanto, pode contribuir para a formação de bolhas ou câmaras de eco, servindo como barreira para a divulgação do conhecimento, devido ao viés ideológico. Em outras palavras, se contrapõem aos objetivos da divulgação científica.

A décima primeira *hashtag* mais utilizada é a #física, com 107 ocorrências, distribuídas nas áreas do conhecimento: Multidisciplinar (67); Ciências Exatas e da Terra (37); Engenharias (2) e Ciências da Saúde (1).

Os *insights* de público demonstraram que o maior alcance dessa *hashtag* foram pessoas entre 18 e 24 anos (76% do público), pessoas entre 25 e 34 anos (16% do público) e pessoas com mais de 35 anos (8% do público) (TIKTOK, 2023). A #física é apontada como extremamente popular no Brasil no momento da elaboração desta pesquisa (fevereiro de 2023) atingindo 104% de pico de audiência nas métricas do TikTok para os últimos 34 dias (TIKTOK, 2023).

Sob os aspectos de indexação social, a #física pode ser considerada um termo geral no contexto do TikTok,

pois é possível pensar em outros termos que poderiam representar especificações para o assunto (e.g. #física-departículas). Quanto aos aspectos linguísticos, possui um erro ortográfico (ausência de sinal gráfico) recorrente na linguagem da *Web*.

É uma *hashtag* interna ao contexto acadêmico, que auxilia a identificar o conteúdo da publicação e, consequentemente, aumenta as chances de alcançar seu público-alvo. No entanto, assim como nas *hashtags* #biologia e #economia é possível compreender que essa *hashtag* é importante para a recuperação da informação, mas deve vir acompanhada de outras mais específicas para contribuir na representação e recuperação de vídeos.

A décima segunda *hashtag* mais utilizada na amostra de pesquisa foi a #econolivia. Possui 101 ocorrências e foi utilizada exclusivamente pela área do conhecimento de Ciências Sociais Aplicadas.

Quanto aos aspectos linguísticos, não se trata de uma abreviação, mas da junção entre o tema principal do perfil (economia) e o nome da produtora de conteúdo (Olivia). Trata-se de uma *hashtag* personalizada, criada pelo canal da @olivia.sensata.

No que diz respeito a popularidade, teve mais de 26 milhões de visualizações e esteve presente em poucas publicações no Brasil (121 ocorrências), entre fevereiro de 2022 a fevereiro de 2023 (TIKTOK, 2023).

Os *insights* de público mostraram que o maior alcance dessa *hashtag* foram pessoas entre 18 e 24 anos (38% do público), pessoas entre 25 e 34 anos (35% do público) e pessoas com mais de 35 anos (27% do público) (TIKTOK, 2023). Esta *hashtag* possui a distribuição mais equilibrada de público por faixa etária. O público não está concentrado majoritariamente entre pessoas de 18 e 24 anos, mas apresenta um quantitativo consistente entre outras faixas etárias.

Foi categorizada no TikTok.Ads como uma *hashtag* ligada ao serviço financeiro, teve em torno de 5 milhões de visualizações, presente em 106 postagens no Brasil, entre os meses de fevereiro de 2022 e fevereiro de 2023 (TIKTOK, 2023).

Observou-se que a #econolivia é um termo de representação relacionado ao perfil que produz o conteúdo no TikTok. É uma *hashtag* externa ao contexto acadêmico, embora funcione como um termo que representa uma especificidade de um perfil, sendo uma *hashtag* personalizada. Por apresentar essa característica, seria um termo de marcação que não oferece suporte à pesquisa de maneira eficaz (GUY; TONKIN, 2006).

Contudo, é possível que essa personalização seja útil ao divulgador científico, pois considerando a pesquisa no TikTok, as *hashtags* são mais um marcador que relaciona o conteúdo com o assunto de economia. A relação entre termos excessivamente personalizados e as parti-

cularidades do TikTok poderão ser aprofundadas em pesquisas futuras.

5 Conclusão

Foi possível identificar estratégias de visibilidade baseada em indexação social para os conteúdos ligados à ciência. Observou-se que existem ocorrências diferentes das *hashtags* nas nove áreas do conhecimento da CAPES e *hashtags* empregadas exclusivamente em determinadas áreas ou criadas para conteúdo, ou perfis.

A análise da indexação social permitiu estabelecer que divulgadores científicos brasileiros utilizam as *hashtags* de forma estratégica para visibilidade. Existem seis tipos de *hashtags* envolvidas nas estratégias de visibilidade do conteúdo: a) internas ao contexto acadêmico (#ciência, #biologia, #economia e #física); b) genéricas e populares (#fy, #fyp); c) vinculadas a uma área do conhecimento (#agro); d) elaboradas para promover o conteúdo educacional (#aprendanotiktok e #curiosidades); e) que identificam ou personalizam a busca por informações veiculadas a um perfil (#cienciaatododia e #econolivia), e; f) de cunho político (#lefttiktok).

Destaca-se que as *hashtags* são parte de um sistema de classificação de conteúdo que considera múltiplas variáveis dentro do TikTok. Desta forma, as *hashtags* sozinhas não garantem a visibilidade e engajamento dos vídeos. Embora a indexação social faça parte do processo, é apenas uma das dimensões de análise considerada pelo algoritmo do TikTok. O algoritmo também considera as interações, tanto aquelas que o usuário faz com o conteúdo, quanto o tipo de conteúdo que o usuário pode produzir.

As *hashtags* são relevantes no processo de classificação no TikTok, com potencial de representar o conteúdo informacional dos vídeos, possibilitando reunir publicações que possuem assuntos relacionados. Em resumo, atuam na dupla função de identificar e reunir publicações em uma coleção.

Destaca-se que a divulgação científica não se destina exclusivamente a um nicho específico da sociedade ou uma faixa etária. Seu objetivo é contribuir com a disseminação de conhecimento a sociedade. Porém, reconhecer uma *hashtag* específica relacionada ao tema e com alcance significativo em um SRSO como TikTok, pode contribuir para a visibilidade de conteúdos científicos na *Web*.

Ao decorrer do processo de análise desta pesquisa, percebeu-se que a indexação social é uma de muitas estratégias e recursos que o algoritmo do TikTok se utiliza para segmentar conteúdo. Por isso, para obter um panorama sobre a visibilidade do conteúdo científico no TikTok é necessário que em trabalhos futuros uma ampliação da amostra e a observação de outros mecanis-

mos alavancam a visibilidade da informação no TikTok. Um exemplo é compreender como o algoritmo se alimenta de outras funcionalidades do SRSO, por meio da trilha sonora escolhida, das interações do usuário na interface *For you* e de outros mecanismos que podem contribuir para que um conteúdo seja considerado relevante ou não ao usuário.

Notas

- [1] O termo Serviços de Redes Sociais *Online* e seu acrônimo SRSO é utilizado para se referir especificamente às plataformas de comunicação disponibilizadas por empresas como Meta Platforms. Inc (Facebook e Instagram) ou ByteDance (TikTok), pois especifica o caráter de ferramenta de comunicação *online*, em que é possível que pessoas conectadas formem Redes Sociais de forma *Online*, ou seja, possibilita que pessoas se conectem *online*.
- [2] Informações sobre os espectadores que assistem a vídeos ligados as *hashtags*
- [3] É importante destacar que não se está afirmando que toda a produção vinculada à *hashtag* no TikTok é composta por postagens de caráter educacional ou científico.
- [4] @bea.zevedo. Disponível em: <https://vm.tiktok.com/ZMYyWWP2U/>. Acesso em: 02 fev. 2023.
- [5] O TikTok Ads é uma ferramenta para gerenciar anúncios no TikTok, que possibilita acesso a métricas de audiência global no TikTok, inclusive das *hashtags*

Referências

- BUENO, W. C. Comunicação científica e divulgação científica: aproximações e rupturas conceituais. **Informação & Informação**, Londrina, v. 15, n. esp., 2010. Disponível em: <https://www.brapci.inf.br/index.php/article/download/14078>. Acesso em: 10 ago. 2022.
- CANDOTTI, E. Ciência na Educação Popular. In: MASSARANI, L.; MOREIRA, I. C.; BRITO, F. (org.). **Ciência e público**: caminhos da divulgação científica no Brasil. Rio de Janeiro: Casa da Ciência, 2002. Disponível em: https://www.museudavida.fiocruz.br/images/Publicacoes_Educacao/PDFs/cienciaepublico.pdf. Acesso: 09 ago. 2022.
- CASTELLS, M. **A sociedade em rede**. 23. ed. São Paulo: Paz e Terra, 2021.
- CATARINO, M. E.; BAPTISTA, A. A. Folksonomias: características das etiquetas na descrição de recursos da Web. **Informação & Informação**, Londrina, v. 14, n. 1 esp., p. 46-67, dez. 2009. Disponível em: <https://brapci.inf.br/index.php/res/download/45658>. Acesso em: 15 jun. 2022
- CECI, L. TikTok - Statistics & Facts. **Statista**, [S. l.], 23 nov. 2022. Disponível em: <https://www.statista.com/topics/6077/tiktok/>. Acesso em: 03 maio 2022.
- GOMES, A. G. *et al.* Análise de Tags na venda de frutas da Amazônia brasileira em rede social. **Informação & Informação**, Londrina, v. 26, n. 3, p. 499–525, 2021. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/44079>. Acesso em: 14 maio 2022.
- GONÇALVES, M. Contribuições das mídias sociais digitais na divulgação científica. In: PINHEIRO, L. V. R.; OLIVEIRA, E. C. P. (org.). **Múltiplas facetas da comunicação e divulgação científica**: transformações em cinco séculos. Brasília: IBICT, 2012. Disponível em: <http://www.ibict.br>. Acesso em: 14 maio 2022.
- GUEDES, R. M.; DIAS, E. J. W. Indexação social: abordagem conceitual. **Revista ACB**: Biblioteconomia em Santa Catarina, Florianópolis, v. 15, n. 1, p. 39- 53, jan./jun. 2010. Disponível em: <https://revista.acb.org.br/racb/article/view/686>. Acesso em: 12 ago. 2022.
- GUY, M.; TONKIN, E. Folksonomies: tidying up tags? **D-Lib Magazine**, [S. l.], v. 12, n. 1, jan. 2006. Disponível em: <https://www.dlib.org/dlib/january06/guy/01guy.html>. Acesso em: 15 ago. 2022.
- HOLSTROM, C. Social Tagging: Organic and Retroactive-Folksonomies. In: ACM/IEEE ON JOINT CONFERENCE ON DIGITAL LIBRARIES, 18., 2018, Fort Worth, TX, USA, **Anais [...]** Fort Worth, TX, USA, 2018. Disponível em: <https://dl-acm-org.ez3.periodicos.capes.gov.br/doi/pdf/10.1145/3197026.3197065>. Acesso em: 30 set. 2022.
- LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. rev. atual. Brasília, DF: Briquet de Lemos/Livros, 2004. Disponível em: <https://bibliotextos.files.wordpress.com/2014/07/livro-indexac3a7c3a3o-e-resumos-teoria-e-prc3a1tica-lancaster.pdf>. Acesso: 29 set. 2022.
- LIMA, G. A. Organização e representação do conhecimento e da informação na web: teorias e técnicas. **Perspectivas em Ciência da Informação**, v. 25, p. 57-97, 2020. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/135734>. Acesso em: 09 set. 2022.
- MATHES, A. **Folksonomies** - cooperative classification and communication through shared metadata. Computer Mediated Communication – LIS590CMC, Urbana : University of Illinois, 2004. Disponível em: <https://adammathes.com/academic/computer-mediated-communication/folksonomies.html>. Acesso em: 17 ago. 2022.
- MUELLER, S. Popularização do conhecimento científico. **DataGramZero**, [S. l.], v. 3, n. 2, 2002. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/5354>. Acesso em: 24 jul. 2022.
- PARISER, E. **O filtro invisível**. [S. l.]: Zara, [2012]. Disponível em: <https://lereumvicio.files.wordpress.com/2016/06/o-filtro-invisivel-eli-pariser.pdf>. Acesso em: 30 set. 2022.
- PAULINO, R. C. R.; VENTURA, M. P. O engajamento no Twitter: métodos de análise para #Somos70porcento. **Cardernos Info**. Santiago, n. 49, p. 51-71, 2021. Disponível em: https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0719-367X2021000200051&lng=pt. Acesso em: 05 dez. 2021.

PINHEIRO, L.V.R. Comunidades científicas e infra-estrutura tecnológica no Brasil para uso de recursos eletrônicos de comunicação e informação na pesquisa. **Ciência da Informação**. Brasília, v.32, n.3, p.62-73, 2003. Disponível em: <https://www.scielo.br/j/ci/a/h6P7hVhMrxRtqTYfRQQGd6t/?format=pdf&lang=pt>. Acesso em: 24 jul. 2022.

SANTAREM SEGUNDO, J. E.; VIDOTTI, S. A. B. G. Representação iterativa e folksonomia assistida para repositórios digitais. **Liinc em Revista**, v. 7, n. 1, p. 283-300, 2011. Disponível em: <http://hdl.handle.net/11449/114930>. Acesso em: 15 ago. 2022

SANTOS, R. F.; CORRÊA, R. F. Conceituando Folksonomia: análise e síntese dos diversos usos do termo na literatura de Ciência da Informação. In: CONGRESSO BRASILEIRO EM ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO, 5., 2017, Belém, **Anais [...]**. Belém: ISKO Brasil, 2017. Disponível em: <https://brapci.inf.br/index.php/res/download/121673>. Acesso em: 09 set. 2022.

TIKTOK. **5 tips for TikTok creators**. [S. l.], 30 jun. 2020. TikTok: Newsroom. Disponível em: <https://newsroom.tiktok.com/en-us/5-tips-for-tiktok-creators>. Acesso em: 15 jan. 2023

TIKTOK. **TikTok dá início ao Programa de Aceleração para apoiar conteúdo educativo no Brasil**. [S. l.], 24 ago. 2021. TikTok: Newsroom. Disponível em: <https://newsroom.tiktok.com/pt-br/aprenda-no-tiktok-educacao-escolar>. Acesso: 13 fev. 2023

TIKTOK. **Segurança e bem-estar dos jovens**. [S. l.], 2022. TikTok: Diretrizes da Comunidade: Segurança de menores. Disponível em: https://www.tiktok.com/community-guidelines?lang=pt_BR#31. Acesso em: 09 set. 2022.

TIKTOK. **Trends**. [S. l.], 2023. TikTok: Ads.TikTok. Disponível em: <https://ads.tiktok.com/business/creativecenter/inspiration/popular/hashtag/pc/en>. Acesso em: 06 fev. 2023.

TORRESI, S. I. C.; PARDINI, VERA, L.; FERREIRA, V. F. Sociedade, divulgação científica e jornalismo científico. **Química Nova**, [S. l.], v. 35, n. 3, p. 447, 2012. Disponível em: <https://doi.org/10.1590/S0100-40422012000300001>. Acesso: 16 ago. 2022.

VIANA, J. M. A.; DAL'EVEDOVE, P. R.; TARTAROTTI, R. C. A inserção da Folksonomia nos Programas de Pós-graduação em Ciência da Informação no Brasil. In: BARROS, T. H. B; TOGNOLI, N. B. **Organização do Conhecimento responsável: promovendo sociedades democráticas e inclusivas**. Belém: EdUFPA, 2019. p. 321-330. Disponível em: <https://isko.org.br/publicacoes/serie-estudos-avancados-em-organizacao-e-representacao-do-conhecimento>. Acesso em: 12 abr. 2022.

Apêndice A – Dados analisados

Tabela 2. *Hashtags mais populares da amostra, ranqueadas de acordo com o total de frequência $n \geq 100$, por área do conhecimento*

#	Hashtags	Ciências Agrárias	Ciências Biológicas	Ciências da Saúde	Ciências Exatas e da Terra	Ciências Humanas	Ciências Sociais Aplicadas	Engenharias	Linguística, Letras e Artes	Multi-disciplinar	Total
1	ciencia	0	47	68	252	18	0	0	0	107	492
2	aprendanotiktok	0	47	1	176	45	65	0	0	0	334
3	fyp	1	0	4	153	79	0	0	16	23	276
4	fy	15	0	0	127	74	0	0	18	30	264
5	curiosidades	0	21	41	0	0	1	0	2	61	126
6	agro	122	0	0	0	0	0	0	0	0	122
7	economia	0	0	0	0	14	98	0	0	0	112
8	biologia	0	97	1	0	0	0	0	0	12	110
9	cienciatododia	0	0	0	0	0	0	0	0	109	109
10	leftiktok	0	0	0	0	108	0	0	0	0	108
11	fisica	0	0	1	37	0	0	2	0	67	107
12	econolivia	0	0	0	0	0	101	0	0	0	101
-	<i>Demais Hashtags</i>	662	540	577	1006	1066	527	176	427	864	5845
-	Total	800	752	693	1751	1404	792	178	463	1273	8106

Fonte: Autores (2023).

Quadro 1. *Características observadas nos termos de indexação social, ranqueadas de acordo com o total de frequência*

Hashtags ¹ (n \geq 100)	Propriedades da indexação						Aspectos linguísticos			
	Termo geral	Termo específico	Termo específico de uma área	Termo composto	Interno ao contexto acadêmico	Externo ao contexto acadêmico	Possui erro de digitação	Linguagem da Web	Abreviação	Representa especificidades do perfil ou área do conhecimento
ciencia	Sim	Não	Não	Não	Sim	Não	Sim	Sim	Não	Não
aprendanotiktok	Não	Sim	Não	Sim	Sim	Sim	Não	Sim	Não	Não
fyp	Sim	Não	Não	Sim	Não	Sim	Não	Sim	Sim	Não
fy	Sim	Não	Não	Sim	Não	Sim	Não	Sim	Sim	Não
curiosidades	Sim	Não	Não	Não	Sim	Não	Não	Não	Não	Não
agro	Não	Não	Sim	Não	Não	Sim	Sim	Sim	Não	Sim
economia	Sim	Não	Não	Não	Sim	Não	Não	Não	Não	Sim
biologia	Sim	Não	Não	Não	Sim	Não	Não	Não	Não	Sim
cienciatododia	Não	Sim	Não	Sim	Sim	Não	Não	Sim	Não	Sim
leftiktok	Não	Sim	Não	Sim	Não	Sm	Não	Sim	Não	Sim
fisica	Sim	Não	Não	Não	Sim	Não	Sim	Não	Não	Sim
econolivia	Não	Sim	Não	Sim	Não	Sim	Não	Sim	Não	Sim

¹ Recorte de hashtags com 100 ou mais ocorrências no universo pesquisado ($n \geq 100$).

Fonte: Autores (2023).

Introdução a práticas de mineração de dados: considerações sobre os dados abertos governamentais em Londrina

Introduction to data mining practices: considerations on open government data in Londrina

Matheus Antunes Palmieri (1), Benjamin Luíz Franklin (2)

(1) Universidade Estadual de Londrina, Campos Universitário, matheus.palmieri@uel.br, (2) belfra@uel.br

Resumo

Este estudo explora o potencial analítico da crescente disponibilidade de documentos públicos online para profissionais da informação e arquivistas. O objetivo é examinar o conjunto documental do portal de transparência da prefeitura de Londrina, convertendo a informação não estruturada em um Modelo Entidade-Relacionamento e, em seguida, em um modelo de rede. A metodologia envolve a aplicação de algoritmos de consolidação, agrupamento e centralidade para oferecer uma visão relacional de atores como Documentos, Secretaria e Credores. A pesquisa baseia-se no arcabouço teórico da Mineração de Dados Públicos e seus recursos analíticos. Os resultados sugerem que a transparência na governança deve estar vinculada ao desenvolvimento de competências analíticas pelos profissionais da informação, considerando a crescente disponibilidade de recursos computacionais e tecnológicos. A ampliação da transparência e a análise de dados públicos são cruciais para a democracia e a prestação de contas. Futuras pesquisas devem focar no estreitamento do diálogo entre a Arquivologia e os métodos analíticos, visando uma crítica mais aprofundada à dinâmica democrática atual. Isso implica na necessidade de maior colaboração entre as áreas e na aplicação de técnicas analíticas para aprimorar a transparência e a eficácia dos processos de governança.

Palavras-chave: Arquivologia. Ciência de dados. Teoria ator-rede. Dados abertos governamentais.

Abstract

This study explores the analytical potential of the increasing availability of public online documents for information professionals and archivists. The objective is to examine the document set of the transparency portal of the city hall of Londrina, converting unstructured information into an Entity-Relationship Model and then into a network model. The methodology involves the application of consolidation, grouping, and centrality algorithms to offer a relational view of actors such as Documents, Secretariat, and Creditors. The research is based on the theoretical framework of Public Data Mining and its analytical resources. The results suggest that transparency in governance should be linked to the development of analytical skills by information professionals, considering the growing availability of computational and technological resources. The expansion of transparency and public data analysis is crucial for democracy and accountability. Future research should focus on narrowing the dialogue between Archivology and analytical methods, aiming at a more in-depth critique of current democratic dynamics. This implies the need for greater collaboration between the fields and the application of analytical techniques to enhance transparency and the effectiveness of governance processes.

Keywords: Archivology, Data science, Actor-network theory, Government open data.

Resumen

Este estudio explora el potencial analítico de la creciente disponibilidad de documentos públicos en línea para profesionales de la información y archivistas. El objetivo es examinar el conjunto documental del portal de transparencia del ayuntamiento de Londrina, convirtiendo la información no estructurada en un Modelo Entidad-Relación y luego en un modelo de red. La metodología implica la aplicación de algoritmos de consolidación, agrupamiento y centralidad para ofrecer una visión relacional de actores como Documentos, Secretaría y Acreedores. La investigación se basa en el marco teórico de la Minería de Datos Públicos y sus recursos analíticos. Los resultados sugieren que la transparencia en la gobernanza debe estar vinculada al desarrollo de competencias analíticas por parte de los profesionales de la información, considerando la creciente disponibilidad de recursos computacionales y tecnológicos. La ampliación de la transparencia y el análisis de datos públicos son cruciales para la democracia y la rendición de cuentas. Las futuras investigaciones deben centrarse en estrechar el diálogo entre la Archivología y los métodos analíticos, buscando una crítica más profunda a la dinámica democrática actual. Esto implica la necesidad de una mayor colaboración entre las áreas y la aplicación de técnicas analíticas para mejorar la transparencia y eficacia de los procesos de gobernanza.

Palabras clave: Arquivología. Ciencia de datos. Teoría actor-red. Datos abiertos gubernamentales.

1 Introdução

A crescente disponibilização de documentos públicos na web oferta um extenso campo exploratório de análise para os profissionais da informação em geral, e os arquivistas em particular. Este trabalho procura evidenciar essa possibilidade ao analisar o conjunto documental do portal de transparência da prefeitura de Londrina, ao transformar a informação não estruturada de seu corpus para o Modelo Entidade-Relacionamento e, recursivamente, em um modelo de rede, visando a aplicação de algoritmos de consolidação, agrupamento e centralidade, ofertando uma visão relacional de alguns de seus atores.

A partir do arcabouço teórico da Mineração de Dados Públicos e seus recursos analíticos, buscamos relacionar atores como: Documentos, Secretaria e Credores, obtendo um quadro de sua dinâmica. Os resultados sugerem que a ampliação da transparência na governança, a partir da ampla divulgação de documentos públicos, deve estar associada à apropriação de competências analíticas pelos profissionais da informação, dada a crescente disponibilidade de recursos computacionais e tecnológicos. Futuras pesquisas, dessa forma, sugerem o estreitamento do diálogo entre a Arquivologia e os métodos analíticos, para realizar a crítica necessária à dinâmica democrática atual.

Buscando “a cada dia praticar a cultura da transparência e do controle social no Município de Londrina” (LONDRINA, 2022, [online]) o portal da transparência do município de Londrina oferece ao público informações detalhadas sobre o orçamento, os gastos públicos, licitações, contratos entre outros e em diferentes áreas, como a saúde, segurança, infraestrutura entre outras. O portal está subordinado as exigências da Lei de Acesso à Informação (Lei 12.527, de 18/11/2011), da Lei Complementar 131, de 27/05/2009 e regulamentações municipais, como descrito no portal. Os dados procuram ser oferecidos em tempo real, respeitando a natureza da informação.

Contudo, a disponibilização de informações governamentais por força da Lei de Acesso à Informação (Lei nº 12.527/2011) e da Lei da Transparência (Lei Complementar 131/2009), se apresenta em diferentes formatos, constituindo “silos de informação” (JACINTO, et al., 2022, p. 40). Os chamados silos de informação correspondem a aglomerados de dados não estruturados em diferentes formatos, este fato atesta limitação no acesso para potencial reuso da informação, uma vez que caberá ao usuário efetuar a limpeza, tratamento e interligações com outras fontes de dados para inferir conhecimentos mais aprofundados a partir da informação disponibilizada em diversos formatos. (JACINTO, et al., 2022).

A *Sunlight Foundation* (2019) disponibiliza dez princípios aplicáveis aos dados abertos governamentais, sendo eles: 1) Completude, 2) Primazia dos dados brutos, 3) Facilidade de acesso, 4) Reutilização, 5) Universalidade, 6) Integridade, 7) Exemplaridade, 8) Melhoria Contínua dos dados, 9) Cooperação e 10) Tecnologia-neutra. Tais princípios são de extrema importância, pois objetivam tornar os dados abertos governamentais mais transparentes, acessíveis e úteis para os cidadãos.

Dentre os dez princípios citados, o da primazia dos dados brutos, da facilidade de acesso e o da integridade, neste contexto da pesquisa, são os princípios essenciais a serem ressaltados visando garantir a necessidade de que os dados estejam disponíveis em formato aberto, estruturado e legível por máquina.

Em outras palavras, o formato aberto se refere a interoperabilidade, reutilização e acessibilidade dos dados governamentais, permitindo que diferentes sistemas e ferramentas possam ler, processar e compartilhar dados de forma mais fácil e eficiente. Já os dados estruturados “[...] tem seu conteúdo interno explicitado por meio de uma semântica que permite que todo seu conteúdo seja interpretado de forma autônoma por máquina” (SANTOS; SANT’ANA, 2015, p. 206). Por fim, o formato será legível por máquina quando disponibilizado em formatos eletrônicos que permitam o processamento automático dos dados. (SUNLIGHT FOUNDATION, 2019)

Isso significa que os dados devem estar estruturados de forma a permitir que sejam facilmente acessados e processados por sistemas computacionais, possibilitando a criação de ferramentas e aplicativos que possam utilizá-los para gerar insights, análises e visualizações.

Neste sentido, utilizaremos o arcabouço conceitual e técnico metodológico da análise da mineração de dados públicos como método aplicável a estes silos de informação.

2 Metodologia

A partir do arcabouço teórico da teoria ator-rede (LATOURETTE, 2012), espera-se sugerir um método de análise de rede aplicado aos dados públicos enquanto técnica de mineração de dados para analisar o conjunto dos dados obtidos, tentando oferecer um recurso para extração de informações úteis e insights fornecendo um mecanismo para tomada de decisões em diversas áreas.

O acesso, eficiência e eficácia dos dados abertos governamentais disponibilizados pelo governo para reutilização e interligação com outras fontes de dados estão relacionados com a disseminação dos dados de forma estruturada e em formatos não proprietários.

A mineração de dados públicos implica no processo de extração de informações úteis de grandes conjuntos de dados disponíveis publicamente e oferecidos pelo Estado a título de transparência a partir de um fluxo de trabalho envolvendo diversas etapas aplicadas aos dados. Essas etapas aplicadas ao ciclo analítico da ciência de dados correspondem a carregar, limpar, transformar, visualizar e modelar (fase exploratória) e comunicar o resultado. (OLIVEIRA; GUERRA; MCDONELL, 2018).

Do ponto de vista social, partindo desse entendimento e o relacionando com a teoria ator-rede (LATOURET, 2012), entendemos que o processo de mineração não se trata apenas de um processo técnico, mas também social para contribuir com o entendimento de como esses dados são produzidos e como as informações são geradas.

Em relação a abordagem, Latour (2012) propõe a escolha de determinar inicialmente o grupo e nível de análise a ser enfatizada ou a adoção dos procedimentos dos atores envolvidos no domínio para o rastreamento de suas pistas na formação e desmantelamento de grupos.

Neste sentido, sendo possível identificar protagonistas do movimento de formação e desmantelamento de grupos através das pistas deixadas pelos atores do domínio em estudo, espera-se contribuir ativamente com o empoderamento da sociedade oferecendo um recurso de fiscalização da transparência governamental. Deste modo, através da análise computacional de conteúdo objetiva-se mapear os relacionamentos de atores envolvidos nas contas pagas do portal da transparência do município de Londrina no período de 01/01/2022 a 09/09/2022.

Partindo do levantamento das notas de pagamento extraídas do Portal da Transparência do Município de Londrina – PR, objetiva-se, também, apresentar uma proposta metodológica que promova a real participação da população na auditoria da transparência governamental para garantir o exercício efetivo da cidadania.

Com este propósito utilizou-se os seguintes softwares livres para garantir a replicabilidade da metodologia: Python (<https://www.python.org/>), LibreOffice Base (<https://pt-br.libreoffice.org/>) e LibreOffice Calculator³, Open Refine (<https://openrefine.org/>) e o principal *software* de visualização e exploração para todos os tipos de gráficos e redes de acesso não proprietário.

2.1 Fonte de coleta de dados

Uma vez que a Universidade Estadual de Londrina se encontra na circunscrição territorial do município de Londrina/PR e um dos eixos das atividades desenvolvidas pela Universidade é a extensão, escolhemos como fonte de coleta de dados o website do portal da transparência do município de Londrina/PR (<https://portal.londrina.pr.gov.br/transparencia>) pretendendo oferecer recurso útil à participação cidadã

na auditoria da transparência do referido Portal, em especial, a população local. Assim, optou-se pela seção de contabilidade e finanças, em virtude da possibilidade de análise exploratória de caráter quantitativo. Para tanto, dentro da categoria de Despesas, foi selecionada a opção de Despesas Pagas. Por entidade foi escolhida a Prefeitura do Município de Londrina durante o exercício de 2022, do período de 01/01/2022 a 09/09/2022.

A escolha de despesas pagas ao invés das despesas empenhadas, enquanto critério de seleção, ocorreu em virtude das despesas pagas representarem um registro concreto dos gastos públicos no período investigado, que apesar de não estarem disponibilizados de forma estruturada poderiam ser submetidos a análise em questão.

Além disso, por ser uma fonte oficial de informações públicas, o portal da transparência permite acessar dados atualizados sobre as despesas do município. Por se tratar de informações governamentais, partiu-se do pressuposto de que os dados íntegros, isto é, precisos e completos, quando analisados e, ainda, as contas pagas estavam disponibilizadas em formato PDF, o que permitiu a conversão para formato TXT do conteúdo sem perda do contexto orgânico do documento.

A opção de despesas pagas em desprovido das contas empenhadas se deu em virtude da suposição de que a população teria mais interesse para consultar a destinação do dinheiro público, isto é, quanto cada credor recebeu. Em outras palavras, a escolha da opção de despesas pagas foi feita com base na relevância dessa informação para a análise em questão, que está focada em entender como os recursos públicos foram destinados no período delimitado através da relação de latência entre os credores e as secretarias.

2.2 Problemas de processamento

Em virtude de problemas de processamento, tais como limitação de *hardware* e *software*, o montante total de documentos foi restringido para permitir o processamento de acordo com a capacidade computacional disponível para as práticas. Não foi possível processar a totalidade das relações esperadas dentro do domínio em estudo em virtude de limitação na disponibilidade de *hardware*, mais especificamente de memória RAM para execução da prática, apesar do emprego de linguagem SQL e de diversas tentativas de por meio do Python, usando a biblioteca Scriptforge, a exportação da totalidade dos dados esperados partindo das consultas de cálculo para visualização do mapeamento das relações de latência entre os credores e as secretarias a partir dos documentos representados em um grafo no *software* Gephi.

2.3 Procedimentos desenvolvidos

Seguindo o ciclo de vida analítico da ciência de dados, começamos pela etapa do carregamento, que implica na obtenção dos dados brutos, no caso, disponíveis em

formato proprietário. Assim, uma vez delimitada a fonte da coleta a ser realizada, empregou-se a linguagem Python para efetuar o download dos documentos, amparando-se nas bibliotecas *time* (<https://docs.python.org/3/library/time.html>), *requests* (<https://pypi.org/project/requests/>) e *multiprocessing* (<https://docs.python.org/3/library/multiprocessing.html>), o que permitiu maior flexibilidade para atender o desafio de estruturação dos dados encontrado na disponibilização dos dados abertos governamentais no Portal da Transparência do Município de Londrina em relação ao tema estudado. Tendo reunido os documentos em formato PDF, efetuou-se uma breve análise objetivando encontrar padrões na apresentação dos dados. Identificamos que os documentos se apresentam como dossiês contendo múltiplos itens documentais separados através das páginas dos PDFs. Assim, foi realizada a divisão dos arquivos PDFs por seu número de páginas para individualização das notas de pagamento. Na sequência, passando para a etapa da transformação, os documentos foram submetidos à conversão de formato de PDF para TXT. Para tanto, também se empregou o uso da linguagem Python, através das bibliotecas *pdfminer*⁹, *io*¹⁰, *glob*¹¹, *os*¹² e *re*¹³, em razão do grande volume de documentos, com o intuito de levá-los ao carregamento no OpenRefine, objetivando a limpeza e tratamento por meio de linguagem GREL e Regex.

Em seguida, empregamos a etapa de limpeza, onde o conteúdo dos documentos já se encontrava em formato TXT. Levamos os dados brutos ao encontro do OpenRefine, onde, por meio do nome do documento com sua relação de página preservou-se a relação orgânica do documento com seu contexto, permitindo assim estabelecer a relação com a triade <e,a,v>, sem perder o vínculo com a origem dos dados. Então, buscamos identificar e estruturar as entidades e os atributos envolvidos a partir de análise exploratória.

Em virtude dos problemas de processamento anteriormente mencionados, a totalidade dos documentos não pôde ser apreciada. Desta forma, os dados foram filtrados para restringir a análise e permitir o seu processamento. Assim, procurou-se mapear a relação de latência existente entre os atores envolvidos no domínio em estudo. Levando os dados estruturados para o LibreOffice Base, efetuou-se o relacionamento entre as entidades, construindo, por meio de consultas, o relacionamento de latência entre as entidades para calcular os *edges* (arestas) e *nodes* (nós) que irão compor o grafo. Com isso, partimos para a etapa de visualização dos dados já estruturados, limpos e filtrados para elaboração da representação gráfica dos dados selecionados concentrados em um grafo utilizando o software livre Gephi. Então, partiu-se para a etapa de modelagem, onde se aplicou o algoritmo de modularidade para calcular a existência de grupos e utilizou-se o algoritmo para visualização geral da rede,

com o objetivo de calcular o grau médio e identificar o peso de cada nó. Em seguida, levando as métricas obtidas a partir dos algoritmos para o LibreOffice Base, identificaram-se os termos com grau mais elevado através de agrupamentos e definiu-se os grupos a partir da relação dos termos candidatos.

Os resultados foram comunicados por meio da apresentação dos grafos, oferecendo o mapeamento da relação de latência dos credores através das secretarias dentro do recorte efetuado para viabilizar o processamento.

2.4 Resultados

Como resultado do download, obteve-se 252 arquivos em formato PDF, dos quais o split alcançou um total de 63.447 documentos, que por sua vez se transformou em número equivalente de arquivos em formato TXT.

Quando carregados no OpenRefine, o posicionamento dos delimitadores dentro do conteúdo dos documentos transcrito em formato TXT apresentou diversidade de padrões, resultando em dificuldade na recuperação da informação, isto é, uma vez que os campos de dados inseridos nos documentos em formato PDF apresentavam diversos posicionamentos semânticos dentro do texto, este fato exigiu o desenvolvimento de diversas expressões em linguagem GREL conjugadas com RegEx para extração das entidades e atributos dentro do corpus do texto. Ao longo da análise exploratória notou-se dados faltantes no histórico de empenho, fato este que compromete o resultado e a confiabilidade das análises, trazendo dúvida no tocante a confiabilidade dos dados da pesquisa e do próprio portal.

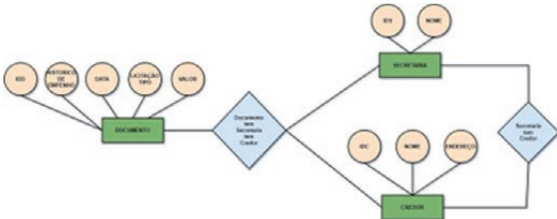
Durante o tratamento e limpeza dos dados extraídos em TXT e carregados no OpenRefine, buscou-se identificar e estruturar as entidades e os atributos envolvidos na análise exploratória. Assim, identificou-se presente nos documentos as seguintes entidades: Fornecedor e Secretaria. Também identificou-se atributos, sendo eles: data, tipo de licitação, endereço, CPF/CNPJ, telefone, histórico de empenho e classificação de despesa.

Por valor, foi considerado o valor total da nota de pagamento descontado a retenção.

Disto resultou um modelo resumido de entidade-relacionamento (MER) para compreensão do domínio em estudo. O modelo de entidade-relacionamento se pauta em três conceitos fundamentais, sendo eles, a entidade, os relacionamentos e os atributos. As entidades são entendidas enquanto objetos e podem se apresentar de formas diferentes e identificáveis, já os relacionamentos são as associações entre as entidades envolvidas em um domínio e por fim, os atributos são as qualidades das entidades.

Assim, o modelo resumido elaborado será apresentado abaixo:

Figura 1. Modelo de Entidade-Relacionamento



Fonte: Elaborado pelos autores.

A entidade documento tem por atributos um ID, data, histórico de empenho, tipo de licitação e valor. Já a entidade “credor”, por se tratar de individualização de pessoa física ou jurídica, se relaciona com um ID único, nome, endereço entre outros atributos entendidos enquanto dados sensíveis e, portanto, foram suprimidos do modelo e da análise de rede.

A partir das entidades observadas nos documentos, identificou-se como atores as secretarias e os credores/fornecedores envolvidos nas notas de pagamento. Assim, procurou-se mapear a relação de latência existente entre os credores e as secretarias, isto é, uma vez que cada nota de pagamento apresenta um único credor/fornecedor e uma única secretaria, procurou-se identificar como os credores/fornecedores se relacionam através das secretarias para mapear esses relacionamentos no período delimitado e oferecer uma visualização gráfica do relacionamento entre os atores do fundo documental em estudo.

O processo de estruturação dos dados para a análise computacional de conteúdo também resultou em diversas expressões para extração da informação nos arquivos TXT. E entendemos que a estruturação dos dados possibilita a realização de observatórios, contribuindo para o desenvolvimento de estudos quantitativos em diversas áreas. Assim, das práticas advindas da aproximação, por meio da análise computacional de conteúdo resultaram grafos, sendo eles:

Figura 2. Mapeamento da relação de latência entre credores e secretarias a partir da seleção de documentos contendo o termo “software” no histórico de empenho.



Fonte: Elaborado pelos autores.

O grafo da figura 2 objetivou, partindo dos credores identificados dentro dos dados extraídos dos PDFs relativos à prestação de serviços de *software* contratados pela prefeitura, identificar padrões e tendências de gastos relacionados a essa área.

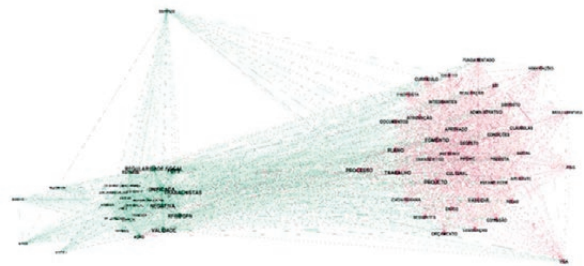
Figura 3. Mapeamento da relação de latência entre credores e secretarias a partir da seleção de documentos contendo o termo “software” e “Impressoras” no histórico de empenho.



Fonte: Elaborado pelos autores.

O grafo da figura 3 intentou identificar possíveis sinergias entre as áreas de software e impressão, possibilitando a otimização de recursos e a redução de custos.

Figura 4. Mapeamento da relação de latência entre termos candidatos a partir da seleção de documentos contendo o termo “Subvenções Sociais” no histórico de empenho.



Fonte: Elaborado pelos autores.

O grafo da figura 4 também se inclinou na identificação de quais termos estão mais relacionados a partir do contexto e entre si para definir quais são mais relevantes para a análise em questão e quais podem ser ignorados. Assim, dentro do conteúdo dos documentos envolvendo as subvenções sociais, o emprego dos algoritmos consolidação, agrupamento e centralidade resultaram em dois grupos distintos, o que sugere

A prática da análise computacional de conteúdo aplicada no domínio em estudo resultou em dois questionamentos,

sendo eles: 1) A análise computacional de conteúdo pode ser utilizada como suporte para a classificação documental na Arquivologia? 2) Quais são as possíveis limitações e desafios da aplicação da análise computacional de conteúdo na Arquivologia?

3 Conclusão

Conforme evidenciado pelo experimento, o tratamento analítico dos silos de informação (JACINTO, *et al.*, 2022) pode contribuir para o processo de transparência governamental. Nas próximas pesquisas buscaremos consolidar as informações sugeridas nos modelos encontrados contribuindo para uma melhor análise quantitativa das relações propostas.

Os estudos e práticas sobre a mineração de dados públicos podem beneficiar o município de Londrina contribuindo com a identificação de problemas, o que pode culminar em soluções mais eficientes para a cidade. Também pode contribuir com a tomada de decisões, oferecendo mais informações a partir de agrupamentos e análises de rede mais aprofundadas. A iniciativa também pode contribuir para o oferecimento de dados em nível de granulometria mais refinado e de forma estruturada, contribuindo com a melhoria da transparência e consequentemente com a participação cidadã na auditoria da transparência do município.

Pode colaborar, ainda, com a economia de recursos a partir dos agrupamentos dos dados abertos governamentais disponibilizados pelo município, o que contribui não apenas para a população, mas também para os gestores públicos na tomada de decisão, promovendo melhoria na eficiência dos serviços prestados pelo município à população e, portanto, contribuindo com o bem-estar social.

A análise computacional de conteúdo por meio da aplicação de algoritmos de consolidação, agrupamento e centralidade pode ser empregada em um conjunto de documentos com a finalidade de identificar palavras-chave e temas recorrentes, o que pode facilitar a categorização de classes e subclasses específicas. Implicando em potencial recurso útil para a classificação documental quando utilizada de forma adequada e em conjunto com as técnicas e metodologias da área.

Em resumo, o tratamento analítico dos silos de informação pode ter um impacto significativo na transparência governamental e na eficiência dos

serviços prestados à população. A mineração de dados públicos pode ajudar a identificar problemas e apoiar a tomada de decisões mais informadas. Além disso, a análise computacional de conteúdo pode ser útil para a classificação documental e a identificação de palavras-chave e temas recorrentes.

Diante desse contexto, é importante que os gestores públicos e pesquisadores continuem a investir em estudos e práticas relacionadas à mineração de dados e análise computacional de conteúdo. Ao fazer isso, podem contribuir para a melhoria da transparência governamental, a eficiência dos serviços prestados à população e o bem-estar social. Além disso, a pesquisa pode levar a novos insights e descobertas que podem ajudar a enfrentar desafios e problemas específicos em diferentes áreas e setores. Em última análise, essas tecnologias e metodologias podem ter um impacto transformador na maneira como a informação é coletada, analisada e utilizada para fins públicos.

Referências

- CAMILLO, C. “Sobre”. **Portal da Prefeitura do Município de Londrina**, 9 de setembro de 2022. Disponível em: <https://portal.londrina.pr.gov.br/index.php/sobre-transparencia>. Acesso em: 20 abr. 2023.
- JACINTO, A. S. *et al.* Aumentando a eficiência da fiscalização social sobre gastos públicos por meio de uma aplicação web baseada em dados abertos. **Ciência, tecnologia e inovação: Geração de emprego e democratização de oportunidades**, Ponta Grossa, p.39 - 44, 2022. Disponível em: <https://www.atenaeditora.com.br/catalogo/ebook/ciencia-tecnologia-e-inovacao-geracao-de-emprego-e-democratizacao-de-oportunidades> Acesso em: 09 dez. 2022.
- LATOUR, B. **Reagregando o social: uma introdução a teoria do ator-rede**. EDUFBA ; EDUSC, 2012.
- OLIVEIRA, P. F.; GUERRA, S.; MCDONNELL, R. **Ciência de Dados com R – Introdução**. Brasília: Editora IBPAD, 2018.
- SUNLIGHT FOUNDATION. **Ten Principles For Opening Up Government Information**. 2017. Disponível em: <https://sunlightfoundation.com/policy/documents/ten-open-data-principles/>. Acesso em: 20 abr. 2023.
- SANTOS, P. L. V. A. C.; SANT'ANA, R. C. G. Dado e granularidade na perspectiva da informação e tecnologia: uma interpretação pela ciência da informação. **Ciência da Informação**, v. 42, n. 2, 2013. DOI: 10.18225/ci.inf.v42i2.1382 Acesso em: 20 abr. 2023.

Investigações acerca da representação da informação em repositórios de dados de pesquisa na Ciência da Informação brasileira

Investigations about the representation of information in research data repositories in Brazilian Information Science

Raimunda Fernanda dos Santos (1), Carla Beatriz Marques Felipe (2)

(1) Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro - RJ/Universidade Estadual de Londrina (UEL), Londrina- PR, fernanda@facc.ufrj.br. (2) Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro-RJ, carlabeatriz@facc.ufrj.br

Resumo

Estuda aspectos relativos à representação da informação em repositórios de dados de pesquisa. Tem como objetivo geral mapear a produção científica nacional sobre representação da informação em Repositórios de Dados de pesquisa no âmbito da Ciência da Informação. Objetiva-se especificamente: investigar os principais enfoques temáticos das produções científicas recuperadas; caracterizar os tipos de representação de dados (descritiva e/ou temática) discutidas na literatura; apontar tendências e perspectivas de estudos relacionados à temática. Utiliza como metodologia as pesquisas bibliográfica, exploratória e descritiva com abordagem qualitativa, tendo como *corpus* os artigos de periódicos indexados na Base de Dados em Ciência da Informação. Como resultados, constata que os artigos estão divididos em quatro grandes eixos temáticos, a saber: apresentação de regras e principais padrões de metadados em Repositórios de Dados de Pesquisa; integração de dados de pesquisa mediante a utilização de estruturas de representação; reflexões acerca da representação e recuperação de dados em domínios específicos; e avaliação da qualidade de metadados em Repositórios de Dados de Pesquisa. Como conclusão, evidencia-se que os aspectos relacionados à representação da informação nos repositórios de dados de pesquisa na Ciência da Informação brasileira vêm ganhando destaque nos últimos anos, especialmente no que concerne à representação descritiva da informação.

Palavras-chave: Representação da Informação; Repositórios de Dados de Pesquisa; Dados de pesquisa.

Abstract

Studies aspects related to the representation of information in research data repositories. General objective is to map the national scientific production on information representation in Research Data Repositories within the scope of Information Science. Specifically aims to: investigate the main thematic approaches of retrieved scientific productions; characterize the types of data representation (descriptive and/or thematic) discussed in the literature; trends and perspectives of studies related to the theme. Uses as a methodology bibliographical, exploratory and descriptive research with a qualitative approach, having as corpus the articles from periodicals indexed in the Information Science Database. Result, it was discovered that the articles are divided into four major thematic axes, namely: presentation of rules and main metadata standards in Research Data Repositories; research data integration through the use of representation structures; reflections on the representation and retrieval of data in specific domains; and evaluation of the quality of metadata in Research Data Repositories. Conclusion, it appears that aspects related to the representation of information in research data repositories in Brazilian Information Science have been gaining prominence in recent years, especially with regard to the descriptive representation of information.

Keywords: Information Representation; Research Data Repositories; Search data.

Resumen

Estudia aspectos relacionados con la representación de la información en los repositorios de datos de investigación. Su objetivo general es mapear la producción científica nacional sobre representación de información en Repositorios de Datos de Investigación en el ámbito de las Ciencias de la Información. Específicamente tiene como objetivo: investigar los principales enfoques temáticos de las producciones científicas recuperadas; caracterizar los tipos de representación de datos (descriptivos y/o temáticos) discutidos en la literatura; señalar tendencias y perspectivas de estudios relacionados con el tema. Utiliza como metodología la investigación bibliográfica, exploratoria y descriptiva con enfoque cualitativo, teniendo como corpus artículos de revistas indexados en la Base de Datos de Ciencias de la Información. Como resultado, se descubrió que los artículos se dividen en cuatro grandes ejes temáticos, a saber: presentación de reglas y principales estándares de metadatos en Repositorios de Datos de Investigación; investigación de integración de datos mediante el uso de estructuras de representación; reflexiones sobre la representación y recuperación de datos en dominios específicos; y evaluar la calidad de los metadatos en los depósitos de datos de investigación. En conclusión, parece que los aspectos relacionados con la representación de la información en los repositorios de datos de investigación en las Ciencias de la

Información brasileñas vienen ganando destaque en los últimos años, especialmente en lo que se refiere a la representación descriptiva de la información.

Palabras clave: Representación de la Información; Repositorios de datos de investigación; Buscar datos.

1 Introdução

A representação consiste em uma operação presente nas práticas sociais e em diferentes ambientes de construção e transfiguração da sociedade. As relações entre a humanidade e o conhecimento evidenciam a necessidade das pessoas em registrar, representar e organizar o conhecimento com vistas à recuperação, preservação, acesso, uso e apropriação da informação.

A dimensão multifacetada da representação desafia aqueles(as) que se propõem a estudá-la no contexto da Ciência da Informação, em especial na Organização e Representação da Informação e do Conhecimento, sob as perspectivas epistemológica e pragmática acerca do acesso às informações a partir dos diferentes dispositivos, plataformas e contextos, como por exemplo os Repositórios de Dados de Pesquisa.

Concorda-se com Felipe e Santos (2022), quando essas autoras afirmam que os dados passam a ser o destaque no fazer científico, sobretudo em ambientes colaborativos, em que o compartilhamento é essencial para o seu reuso. Esses dados devem ser representados a partir de um conjunto de operações que contemplam a descrição física (representação descritiva) e a representação do seu conteúdo (tratamento temático) para fins de sua recuperação, acesso, uso e reuso. Essa representação é realizada nos metadados do Repositório de Dados de Pesquisa, os quais assumem o papel de substituto dos dados, necessitando ser bem estruturados para atender à descrição, referência, significação, uso e reuso de dados de pesquisa (COYLE, 2005).

Com o advento de aplicações *web* e de ambientes como os repositórios de dados de pesquisa que facilitam as atividades de organização e representação de dados por meio da Folksonomia surge a descentralização no processo de representação, haja vista que quem representa os dados (no caso dos repositórios de dados de pesquisa), em linhas gerais, são os seus produtores, com diversos graus de subjetividade e envolvimento com as pesquisas desenvolvidas, tanto de conhecimento como por experiência ou sentimento. Tal fato pode apresentar implicações positivas e negativas no processo de representação e recuperação dos dados de pesquisa.

Nessa perspectiva, a Folksonomia é definida como o resultado do processo de etiquetagem livre (atribuição de tags, etiquetas, palavras-chave) por sujeitos (humanos ou robôs) em ambientes digitais colaborativos para a representação e recuperação de dados e objetos informacionais como áudios, textos,

imagens, vídeos, etc. (SANTOS; CORRÊA, 2018). Essa prática é realizada sem o auxílio de instrumentos de controle terminológico (como por exemplo tesouros, taxonomias, ontologias) e, embora tenha baixo grau de formalização semântica, representa discursos instituídos por sujeitos em diferentes dispositivos, plataformas e contextos.

Com isso, faz-se necessário depreender de que maneira as práticas de representação da informação em Repositórios de Dados de Pesquisa têm sido discutidas na literatura. Para tanto, surge o seguinte questionamento: qual o *status* das pesquisas sobre Representação da Informação em Repositórios de Dados de Pesquisa na Ciência da Informação brasileira?

Visando responder esse questionamento, o objetivo geral desta pesquisa consiste em mapear a produção científica nacional sobre as práticas de representação da informação em repositórios de dados de pesquisa no âmbito da Ciência da Informação do Brasil. Objetiva-se especificamente: investigar os principais enfoques temáticos das produções científicas recuperadas; caracterizar os tipos de representação de dados (descritiva e/ou temática) discutidas na literatura; apontar tendências e perspectivas de estudos relacionados à temática.

A relevância desta pesquisa decorre, em linhas gerais, da ausência de estudos que se dedicam a investigar como os(as) pesquisadores(as) da área de Ciência da Informação têm discutido as questões concernentes à representação de dados em infraestruturas destinadas ao seu armazenamento, descrição, preservação, acesso, uso e reuso. A representação de dados de pesquisa se configura como uma atividade imprescindível para o efetivo acesso do conhecimento científico por usuários e pesquisadores de diversas áreas do conhecimento em âmbito acadêmico, institucional e social, sobretudo em tempos de alta incidência de *fake news*.

Esta pesquisa se propõe a analisar o cenário nacional, uma vez que Souza et al (2022) em sua pesquisa já afirmam a inexistência de relatos de pesquisa e de experiências relacionadas à representação de dados de pesquisa no contexto da Biblioteconomia e da Ciência da Informação internacional, destacando a importância dos estudos da catalogação de dados de pesquisa. Com isso, este estudo pretende analisar o cenário nacional das produções científicas sobre a representação de dados de pesquisa não apenas na perspectiva da representação descritiva (catalogação), mas tentando identificar estudos cujo foco também seja as operações de representação temática de dados nesses ambientes.

Para dar continuidade a essas considerações, a seção a seguir discorre questões relativas aos dados de pesquisa.

2 Dados de pesquisa

Questões relacionadas à gestão e reuso de dados de pesquisa vêm ganhando destaque no âmbito científico, uma vez que os dados são insumos centrais no desenvolvimento das investigações. Segundo Borgman (2012) os dados são considerados como molas propulsoras para o desenvolvimento da pesquisa em qualquer área do conhecimento. Seu formato dependerá da sua origem, finalidade e área do conhecimento. Como exemplo, podemos citar uma Pedra que pode ser dado tanto para a História quanto para a Geologia. A Fotografia, por exemplo, pode ser dado para Antropologia, História, Medicina, Biodiversidade, entre outras disciplinas. A OECD (2007, p.13, tradução nossa) define dados de pesquisa como.

registros factuais (números, registros textuais, imagens e sons) usados como fontes primárias para pesquisa científica, e que são comumente aceitos na comunidade científica como necessários para validar resultados da pesquisa.

Nesse sentido, é possível inferir que os dados de pesquisa são fundamentais para legitimar a pesquisa, pois são coletados por meio de um método específico, para um fim específico e são validados por meio de suas análises. Bertin, Visoli e Druker (2017, p. 38) discorrem que “Dados de pesquisa, por sua vez – e também de modo simplificado –, são todo o tipo de registro produzido, compilado ou utilizado no decorrer da pesquisa”. Isto é, tudo aquilo que é gerado e empregado no âmbito da pesquisa são dados de pesquisa.

Por sua vez, Sales *et. al.* (2019, p.306) descrevem dos dados evidenciando a sua importância para a ciência “dados científicos são ativos informacionais imprescindíveis para o progresso da ciência e para a viabilização de novas descobertas que vão das ciências exatas à humanidades, arte e cultura”. Dessa forma, os dados tornam-se protagonistas dentro da ciência, possibilitando o surgimento de novos conhecimentos.

Os dados de pesquisa são importantes para o avanço do conhecimento científico, com isso a demanda por uma abertura de dados com vistas a sua reutilização vem crescendo e, conseqüentemente, surgem os ambientes de publicação de dados. Um dos ambientes que merecem a atenção do pesquisador é o repositório de dados de pesquisa, estrutura voltada para o compartilhamento, preservação e gestão desses tipos de dados.

2.1 Repositórios de dados de pesquisa

As mudanças na ciência, alinhadas ao grande uso das tecnologias de informação favorecem o surgimento do movimento de Ciência aberta, cujo foco é facilitar o acesso à ciência para a cientistas e a sociedade. Na ciência aberta, prega-se que, se a pesquisa foi realizada com investimento público, logo essa deve ser aberta para quem quiser utilizá-la.

Para tal, existem aspectos que visam facilitar seu progresso. São enfoques desenvolvidos pela ciência aberta o acesso aberto às publicações científicas, educação aberta, elaboração de ferramentas e materiais científicos abertos, ciência cidadã, cadernos de pesquisa abertos e dados abertos de pesquisa.

Nesse contexto, a abertura dos dados faz com que a ciência avance sem necessariamente reproduzir os custos, pois a disponibilização desses dados garante e evita a duplicação de esforços.

Com relação à reprodução de pesquisas por meio dos dados abertos, Curty (2017, p.2) afirma que as tecnologias impulsionam essa ação, “ferramentas computacionais avançadas para compartilhamento e distribuição de dados estão pavimentando o caminho para uma melhor reprodutibilidade nas investigações científicas”. Ou seja, as ações de abertura dos dados, contribuem para a criação de ambientes colaborativos que favorecem a reutilização dos dados nas pesquisas. Nesse contexto, surgem os repositórios de dados de pesquisa.

^[1]O OpenAIRE (2018), define repositório de dados como “um arquivo digital que coleta e exibe conjuntos de dados e seus metadados”. Entende-se que, diferentemente dos outros produtos oriundos da pesquisa como artigos de periódicos, os dados precisam de informações complementares para seu entendimento.

Para Tartarotti, Dal'Evedove Fujita (2019), os repositórios de dados de pesquisa englobam os mais variados tipos de dados oriundos das pesquisas, organizados com vistas ao uso pela comunidade interessada. Assim, esses sistemas podem estar ligados a instituições - como Universidades e centros de pesquisa - ou grupos de pesquisa.

Monteiro *et al.* (2017) ressaltam as seguintes funções dos repositórios de dados: “[...] armazenar, representar, gerenciar, disseminar, disponibilizar, e preservar dados neles depositados”. Desse modo, o repositório torna-se uma importante ferramenta não só para a preservação dos dados, mas também para o seu acesso, uso e reuso.

Cabe citar, que não existe apenas os repositórios de dados de pesquisa como ferramentas de publicação dos dados, tem-se ainda os artigos de dados e os periódicos de dados, porém em comparação com essas duas ferramentas, os repositórios estão mais avançados e

existem em larga escala, englobando os dados de pesquisa dos mais variados domínios, além de oferecer inúmeros benefícios para a comunidade científica.

Para Sayão e Sales (2016), além de proverem o reuso dos dados, os repositórios apresentam inúmeros benefícios, dentre eles: a visibilidade dos dados, crédito ao autor dos dados, preservação digital, memória científica, transparência, entre outros. Portanto, os repositórios tornam-se espaços nos quais o produtor do dado pode ganhar destaque, além de receber o devido crédito por seu trabalho. Também proporcionam a preservação dos dados por um longo período, tornando o acesso valioso, favorecendo a conservação da memória científica.

Para que os dados de pesquisa possam ser recuperados, acessados, usados e reutilizados, faz-se necessária a sua representação em sistemas como os repositórios. Nesse sentido, a seguir são discutidas questões concernentes à representação da informação.

3 Representação da Informação

Com o advento das tecnologias digitais e as novas possibilidades de organização e tratamento de recursos informacionais, a organização e a representação da informação se ampliam cada vez mais não apenas como operações, mas como campo que contempla estudos epistemológicos e pragmáticos com foco no acesso, uso e apropriação do conhecimento produzido e registrado.

Enquanto processo cognitivo a representação da informação possui um alto nível de complexidade, uma vez que o ato de representar está atrelado ao conceito de substituição, ou seja, de “estar no lugar de outra coisa” (ALVARENGA, 2003, p. 18). Para a autora, quando se representa algo é criada uma relação entre o que se apresenta e o signo ou símbolo (seja sonoro, imagético, escrito, etc.) em um ato de substituição utilizando elementos simbólicos como palavras, números, esquemas, figuras, etc. Desse modo, um dado ou recurso informacional é representado por meio da linguagem para fins de sua organização em um determinado contexto.

A representação da informação contempla processos, métodos e técnicas voltados para a descrição física (representação descritiva) e de conteúdo (representação temática) dos recursos. A representação descritiva tem como foco os dados físicos extraídos dos itens informacionais, contemplando aspectos mais objetivos que representam as suas estruturas por meio de atividades como a catalogação. Por outro lado, na representação temática o seu foco consiste no conteúdo (assunto) desses itens, caracterizando-o sob o ponto de vista do seu assunto por meio de atividades como a indexação (DIAS; NAVES, 2007).

Para Gordana e Dragan (2017), a catalogação contribui para a transparência da pesquisa, favorecendo a descoberta de dados, a reprodutibilidade dos mesmos e a averiguação da pesquisa publicada. Essa operação é realizada nos metadados de sistemas de informação.

Os metadados são essenciais para a representação e recuperação dos dados de pesquisa. De acordo com Alves (2010, p.47):

Metadados são atributos que representam uma entidade (objeto do mundo real) em um sistema de informação. Em outras palavras, são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades; são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação.

Assim, os metadados são elementos importantes para identificar e descrever o conteúdo de um dado ou recurso informacional, além de representá-lo e localizá-lo, viabilizando a sua busca e recuperação na *web*.

Para uma melhor representação e recuperação de um dado ou recurso informacional, são necessárias estruturas com formatos, padrões e normas técnicas em sistemas de informação, essas estruturas são denominadas como padrões de metadados.

De acordo com Alves (2010), os padrões de metadados viabilizam a troca de informações entre sistemas que utilizam os mesmos padrões ou padrões diferentes. Essa troca é denominada de interoperabilidade e possibilita que diferentes instituições com diferentes *softwares* possam comutar os seus dados, caso as informações estejam bem estruturadas e representadas nesses ambientes por meio da catalogação.

Segundo Souza et al. (2022), no contexto da Ciência Aberta, a catalogação de *datasets* (coleções de dados estruturados) deve ser implementada em consonância com os princípios FAIR (*Findable, Accessible, Interoperable, Reusable*), objetivando que os metadados e os dados sejam representados com padrão de qualidade para fins de sua recuperação, acesso, uso e reuso.

Henning et al. (2019, p. 178-179) apresenta um quadro (Anexo A) com os princípios FAIR e seus elementos. Esses elementos contemplam regras que envolvem os metadados visando a encontrabilidade, acessibilidade, interoperabilidade e reusabilidade. No quadro apresentado por esses autores também é possível visualizar elementos dos princípios FAIR que estão diretamente associados à atividade de representação de dados e que contribuem para a qualidade do seu tratamento.

De acordo com Souza et al (2022), em 2018 a *European Commission Expert Group on FAIR Data* publicou o documento intitulado “*Turning FAIR reality*” como um instrumento que visa esclarecer esses

princípios e contribuir para a sua aplicabilidade em diferentes contextos. A partir disso, outras iniciativas nessa seara foram desenvolvidas e outras propostas se encontram em fase de desenvolvimento.

Desse modo, a representação de dados de pesquisa com o auxílio de diretrizes, regras ou padrões específicos faz-se necessária para o acesso, uso e reuso desses dados.

Diante do exposto, torna-se importante investigar o *status* das pesquisas acerca do processo de representação de dados em repositórios de dados de pesquisa.

A seguir são apresentados os procedimentos metodológicos empreendidos neste estudo.

4 Procedimentos metodológicos

Visando atender o seu objetivo geral, foram realizadas as pesquisas bibliográfica, exploratória e descritiva com abordagem quali-quantitativa para a recuperação, análise e mapeamento da produção científica nacional sobre representação da informação em Repositórios de Dados de Pesquisa.

As buscas dessas produções científicas foram realizadas na Base de Dados em Ciência da Informação (BRAPCI) no mês de março de 2023. Essa fonte de informação foi selecionada nesta pesquisa devido à possibilidade de acesso aos artigos publicados nos principais periódicos e eventos científicos nacionais na área de Ciência da Informação do Brasil em diferentes modalidades: artigos originais, artigos de revisão, relatos de pesquisa, por exemplo.

Para a recuperação das produções científicas sobre o tema não foi delimitado o recorte temporal, com vistas a recuperar todas as pesquisas publicadas acerca da temática. Para tanto, foram empregadas as seguintes restrições lógicas no processo de busca na BRAPCI: “Representação da informação AND Repositório de Dados” e “Representação da informação AND Dados de Pesquisa”.

As combinações com o operador booleano “AND”, bem como a quantidade de termos selecionados, foram aplicados na base de dados com o objetivo de restringir logicamente as buscas, possibilitando a recuperação de produções científicas que discorrem especificamente sobre a Representação da Informação em Repositórios de Dados de Pesquisa.

Após a realização das buscas verificou-se que foram recuperadas 49 (quarenta e nove) produções científicas, porém após a realização de uma leitura técnica dos trabalhos recuperados e a exclusão de trabalhos duplicados ou que não possuíam articulação direta com o enfoque temático desta pesquisa, o *corpus* final para

a realização das análises deste estudo foi constituído de oito artigos de periódicos.

É importante ressaltar que, embora uma quantidade significativa de produções científicas recuperadas apresentassem pelo menos um dos termos pesquisados em seu conteúdo, observou-se que do total citado anteriormente, seis trabalhos eram duplicados e 35 (trinta e cinco) artigos científicos não se articulavam diretamente com esta investigação, pois seu objeto de estudo estava voltado para ambientes como Repositórios Institucionais, bases de dados governamentais, catálogos eletrônicos, bibliotecas digitais, além de discutir especificamente questões voltadas para dados governamentais abertos, curadoria de dados, *e-Science* e *Data Science*.

Ademais, na maioria das produções científicas recuperadas, um dos termos presentes nas expressões de busca (representação da informação, repositório de dados, dados de pesquisa) aparecia apenas uma vez no texto, sem fazer a conexão entre a Representação da Informação e os Repositórios de Dados de Pesquisa.

Como procedimento para a realização da análise dessas pesquisas foi utilizada a análise de conteúdo por meio de uma leitura técnica dos trabalhos recuperados, possibilitando tornar compreensíveis os principais aspectos enfatizados nos estudos que interligam ambas as temáticas. Além disso, foi possível caracterizar os tipos de representação de dados (descritiva e/ou temática) discutidas na literatura; apontar tendências e perspectivas de estudos relacionados à temática.

A leitura técnica das produções científicas foi baseada na norma da Associação Brasileira de Normas Técnicas (ABNT) NBR 12676, que trata especificamente sobre métodos de análise de documentos e determinação de seus assuntos. Para tanto, foram considerados os seguintes elementos informativos nas produções científicas: título e subtítulo; resumo; palavras-chave; introdução; ilustrações; diagramas; tabelas e explicações; resultados; considerações finais e referências.

A partir dessa lógica, a seguir são apresentados os principais resultados da pesquisa.

5 Análise e discussão dos resultados

Após as análises realizadas, verificou-se que as oito produções científicas que se constituíram como *corpus* final desta pesquisa estão distribuídas no intervalo entre 2017 e 2022, sendo 2022 o ano que apresentou maior quantidade de publicações (três trabalhos) sobre representação de dados de pesquisa, conforme é possível visualizar no Quadro 1 (Apêndice A).

No Quadro supracitado é possível verificar uma quantidade significativa de produções científicas publicadas

sobre a temática no último ano, o que evidencia uma tendência de maior preocupação dos pesquisadores da área de Ciência da Informação acerca da descrição e estruturação de dados de pesquisa, mediante a análise, a identificação e/ou a proposição de padrões de metadados para a representação, recuperação e integração de dados.

De acordo com a análise de conteúdo realizada nas publicações listadas, emergiram quatro enfoques de estudos (Quadro 2):

Quadro 2. *Enfoques temáticos das produções científicas analisadas*

<i>Enfoques temáticos</i>	<i>Autores(as)</i>
Apresentação de regras e principais padrões de metadados em Repositórios de Dados de Pesquisa	Pinto et al (2017) Sanchez; Silva e Vechiato (2018) Souza et al (2022)
Integração de dados de pesquisa mediante a utilização de estruturas de representação	Tadini e Santarém Segundo (2019) Torino; Coneglian; Vidotti (2020)
Reflexões acerca da representação e recuperação de dados em domínios específicos	Reyes; Sant'Ana; Ramalho (2017) Viola e Sales (2022)
Avaliação da qualidade de metadados em Repositórios de Dados de Pesquisa	Felipe e Santos (2022)

Fonte: Elaborada pelas autoras.

Conforme é possível visualizar no Quadro anterior, a maioria das produções científicas analisadas têm como foco apresentar “regras e principais padrões de metadados em Repositórios de Dados de Pesquisa”, contemplando orientações para a representação descritiva e temática; instruções para o preenchimento de formulários de catalogação; orientações para uso de vocabulários específicos com base nos princípios FAIR; recomendações para a revisão na catalogação. Tais contribuições objetivam padronizar o processo de representação da informação, bem como contribuir para a geração de metadados qualificados. As pesquisas de Pinto et al (2017); Sanchez, Silva e Vechiato (2018) e Souza et al (2022) contemplam questões relativas a esse enfoque temático.

Os principais padrões de metadados citados pelos autores apresentados anteriormente foram: *Dublin Core, Data Documentation Initiative, ISO 19115 – Geographic Information – Metadata*, seguidos de *Life Science Metadata, ECHO Metadata Standard, Jakarta Lucene Search and Retrieval Protocol, ANZLIC Metadata*.

Os padrões de metadados apresentam um conjunto de elementos que auxiliam na veracidade das informações relativas aos dados a serem descritos em um sistema. De acordo com Sanchez; Silva e Vechiato (2018) um repositório de dados de pesquisa pode implementar variados tipos de padrões de metadados visando contemplar diferentes domínios do conhecimento.

Outro enfoque temático identificado na literatura foi a “*integração de dados de pesquisa mediante a utilização de estruturas de representação*”. Esses estudos buscam discutir as possibilidades de integração de dados e metadados de diferentes tipos de repositórios, conforme propõem as técnicas de *Linked Data*, contribuindo para o aumento do impacto das pesquisas científicas. Iniciativas como essa demandam que os profissionais da informação envolvidos nas práticas de representação da informação em repositórios de dados de pesquisa possuam conhecimentos relacionados à arquitetura da informação, padrões e protocolos de comunicação, interoperabilidade e tecnologias semânticas. Esses e outros aspectos são destacados nos estudos de Tadini e Santarém Segundo (2019); Torino, Coneglian e Vidotti.

Também foram identificados na literatura nacional estudos cujo foco consiste em apresentar “*reflexões acerca da representação e recuperação de dados em domínios específicos*” mediante uma análise da literatura, abordando noções como relevância, importância da curadoria digital e reuso de dados. Essas reflexões podem ser encontradas nas pesquisas de Reyes, Santana e Ramalho (2017); Viola e Sales (2022).

Há ainda na literatura analisada a preocupação de pesquisadores no que concerne à “*avaliação de metadados em Repositórios de Dados de Pesquisa*”. Estudos com esse enfoque temático contribuem para apontar eventuais lacunas existentes na representação dos dados nesses ambientes, além de propor melhorias para a descrição de conjunto de dados. A pesquisa das autoras Felipe e Santos (2022) apresenta análises e contribuições relacionadas a essa perspectiva.

Diante do exposto, foram observados quatro enfoques de estudos na literatura nacional sobre representação da informação em repositórios de dados de pesquisa. Assim, é possível constatar a riqueza de possibilidades de investigações oferecidas pelo tema, visando incluir diferentes perspectivas futuras no campo da Organização e Representação da Informação e do Conhecimento.

Considerando que esta investigação tem como foco a representação da informação em repositórios de dados de pesquisa, foi possível identificar ainda quais os repositórios analisados ou citados nas produções científicas recuperadas, são eles: Repositório de Dados de Pesquisa da Embrapa (Redape), Repositório de Dados

de Pesquisas em Biodiversidade (PPBio), Repositório de Dados Espaciais.

Além disso, é importante enfatizar que o diretório Re3Data consiste em uma ferramenta de acesso aberto que fornece uma visão panorâmica dos repositórios nacionais e internacionais em diferentes áreas do conhecimento.

Essa fonte de informação também pode ser consultada por pesquisadores e profissionais da informação com vistas a compreender ou analisar diferentes sistemas e vocabulários na perspectiva da organização e representação da informação e do conhecimento.

Por fim, foi possível identificar que os estudos relacionados à representação da informação em repositórios de dados de pesquisa centram-se, sobretudo, no campo da representação descritiva (catalogação) de dados a partir de padrões de metadados específicos. Embora esses ambientes possuam regras e padrões de metadados, o processo de catalogação de dados tem sido realizado, em linhas gerais, pelo autor do *dataset* ou outra pessoa de sua responsabilidade mediante o preenchimento dos elementos de metadados recomendados pelo sistema, caracterizando assim como um processo colaborativo que utiliza a Folksonomia.

6 Conclusão

Com o estabelecimento da cultura de acesso aberto aos dados, surge a necessidade da criação de ambientes que favoreçam a disponibilidade e compartilhamento de dados de pesquisa, visando o reuso por parte dos pares interessados. Logo, tem-se os repositórios de dados de pesquisa, ambientes digitais desenvolvidos para armazenar, disseminar e preservar os dados de pesquisa. Como ambientes de compartilhamento de informação, necessitam de uma organização e com isso fazem uso da representação da informação.

Após investigação no cenário brasileiro, fica evidente que questões ligadas à representação da informação nos repositórios de dados de pesquisa na Ciência da Informação brasileira vêm ganhando destaque nos últimos anos. Isso demonstra que a gestão dos repositórios de dados não cabe somente aos profissionais de Tecnologia da Informação e Comunicação, mas que deve ser desenvolvida em consonância com o profissional Bibliotecário, com domínio de processos, produtos, instrumentos e serviços da organização e representação da informação e do conhecimento para fins de recuperação da informação.

Observou-se ainda que os aspectos relativos à representação descritiva estão mais centralizados nas pesquisas sobre a representação da informação em repositório de dados. Além disso, os enfoques destacados na produção recuperada, comprovam que a literatura vem se desenvolvendo de acordo com os princípios FAIR, os

quais apresentam ênfase para o uso dos metadados visando a representação, recuperação e reuso da informação contida nos repositórios de dados de pesquisa.

Para pesquisas futuras, tem-se como objetivo: ampliar o escopo da análise (incluindo termos relacionados à representação da informação e contemplando a literatura internacional sobre o tema); aprofundar nas questões relacionadas à representação temática nesses repositórios, incluindo sugestões para o uso e a aplicação da Folksonomia Assistida com vistas a auxiliar o autor do *dataset* ou outra pessoa de sua responsabilidade no processo de indexação com vistas a produção significativa de metadados semânticos nesse contexto.

Notas

[1] Documento online, não paginado.

Referências

- ALVARENGA, Lúcia. Representação do Conhecimento na perspectiva da Ciência da Informação em tempo e espaços digitais. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v.1, n.15, p.18-40, 2003.
- ALVES, R. C. V. **Metadados como elementos do processo de catalogação**. 2010. 134 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- BERTIN, Patrícia Rocha Bello; VISOLI, Marcos Cezar; DRUCKER, Debora Pignatari. A gestão de dados de pesquisa no contexto da e-science: benefícios, desafios e oportunidades para organizações de P&D. **Ponto de Acesso**, Salvador, v.11, n.2, p. 34-48, ago. 2017. Disponível em: <https://portalseer.ufba.br/index.php/revistaici/article/view/21449>. Acesso em 18 fev. 2023
- BORGMAN, Christine L. The conundrum of sharing research data. **Journal of the Association for Information Science and Technology**, v. 63, n.6, p.10-59-1078, jun. 2012. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1002/asi.22634>. Acesso em: 02 ago. 2020.
- CURTY, Renata Gonçalves. O paradigma da publicação de dados e suas diferentes abordagens. In: XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB, 18, 2017, Marília. **Anais [...]**. Marília: UNESP, 2017.p. 1-20. Disponível em: <https://brapci.inf.br/index.php/res/v/105144>. Acesso em: 19 fev. 2023.
- DIAS, E.W.; NAVES, M.M.L. **Análise de assunto**: teoria e prática. Brasília: Thesaurus, 2007.
- FELIPE, C. B. M.; SANTOS, R. F. dos. Avaliação de metadados em Repositórios de Dados de Pesquisa sobre biodiversidade. **Em Questão**, Porto Alegre, v. 8, n.3, p. 1-19, jul./set. 2022.
- GORDANA, R.; DRAGAN, I. Cataloguing dataset inLibrary Information Systems using the MARC 21 format. In: INTERNATIONAL CONFERENCE ON INFORMATION

- SOCIETY AND TECHNOLOGY, 7, 2017, Kopaonik. **Proceedings** [...] Belgrade: Society for Information Systems and Computer Networks Belgrade, 2017. P.395-399.
- HENNING, P. C. Desmistificando os princípios FAIR: conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados FAIR. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, João Pessoa, v.14, n.5, p.175-192, 2019.
- Monteiro, Elizabete Cristina de Souza de Aguiar et al. Repositório de Dados Científicos: Aspectos sobre Privacidade de Dados. *In: I WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA*, 1, 2017, Florianópolis, **Anais** [...]. Florianópolis, UFSC, 2017.
- OPEN AIRE – Open Access Infrastructure for Research in Europe. FAQ (What are repositorie?). 2018. Disponível em: <https://www.openaire.eu/where-can-i-read-more-about-fp7>. Acesso em: 02 mar. 2023.
- ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO – OCDE. Principles and guidelines for access to research data from public data. 2007. Disponível em: <https://www.oecd.org/sti/inno/38500813.pdf>. Acesso em: 20 fev. 2022.
- PINTO, et al. Descrição dos dados da pesquisa geoespacial: a experiência da Embrapa. **Reciis: Rev. Eletron. Comun. Inf. Inov. Saúde**, [S.l.], v.11, n.1, p.1-16, 2017.
- REYES, J. T. C.; SANT'ANA, R. C. G.; RAMALHO, R. A. S. A produção científica sobre representação de dados dentro da área temática da agricultura. **Informação e Informação**, Londrina, v. 22, n. 2, p. 467-480, 2017.
- SALES, Luana Farias *et al.* Competências dos bibliotecários na gestão dos dados de pesquisa. **Informação & Informação**, Brasília, DF, v.48 n.3 (Supl.), p.303-313, set./dez. 2019. Disponível em: <https://www.arca.fiocruz.br/handle/icict/43074>. Acesso em: 15 mar. 2023
- SANCHEZ, F. A.; SILVA, N. B. P.; VECHIATO, F. L. Padrões de metadados para a representação e organização da informação em repositórios de dados de pesquisa. **Informação & Tecnologia**, [S. l.], v. 5, n.1, p. 37-51, 2018.
- SANTOS, R. F. dos; CORRÊA, R. F. Análise das definições de Folksonomia: em busca de uma síntese. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 23, n.2, p. 1-32, 2018.
- SAYÃO, Luís Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, Londrina, v. 21, n. 2, p. 90 – 115, maio/ago., 2016.
- SOUZA, I. F. S. et al. Catalogação de Datasets ômicos no Repositório de Dados de Pesquisa da Embrapa. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 18, p. 01-28, 2022.
- TADINI, A. V. W. SANTARÉM SEGUNDO, J. E. Modelos semânticos para dados bibliográficos de publicações científicas disponibilizados como Linked Data. **Informação em Pauta**, [S.l.], v. 4, n.2, p. 2019.
- TARTAROTTI, R. C. D.; DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Biblioteconomia de dados em repositórios de pesquisa: perspectivas para a atuação bibliotecária. **Informação & Informação**, v. 24, n. 3, p. 207-226, 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/134221>. Acesso em: 03 mar. 2023.
- TORINO, E.; CONEGLIAN, C. S.; VIDOTTI, S. A. B. G. Estruturas de representação para reuso de dados no contexto da ecologia de pesquisa: CRIS institucional. **Informação & Informação**, Londrina, v. 5, n.3, p. 1-27, jul/set. 2020.
- VIOLA, C. M. M.; SALES, L. F. A “alquimia” da relevância, o reuso dos dados e seu impacto no desenvolvimento sustentável: contribuições à agenda 2030. **Encontros Bibli: Rev. Eletr. Bibl. E Ci.**, [S. l.], v. 27, n.1, p. 1-21, 2022.

Apêndice A

Quadro 1. Produções científicas analisadas

<i>Título do trabalho</i>	<i>Autores e ano de publicação</i>
Descrição dos dados da pesquisa geoespacial: a experiência da Embrapa	Pinto et al (2017)
A produção científica sobre representação de dados dentro da área temática da agricultura	Reyes; Sant'Ana; Ramalho (2017)
Padrões de metadados para a organização da informação em Repositórios de dados de pesquisa	Sanchez; Silva; Vechiato (2018)
Modelos semânticos para dados bibliográficos de publicações científicas disponibilizados como Linked Data	Tadini e Santarém Segundo (2019)
Estruturas de representação para reuso de dados no contexto da ecologia de pesquisa: CRIS institucional	Torino; Coneglian; Vidotti (2020)
Avaliação de metadados em repositórios de dados de pesquisa sobre biodiversidade	Felipe e Santos (2022)
A “alquimia” da relevância, o reuso dos dados e o seu impacto no desenvolvimento sustentável: contribuições à agenda 2023	Viola e Sales (2022)
Catálogo de datasets ômicos no Repositório de Dados de Pesquisa da Embrapa	Souza et al (2022)

Fonte: Elaborado pelas autoras.

Anexo A

Quadro 2. *Princípios FAIR e seus elementos*

<i>Princípio F - Findable (Encontrabilidade) e seus elementos</i>
F1. Os metadados e dados devem ter identificadores globais, persistentes e identificáveis
F2. Os dados devem ser descritos com metadados enriquecidos
F3. Os metadados devem incluir claramente e explicitamente os identificadores dos dados que descrevem
<i>Princípio A - Accessible (Acessibilidade) e seus elementos</i>
A1. Metadados e dados devem ser recuperáveis pelos seus identificadores usando protocolo de comunicação padronizado
A1.1 O protocolo deve ser aberto, gratuito e universalmente implementável
A1.2 O protocolo deve permitir procedimentos de autenticação e autorização, quando necessário
A2. Metadados devem ser acessíveis, mesmo quando os dados não estão mais disponíveis
<i>Princípio I - Interoperable (Interoperabilidade) e seus elementos</i>
I1. Metadados e dados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento
I2. Metadados e dados devem usar vocabulários que assegurem os princípios FAIR
I3. Metadados e dados devem incluir referências qualificadas para outros metadados e dados
<i>Princípio R - Reusable (Reusabilidade) e seus elementos</i>
R1. Metadados e dados são descritos com uma pluralidade de atributos precisos e relevantes
R1.1. Metadados e dados devem ser disponibilizados com licenças de uso claras e acessíveis
R.1.2. Metadados e dados devem estar associados à sua proveniência
R.1.3. Metadados e dados devem estar alinhados com padrões relevantes ao seu domínio.

Fonte: Henning et al (2019, p. 178-179).

Knowledge Organization no processo de FAIRificação de *datasets*: estruturando a semântica e interligando as notícias do banco de dados de periódicos musicais oitocentistas

Knowledge Organization on datasets FAIRification process: data semantics and linked data in music in 19th century periodicals database system

Cláudio José Silva Ribeiro (1) ¹, Cintia Sales de Sousa (2), Ciro Ítalo Tertulino (3), Ilvana Sousa do Amaral (4), Marlus José Soares dos Santos (5), Suelen de Mendonça Soares Cóquero (6), Martha Tupinambá de Ulhôa (7)

(1) Prof. Dr. - Unirio/DPTD e PPGb, Av. Pasteur, 458, Rio de Janeiro, claudio.ribeiro@unirio.br (2) Unirio/PPGB, cintiasales10@gmail.com (3) ciro.italo@hotmail.com (4) ilvana.amaral@edu.unirio.br (5) marlujss@gmail.com (6) suelencoquero@id.uff.br (7) Prof. Dra. - Unirio/PPGM, mulhoa@unirio.br

Resumo

A ciência orientada aos dados de pesquisa está em processo de consolidação, envolvendo a atuação de agências de fomento, instituições de pesquisa e universidades. Este artigo tem como objetivo demonstrar a aplicação prática do processo de FAIRificação na base Música em Periódicos Oitocentistas, a partir do conceito de Ciência Aberta e da noção de interoperabilidade semântica. A presente pesquisa é um estudo de caso de natureza aplicada e exploratória em relação ao seu objetivo, utilizando, como procedimento técnico-metodológico, a revisão de literatura na Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapci), além de recorte e seleção de *datasets* na fase de pré-FAIRificação. Como resultado, foi possível propor um modelo semântico presente no objeto de estudo, baseado em vocabulários que descrevem recursos eletrônicos. Conclui-se que o processo de transformação de dados de pesquisa no sentido de torná-los FAIR ainda é uma iniciativa difícil de ser conduzida. Além disso, foi percebido que existem pontos de convergência entre o processo de FAIRificação e a preparação para uso dos dados ligados.

Palavras-chave: Princípios FAIR. FAIRificação. Interoperabilidade Semântica. Dados de Pesquisa. Ciência Aberta. Música. Oitocentistas.

Abstract

Data oriented-science is in a consolidation process that involves research funding agencies, research institutions and universities. This article aims to demonstrate the application of the FAIRification process in the Music in Nineteenth Century Periodicals database system, taking into account the concept of Open Science and the notion of semantic interoperability. This is an applied case study research with an exploratory relation to its objective, using literature reviews of the Reference Database of Journal Articles in Information Science (Brapci) as a technical-methodological procedure, as well as clipping and selection of datasets in the pre-FAIRification phase. As a result, it was possible to propose a semantic model for the dataset selected as object of study, using some vocabularies that describe electronic resources. It is concluded that, in an effort to turn these informations into FAIR data, this research data transformation process is still hard. In addition, it has been perceived that there are some points of convergence between the FAIRification processes and the processes aiming the preparation to use linked data.

Keywords: Fair Principles. FAIRification process. Semantic Interoperability. Research data. Open Science. Music. Nineteenth Century Periodicals.

Resumen

Las dificultades en el proceso de implantación del uso compartido de los datos, sobre todo con la práctica de los principios FAIR, permean diversas áreas del conocimiento. La ciencia orientada a datos de investigación se encuentra en proceso de consolidación, incluyendo la actuación de agencias de fomento, instituciones de investigación y universidades. Este informe se basa en un estudio sobre el intercambio, el uso y la reutilización de datos, vuelto hacia el área de la Musicología. Tiene por objetivo demostrar en la práctica la aplicación de los principios FAIR sobre la base Música en Periódicos Oitocentistas, bajo el concepto de Ciencia Abierta y so la noción de interoperabilidad semántica. La presente investigación es un estudio de caso caracterizado por su naturaleza aplicada y exploratoria con respecto a su objetivo, utilizando como procedimiento técnico y metodológico la revisión de la literatura en la Base de Datos Referencial de Artículos de Periódicos en Ciencia de la Información (BRAPCI - brapci.inf.br). Está caracterizado por su naturaleza cualitativa para el análisis de datos, con la exploración bibliográfica y documental, en sitios del área de Musicología y en el buscador Google Académico para recorte y de la selección de conjuntos de datos en el nivel de preproceso FAIR. Como consecuencia, se pudo proponer un modelo semántico presente en el objeto de investigación, basándose en vocabularios que describen recursos

electrónicos. Se deduce que, para volverlos FAIR, el proceso de transformación de datos de investigación todavía es una iniciativa difícil de conducirse. Además de eso, se percibieron puntos de convergencia entre los procesos de FAIRificación y la preparación para el uso de datos relacionados.

Palabras clave: Principios FAIR. FAIRificación. Interoperabilidad Semántica. Datos de investigación. Ciencia Abierta. Base musical de publicaciones periódicas del siglo XIX.

1 Introdução

Vivemos numa conjuntura em que as práticas do fazer científico estão em constante mudança gerando novas formas de produção do conhecimento, contribuindo para o progresso da ciência. Este cenário de transformações resultou em um aumento do quantitativo de dados de pesquisa e no interesse do compartilhamento desses dados de maneira que seja possível a sua reutilização por outros pesquisadores.

As diversas possibilidades de acesso aos dados de pesquisa têm impulsionado a concepção de políticas e práticas visando à sua abertura e compartilhamento, intensificadas pelo desenvolvimento e a emergência de novas tecnologias da informação e comunicação. Neste contexto, são ampliadas as possibilidades de reuso dos dados em novas pesquisas e de reprodutibilidade de experimentos para verificação e validação da qualidade dos resultados da pesquisa científica (JORGE; ALBAGLI, 2018).

A ciência orientada aos dados de pesquisa está em processo de consolidação, envolvendo a atuação de agências de fomento, instituições de pesquisa e universidades. No Brasil, há avanços nesta prática, como o Plano de Dados Abertos 2017-2018, promovido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), a exigência de um Plano de Gestão dos Dados pela Fundação de Apoio à Pesquisa do Estado de São Paulo (Fapesp) na submissão de projetos de pesquisa e a implantação de repositórios de dados de pesquisa, como o CarpeDIEN (IEN), o Redape (Embrapa), o Arca Dados (Fiocruz), o PPBiodata (INPA) e o Ibict Dataverse.

O contexto científico atual tem resultado em uma maior valorização da disponibilidade de dados científicos, em que a transparência e a reprodutibilidade em pesquisas são aspectos preponderantes (CURTY; AVENTURIER, 2017). No entanto, observa-se um debate entre os pesquisadores sobre uma “crise da ciência” ou “crise da reprodutibilidade”, com destaque em questões que abarcam problemas no reuso dos objetos, o que acarreta dificuldade em reproduzir as pesquisas científicas.

Em um estudo recente, Fanelli (2018) apresenta um levantamento feito na Base de Dados *Web of Science* sobre a ocorrência, nos artigos indexados, de termos relacionados com a crise de reprodutibilidade. Os resultados mostraram um crescimento no número de artigos, o que endossa que a crise aumentou nos últimos

anos. Outras pesquisas também apontaram que existe uma dificuldade em reproduzir experimentos pelos pesquisadores, como relatado nos trabalhos de Baker (2016) e Nüst *et al.* (2018).

A fim de facilitar o compartilhamento, o uso e o reuso dos dados de pesquisa, os princípios FAIR surgem com a intenção de possibilitar que os dados sejam mais facilmente localizáveis, acessíveis, interoperáveis e reusáveis, oferecendo recursos para o seu gerenciamento. Baseados em um conjunto de princípios, visam à maior utilidade dos dados, de forma a ampliar sua reutilização por seres humanos e máquinas e colaborar na intermediação entre o gerador dos dados e o interessado no uso dos mesmos (KALINAUSKAITĖ, 2017).

Este artigo tem como objetivo demonstrar a aplicação prática do processo de FAIRificação na base Música em Periódicos Oitocentistas, a partir do conceito de Ciência Aberta e da noção de interoperabilidade semântica. Além disso, apresenta os princípios FAIR como solução para os problemas relacionados aos dados de pesquisa, com a finalidade de estabelecer melhorias na gestão, compartilhamento e reuso desses dados.

2 Open Science e o compartilhamento de dados

As mudanças nos modelos de produção e comunicação do conhecimento têm gerado transformações nas atividades científicas, constituindo uma nova ciência colaborativa e que envolve a sociedade nas suas práticas, democratizando o acesso e o uso das informações científicas e priorizando o compartilhamento e o reuso dessas informações. Neste contexto de expansão das fronteiras científicas e de ampliação da participação social, surge a Ciência Aberta, que é um conceito baseado em um conjunto de práticas e iniciativas científicas que se desenvolvem de modo aberto, representando um novo paradigma na comunidade científica.

A Ciência Aberta vai além dos moldes de pesquisa tradicional e do acesso às publicações e aos dados de pesquisa. Compreende que todo o processo de pesquisa em si seja aberto, com inclusão de seus métodos de análise e códigos de programas, ferramentas, entre outros materiais, de modo a contribuir para a colaboração, o compartilhamento e a reutilização. Dessa forma, expressa o desenvolvimento de uma ciência que abarca várias práticas e ações, como: acesso aberto,

dados abertos, ciência cidadã, revisão por pares aberta, códigos abertos, cadernos de laboratório abertos, recursos educacionais abertos e redes sociais científicas (PACKER; SANTOS, 2019; PORTUGAL, 2016).

Mundialmente, diversas instituições e pesquisadores têm se voltado ao fomento do acesso aberto e da abertura dos dados gerados pelas pesquisas científicas. Algumas ações proporcionaram esse alcance global, colaborando para a conscientização e a adoção desta nova filosofia. A *Open Archives Initiative* (OAI) surgiu como uma das primeiras iniciativas sobre acesso aberto e originou importantes declarações, como a Declaração de Budapeste, a Declaração de Bethesda e a Declaração de Berlim (SILVA; ALCARÁ, 2009).

Embora o Brasil ainda não possua uma política pública orientada para a Ciência Aberta, podemos citar algumas iniciativas em direção ao fortalecimento da abertura e ao compartilhamento dos dados governamentais, na promoção do Governo Aberto e do Acesso Aberto. Conhecida como a Lei da Transparência ou Lei de Acesso à Informação (LAI), a Lei 12.527/2011 assegura o direito fundamental de acesso à informação de maneira transparente, clara e em linguagem de fácil compreensão, preconizando que a publicidade dos dados e das informações é a regra e o sigilo é a exceção (SANTOS; ALMEIDA; HENNING, 2017).

Também, em 2011, foi lançado o Plano de Ação Nacional sobre Governo Aberto, com diretrizes de abertura de dados para promover a transparência e o acesso às informações públicas. Nesse mesmo ano, o Brasil passou a ser membro do *Open Government Partnership* (OGP), juntamente com outros 70 países, o qual visa incentivar a transparência dos governos e as práticas de compartilhamento de dados gerados pela administração pública. A filiação do Brasil ao OGP possibilitou o amadurecimento dos processos para disseminação, compartilhamento e reúso das informações produzidas na esfera governamental (SANTOS; ALMEIDA; HENNING, 2017; PACKER; SANTOS, 2019; RIBEIRO; VELHO, 2021).

Posteriormente, em 2012, a Infraestrutura Nacional de Dados Abertos (INDA) foi criada, com a finalidade de atender às questões estabelecidas na LAI e no Plano de Ação Nacional sobre Governo Aberto 2018-2020 (SALES *et al.*, 2020). A INDA baseia-se em um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos, em conformidade com o disposto no Projeto de Interoperabilidade no Governo – ePING (BRASIL, 2019). Amparado pelas ações relacionadas ao INDA, o Portal de Dados Abertos Brasileiros surgiu como uma iniciativa da parceria do Governo Aberto,

disponibilizando os conjuntos de dados sobre as ações de governo (RIBEIRO; VELHO, 2021).

Em todo o mundo são registradas diversas discussões que tratam do acesso e da abertura dos dados a fim de que eles sejam compartilhados e alcancem o máximo potencial de seu reúso. Os esforços brasileiros já realizados indicam que o país está caminhando progressivamente na perspectiva de promover o acesso livre, aberto e o compartilhamento dos dados governamentais. No entanto, diante da diversidade e complexidade dos dados, é necessária a adoção de elementos que melhorem a qualidade dos dados e facilitem a sua gestão. Assim, apresentaremos na próxima seção a aplicação dos princípios *FAIR*, que foram criados como forma de minimizar as dificuldades encontradas na criação e no gerenciamento dos dados.

3 Princípios *FAIR*

Com o intuito de solucionar problemas no tratamento dos inúmeros dados de pesquisa gerados, surgiram os princípios *FAIR* (*Findable, Accessible, Interoperable, Reusable*), como proposta de estabelecer melhores práticas de gestão, compartilhamento e reúso dos dados de pesquisa. Os princípios *FAIR* foram elaborados em 2014, como resultado de uma conferência denominada *Jointly designing a data FAIRPORT*.

Nessa conferência, um grupo de especialistas de várias áreas do conhecimento, preocupados com a utilização e reutilização de dados, discutiram e estabeleceram os princípios *FAIR* no contexto da *E-science* (HENNING *et al.*, 2018). O principal foco era gerar instrumentos capazes de orientar a gestão adequada dos dados e garantir a sua preservação, visibilidade e permanência para as gerações futuras. Em pouco tempo, os princípios *FAIR* foram aceitos e implementados por editores, universidades, pesquisadores, agências de financiamento, bibliotecários e todas as instituições que os viram como solução para os dados poderem ser reusados por homens e por máquinas.

A referida conferência resultou na organização de um guia dos princípios, como forma de nortear sua aplicação e sua utilidade (WILKINSON *et al.*, 2017). Partindo de cada uma das letras do acrônimo *FAIR*, tem-se a adoção de um conjunto de especificações que demandam um domínio de conhecimentos técnicos e estruturais distintos.

Em *Findable* (F), há quatro subprincípios que tratam dos elementos que os dados e metadados devem possuir para tornar o dado localizável. São eles:

- (F1) - Os dados e metadados devem ter identificadores globais, persistentes e identificáveis;

- (F2) - Os dados devem ser descritos com metadados enriquecidos;
- (F3) - Os metadados devem incluir clara e explicitamente identificadores dos dados que descrevem;
- (F4) - Os dados e metadados devem ser registrados ou indexados com recursos que ofereçam capacidade de busca.

Em *Accessible* (A), dois subprincípios são identificados como fundamentais para tornar os dados e metadados acessíveis:

- (A1) – Os conjuntos de dados e metadados devem possuir identificador que utilize protocolo de comunicação padronizado, para facilitar a recuperação;
- (A1.1) - O protocolo deve ser aberto, gratuito e universalmente implementável;
- (A1.2) - O protocolo deve permitir procedimentos de autenticação e autorização, caso seja necessário;
- (A2) - Garantir a acessibilidade aos metadados, mesmo quando os dados não estiverem mais disponíveis.

Interoperable (I) é o princípio que trata sobre a importância de os dados e metadados estarem em conformidade com formatos e padrões reconhecidos, objetivando o intercâmbio desses elementos entre sistemas. Possui três subprincípios:

- (I1) – Os dados e metadados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento;
- (I2) - Os dados e metadados devem usar vocabulários que seguem os princípios *FAIR*;
- (I3) - Os dados e metadados devem incluir referências qualificadas para outros dados e metadados, possibilitando a conexão semântica entre eles.

Reusable (R) destaca a necessidade de os dados e metadados possuírem documentação apropriada para possibilitar o seu reuso. Apresenta o seguinte subprincípio:

- (R1) – Os dados e metadados são descritos com uma pluralidade de atributos precisos e relevantes;
- (R1.1) – Os dados e metadados devem ser disponibilizados com licenças de uso claras e acessíveis;
- (R1.2) – Os dados e metadados devem estar associados à sua proveniência;

- (R1.3) – Os dados e metadados devem estar alinhados com padrões relevantes do seu domínio.

Como mencionado anteriormente, a evolução tecnológica alavancou o movimento da Ciência Aberta e com isso tornou-se cada vez mais comum e benéfico o compartilhamento de dados e informações. Nesse sentido, os princípios *FAIR* surgiram como solução para nortear boas práticas de gestão de dados científicos, pois envolvem ideias de como melhorar a qualidade dos dados e, conseqüentemente, seu reuso por homens e máquinas.

3.1 FAIRificação e interoperabilidade semântica

Tomando por base o contexto do compartilhamento de dados nos esforços para o acesso aberto à produção em ciência e tecnologia, o Programa de Pós-Graduação em Biblioteconomia da Unirio (PPGB) estruturou uma disciplina para elucidar e debater o compartilhamento de dados com o uso dos princípios *FAIR*, por meio dos processos de transformação de dados e metadados com o apoio de modelos semânticos (processo de FAIRificação).

Com um debate sobre o movimento da Ciência Aberta e iniciativas em curso para a sua implementação, além da noção de interoperabilidade, trazida pela Web de Dados (*Linked Open Data Cloud*), a disciplina destacou o processo de FAIRificação com base em Jacobsen *et al.* (2020), no conteúdo do curso *FAIRification Process* (UNIVERSITY OF TWENTE, 2022) e no quadro de padrões semânticos no domínio das GLAM (quadro 1), para discutir o uso e a elaboração de modelos semânticos em um estudo de caso prático sobre os periódicos musicais oitocentistas.

Quadro 1. Extrato das iniciativas e padrões semânticos no domínio das GLAM^[2]

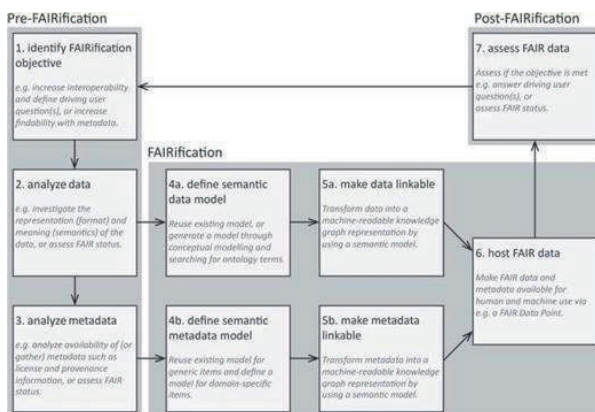
<i>Formalismo/Modelo Semântico</i>	<i>Escopo</i>
FRBR(er) ¹ ; FRBR com RDF (FRBRCore) ² ; FRBRoo ³ ; FRBR OWL2 DL ⁴ ; BIBFRAME ⁵ ; FaBIO ⁶ ; DataCite ⁷ ; BIR0 ⁸ ; BIBO ⁹ ; SKOS ¹⁰ ; DC ¹¹	Registros Bibliográficos
PROV-DM ¹²	Registro para Proveniência

Siglas/Acrônimos
¹ <i>Functional Requirements for Bibliographic Records – entity-relationship model.</i>
² <i>FRBR com Resource Description Framework.</i>
³ <i>FRBR Object Oriented.</i>
⁴ <i>FRBR in OWL2 DL Ontology.</i>
⁵ <i>Bibliographic Framework as a Web of Data – Linked Data Model and Supporting Services.</i>
⁶ <i>FRBR – aligned Bibliographic Ontology.</i>
⁷ <i>DataCite Metadata Schema.</i>
⁸ <i>Bibliographic Reference Ontology.</i>
⁹ <i>Bibliographic Ontology.</i>
¹⁰ <i>Simple Knowledge Organization System.</i>
¹¹ <i>Dublin Core.</i>
¹² <i>Provenance Data Model.</i>

Fonte: Adaptado de Ribeiro, Santos e Moreira (2020).

A figura 1, a seguir, apresenta o fluxo de FAIRificação trabalhado.

Figura 1. Fluxo de FAIRificação



Fonte: Jacobsen *et al.* (2020).

4 Música em periódicos oitocentistas

Conforme descrito no sítio do projeto Música em Periódicos Oitocentistas, os conjuntos de registros disponíveis compõem um banco de dados online contendo notícias e comentários sobre música em periódicos do século XIX [3]. Iniciado em 2002, tem servido de suporte para pesquisas sobre gêneros musicais, bem como sobre músicos, a imprensa musical, o teatro musical e a ópera, em periódicos brasileiros dos anos 1800.

A carga de registros é feita manualmente por pesquisadores que participam de projetos de pesquisa associados. Inicialmente, a consulta às fontes de informação (periódicos microfilmados) era feita na sala de obras raras da Biblioteca Nacional. Em julho de

2012, com a inauguração da Hemeroteca Digital Brasileira, o acesso passou a estar disponível de forma online.

O acesso na Hemeroteca Digital pode ser feito por título, período, edição, local de publicação e palavra(s)-chave. Os resultados da busca são apresentados sob a forma de arquivos em imagem e isso dificulta a captura automática. Assim, a extração e carga de dados ocorrem artesanalmente no banco de dados, sendo as informações digitadas em tabelas que incluem o título do periódico, local (cidade e estado), número da edição, número da página, data de publicação e conteúdo.

A consulta ao sítio que contém o banco de dados pode ser feita por todos esses parâmetros, além de palavras ou expressões em um campo de consulta livre. Seu aproveitamento por outros pesquisadores é muito precário ainda. Os conjuntos de dados precisam ser disponibilizados de forma adequada aos critérios e padrões de compartilhamento de dados em rede, incluindo a possibilidade de interoperabilidade. Assim, investigadores em qualquer parte do mundo terão acesso irrestrito a esses dados, a fim de reutilizá-los em suas pesquisas.

5 Procedimentos metodológicos

A presente pesquisa é de natureza aplicada, pois visa “contribuir para fins práticos mais ou menos imediatos, buscando soluções para problemas concretos” (CERVO; BERVIAN; SILVA, 2007, p. 72). Em relação ao objetivo, caracteriza-se como exploratória, uma vez que trabalha com o “levantamento da presença das variáveis e da sua caracterização quantitativa ou qualitativa” (KOCHE, 2015, p. 126). É considerada um estudo de caso, pois se concentra no estudo de um caso particular, no qual o objeto da pesquisa é uma unidade que se analisa profundamente (SEVERINO, 2007; TRIVIÑOS, 1992).

Com o objetivo de identificar os trabalhos correlatos e a natureza das iniciativas existentes na literatura científica, foi realizada uma revisão de literatura na Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapi), a partir da execução da seguinte estratégia de pesquisa: busca pelo termo *FAIR* em Todos os Campos, resultando na identificação de iniciativas *FAIR* voltadas para o planejamento visando à FAIRificação (71%) e para a execução da FAIRificação (29%). Esta base foi selecionada por ser referência para a área da Biblioteconomia e ter como finalidade subsidiar os estudos no âmbito da Ciência da Informação.

Como etapas do planejamento de FAIRificação, utilizou-se o fluxo proposto por Jacobsen *et al.* (2020), aplicado a um conjunto de dados composto por 32

notícias oriundas da base Música em Periódicos Oitocentistas que atendiam a critérios preestabelecidos.

As notícias selecionadas atendiam aos seguintes critérios: estarem referenciadas por um localizador e/ou identificador (URL e/ou URI) [4]; e possuir gêneros musicais atendidos por verbetes na Wikipedia e entradas na Wikidata. Esses critérios foram estabelecidos para que pudéssemos obter um alinhamento com os princípios *FAIR* (conforme seção 3), além de permitir conexões com termos já disponíveis na grande rede. O quadro 2 (apêndice A) apresenta um extrato com 10 notícias dentre as selecionadas.

Com o desenvolvimento do trabalho houve necessidade de ajustes na metodologia. Essas mudanças foram apresentadas junto aos resultados alcançados com o intuito de melhorar sua compreensão.

6 Resultados

O estudo de caso foi desenvolvido segundo o fluxo apresentado na figura 1, mas com resultados práticos decorrentes da fase de pré-FAIRificação.

6.1 A fase de pré-FAIRificação: recorte e seleção de datasets

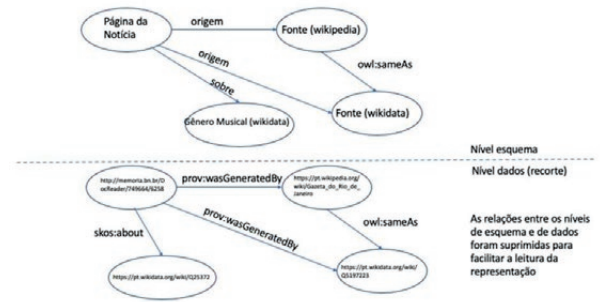
Para o planejamento do processo de FAIRificação, foram realizadas buscas no sítio do projeto alvo e extração manual de conteúdo para selecionar os registros que seriam utilizados no estudo de caso.

Após debate realizado pelos membros da equipe de investigação, decidiu-se recortar um pouco mais o escopo dos registros recuperados, por meio da fixação de um periódico específico, conforme quadro 3 (apêndice B). Esta orientação permitiu que pudéssemos planejar a publicação dos dados, agora segmentados por um título de periódico específico, pois a proposta inicial era tornar disponíveis grupos de dados que possibilitassem o reúso por outros projetos.

Em seguida, procurou-se identificar vocabulários que pudessem descrever semanticamente as relações (conforme seção 3.1). No contexto desta pesquisa, foram selecionados os vocabulários *Dublin Core* (DC), *Simple Knowledge Organization System* (SKOS), *Provenance Data Model* (PROV-DM) e *Web Ontology Language* (OWL). Para o recorte de dados neste estudo de caso, foram selecionados os predicados *prov:wasGeneratedBy*, *skos:about* e *owl:sameAs*.

A figura 2 apresenta o grafo *RDFSchema* e uma instância RDF gerada para ilustrar as conexões.

Figura 2. Grafo *RDFSchema/RDF*



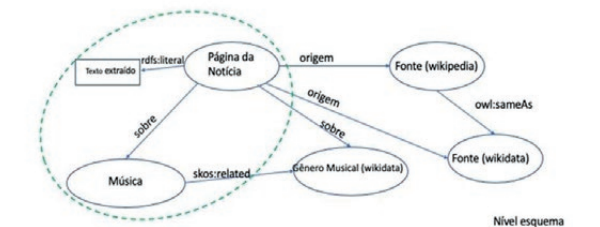
Fonte: Dados da pesquisa, 2023.

6.2 Preparação para as fases de FAIRificação e pós-FAIRificação

Com o recorte definido, o modelo semântico ajustado e os vocabulários escolhidos, o estudo de caso encontra-se pronto para ser testado em plataforma para compartilhamento. As próximas etapas do projeto serão a carga, em plataforma para compartilhamento, das triplas geradas (sujeito e objeto – conforme quadros 2 e 3, apêndices A e B, respectivamente), mas incrementadas semanticamente com os predicados, utilizando-se os vocabulários apontados anteriormente (DC, SKOS, PROV-DM e OWL), de acordo com a representação gráfica trazida pela figura 2. São etapas que atenderão às atividades de *Host* e *Access FAIRData*, presentes na figura 1 (Fluxo FAIRificação).

Ademais, já se vislumbrou a possibilidade de incrementar as conexões semânticas com outros dados ligados. Criar uma conexão com características da música e, também, trazer os trechos de texto que estão armazenados no banco de dados, podem ser atividades futuras para possíveis integrações com outras estruturas semânticas, como, por exemplo, a *Musicontology*, suas classes e propriedades, envolvendo artistas e demais músicos.

Figura 3. Grafo *RDFSchema* para a próxima implementação



Fonte: Dados da pesquisa, 2023.

7 Considerações finais

O processo de transformação de dados de pesquisa para torná-los *FAIR* ainda é uma iniciativa difícil de ser conduzida. A reunião dos dados apresentados na fase de revisão de literatura convalidou essa percepção, pois a maior parte das publicações recuperadas no contexto desta pesquisa (cerca de 70%) explorou o processo de planejamento e fundamentação dos princípios *FAIR*.

A noção de conjunto de dados é item relevante na disponibilização para o reuso. Delimitar essa coleção quando se trabalha com os registros armazenados em um banco de dados relacional é tarefa essencial. Essa discussão levou a equipe da pesquisa a ajustar a estratégia para transformação dos dados, coletando e disponibilizando segundo assuntos específicos e de interesse para a comunidade envolvida no uso dos dados.

Adicionalmente, foi percebido neste estudo que existem pontos de convergência entre as atividades presentes no processo de FAIRificação e a preparação para uso de dados ligados. A execução de etapas do fluxo apresentado na figura 1 para tornar ligáveis os dados ligáveis (etapas 5a e 5b) se alinhou com as melhores práticas para consumo e publicação de dados na Web (RAUTENBERG *et al.*, 2018). Esse alinhamento pode impulsionar os processos de disponibilização dos dados usando *FAIR*, na medida em que o uso de dados ligados é suportado por extensa literatura nacional.

Notas

- [1] O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq.
- [2] Acrônimo de *Galleries, Libraries, Archives and Museums*.
- [3] Disponível em: <http://musica-sec-xix.unirio.br/>.
- [4] Na ausência de URI, optou-se por usar URL como PID (*Persistent Identifier*) apenas para demonstrar a aplicabilidade do processo no estudo de caso.

Referências

- BAKER, M. Is there a reproducibility crisis? *Nature*, v. 533, p. 452, 26 maio 2016. Disponível em: <https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970>. Acesso em: 22 nov. 2022.
- BRASIL. **Infraestrutura nacional de dados abertos**. 05 dez. 2019. Disponível em: <https://www.gov.br/governodigital/pt-br/dados-abertos/infraestrutura-nacional-de-dados-abertos>. Acesso em: 10 dez. 2022.
- CERVO, A. L.; BERVIAN, P. A.; SILVA, R. Técnicas de coleta de dados. In: CERVO, A. L.; BERVIAN, P. A.; SILVA, R. **Metodologia científica**. 6. ed. São Paulo: Pearson Prentice Hall, 2007. p. 69-82.

CURTY, R. G.; AVENTURIER, P. O paradigma da publicação de dados e suas diferentes abordagens. **Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação**, n. XVIII ENANCIB, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/105144>. Acesso em: 24 nov. 2022.

FANELLI, D. Opinion: Is science really facing a reproducibility crisis, and do we need it to? **Proceedings of the National Academy of Sciences of the United States of America**, v. 115, n. 11, p. 2628-2631, mar. 2018. Disponível em: <https://www.pnas.org/doi/10.1073/pnas.1708272114>. Acesso em: 24 nov. 2022.

HENNING, P. C. *et al.* Desmistificando os princípios *FAIR*: conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados *FAIR*. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 11, n. 1, p. 1-23, 2008. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/458>. Acesso em: 17 out. 2022.

JACOBSEN, A. *et al.* A generic workflow for the Data FAIRification Process. **Data Intelligence**, Cambridge, v. 2, n. 1-2, p. 56-65, 2020. Disponível em: https://doi.org/10.1162/dint_a_00028. Acesso em: 22 mar. 2023.

JORGE, V. A.; ALBAGLI, S. Compartilhamento de dados de pesquisa em saúde: iniciativas do National Institutes of Health (NIH). **Reciis**, v. 12, n. 4., out./dez. 2018. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/1499>. Acesso em: 24 nov. 2022.

KALINAUSKAITĖ, D. To be findable, accessible, interoperable and reusable: language data and technology infrastructure for supporting the FAIR data approach. In: CEUR Workshop Proceedings, Kaunas, 2017. **ICYRIME 2017: proceedings of the symposium for young researchers in informatics, mathematics and engineering**. Aachen: CEUR-WS, 2017. Disponível em: <https://www.vdu.lt/cris/entities/publication/6138a749-b4c6-48b0-9020-b2b97ef1dd91/details>. Acesso em: 09 dez. 2022.

KOCHE, J. C. O fluxograma da pesquisa científica. In: KOCHE, J. C. **Fundamentos de metodologia científica: teoria da ciência e iniciação à pesquisa**. 34. ed. Petrópolis: Vozes, 2015. p. 121-136.

NÜST, D. *et al.* Reproducible research and GIScience: an evaluation using AGILE conference papers. **PeerJ**, v. 6, e5072, 2018. Disponível em: <https://peerj.com/articles/5072/>. Acesso em: 11 dez. 2022.

PACKER, A. L.; SANTOS, S. **Ciência aberta e o novo modus operandi de comunicar pesquisa – parte I**. ago. 2019. Disponível em: <https://blog.scielo.org/blog/2019/08/01/ciencia-aberta-e-o-novo-modus-operandi-de-comunicar-pesquisa-parte-i/#.Y5Rn433MLIV>. Acesso em: 10 dez. 2022.

PORTUGAL. Ministério da Ciência, Tecnologia e Ensino Superior. **Ciência aberta**. 2016. Disponível em: <https://www.ciencia-aberta.pt/sobre-ciencia-aberta>. Acesso em: 11 dez. 2022.

- RAUTENBERG, S. *et al.* **Guia prático para publicação de dados abertos conectados na web.** Curitiba: Appris, 2018.
- RIBEIRO, C. J. S.; SANTOS, L. O. B. S.; MOREIRA, J. L. R. Princípios *FAIR* e interoperabilidade: reprodutibilidade com o uso de modelagem conceitual orientada a ontologias. **Informação & Informação**, v. 25, n. 3, p. 28-51, out. 2020. Disponível em: <https://www.uel.br/revistas/uel/index.php/informacao/article/view/41874>. Acesso em: 22 mar. 2023.
- RIBEIRO, C. J. S.; VELHO, A. C. M. Princípios *FAIR* e a gestão de bases governamentais: análise do compartilhamento de dados de registros civis por meio da iniciativa GovData. In: SALES, L. F. *et al.* (org.). **Princípios *FAIR* aplicados à gestão de dados de pesquisa.** Rio de Janeiro: Ibict, 2021. Disponível em: <https://ridi.ibict.br/handle/123456789/1182>. Acesso em: 11 dez. 2022.
- SALES, L. F. *et al.* GO FAIR Brazil: a challenge for brazilian data science. **Data Intelligence**, v. 2, n. 1-2, p. 238-245, jan. 2020. Disponível em: <https://direct.mit.edu/dint/article/2/1-2/238/10004/GO-FAIR-Brazil-A-Challenge-for-Brazilian-Data>. Acesso em: 10 dez. 2022.
- SANTOS, P. X.; ALMEIDA, B. A.; HENNING, P. (org.). **Livro verde - ciência aberta e dados abertos:** mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional. Rio de Janeiro: Fiocruz, 2017. Disponível em: <https://www.arca.fiocruz.br/handle/icict/24117>. Acesso em: 10 dez. 2022.
- SEVERINO, A. J. **Metodologia do trabalho científico.** 23. ed. rev. e atual. São Paulo: Cortez, 2007.
- SILVA, T. E.; ALCARÁ, A. R. Acesso aberto à informação científica: políticas e iniciativas governamentais. **Informação & Informação**, v. 14, n. 2, p. 100-116, jul./dez. 2009. Disponível em: https://brapci.inf.br/_repositorio/2010/05/pdf_5d9c4687ca_010248.pdf. Acesso em: 23 nov. 2022.
- TRIVIÑOS, A. N. S. **Introdução à pesquisa de ciências sociais:** a pesquisa qualitativa em educação. São Paulo: Atlas, 1992.
- UNIVERSITY OF TWENTE. **FAIR principles and the FAIRification process (2021-2B) course.** 2022.
- WILKINSON, M. D. *et al.* **A design framework and exemplar metrics for FAIRness.** nov. 2017. Disponível em: <https://www.biorxiv.org/content/biorxiv/early/2017/11/29/225490.full.pdf>. Acesso em: 22 mar. 2022.

Apêndice A – Quadro 2. Extrato da relação de dados para composição de atributos em triplas RDF

<i>Sujeito</i>	<i>Objeto</i>
http://memoria.bn.br/DocReader/749664/6258	https://pt.wikipedia.org/wiki/Gazeta_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q5197223
http://memoria.bn.br/DocReader/094170_01/4569	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_01/42876	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_01/45836	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_01/46037	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_02/14181	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_02/15887	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_02/16135	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_02/24875	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643
http://memoria.bn.br/DocReader/094170_02/25025	https://pt.wikipedia.org/wiki/Diário_do_Rio_de_Janeiro https://www.wikidata.org/wiki/Q10267643

Fonte: Dados da pesquisa, 2023.

Apêndice B – Quadro 3. Extrato da relação de dados para composição de atributos em triplas RDF – recortado para periódico específico

<i>Sujeito</i>	<i>Objeto</i>
http://memoria.bn.br/DocReader/706906/10 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982
http://memoria.bn.br/DocReader/706906/26 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982
http://memoria.bn.br/DocReader/706906/30 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982
http://memoria.bn.br/DocReader/706906/38 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982
http://memoria.bn.br/DocReader/706906/39 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982
http://memoria.bn.br/DocReader/706906/37 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982
http://memoria.bn.br/DocReader/706906/43 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982
http://memoria.bn.br/DocReader/706906/46 (A Marmota na Corte (RJ) - 1849 a 1853)	https://pt.wikipedia.org/wiki/A_Marmota https://www.wikidata.org/wiki/Q16142982

Fonte: Dados da pesquisa, 2023.

Levantamento e tratamento dos dados de pesquisa da Ciência da Informação à luz dos dados abertos da CAPES

Survey and treatment of Information Science research data in the light of CAPES open data

Francis Bento Marques (1), Benildes Coura Moreira dos Santos Maculan (2), Renato Rocha Souza (3)

(1) Universidade Federal de Minas Gerais, Av. Pres. Antônio Carlos, 6627 - Pampulha, Belo Horizonte - MG, 31270-901, fbmarques@gmail.com. (2) benildes@gmail.com, (3) rsouzaufmg@gmail.com

Resumo

Este estudo debate a temática dos Dados Governamentais Abertos, com foco nos dados que representam áreas científicas. O objetivo é apresentar a estruturação dos dados de pesquisa da área da Ciência da Informação, à luz dos dados abertos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Do ponto de vista metodológico, caracteriza-se a pesquisa como descritiva e quantitativa, tendo como procedimentos técnicos a aplicação de uma pesquisa documental, dividida em duas etapas: exploração de dados (levantamento) e preparação de dados (ajustes de inconsistência nos dados). Para execução da parte técnica da pesquisa, recorreu-se ao uso dos softwares Open Refine e do Calc do Libreoffice, do ambiente Google Colaboratory, da linguagem de programação Python e de quatro bibliotecas de visualização de dados. Os resultados revelam a existência de problemas relacionados à digitação, à falta de padronização, à ausência e à duplicidade de dados inseridos nos respectivos campos de registro, não sendo identificado se os problemas ocorrem no preenchimento dos campos pelas instituições ou se estão relacionados ao próprio funcionamento do sistema da Agência. Com isso, concluiu-se que são necessárias ações de melhorias na estruturação e disponibilização dos dados abertos referentes à Ciência da Informação, tais como a criação de campos obrigatórios pela Agência e maior detalhamento e uniformidade no preenchimento dos dados pelos programas de Pós-Graduação.

Palavras-chave: Dados Abertos Governamentais. Dados de Pesquisa. Dados Abertos da Capes. Ciência da Informação.

Abstract

This study discusses the theme of Open Government Data, with a focus on data representing scientific areas. The objective is to present the structuring of research data in the field of Information Science, in light of the open data from the Coordination for the Improvement of Higher Education Personnel. From a methodological point of view, the research is characterized as descriptive and quantitative, with technical procedures involving the application of a documentary research, divided into two stages: data exploration (survey) and data preparation (adjustment of data inconsistencies). For the execution of the technical part of the research, we resorted to the use of Open Refine and LibreOffice Calc software, the Google Colaboratory environment, the Python programming language, and four data visualization libraries. The results reveal the existence of problems related to typing, lack of standardization, absence, and duplication of data entered in the respective registration fields, with it not being identified whether the problems occur in the filling of fields by institutions or if they are related to the functioning of the Agency's system. Consequently, it was concluded that actions for improvement in the structuring and availability of open data related to Information Science are necessary, such as the creation of mandatory fields by the Agency and greater detail and uniformity in the filling of data by postgraduate programs.

Keywords: Open Government Data. Search Data. Capes Open Data. Information Science.

Resumen

Este estudio aborda el tema de los Datos Gubernamentales Abiertos, con enfoque en los datos que representan áreas científicas. El objetivo es presentar la estructuración de los datos de investigación en el campo de la Ciencia de la Información, a la luz de los datos abiertos de la Coordinación de Mejoramiento del Personal de Nivel Superior. Desde un punto de vista metodológico, la investigación se caracteriza como descriptiva y cuantitativa, con procedimientos técnicos que incluyen la aplicación de una investigación documental, dividida en dos etapas: exploración de datos (levantamiento) y preparación de datos (ajustes de inconsistencia en los datos). Para la ejecución de la parte técnica de la investigación, recurrimos al uso del software Open Refine y Calc de LibreOffice, el entorno de Google Colaboratory, el lenguaje de programación Python y cuatro bibliotecas de visualización de datos. Los resultados revelan la existencia de problemas relacionados con la escritura, la falta de estandarización, la ausencia y la duplicidad de datos ingresados en los respectivos campos de registro, sin identificar si los problemas ocurren en el llenado de campos por parte de las instituciones o si están relacionados con el funcionamiento del sistema de la Agencia. Con esto, se concluyó que son necesarias acciones de mejora en la estructuración y disponibilidad de los datos abiertos relacionados con la Ciencia de la Información, como la creación de campos obligatorios por parte de la Agencia y un mayor detalle y uniformidad en el llenado de datos por parte de los programas de posgrado.

Palabras clave: Datos Gubernamentales Abiertos. Datos de Investigación. Datos Abiertos de Capes. Ciencia de la Información

1 Introdução

A disciplina Organização do Conhecimento tem como foco e objeto a construção planejada de uma estrutura (organização) de modo que aquilo que é conhecido e está registrado (conhecimento) possa ser “compartilhado por meio de nossa capacidade de lidar com as ‘coisas deste mundo’ e pelo uso de nossas habilidades linguísticas para expressar nossa experiência e percepções” (DAHLBERG, 2006, p. 12). Além da forma linguística, segundo a autora, esse conhecimento pode ser comunicado por outros diversos modos de representação, seja por palavras ou sinais, pois é preciso uma forma de representação para a “verificação da realidade, a fim de examinar sua veracidade e torná-lo intersubjetivo, ou seja, o mais objetivo possível, pelo menos por um certo tempo (DAHLBERG, 2006, p. 12). Essa representação pode estar em distintos graus de complexidade.

A autora destaca que o objeto da organização do conhecimento engloba apreender os referentes no mundo, real ou abstrato, organizando-os em conformidade com um plano, representando-os e agrupando-os de modo que possam ser reconhecidos e que novos conhecidos possam ser inferidos a partir deles (DAHLBERG, 1978; 2006). A pesquisadora ainda relata que uma das aplicações da organização do conhecimento é o mapeamento de unidades de conhecimento a partir de visualizações de relações encontradas entre essas unidades por meio de gráficos.

Hjørland (2008) descreve sete abordagens de organização do conhecimento propostas por Broughton et al. (2005), dentre as quais destaca-se as abordagens bibliométricas, que estão alinhadas com o contexto deste estudo. Nas abordagens bibliométricas, segundo o autor, faz-se uso de referências bibliográficas para organizar redes de artigos ou de análise de citação, buscando criar e “interpretar mapas bibliométricos como estruturas de campos de pesquisa” (HJØRLAND, 2008, p. 94). O autor destaca que os estudos bibliométricos estão mais próximos da organização social do conhecimento, sendo que oferecem informações referenciais adicionais, que são úteis por si só.

No contexto dos estudos bibliométricos, a produção científica brasileira gerada como resultado de pesquisas vem sendo analisada por instrumentos de avaliação da ciência desenvolvidos pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Esses estudos são, segundo Hjørland (2002), um dos instrumentos metodológicos para analisar uma área do conhecimento, quando se usa os princípios da Teoria Analítica do Domínio. Hjørland (2002, p. 260) afirma que é o campo mais amplo de uma área “que desenvolva e compartilhe conceitos, termos e conhecimentos comuns [...] em diferentes camadas de generalização” e que um segundo passo seria “investigar a natureza e a

estrutura do conhecimento e da comunicação no nível de especialização escolhido”. Essas análises metódicas sobre um domínio permitem perceber os aspectos mais relevantes “de sua organização, de sua formação, de seu desenvolvimento, de seu funcionamento e de seus produtos intelectuais” (TESSER, 1994, p. 92), trazendo como resultado informações estruturadas sobre o domínio. Essas informações estruturadas são o foco nos estudos na área da Ciência da Informação, pois, como afirma Rodrigues (2019), a informação é seu objeto de estudo e está sempre em evidência nas suas pesquisas acadêmicas, seja por uma perspectiva filosófica (BROOKES, 1980), epistemológica (SARACEVIC, 1970), fundamentos teóricos (CAPURRO, 1992), interdisciplinaridade (SARACEVIC, 1992), ou da informação como coisa (BUCKLAND, 1991).

Originalmente denominada por Hulme, em 1923, como bibliografia estatística, o termo “bibliometria” foi citado por Paul Otlet, em 1934, no seu livro *Traité de Documentation* (FONSECA, 1973), como parte mensurada ou quantificada da Bibliologia, e mais tarde popularizado por Allan Pritchard, em 1969, no seu livro *Concepts of Information Retrieval* (PAO, 1989). Os estudos bibliométricos têm como objeto os aspectos quantitativos da produção científica, agregando elementos tais como a disseminação e o uso da informação publicada, assim como da sua validade (TAGUESUTCLIFFE, 1992), estabelecendo, por exemplo, indicadores de produção científica de uma instituição ou área do conhecimento em relação ao país.

Desde a origem da Ciência da Informação como disciplina, a bibliometria se instalou não como um campo de pesquisa, mas como uma técnica ou instrumental para medir a produção científica, pois a área teve início, basicamente, para “lidar com o grande volume e a diversificação de informações registradas em variadas formas, com vistas à sua mais ampla difusão [...]. Fortemente influenciada pelas ciências empíricas [...]” (CARDOSO, 1996, p. 73)

Muitas vezes, a coleta de dados para a análise bibliométrica é realizada por meio de revisões bibliográficas ou revisões sistemáticas da literatura. Porém, outra alternativa para a coleta de dados é utilizar dados abertos, uma vez que há várias iniciativas, no mundo todo, para a abertura de dados, sobretudo no eixo público (dados governamentais), tendo em vista dar maior transparência e acesso às informações, assim como promover participação social nas gestões. Algumas vezes, o que impede o uso desses dados nos estudos bibliométricos se refere à falta de letramento de dados, pois a coleta desses dados exige conhecimentos técnicos específicos para o acesso, o uso, o tratamento e a apresentação dos dados já tratados e estruturados (BUZATO; TORRES, 2019), de modo que seja possível analisá-los e inferir sobre eles.

O movimento crescente de Dados Governamentais Abertos (DGA), do inglês Open Government Partnership (OGP), vem encorajando as organizações governamentais a disponibilizarem seus dados de forma pública, com o objetivo de aumentar a transparência (ATTARD et al., 2015). O assunto vem sendo bastante relatado na literatura, considerando a contribuição para a consolidação da transparência e da responsabilização dos governos, na medida em que propicia o acesso para a população de um conjunto de dados que foi produzido na esfera pública e, assim, possibilita a geração de valor público, que acaba fortalecendo o relacionamento entre governo e cidadão. No entanto, o uso escasso dos DGA é que tem impedido que muitas iniciativas alcancem seu potencial máximo (MARTIN, 2014; RUIJER; MARTINIUS, 2017).

De acordo com Graves e Hendler (2014), vários estudos sobre os DGA mostram que o uso e o entendimento dos dados requerem conhecimento mais profundo de tecnologias de gerenciamento de dados e de técnicas estatísticas. As habilidades requeridas para tratamento desses dados não estão descritas nos currículos de todos os cidadãos. Desse modo, treinamentos são necessários para melhorarem as práticas dos usuários e, portanto, aumentarem o uso dos DGA (GASCÓ-HERNÁNDEZ et al., 2018).

No entendimento de Osagie et al. (2017), o uso dos dados por pessoas sem conhecimento profundo em tecnologia fez com que a União Europeia desenvolvesse o projeto chamado de Raising Open User-friendly Transparency-Enabling Technologies for Public Administration (ROUTE-TO-PA). Os autores afirmam que o projeto foi financiado como parte do Horizonte 2020 da Comissão Europeia, com o intuito de melhorar a transparência governamental, projetando e desenvolvendo tecnologias que simplificam e aumentam o acesso aos dados. Com isso, defendem os autores, os cidadãos não conhecedores de tecnologia podem envolver-se em debates baseados em dados sobre questões sociais.

A dificuldade no uso das plataformas e conjuntos de dados disponibilizados vem criando uma lacuna entre produtores e usuários de dados. Algumas ferramentas estão sendo desenvolvidas para resolver esse problema e acrescentar aos dados uma visualização que possa ajudar, aproximar e envolver diferentes tipos de usuários interessados nos dados (BARCELLOS et al., 2017; GRAVES; HENDLER, 2014). A visualização de dados pode ajudar os governos a completarem os seus objetivos para com os DGA. Dessa forma, o desejo governamental é entregar uma ferramenta que possa ajudar os usuários a compreenderem os dados e também a se comunicarem de modo mais simples e objetivo com diversos públicos (BRUGGER et al., 2016; GRAVES; HENDLER, 2014).

No Brasil, a CAPES é uma das instituições preocupadas com a pesquisa e tem realizado esforços para oportunizar transparência aos dados gerados nas pesquisas científicas de diferentes áreas do conhecimento. Em 2017, a instituição disponibilizou o seu Plano de Dados Abertos (PDA), conforme orienta o Decreto nº 8.777/2016, que tem o intuito de publicizar as ações e estratégias organizacionais que nortearão as atividades de implementação e promoção da abertura de dados. A priorização dos conjuntos de dados a serem disponibilizados foi definida considerando a relevância das informações para o cidadão e o grau de maturidade dos conjuntos de dados dentro da instituição (CAPES, 2022).

Mesmo com o PDA, ainda há necessidade de estudos que investiguem, de modo mais pontual e profundo, a estruturação e a disponibilização dos dados de pesquisa, considerando, sobretudo, áreas específicas de conhecimento, como a Ciência da Informação. Nesse sentido, o problema deste estudo se baseia na questão de como os dados de pesquisa da Ciência da Informação estão sendo disponibilizados pela Capes. Assim, objetiva-se apresentar a estruturação dos dados de pesquisa da área da Ciência da Informação, à luz dos dados abertos da Capes.

O presente estudo é parte integrante de uma tese de doutorado em desenvolvimento e em processo de qualificação. Nesta ocasião, são apresentados os resultados das duas primeiras etapas, de análise exploratória dos dados coletados e de preparação dos dados (com o tratamento dos dados), respectivamente, e se projeta uma terceira etapa como estudo futuro para a finalização da tese.

2 Dados Governamentais Abertos e sua reutilização

O movimento dos DGA foi lançado em 2011 por oito governos fundadores, constituídos por Brasil, Indonésia, México, Noruega, Filipinas, África do Sul, Reino Unido e Estados Unidos. Ao longo dos anos, o movimento cresceu para mais de 75 países membros. Essa iniciativa internacional vem encorajando reformas governamentais nas áreas de transparência, prestação de contas, participação e tecnologia na governança. Vem também promovendo a participação e elaboração das normas tecnológicas para ajudar os governos locais a se tornarem mais transparentes, responsáveis e receptivos para com os cidadãos, exigindo compromissos voluntários e colaboração com a sociedade civil no desenvolvimento de planos de ação nacionais (BERTIN, 2019).

Nesse contexto, as autoridades governamentais desenvolvem planos de ação e identificam quais são os dados que podem ou devem ter acesso público. Uma das maneiras pelas quais os DGA esperam contribuir para um governo mais aberto e responsável é socializando

as normas de participação cívica dentro das instituições governamentais, por meio de várias interações com a sociedade civil, com base em políticas nacionais. Logo, os dados governamentais tornam-se abertos e, portanto, autorizados para reutilização por agentes públicos e privados (ATTARD et al., 2015; GRAVES; HENDLER, 2014; MARTIN, 2014).

Para Freitas et al. (2018), alguns conceitos como dados abertos, governo aberto, transparência e até a participação da população são acessíveis por meio do uso de tecnologias. Com essas tecnologias, segundo os autores citados, demonstra-se a necessidade da construção de ferramentas para auxiliar os cidadãos no acesso, compreensão e manipulação dos dados disponibilizados pelo governo.

O resultado desse processo será o fomento à reutilização dos dados, que podem servir a vários propósitos, dentre os quais se destacam a exploração, análise e apresentação de dados públicos pelos cidadãos. O reuso dos dados é um elemento importante para os cidadãos, que passam a fiscalizar o governo e podem até responsabilizar as autoridades (CUCCINIELLO; NASSI; VALOTTI, 2012).

A reutilização dos DGA na pesquisa científica, no âmbito dos conjuntos de dados públicos, tem uma aceitação relevante e é altamente benéfica para a sociedade (ZIMMERMANN; PUCHAR, 2015). Além disso, os DGA podem ser usados pelo setor privado para o desenvolvimento e a criação de novos produtos e serviços com fins lucrativos. Dessa maneira, tem-se a expectativa de que a reutilização comercial dos DGA possa gerar impacto econômico na sociedade, ao impulsionar a inovação, criar empregos e gerar riqueza (YU; ROBINSON, 2011).

Kassen (2021) afirma que podem existir alguns obstáculos tecnológicos para que os usuários possam reutilizar os dados abertos e aproveitá-los em potencial, tais como: falta de conhecimento tecnológico, acesso a poucos equipamentos necessários ou o acesso aos dados incompletos. Outro problema, segundo o autor, é que alguns dos conjuntos de dados podem conter formatos de extensão exclusivos, sendo lidos apenas por softwares proprietários. Outras questões problemáticas pontuadas por Kassen (2021) é o custo desses softwares e os dispositivos exclusivos ou equipamentos de última geração, que podem ser inacessíveis para a maioria dos usuários que desejam trabalhar com esse tipo de dados.

Na maioria das vezes, as barreiras para o uso e o reuso dos dados abertos podem ser agrupadas em duas grandes categorias: as de implementação e as de uso (MARTIN, 2014). Em ambos os casos, o estudo mencionado aponta a necessidade da criação de portais para uma maior abertura dos dados com a utilização de ferramentas que facilitem o uso e também que forneçam

instruções de como utilizar, além da possibilidade de integrar esses dados com outros dados e/ou informações. Desse modo, Martin (2014) conclui que o desenvolvimento de plataformas para o reuso e a integração com outros dados apresenta-se muito viável e importante para toda a comunidade.

3 Metodologia do estudo

Quanto aos propósitos, este estudo constitui-se como uma pesquisa descritiva e quantitativa. Do ponto de vista técnico, tem-se a aplicação de uma pesquisa documental, realizada nos conjuntos de dados abertos disponibilizados pela Capes, considerando, como recorte, os dados registrados e que representam a área da Ciência da Informação brasileira.

Importante frisar que, em seu sítio, a Capes disponibiliza 63 conjuntos de dados, com os temas: Acessos ao Portal de Periódicos; Bolsas e Auxílios; Avaliação da Pós-Graduação *Stricto Sensu* e Servidores; Contratos; Dotação e Execução Orçamentária da Capes.

No âmbito de cada tema, existem diversos conjuntos de dados, que são disponibilizados em diferentes formatos de arquivos para download, como: CSV, XLS, RDF, TTL, HTML e PDF. O conjunto de dados trabalhados nesta pesquisa é o de Avaliação da Pós-Graduação *Stricto Sensu*, composto por 37 *datasets*, sendo selecionados para análise os que contêm dados referentes às teses e dissertações desenvolvidas no Brasil e aos docentes e discentes. Em todos os casos, foram considerados os estudos defendidos no intervalo de 2004 a 2021 e cujos arquivos se encontravam disponibilizados no formato CSV.

Para cobrir esse período, foram utilizados os quatro Catálogos de Teses e Dissertações (CTD) disponibilizados pela Capes. O primeiro CTD refere-se às teses e dissertações defendidas de 1987 a 2012. O PDA da Capes disponibiliza 25 arquivos, cada um possui 42 variáveis, contendo dados como resumo, palavras-chave, nome do discente, nome da instituição de ensino, unidade federativa, entre outros (CAPES, 2022).

O segundo CTD abrange os estudos desenvolvidos de 2013 a 2016, com cinco arquivos, cujo número de variáveis passou para 56 e foram adicionados códigos de identificação para instituição, discente, docente, data de entrada e saída do programa, o endereço para o texto completo da dissertação ou tese na Plataforma Sucupira, dentre outros.

No terceiro CTD, há quatro arquivos que compreendem os anos de 2017 a 2020 e são compostos por 58 variáveis, sendo adicionados mais dois campos em relação ao segundo CTD, que são `IN_TCC_COM_VINCULO_PRODUCAO` e `ID_ADD_PRODUCAO_VINCULO_CT`, que mostram se houve estudo vinculado a essa produção e, em caso

positivo, qual é a sua identificação em outra base de dados.

No quarto e último CTD, aparecem, apenas, os dados de 2021, pois até o momento atual da pesquisa (março de 2023), os dados disponibilizados contêm informações sobre a Pós-Graduação do ano base 2021, com variáveis que se assemelham ao terceiro CTD (CAPES, 2022).

Para a parte empírica da pesquisa, sobretudo no que tange à coleta de dados, foi feito o uso de dois softwares: *Open Refine* (versão 3.6.1) e *Calc*, do *Libre Office* (versão 7.4.1.2). Ao manusear os dados coletados e sua importação para planilhas eletrônicas, optou-se por utilizar o ambiente *Google Colaboratory*, com o uso da linguagem de programação *Python*, na versão 3.10, e as bibliotecas *Pandas*, na versão 1.4.2, *Seaborn*, na versão 0.11.2, *Matplotlib*, versão 3.5.0 e *Numpy*, versão 1.23.0. O uso dessas ferramentas foi necessário, a fim de garantir o entendimento dos dados e a preparação deles.

A princípio, foi executado o download de 46 arquivos em formato CSV, totalizando 7,5 GB. Com o auxílio do software *Open Refine*, cada arquivo foi aberto. Em seguida, procedeu-se à seleção ou filtro do campo `CD_AREA_CONHECIMENTO`, em específico para o código 60700009, que se refere à área da Ciência da Informação, conforme estruturação do PDA.

No desenvolvimento desta pesquisa, optou-se em agrupar as atividades em duas etapas: 1 – análise exploratória e 2 - preparação dos dados. O quadro 1 apresenta as atividades de cada etapa e as variáveis ou elementos que foram abordados em cada uma.

Quadro 1. *Etapas da pesquisa*

<i>Nome da etapa</i>	<i>Atividades realizadas</i>	<i>Elementos ou variáveis abordados</i>
Análise exploratória	- Levantamento dos dados no site da Capes; 2 – Registro dos dados brutos em planilha; 3 – Geração de arquivo em formato CSV; 4 – Mapeamento de campos e variáveis.	Ano, área de conhecimento, programa, instituição, Estado da instituição, status jurídico, área de concentração, linhas de pesquisa, tipo de produção, <i>abstract</i> , projeto, dados do docente e dados do discente
Preparação dos dados	1 – Verificação minuciosa dos dados; 2 – Localização	Os mesmos da etapa anterior.

<i>Nome da etapa</i>	<i>Atividades realizadas</i>	<i>Elementos ou variáveis abordados</i>
	de inconsistências; 3 – Correção de inconsistências (contato com instituições e consultas aos repositórios); 4 – Normalização dos dados.	

Fonte: Autores.

Importante destacar que, em ambas as etapas, os dados coletados foram armazenados em um arquivo em formato CSV e submetidos a técnicas de mensuração, comparação e ajustes.

4 Resultados e discussão

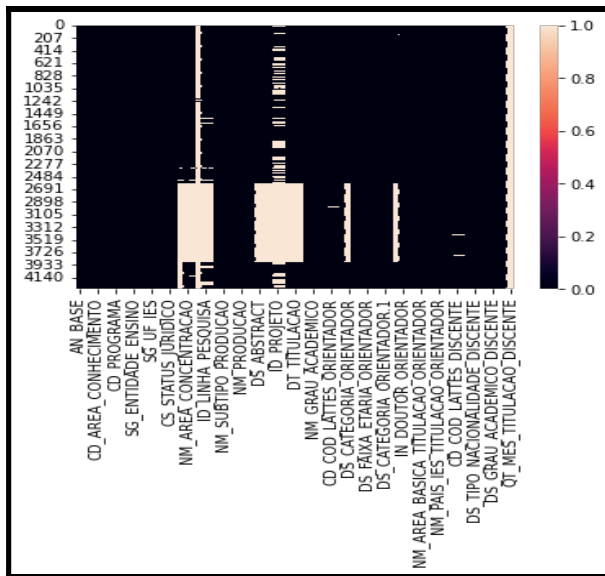
Aqui estão apresentados os resultados das duas etapas desenvolvidas nesta fase da pesquisa.

4.1 Primeira etapa: análise exploratória dos dados

Entre os anos de 2004 a 2012 e 2013 a 2021, foi constatada uma diferença entre os dados disponibilizados pela Capes e também na forma como os mesmos estão dispostos. No primeiro conjunto, verificou-se a existência de 41 colunas e, no segundo, 56. Há 16 colunas do primeiro conjunto que não aparecem no segundo, contendo campos com parte do número de CPF do discente e do docente, que aparentam ser um tipo de chave ou código para identificar os mesmos dados, e também o nome e o documento de mais três orientadores.

No segundo conjunto, encontraram-se 15 campos a mais e se forem desconsiderados os documentos e nomes dos três orientadores, o resultado seria de 21 campos. Entre os principais destaques, têm-se os campos que são códigos únicos para a Capes; se um discente se tornar docente, ficará com o mesmo código no sistema, e isso facilita a busca por uma pessoa, eliminando duplicidade. Foram adicionados campos para o “abstract” e as “keywords”, que não estavam contidos no primeiro conjunto. Algumas informações do docente, como tipo de vínculo com a instituição, a categoria do mesmo; e se este participou da banca de defesa também foram identificadas. Após a análise dos conjuntos, procedeu-se à junção deles, cujos resultados indicaram a quantidade de 4329 registros. Assim, foi gerado um arquivo em formato CSV, com as informações separadas por colunas. A Figura 1 apresenta o resultado dos dados processados, durante a análise exploratória, no ambiente Google Colab.

Figura 1. Variáveis mapeadas dos conjuntos de dados dos CTD da Capes



Fonte: Autores.

Na Figura 1, podem-se verificar alguns espaços vazios, os quais fazem menção aos campos sem informações, no conjunto de dados disponibilizados pela Capes, o que revela inconsistência no momento do preenchimento desses campos. Os campos referentes ao “ano”, “grande área do conhecimento”, “área do conhecimento”, “área de avaliação”, “código” e “nome do programa”, “sigla” e “nome da instituição”, “nome da região” onde se encontra o programa e o “estado”, “tipo de produção”, “nome da produção”, “resumo”, “palavras-chave” e “nome do docente” e “discente” foram acoplados de modo adequado, não resultando problemas de visualização. O “nome da área de concentração” e o “ID” (coluna que identifica cada campo como único no conjunto) da mesma não estavam no primeiro conjunto. Assim, foi adicionado o código 99999 e o nome de não informado para os campos vazios. Para os campos “nome da cidade” e “nome do estado onde se localiza a instituição”, foram acrescentadas funções para inserir os dados aos campos faltantes. Um ajuste também foi efetuado no campo “modalidade do programa”, sendo criada uma função para inserir os dados após verificação na Plataforma Sucupira.

A disposição dos dados no gráfico 1 confirma a importância do levantamento dos dados e a sua exposição ilustrativa, facilitando, assim, a comparação dos dados e a identificação de inconsistência. Isso está de acordo com o que apontaram Graves e Hendler (2014), ao afirmarem que o uso de tecnologias de gerenciamento de dados e técnicas estatísticas é ponto crucial na obtenção, no tratamento e na visualização dos dados, além de requerer um profundo conhecimento no assunto em que se deseja pesquisar.

O próximo passo foi inserir os dados dos docentes. Para o conjunto de dados referentes ao intervalo de 2013 a 2021, existe o “código do docente”. Então, foram abertos os arquivos por ano, e foi criada uma função para buscar alguns dados como: “categoria”, “ano de nascimento”, “faixa etária do orientador no ato da defesa”, “nacionalidade”, “país de origem”, “tipo de vínculo”, “se é doutor” e “informações da última titulação” (ano, grau, área básica, instituição de vínculo e país da instituição). Para o conjunto de dados dos anos de 2004 a 2012, como não existe o “ID”, foi feita uma busca pelos dados contidos no campo “DocumentoOrientador_1” e as informações foram validadas pelo nome do orientador.

Para os discentes, foram adicionadas as seguintes informações: “nome do país do discente”, “tipo de nacionalidade” (brasileiro ou estrangeiro), “ano de nascimento”, “faixa etária no ato da defesa”, “grau acadêmico”, “data de matrícula” e “situação de vínculo”. Semelhante aos docentes, para os anos que não existem “ID”, foram utilizados os campos “DocumentoDiscente” e o nome do mesmo, cujos dados também foram validados ao serem acoplados na planilha com os dados da produção.

Foram adicionados campos de forma manual na planilha, referentes aos links para o Currículo Lattes dos docentes e discentes, e um campo para o cálculo entre a data da situação e a data de matrícula, em meses, demonstrando, assim, o tempo utilizado pelo discente para a titulação. Com isso, a planilha final ficou com 73 campos.

Ao finalizar a etapa de levantamento e exploração dos dados, novamente, reforça-se a importância dessa atividade, como revelado por Graves e Hendler (2014). E, como também demonstrado por Sievert e Shirley (2014), para a interpretação dos resultados é aconselhado o uso de ferramentas de armazenamento, visualização e gestão de dados, pois permite que os usuários observem com mais atenção o que está sendo extraído dos dados.

4.2 Segunda etapa: preparação dos dados

Durante o levantamento dos dados, na primeira etapa, foi anotado aquilo que deveria ser ajustado, como correção de nomes, siglas, valores, inserção de informações em campos e exclusão de valores e colunas. As informações disponibilizadas nos dados abertos da Capes são inseridas pelos Programas de Pós-Graduação, e, por esse motivo, algumas informações acabam sendo inseridas com algum tipo de inconsistência, seja na duplicação ou ausência de dados, seja na digitação ou formato dos dados registrados em seus respectivos campos.

Nos campos onde se encontram o “nome” e a “sigla da instituição”, continham nomes escritos de forma diferente para a sigla, como: UNESP/MAR e UNESP-

MAR, UFPB-JP e UFPB/J.P. Assim, tais siglas foram padronizadas do seguinte formato: UNESP-MAR e UFPB-JP. Já no campo “nome da instituição”, foram encontradas situações como: UNIVERSIDADE FEDERAL DA PARAÍBA (JOÃO PESSOA) e UNIVERSIDADE FEDERAL DA PARAÍBA/JOÃO PESSOA, sendo necessário, também, realizar uma normalização. Foi constatado que não há um código de identificação para a instituição, o que dificultou a comparação, pois a instituição deve ser representada por seu nome. Desse modo, os campos deveriam ter a mesma informação.

Para o campo que contém o nome do município onde se encontra a instituição, o nome da cidade do Rio de Janeiro estava registrado dessa forma: ‘Rio de Janeiro’, com um espaço após a letra ‘o’, o que mostrava na contagem de cidades uma instituição a mais. Esse campo também não possui ID.

No campo “nome da produção”, representado pelo título da produção, foram encontrados dois estudos com o mesmo nome, ambos da mesma instituição, a FUMEC. Ao avaliar os estudos, foi possível perceber que o mesmo foi lançado na plataforma duas vezes, o que mostra um problema no sistema da Capes, que gera campos distintos para as mesmas informações, não havendo restrições para duplicidades. Percebeu-se, também, que o campo com o código do programa era diferente. Assim, foi feita uma consulta à Plataforma Sucupira, sendo verificado que um dos programas se encontra desativado. Portanto, outro problema encontrado foi a possibilidade de cadastrar um estudo em um programa desativado. Assim, foram excluídos os dois trabalhos com código de programa desativado.

Para o campo “palavras-chave”, foram encontrados 4328 estudos, sendo que 4315 eram únicos (não duplicados). Logo, identificaram-se 13 ocorrências de registros com as mesmas palavras-chave. Evidenciou-se, assim, que foram cadastradas as mesmas palavras-chave para dois estudos em dois programas diferentes. Para resolver essa inconsistência, foi procurado o estudo nos repositórios das instituições UNESP-MAR e UNB, cujas palavras-chave foram ajustadas no *dataframe*. No Programa Ciência da Informação da Universidade Federal da Paraíba/João Pessoa, foram encontrados vários registros com a frase “A SER PREENCHIDO”, no campo onde deveriam estar as palavras-chave. Para esses casos, recorreu-se ao repositório institucional, devido à necessidade de recuperar os estudos para identificar e inserir os dados corretos no *dataframe*.

Foi realizada, também, uma contagem das palavras no campo “palavras-chave”, sendo localizados alguns campos com cinco, sete ou dez palavras. A fim de solucionar esse problema, também foi necessário localizar os nove estudos no repositório da instituição, e ajustar os dados no *dataframe*. Após essas validações,

foi verificado que as palavras-chave eram separadas por vários tipos de caracteres e/ou números. Assim, procedeu-se à normalização e todas as palavras-chave foram separadas pelo caractere *underline* (), com o objetivo de facilitar a abertura do arquivo CSV como planilha para possíveis validações.

No campo “resumo”, foram encontrados 13 registros com o mesmo valor, nove deles com o termo “A SER PREENCHIDO” e nove com a palavra “FALTA”. Para esses registros, como realizado nos casos anteriores, recorreu-se aos repositórios das instituições, para inserção dos dados no *dataframe*. O campo que contém o endereço do estudo na Plataforma Sucupira apresentava todas as letras em maiúsculo no conjunto de dados de 2013 a 2017. O campo foi também normalizado para que fosse possível acessar o endereço na internet. Para os campos “ID” e “nome do projeto”, há poucas informações, sendo 2396 campos nulos, o que correspondeu a mais da metade dos dados.

Os campos “data de matrícula” e “titulação” possuíam diferentes formatos de data, sendo necessário normalizá-los, criando funções para unificar a forma como as datas são demonstradas e, assim, fosse possível realizar cálculos com as mesmas. Alguns campos, como o “número de páginas do trabalho”, não registravam informações coerentes, por exemplo, em alguns casos, o valor era zero ou um. Foram localizados vários registros nessa situação, o que dificultou, assim, realizar uma comparação por páginas entre pesquisas de Mestrado e/ou de Doutorado.

No conjunto de dados referente aos docentes, havia alguns campos com informações que não pareciam precisas, como é o caso do “ano de nascimento do orientador” com um registro datado de 1900. Para essa situação, foram solicitadas informações sobre o professor orientador junto ao Programa, que não soube informar, por se tratar de uma professora estrangeira (espanhola). Na tentativa de resolver essa inconsistência, após uma busca avançada no Google, foi encontrada a instituição de vínculo da professora e o seu e-mail. Mediante mensagem encaminhada para o contato encontrado, a resposta foi devolvida com sucesso e o dado gravado no *dataframe*.

Os campos “sigla” e “nome da instituição da última titulação do orientador” também apresentam problemas, não havendo uniformidade para o cadastro dessas informações. Desse modo, o cadastramento realizado pelos programas dificulta a compreensão, com vários campos duplicados para a sigla e com nomes diferentes. A princípio, eram 127 siglas únicas e 141 nomes únicos de instituições; após a comparação e a normalização, o conjunto foi reduzido a 115 siglas únicas referentes aos dois campos.

O conjunto de dados referente aos discentes contém poucas informações únicas, como o “ano de nascimen-

to” e a “faixa etária”, no momento da defesa. Dois campos que foram utilizados são os da “data de matrícula” e da “titulação do discente”; neste último, constam datas em diferentes formatos, sendo necessário, portanto, realizar ajustes. A maior dificuldade é a da confiança nos dados informados pelo Programa quanto às datas, pois, após a normalização, foi subtraída a quantidade de meses que o discente usou para a realização do Mestrado e/ou Doutorado. O resultado dessa intervenção foram valores que, possivelmente, não refletem a realidade. Como exemplos dessa ocorrência, citam-se casos de menos de dez meses para a realização de Doutorado e Mestrado e casos de 120 meses para realização de Mestrado. Como mencionado, foi enviada uma solicitação de esclarecimentos para as universidades, a partir de datas que pareciam estar distorcidas, como o caso de conclusão de cursos com menos de 12 meses e para a confirmação de dados maiores que 60 meses, respostas ainda não devolvidas pelas instituições.

A falta de clareza e de padronização e/ou a ausência de dados e a incompletude deles são problemas que geram prejuízos não apenas para a área de conhecimento que está sendo analisada, mas também para o desenvolvimento da ciência, e, por conseguinte, acarretam prejuízos para toda a sociedade. Perovich, Wylie e BonGiovanni (2020) corroboram essa reflexão, destacando que o mapeamento coerente dos dados de pesquisa pode fundamentar a tomada de decisões e melhorar a comunicação em uma ampla variedade de setores, incluindo saúde, finanças, meio ambiente, educação e políticas públicas. Desse modo, destacam os autores, disponibilizar, com consistência, os dados abertos das instituições é fundamental para promover a transparência, a responsabilidade e a compreensão das informações científicas, ajudando as pessoas a se beneficiarem com as conquistas e a monitorarem os avanços da ciência e das instituições.

Os resultados ora alcançados também vão ao encontro do que defenderam Barcellos *et al.* (2017) a respeito do distanciamento que se estabelece entre instituições e sociedade, quando as informações de acesso público não são apresentadas com coerência e nem representam a realidade institucional. De fato, além de não haver o reconhecimento dos investimentos públicos, segundo Kassen (2021), as inconsistências dos dados impedirão o seu reuso, visto que os interessados não terão confiança no trabalho institucional, gerando falta de aplicação das pesquisas e aumento de custos.

A área específica da Ciência da Informação necessita de reforço em suas estratégias de divulgação, tendo em vista adquirir maior relevância, reconhecimento e valorização. Portanto, a disponibilização inadequada dos dados dessa área pode trazer prejuízos quanto a sua visibilidade, sobretudo internacional. Como foi descoberto no estudo de Araújo e Valentim (2019, p. 254),

percebe-se que a pesquisa brasileira em Ciência da Informação “[...] tem evoluído ao longo dos anos, embora ainda carece de visibilidade internacional [...]”, ação essa que pode ser desenvolvida pelos programas de Pós-Graduação, no sentido de fortalecer a “[...] divulgação do pensamento científico em Ciência da Informação no contexto internacional”.

Quanto à necessidade de atenção no cadastramento dos dados, é importante reforçar a estratégia de se oferecer treinamentos para os programas, sobretudo quanto ao uso das ferramentas tecnológicas e a necessidade de normalização dos dados registrados (MARTIN, 2014). É provável que essa iniciativa enalteça o papel da curadoria dos dados de pesquisa, processo esse que, segundo Resende e Bax (2020), não pode limitar-se apenas à preservação, mas, principalmente, à gestão desses dados, com o propósito de ampliar sua utilização e, dessa forma, poupar recursos e contribuir para o avanço de novas descobertas científicas na área da Ciência da Informação.

5 Considerações finais

Neste estudo a dimensão aplicada da Organização do Conhecimento se evidencia no processo de modelagem que envolve a preparação, armazenagem, organização e estruturação de dados contidos em documentos (dados da Capes) que ainda não receberam tratamento científico, tratando-se de uma pesquisa documental. Inserida no campo da Ciência da Informação, este estudo passa pelos processos de coleta, tratamento e organização dos dados que foram registrados e são passíveis de transmissão, uma vez que se propõe apresentá-los de modo estruturado e com significado dentro do contexto trabalhado.

Considerando o problema e o objetivo definidos para este estudo, foram executadas atividades iniciais de análise exploratória de dados (primeira etapa), quando houve o levantamento de dados no âmbito do PDA da Capes, e a preparação dos dados (segunda etapa), com o seu tratamento para eliminação de inconsistências.

Neste estudo foi apresentada a estruturação dos dados de pesquisa gerados na área da Ciência da Informação. A verificação dos dados disponibilizados pela Capes evidenciou problemas relacionados à digitação, à falta de padronização, à ausência de registros e também de duplicidade de dados inseridos nos respectivos campos de registro. Verificou-se que a Capes não informa quais campos são obrigatórios na coleta das informações. Com isso, não há como constatar a origem do problema, ou seja, se está no preenchimento dos campos pelas instituições ou se está no próprio sistema da Capes.

A partir das inconsistências localizadas, torna-se possível tecer algumas primeiras proposições de melhorias.

Tais sugestões são indicadas tanto para a Capes quanto para os responsáveis pelos cadastros, como os programas, por se tratar de um trabalho colaborativo, que demanda participação de todas as instituições responsáveis pelas pesquisas.

Para a Capes, sugere-se a criação de códigos únicos e IDs específicos para cidades, estados, instituições, nacionalidade e países, com a finalidade de impedir equívocos ou esquecimentos durante o preenchimento dos dados no sistema, com restrições de obrigatoriedade. Também poderiam ser criados campos obrigatórios com valores mínimos para inserção no sistema, como palavras-chave, resumo, *abstract* e *keywords*, de modo que os programas não deixem de inserir os dados solicitados. Outra sugestão para evitar erros seria modificar o campo destinado ao registro sobre a banca de avaliação do estudo. Nesse caso, o campo deixaria de ser livre (com registro dos nomes dos participantes) e passaria a conter um cadastro único, com uma chave, como o CPF, o que evitaria, assim, o registro de informações duplicadas.

Para os programas, sugere-se um trabalho mais cauteloso, com o uso de ferramentas específicas para inserção dos dados, como também a oferta de treinamento para os operadores, além de mais atenção no momento de se fornecer as informações. A demora ou ausência no fornecimento de informações foram constatadas, no decorrer desta pesquisa, quando foram solicitados esclarecimentos, por e-mail, sobre as inconsistências constatadas. Em alguns casos, foi preciso recorrer à ferramenta <https://falabr.cgu.gov.br/> para se conseguir algum retorno sobre o que foi indagado.

Os resultados das duas etapas executadas neste estudo possibilitam a execução da próxima etapa, indicada como estudo futuro. Nela, pretende-se trabalhar a visualização dos dados, finalizando este estudo bibliométrico, que apresente, de modo visual, os dados de pesquisa da Ciência da Informação à luz dos dados abertos da Capes. Por fim, destaca-se que mesmo com os dados brutos, sem qualquer tipo de intervenção, já é possível desenvolver análises quantitativas em vários campos, como por universidade, por região, por estado, por ano, dentre outras possibilidades.

Referências

ARAÚJO, Carlos Alberto Ávila; VALENTIM, Marta Lígia Pomim. A Ciência da Informação no Brasil: mapeamento da pesquisa e cenário institucional. **Bibliotecas e Anales de Investigación**, [S. l.], v. 15, n. 2, p. 232-259, jul./dez. 2019.

ATTARD, Judie *et al.* A systematic review of open government data initiatives. **Government Information Quarterly**, [S. l.], v. 32, n. 4, p. 399-418, abr. 2015.

BARCELLOS, Raissa *et al.* Transparency in practice: using visualization to enhance the interpretability of open data. In: ANNUAL INTERNATIONAL CONFERENCE ON

DIGITAL GOVERNMENT RESEARCH, 18., 2017, [S. l.]. **Proceedings** [...]. [S. l.]: [S. n.], 2017. p. 139-148.

BERTIN, Patrícia Rocha Bello *et al.* A parceria para Governo Aberto como plataforma para o avanço da Ciência Aberta no Brasil. **Transinformação**, Campinas, v. 31, 2019.

BROOKES, B.C. The foundations of information science. Part I. Philosophical aspects. **Journal of Information Science**, [S. l.], v. 2, p. 125-133, 1980.

BROUGHTON, Vanda *et al.* Knowledge organisation: report of working group 7. In: KAJBERG, L.; LØRRING, L. (eds.). **European curriculum reflections on education in library and information science**. Copenhagen: Royal School of Library and Information Science, 2005.

BRUGGER, Jérôme *et al.* Current barriers to open government data use and visualization by political intermediaries. In: CONFERENCE FOR E-DEMOCRACY AND OPEN GOVERNMENT (CeDEM), 1., 2016, [S. l.]. **Proceedings** [...]. [S. l.]: [S. n.], 2016. p. 219-229.

BUCKLAND, M.K. Information as a thing. **Journal of the American Society for Information Science (JASIS)**, [S. l.], v. 45, n. 5, p. 351-360, 1991.

BUZATO, M. E. K.; TORRES, C. C. Dadificação e transdisciplinaridade nos estudos do letramento: o jornalismo de dados como contexto de pesquisa. **Revista da Anpoll**, Brasília-DF, v. 1, n. 49, p. 128-141, 2019.

CAPURRO, Rafael. What is information science for? A philosophical reflection. In: VAKKARI, Pertti; CRONIN, Blaise (eds.). **Conceptions of Library and Information Science: historical, empirical and theoretical perspectives**. London: Taylor Graham, 1992. p. 82-96.

CARDOSO, Ana Maria Pereira. Pós-Modernidade e informação: conceitos complementares? **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 63-79, jan./jul. 1996.

COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Plano de dados abertos**. Brasília, set. 2022. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/documentos/PlanodeDadosAbertosdaCAPES202022.pdf>. Acesso em: 28 fev. 2023.

CUCCINIELLO, Maria; NASI, Greta; VALOTTI, Giovanni. Assessing transparency in government: rhetoric, reality and desire. In: HAWAII INTERNATIONAL CONFERENCE ON SYSTEM SCIENCES, 45., 2012, [S. l.]. **Proceedings** [...]. [S. l.]: [S. n.], 2012. p. 2451-2461.

DAHLBERG, Ingetraut. A referent-oriented, analytical concept theory of Interconcept. **International Classification**, Frankfurt, v. 5, n. 3, p. 122-151, 1978.

DAHLBERG, Ingetraut. Knowledge organization: a new science? **Knowledge Organization**, [S. l.], v. 33, n. 1, p. 11-19, 2006.

FONSECA, E. N. da. Bibliografia estatística e bibliometria: uma reivindicação de prioridades. **Ciência da Informação**, Brasília, v. 2, n. 1, p. 5-7, 1973.

FREITAS, José Antônio de Carvalho *et al.* O ecossistema de dados abertos do governo federal: composição e desafios. **Ciência da Informação**, Brasília, v. 47, n. 2, 2018.

GASCÓ-HERNÁNDEZ, Mila *et al.* Promoting the use of open government data: cases of training and engagement. **Government Information Quarterly**, [S.l.], v. 35, n. 2, p. 233-242, jan./jun. 2018.

GRAVES, Álvaro; HENDLER, James. A study on the use of visualizations for open government data. **Information Polity**, [S.l.], v. 19, n. 1-2, p. 73-91, jan./fev. 2014.

HJØRLAND, Birger. Epistemology and the Socio-Cognitive perspective in Information Science. **Journal of the American Society for Information Science and Technology**, [S.l.], v. 53, n. 4, p. 257-270, Feb. 2002.

HJØRLAND, Birger. What is knowledge organization (KO)? **Knowledge Organization**, [S.l.], v. 35, n. 3/2, p. 86-111, 2008.

KASSEN, Maxat. Understanding motivations of citizens to reuse open data: open government data as a philanthropic movement. **Innovation**, [S.l.], v. 23, n. 1, p. 44-70, jan./jun. 2021.

MARTIN, Chris. Barriers to the open government data agenda: Taking a multi-level perspective. **Policy & Internet**, [S.l.], v. 6, n. 3, p. 217-240, mar. 2014.

OSAGIE, Edobor *et al.* Usability evaluation of an open data platform. In: ANNUAL INTERNATIONAL CONFERENCE ON DIGITAL GOVERNMENT RESEARCH, 18., 2017, [S.l.]. **Proceedings** [...]. [S.l.]: [S.n.], 2017. p. 495-504.

PAO, M. L. **Concepts of information retrieval**. Englewood, Colorado: Libraries Unlimited Inc., 1989.

PEROVICH, Laura; WYLIE, Sara Ann; BONGIOVANNI, Roseann. Chemicals in the Creek: designing a situated data physicalization of open government data with the community. **Transactions on Visualization and Computer Graphics**, [S.l.], v. 27, n. 2, p. 913-923, jul./dez. 2020.

RESENDE, Liliane Chaves de; BAX, Marcello Peixoto. A curadoria de dados científicos na Ciência da Informação: levantamento do cenário nacional. **AtoZ: novas práticas em informação e conhecimento**, Curitiba, v. 9, n. 1, p. 94-110, jan./jun. 2020.

RODRIGUES, Adriana Alves. **Visualização de dados no cenário da Data Science: práticas de laboratórios de**

inovação guiados por dados. 2019. 264f. Tese (Doutorado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação do Centro de Ciências Sociais Aplicadas da Universidade Federal da Paraíba. João Pessoa: UFPB/CCSA, 2019.

RUIJER, Erna; MARTINIUS, Evelijn. Researching the democratic impact of open government data: a systematic literature review. **Information Polity**, [S.l.], v. 22, n. 4, p. 233-250, abr. 2017.

SARACEVIC, T. Information Science: origin, evolution and relations. In: VAKKARI, Pertti, CRONIN, B. (eds.). **Conceptions of Library and Information Science; historical, empirical and theoretical perspectives. THE INTERNATIONAL CONFERENCE FOR THE CELEBRATION OF 20TH ANNIVERSARY OF THE DEPARTMENT OF INFORMATION STUDIES OF UNIVERSITY OF TAMPERE, 1991, Proceedings** [...] Finland. London, Los Angeles: Taylor Graham, 1992. p. 5-27.

SARACEVIC, Tefko. **Introduction to Information Science**. Chatham, Nova Jersey: R.R. Bowker Company, 1970.

SIEVERT, Carson; SHIRLEY, Kenneth. LDavis: A method for visualizing and interpreting topics. In: WORKSHOP ON INTERACTIVE LANGUAGE LEARNING, VISUALIZATION, AND INTERFACES, June 27, 2014, Baltimore, Maryland, USA. **Proceedings** [...] Baltimore, Maryland, USA: Association for Computational Linguistics, 2014. p. 63-70.

TAGUE-SUTCLIFFE, J. An introduction to informetrics. **Information Processing & Management**, [S.l.], v. 28, n. 1, p. 1-3, 1992.

TESSER, Gelson João. Principais linhas epistemológicas contemporâneas. **Educar**, Curitiba, v. 10, p. 91-98, dez. 1994.

YU, Harlan; ROBINSON, David. The new ambiguity of open government. **UCLA L. Rev. Discourse**, [S.l.], v. 59, n. 1, p. 178, jan./jun. 2011.

ZIMMERMANN, Hans-Dieter; PUCIHAR, Andreja. Open innovation, open data and new business models. In: I OPEN DATA AND NEW BUSINESS MODELS, 1., 2015, [S.l.]. **Proceedings** [...]. [S.l.]: [S.n.], 2015. p. 1-5.

O lugar da normalização na organização do conhecimento: análise da coleta de referências em teses e dissertação de um programa de pós-graduação em Ciência da Informação

The place of normalization in the organization of knowledge: analysis of the collection of references in theses and dissertations of a Graduate Program in Information Science

Michely Jabala Mamede Vogel

Universidade Federal Fluminense, Rua Lara Vilela, 126 Niterói/RJ, michelyvogel@id.uff.br

Resumo

A normalização é função da Biblioteconomia e Ciência da Informação e impacta na Organização e Recuperação da Informação e nos Estudos Métricos da Informação, especialmente na análise de citações. Objetiva-se levantar os impactos que a falta de normalização pode gerar nos processos de organização do conhecimento, em especial no que tange à análise de citações, por meio da descrição do processo de coleta das referências utilizadas nas teses e dissertações do programa de pós-graduação em Ciência da Informação da Universidade Federal Fluminense, defendidas entre 2011 e 2020. Como aporte metodológico, recorreu-se a um estudo descritivo, à pesquisa bibliográfica para identificação de conceitos e à análise documental. Os resultados apontam os desafios da falta de normalização, como ausência de metadados voltados às referências, falta de ferramentas para sua exportação e problemas de identificação e localização das referências utilizadas nas teses e dissertações analisadas. Conclui-se que é preciso investir em ensino de normalização, necessita-se de maiores investimentos nas instituições de pesquisa pública para aquisição ou desenvolvimento de ferramentas de análise bibliométrica.

Palavras-chave: Normalização. Ciência da Informação e Biblioteconomia. Organização do Conhecimento.

Abstract

Normalization is a function of Library and Information Science and impacts on Information Organization and Retrieval and on Information Metrics Studies, especially on citation analysis. The objective is to raise the impacts that the lack of normalization can generate in the processes of knowledge organization, especially about the analysis of citations, through the description of the process of collecting references used in theses and dissertations from graduate program in Information Science of Universidade Federal Fluminense, defended between 2011 and 2020. As a methodological contribution, a descriptive study was used, bibliographic research to identify concepts and document analysis as well. The results point out the challenges of the lack of normalization, such as the absence of metadata related to the references, lack of tools for its export and problems of identification and location of the references used in the analyzed theses and dissertations. It is concluded that it is necessary to invest in normalization teaching, greater investments are needed in public research institutions for the acquisition or development of bibliometric analysis tools.

Keywords: Normalization. Library and Information Science. Knowledge Organization.

Resumen

La normalización es una función de la bibliotecología y la ciencia de la información e impacta en la organización y recuperación de la información y en los estudios de métricas de la información, especialmente en el análisis de citas. El objetivo es plantear los impactos que la falta de estandarización puede generar en los procesos de organización del conocimiento, especialmente en lo que se refiere al análisis de citas, a través de la descripción del proceso de recolección de referencias utilizadas en tesis y disertaciones del programa de posgrado en Ciencias de la Información de la Universidad Federal Fluminense, defendida entre 2011 y 2020. Como aporte metodológico se utilizó un estudio descriptivo, investigación bibliográfica para identificar conceptos y análisis de documentos. Los resultados apuntan los desafíos de la falta de estandarización, como la ausencia de metadatos relacionados con las referencias, falta de herramientas para su exportación y problemas de identificación y localización de las referencias utilizadas en las tesis y disertaciones analizadas. Se concluye que es necesario invertir en la normalización de la docencia, se necesitan mayores inversiones en las instituciones públicas de investigación para la adquisición o desarrollo de herramientas de análisis bibliométrico.

Palabras clave: Normalización. Ciencias de la Información y Bibliotecología. Organización del conocimiento.

1 Introdução

As teses e dissertações defendidas num programa de pós-graduação (PPG) constituem um corpus de interesse para conhecer suas bases intelectuais. Elas representam o processo de evolução e concepção de um pesquisador, estão presentes no fluxo da comunicação científica e são consideradas como fontes de informação primária. Sua importância reside no fato de trazerem, via de regra, contribuições inovadoras, ocupando um relevante papel na produção científica (WITTER; PÉCORA, 1997). Ademais, para Grácio e Oliveira (2017), o conjunto de referências citadas na literatura científica pode representar uma comunidade discursiva.

Atualmente, a maioria das universidades brasileiras incentiva a divulgação das pesquisas produzidas na instituição, apesar de não serem publicadas por casas editoriais, como livros e periódicos. Se antes as teses e dissertações ficavam ao alcance de um número restrito de pessoas, hoje, com as possibilidades da publicação eletrônica, seu acesso e disponibilização por meios digitais rompeu essa barreira sendo, inclusive, item considerado na avaliação dos PPG pela Coordenação para o Aperfeiçoamento do Pessoal de Nível Superior (CAPES). Para tanto, foram desenvolvidos páginas, repositórios e bibliotecas digitais para reuni-las, como, por exemplo, a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), coordenada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a *Networked Digital Library of Theses and Dissertations* (NDLTD), em âmbito internacional, patrocinada pela UNESCO, bem como os repositórios institucionais de universidades. Assim, as teses e dissertações podem ser acessadas como fonte de informação e podem ser utilizadas para estimular novas pesquisas e para aperfeiçoar outras já existentes.

Ao analisar as referências num conjunto de 145 teses e dissertação do Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense (PPGCI/UFF), referentes a um período de dez anos, foram identificados diversos desafios para a área da Organização do conhecimento. Trata-se de questões relacionadas ao acesso ao material e à forma como eles são normalizados em sua versão final.

Diante de tal coleta, e partindo do pressuposto que boa parte dos mestrados e doutorandos, assim como de orientadores e membros de bancas de Ciência da Informação ou são bibliotecários de formação ou convivem de perto com tais profissionais, aponta-se o seguinte problema: por que a normalização de referências bibliográficas, apesar de ser uma área da Ciência da Informação e da Biblioteconomia, não é seguida à risca nos produtos dos PPG dessa área?

Na tentativa de aproximação com tal questão, o objetivo proposto é levantar os impactos que a falta de normalização pode gerar nos processos de organização

do conhecimento, em especial no que tange à análise de citações. Para tanto, será descrito o processo de coleta das referências utilizadas nas teses e dissertações PPGCI/UFF defendidas entre 2011 e 2020, destacando os maiores desafios encontrados.

Como aporte metodológico, recorreu-se a um estudo descritivo, ao relatar os passos de uma pesquisa realizada entre 2021 e 2022, à pesquisa bibliográfica para identificação de conceitos e à análise documental. Esta “compreende a identificação, a verificação e a apreciação de documentos para determinado fim” (MOREIRA, 2006, p. 271), sendo na maioria das vezes qualitativa, ao verificar o conteúdo do material analisado. Dessa forma, ao se observar tanto o processo de coleta como as referências coletadas, procede-se a uma análise de tais fontes como documentos (e não como bibliográficas).

Após esta introdução apresentamos uma revisão sobre normalização, e uma revisão sobre análise de citações e sua utilização para mapeamento de teses e dissertações, descreveremos a metodologia e o processo utilizado para recuperar as referências das teses e dissertações do PPG em questão, discutiremos a questão da normalização como questão de organização da informação e apresentaremos nossas conclusões.

2 A Normalização

Segundo o *Diccionario de Organización del Conocimiento*, Normalização é:

Submissão de um conjunto de fenômenos, situações ou objetos a um modelo, regulamento ou padrão. A padronização se expressa atualmente em toda a gama de atividades humanas especializadas (ciência, técnica, economia, comércio, indústria, etc.), e busca a padronização, simplificação e coordenação de elementos. Consequentemente, isso permite a redução de custos e tempo. (BARITÉ et al, 2015, p. 111, tradução nossa).

De acordo com Meadows (1999), a falta de normalização bibliográfica gerou documentos incompletos, que acabavam por omitir dados importantes das pesquisas, e que muitas vezes geravam a importância de localizar informações e identificar pesquisas em desenvolvimento. Por isso, a “a normalização da produção científica é essencial por dinamizar a divulgação do que é gerado pela ciência” (CRESPO; RODRIGUES, 2011, p. 38)

Segundo Sperandio (2019) os trabalhos de conclusão de curso (TCC) no âmbito da graduação, as monografias nos cursos de especialização, as dissertações no mestrado e as teses no doutorado ou pós-doutorado e ainda os artigos de periódicos são os tipos de trabalho que no Brasil devem ser normalizados segundo as normas da Associação Brasileira de Normas Técnicas (ABNT). Nos programas de pós-graduação em Ciência da Informação e nos cursos de graduação em

Biblioteconomia e Documentação nacionais essa costuma ser a regra.

Embora as normas sejam de ordem orientativa, sua aplicação garante a organização do trabalho, facilita sua recuperação, sua avaliação e seu mapeamento bibliométrico.

Historicamente, é possível encontrar aplicação de tentativas de normalização desde a pré-história, mas foi durante a revolução industrial que ela “começou a ser considerada importante, principalmente com a criação de sistemas métricos materializados, como o metro e o quilograma (SILVA, 2003 *apud* SPERANDIO, 2019, p. 70). Para Crespo e Rodrigues (2011), a aplicação de normas nos trabalhos científicos é responsável por facilitar, por um lado, sua divulgação posterior, por adequar-se prontamente aos padrões e exigências de periódicos, repositórios e bases de dados, e por outro facilitar as trocas no interior das comunidades científicas, uma vez que documentos normalizados se enquadram nos critérios de cientificidade e qualidade exigidos no fazer científico.

Para Curti e Boccato (2005, p. 96), a normalização “facilita as operações documentais e diminui o custo e o tempo necessário para realizá-las, viabilizando o intercâmbio e a recuperação de informações”. Ou seja, pode ser trabalhoso garantir aplicar as normas ao se redigir um trabalho, no entanto, isso deixa de ser necessário a cada vez que um trabalho for ser citado ou incluído numa base de dados, ganhando tempo nestas etapas que se repetem, em geral, mais de uma vez.

Ao trabalhar com a comunicação científica, vários são os desafios encontrados para coletar e organizar dados na ótica dos estudos bibliométricos: fontes variadas, tipologias diversas, idiomas diferentes. Na área dos estudos métricos da informação, a literatura identifica problemas: esta área já enfrenta a questão de lidar com resultados vastos muitas vezes sem a infraestrutura tecnológica adequada, e tal desafio é potencializado pela “baixa qualidade da normalização documentária das citações e referências presentes em publicações de cunho científico, como teses, dissertações, artigos etc”. (SOUSA, 2020, p. 735). De acordo com Sousa (2020), pesquisadores que atuam com Bibliometria e análise de citação indicam justamente a ausência de normalização como umas das principais dificuldades para o bom andamento das análises bibliométricas. Como consequências, são apontados equívocos na mensuração das citações e das frentes de pesquisa, como por exemplo: erros na atribuição de autoria, dispersão de contagem de citações, levando um autor a receber menos ou mais indicações do que o correto, e a geração de classificações falsas tanto de frentes de pesquisa como de pesquisadores outliers.

Corroborando e complementando a questão, Silva, Hayashi e Hayashi (2011, p. 125), identificaram como problemas para a coleta de informações bibliométricas a

ausência de consistência nas bases de dados, como falta de estrutura, problemas de atualização e registros incoerentes, ausência de recursos voltados aos mapeamentos bibliométricos em softwares, como por exemplo não oferecer as referências de cada trabalho em formato exportável e editável, mas somente em pdf assim como a ausência de informações relevantes em documentos, como palavras-chave, datas, entre outros.

Defende-se que tais questões são de ordem da organização de conhecimento: metadados disponíveis, formatos padronizados de preenchimentos e respeito aos conteúdos obrigatórios nas referências. Como alerta Sousa (2020, p. 748), tais inconsistências “podem implicar em erros que afetarão os resultados e colocarão em xeque a credibilidade do estudo e, também, a do pesquisador”. Por isso, nos passos metodológicos das pesquisas bibliométricas dever-se-ia seguir o mesmo rigor típico da Organização da Informação nos processos de representação documental. Souza (2020) ainda destaca que regras de catalogação, bases de dados de bibliotecas e análise documentos são fundamentais tanto para a “confirmação e uniformização de elementos e dados, mas também como aporte teórico para a resolução de desafios que surgiram durante a pesquisa.” (SOUSA, 2020, p. 749).

Afinal, conforme asseverado por Lubisco (2008), é justamente a normalização adequada dos trabalhos que oferece as condições ideais para sua indexação e recuperação com qualidade, promovendo dessa forma os processos da comunicação científica. Trata-se de interesse duplo dos pesquisadores: por um lado a certeza de que suas publicações terão condições de serem inseridas em bases de dados de qualidade e por outro a possibilidade de enriquecimento de seu currículo (LUBISCO, 2008).

Como lembra Sant’Anna (2019), o bibliotecário é o profissional competente para a realização da normalização bibliográfica, dentre tantas outras possibilidades de trabalho. Com tal padronização, assegura-se conforto e comodidade nos processos de acesso, transferência e uso da informação científica tanto para usuários como pesquisadores, especialmente da análise de citações.

Do ponto de vista das habilidades e competências necessárias ao exercício da normalização, é possível identificar no fazer da Biblioteconomia e Ciência da Informação as disciplinas chave, conforme o quadro 1.

Quadro 1. *Principais disciplinas que contribuem com o exercício da normalização bibliográfica*

<i>Nome da Disciplina</i>	<i>Principal objetivo</i>
Normalização da informação	Capacitar o aluno a reconhecer os diversos documentos bibliográficos e as técnicas e métodos de normalização propostos por órgãos normalizadores, como a ABNT.

<i>Nome da Disciplina</i>	<i>Principal objetivo</i>
Fontes de informação	Despertar a capacidade de reconhecimento e identificação das características principais das diversas fontes de informação que registram conhecimento.
Representação descritiva	Demonstrar o uso das técnicas de catalogação, tendo em vista o registro de documentos em acervos bibliográficos nos mais diferenciados suportes.
Formação e desenvolvimento de coleções	Demonstrar o uso das técnicas de catalogação, tendo em vista o registro de documentos em acervos bibliográficos nos mais diferenciados suportes.
Editoração	Caracterizar os aspectos culturais, industriais e comerciais do mercado editorial; compreender os diversos elementos relacionados ao processo editorial ao longo da história e a importância do bibliotecário na divulgação e difusão do conhecimento.
Métodos e técnicas de pesquisa	Levar os alunos a conhecerem o caminho de construção de um objeto de pesquisa e a terem uma visão geral dos métodos de preparação de trabalhos/projetos de pesquisa científica.
Serviços de recuperação da informação	Recuperação da informação: conceituação e objetivos. Atuação profissional de referência. Processo de referência. Estratégia de busca.

Fonte: Sant'Anna (2019, p. 226.)

Quase todas as disciplinas referem-se ao campo da Organização do Conhecimento, ou nela encontram suporte para se desenvolverem. Por isso, defende-se neste trabalho que a normalização deva ser preocupação da Organização do Conhecimento.

3 Análise de Citações de Teses e Dissertações

A literatura científica e as atividades dos cientistas constituem um interessante objeto de estudo para a Ciência da Informação. Estudar as bases intelectuais de um pesquisador ou de um grupo de pesquisadores permite identificar suas influências e bases teóricas, epistemológicas e metodológicas (ARAÚJO; BUFREM, 2021). Este tipo de estudo envolve a análise e descrição de um conjunto de estudo composto pelas referências citadas pelos autores, oferecendo um panorama das bases teóricas e suas implicações GRÁCIO; OLIVEIRA, 2017).

A análise de citações é reconhecida como método efetivo de identificar as influências intelectuais de um campo, uma vez que permite estabelecer o caminho para associar um estudo a um autor mostrando evidências de

conexões de um campo e mesmo fora dele (SMIRAGLIA, 2014), e seu objeto são as citações e as referências registradas nos textos científicos para fins de avaliação das atividades e práticas de produção, comunicação e uso do conhecimento (SILVEIRA: CAREGNATO, 2017). Trata-se de uma técnica de coleta e análise de dados que visa identificar fenômenos na comunicação científica, tais como o impacto de publicações, padrões de uso de documentos por pesquisadores e redes cocitação. Por sua natureza dinâmica e social e, especialmente, por se valer da literatura científica (HJØRLAND, 2013) permite identificar os grupos de cientistas, suas publicações, os autores com mais impacto, seus paradigmas e procedimentos metodológicos (GLÄNZEL, 2003; GRÁCIO, 2016).

A análise de citações permite gerar alguns produtos bibliométricos (GUEDES, 2012):

- Frente de pesquisa, representada por um conjunto de autores que se citam na literatura recente;
- Fator de impacto ou de imediatismo, calculada a partir da concentração de citações a um determinado artigo, em documentos publicados nos últimos quinze anos;
- Acoplamento bibliográfico, procedimento de análise retrospectiva de citações que consiste na união de artigos citantes, citando o(s) mesmo(s) documento(s);
- Análise de cocitação, procedimento de análise prospectiva de citações, que consiste na união de artigos, citados pelo(s) mesmo(s) documento(s);
- Obsolescência da literatura, estimada pela análise do declínio do uso da literatura, em um período determinado;
- Vida média, calculada a partir da razão de obsolescência e da razão de crescimento, de um determinado corpus da literatura;
- Lei do Elitismo, identificada a partir das citações em uma área de assunto, buscando demonstrar que somente uma parte da literatura existente é estreitamente entrelaçada e identifica, e esse pequeno grupo de artigos a frente de pesquisa de uma área do conhecimento.

Inicialmente voltada à observação da distribuição de artigos e o impacto de periódicos em determinada área (WALTMAN, 2016), a análise de citação passou a ser utilizada também para medir a influência de pesquisadores para formação de redes de colaboração científica e mapeamento de tendências de pesquisa (HOU, 2017) e a utilizar outras fontes, incluindo teses e dissertações para identificar padrões de citação e as características dos documentos citados por

pesquisadores de diferentes áreas do conhecimento (MILLER, 2011).

Conforme levantamento feito na BRAPCI (Base de Dados Brasileira de Periódicos em Ciência da Informação), a análise de citações aplicada a teses e dissertações em com PPG em Ciência da Informação (PPGCI) foi encontrada em três trabalhos: Rodrigues (1982) fez suas análises em cima das dissertações do PPGCI do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) defendidas entre 1972 e 1979, buscando identificar a distribuição e incidência dos autores citados, vida média da literatura, frente e elite de pesquisa, autores mais citados e tipologias mais utilizadas. Menezes e Oddone (2014) observaram as citações de diversos PPGCI cujas teses e dissertações defendidas entre 2007 e 2009 constavam da BDTD, buscando identificar a citação feitas a livros. Custódio, Lima e Bufrem (2017) observaram especificamente as redes de cocitações a partir das referências de teses e dissertações do PPGCI da Universidade de São Paulo (USP), defendidas entre 2013 e 2015. No entanto, pouco é apresentado em relação aos passos da efetiva coleta das referências e sua sistematização, especialmente dos potenciais problemas que podem impactar na qualidade dos dados.

4 Metodologia

Este é um estudo descritivo, que buscou analisar aspectos da normalização de referências nas teses e dissertações do PPGCI/UFF. Observando-as como documentos (pesquisa documental) e não como fonte bibliográfica (que caracterizaria o estudo bibliográfico).

A busca foi feita no repositório institucional da UFF (RI/UFF) e na BDTD. Embora as teses e dissertações tenham espaço garantido no RI/UFF, sua busca não se mostrou intuitiva. Dessa forma, um primeiro passo foi consultar a secretaria do PPGCI/UFF em busca dos títulos defendidos entre 2011 e 2020 e que representam os resultados dos dez primeiros anos de funcionamento do PPG. Vale lembrar que o PPGCI/UFF iniciou suas atividades no ano de 2009, tendo as primeiras defesas de dissertações de mestrado em 2011, e abriu o doutorado no ano de 2015, apresentando as primeiras defesas em 2019.

Nem na BDTD nem no RI/UFF há atualmente formato para exportação dos metadados e não havia dados específicos das citações e referências em cada trabalho. Diante disso, não houve outra solução a não ser coletar manualmente as referências de cada uma das teses e dissertações. A coleta foi realizada de outubro de 2021 até março de 2022 e contou com o apoio de um doutorando, cinco mestrandos e dez graduandos.

As teses e dissertações foram todas reunidas em formato pdf (único disponível) numa pasta online compartilhada com os participantes (*Google Drive*), seus títulos foram codificados, e um formulário foi desenvolvido na ferramenta *Google Forms* para a coleta de cada referência, com os seguintes campos: Nome do responsável pela coleta, Código do trabalho; Ano do trabalho; Tipo do trabalho (tese ou dissertação); Referência da citação coletada; Tipo da obra citada (artigo de jornal, artigo de periódico, capítulo de livro, dicionário, jornal, livro, periódico, regulamento/lei, sites, tese/dissertação/TCC, trabalho de evento, outro); Ano da citação; Idioma; Tipo de autoria (pessoa, instituição, mista, desconhecida); Quantidade de autores; Autores citados. A vantagem da ferramenta *Forms* é a geração automática de uma planilha, a partir da qual os dados podem ser analisados e contados de maneira mais ágil. Um grupo na rede *Whatsapp* foi criado para solucionar de forma mais ágil dúvidas dos participantes.

5 Os desafios da coleta de referências: o caso do PPGCI/UFF

Quando saímos do universo dos artigos de periódicos, a questão da infraestrutura de pesquisa voltada a estudos métricos da informação se torna evidente. Quase todo o trabalho de coleta depende de ações manuais. Isso, somado a falta de recursos para investimentos em soluções de tecnologia adequadas para análise nas universidades públicas, torna o trabalho demorado e desafiador.

Desafio 1: acesso ao material e às suas referências

Na BDTD, apenas dez trabalhos foram encontrados, indicando falta de atualização da base de dados. No RI/UFF, encontramos todas as obras. Foram coletados 145 trabalhos, sendo 131 dissertações e 14 teses. Foi necessário acessar cada trabalho e copiar uma a uma cada referência, sempre que possível, utilizando os recursos de copiar e colar existentes nos editores de texto, formulários e arquivos em pdf. No entanto, havia trabalhos em pdf protegidos, isto é, que não permitem o recurso de copiar, e dessa forma houve casos durante a coleta em que foi necessário escrever toda a citação para registrá-la no formulário. Ao todo, foram coletadas 17.193 referências (14.125 das dissertações e 3.068 das teses), conforme apresentado na Tabela 1.

Interessante notar que a média de referências utilizadas por dissertação varia de 94 a 129 trabalhos e nas teses é de 211 a 241. Nota-se que apesar de seu menor tempo de existência no programa, as teses possuem uma alta fundamentação de referências, quantitativamente se sobressaindo quando comparado às dissertações – que apesar de possuírem um alto índice também, se justifica pela divisão de tempo de pesquisa.

Tabela 1. Resultado da coleta de referências por dissertações (Diss.) e Teses

ANO	Defesas		Quantidade de referências		Média por trabalho	
	Diss.	Teses	Diss.	Teses	Diss.	Teses
2011	14	-	1329	-	94,9	-
2012	14	-	1708	-	122,0	-
2013	16	-	1561	-	97,6	-
2014	16	-	1872	-	117,0	-
2015	14	-	1518	-	108,4	-
2016	12	-	1168	-	97,3	-
2017	9	-	851	-	94,6	-
2018	8	2	776	470	97,0	235,0
2019	13	10	1679	2115	129,2	211,5
2020	15	2	1663	483	110,9	241,5
TOTAL	131	14	14125	3068	107,8	219,1

Fonte: Dados da pesquisa, 2022.

Desafio 2: normalização em si.

Embora no conjunto analisado seja regra a apresentação das referências segundo a norma ABNT 6023 (2018), muitas delas foram encontradas fora do padrão de normalização. Entre os erros encontrados, pode-se mencionar capítulos de livros citados com se fossem artigos, autores institucionais confundidos com editoras e trabalhos de eventos ora citados como artigos de periódicos ora como capítulos de livros. Isso exigiu, por parte dos envolvidos na coleta, intenso trabalho de pesquisa para verificar a correção dos dados. Vale dizer que muitas vezes os trabalhos eram completamente diferentes do que aquilo que estava citado. Nesse sentido, o grupo de *Whatsapp* foi essencial para que as decisões fossem tomadas. Houve casos, inclusive, em que não foi possível identificar os autores como pessoa ou instituição (224 casos) e nem ao mesmo saber a que tipo de fonte se refere o documento (seria um texto impresso, uma base de dados que não existe mais, outro tipo de recurso?). Estes casos, 79 referências, foram categorizados como fontes de acesso restrito. Enfatizamos, todavia, que todo esforço possível foi realizado para identificação das fontes e nas planilhas o desvio era corrigido de modo a constar a identificação correta da fonte. Outra questão é função da referência: se a ideia é mencionar o trabalho na pesquisa para que os futuros leitores possam consultá-los para saber mais sobre o assunto para dirimir dúvidas que ficaram na leitura, a citação a documentos restritos torna-se um desafio. É preciso refletir sobre o que significa fundamentar uma pesquisa, e possivelmente diferenciar bibliografia de apoio a essa fundamentação da bibliografia utilizada como objeto.

Além disso, pode-se notar como os mesmos autores são referenciados de formas distintas. Inclusive de documentos institucionais. Isto certamente gerou uma dispersão de difícil controle, mesmo que o processo fosse feito por máquina. Vale ressaltar que os pesquisadores da instituição em que se localiza o PPGCI analisado têm a sua disposição serviço de normalização por meio de suas bibliotecas, além de entender que tal operação é exigência na apresentação das teses e dissertações. De modo ainda não sistemático, cursos de gerenciadores de referências, como Zotero, tem sido oferecidos no âmbito da UFF, mas ainda com tímida adesão do corpo discente e docente da UFF. De maneira geral, na literatura de Biblioteconomia e Ciência da Informação também é tímida a discussão sobre essas ferramentas, como apontado por Quadros e Vogel (2022). Em tal estudo, as autoras apontam que a maior parte dos trabalhos identificados na BRAPCI sobre gerenciadores de referências na verdade os utilizam como ferramentas para estudos métricos e não para discutir a importância da normalização nem para ensinar ou divulgar seu uso durante a realização de trabalhos acadêmicos.

6 Conclusão

Compreende-se que ainda é preciso investir em ensino de normalização para que os pesquisadores de mestrado e doutorado compreendam a importância das referências não apenas para a checagem das fontes de informação, mas também para o mapeamento da produção de conhecimento.

Tal cenário, somado ao desenvolvimento de repositórios institucionais e bases de dados diversificados e preparados para análises métricas, bem como maiores investimentos nas instituições de pesquisa pública para aquisição de ferramentas de análise ou para o desenvolvimento de soluções abertas dessa natureza resultará em estudos bibliométricos mais abrangentes e mais próximos da realidade nacional.

Enquanto não houver uma infraestrutura adequada para coleta dos dados, possivelmente muitas informações ficarão à margem dos estudos métricos. Continuaremos a depender das informações geradas por bases de dados cujos interesses não necessariamente convergem com as necessidades de pesquisa brasileiras, como Web of Science e Scopus. Nesse sentido, iniciativas como a da BRAPCI devem ser encorajadas e ampliadas a outras áreas do conhecimento e a outros tipos de comunicação científica (eventos, teses e dissertações, TCCs entre outros).

Considera-se que é preciso investir em comunicação institucional, considerando que as instituições de pesquisas contam com bibliotecas e que elas, via de regra, oferecem o serviço de normalização, havendo, possivelmente, desconhecimento deste por parte dos

pesquisadores. Uma possível recomendação seria tornar obrigatória uma revisão de normalização após a defesa, garantindo que a tese ou dissertação fosse para o repositório ou biblioteca já dentro dos padrões estabelecidos, e por outro lado pedir que tais repositórios coloque as referências em metadado adequado e exportável, para facilitar a coleta.

Ressalta-se por fim, o desenvolvimento dos gerenciadores de referências, como Zotero, Mendeley, EndNote, entre outros. Tais ferramentas podem ser conhecidas de maneira autodidata, mas é interessante ofertar cada vez mais espaços nas instituições para seu aprendizado. No entanto, é preciso que o pesquisador saiba indicar ao software o tipo de referência, caso contrário este criará uma perfeita referência de artigo de periódico, ao que na verdade era um capítulo de livro. Por isso, novamente, reforça-se a necessidade de disciplinas e orientação que de fato ensino aos pesquisadores, especialmente aos da Biblioteconomia e Ciência da Informação a relevância, o papel e como aplicar a normalização.

Referências

- ABNT (ASSOCIAÇÃO BRASILEIRA DE NORMAS E TÉCNICAS). **NBR 6023** – Informação e Documentação – Referências – Elaboração. Rio de Janeiro, ABNT, 2018.
- ARAÚJO, P. C.; BUFREM, L. S. The intellectual foundation of the elite of Brazilian researches on knowledge organization domain. **Transinformação**, Campinas, v. 33, e200068, 2021. DOI: <https://doi.org/10.1590/2318-0889202133e200068>.
- BARITÉ, M. et al. **Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología**. 6.a ed. corregida y aumentada. Montevideo: csic, 2015.
- CRESPO, I. M.; RODRIGUES, A. V. F. Normas técnicas e comunicação científica: enfoque no meio acadêmico. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 9, n. 1, p. 35-55, jul./dez. 2011. DOI: <http://doi.org/10.20396/rdbci.v9i1.1918>.
- CURTY, M. G.; BOCCATO, V. R. C. O artigo científico como forma de comunicação do conhecimento na área de Ciência da Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 10, n. 1, p. 94-107, jan./jun. 2005. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/download/305/108>. Acesso em: 15 ago.2019.
- CUSTÓDIO, P. G. G. R.; LIMA, L. S.; BUFREM, L. S. Citações e cocitações nas teses de ciência da informação da USP. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais ...**, Marília: UNESP, 2017. Disponível em: <https://brapci.inf.br/index.php/res/download/125157>. Acesso em: 25 fev. 2022.
- GLÄNZEL, W. **Bibliometrics as a research field: a course on theory and application of bibliometric indicators**. Bélgica, 2003. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.5311&rep=rep1&type=pdf>. Acesso em: 25 fev. 2022.
- GRÁCIO, M. C. C. Acomplamento bibliográfico e análise de cocitação: revisão teórico-conceitual. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 21, n. 47, p. 82-99, 2016. DOI: <https://doi.org/10.5007/1518-2924.2016v21n47p82>
- GRÁCIO, M. C. C.; OLIVEIRA, E. F. T. A pesquisa brasileira em estudos métricos da informação: proximidade entre pesquisadores de destaque e áreas afins. **Informação & Sociedade: Estudos**, v. 27, n. 2, p. 105-116, 2017. DOI: <https://doi.org/10.22478/ufpb.1809-4783.2017v27n2.32483>
- GUEDES, V. L. da S. A bibliometria e a gestão da informação e do conhecimento científico e tecnológico: uma revisão da literatura. **PontodeAcesso**, [S. l.], v. 6, n. 2, p. 74-109, 2012. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/5695>. Acesso em: 26 maio. 2023.
- HJØRLAND, B. Citation analysis: a social and dynamic approach to knowledge organization. **Information Processing and Management**, v. 49, n. 6, p. 1313-1325, 2013. DOI: <https://doi.org/10.1016/j.ipm.2013.07.001>.
- HOU, J. Exploration into the evolution and historical roots of citation analysis by referenced publication year spectroscopy. **Scientometrics**, Dordrecht, v. 110, n. 3, p. 1437-1452, 2017. DOI: <https://doi.org/10.1007/s11192-016-2206-9>.
- MENEZES, V. S.; ODDONE, N. E. Os livros nas teses da ciência da informação: um estudo de citação (2007-2009). **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 7, n. 1, 2014. Disponível em: <https://brapci.inf.br/index.php/res/download/121598>. Acesso em: 25 fev. 2022.
- LUBISCO, N. M. L. **Manual de estilo acadêmico: monografias, dissertações e teses**. 4. ed. Salvador: EDUFBA, 2008.
- MEADOWS, A. J. **Comunicação científica**. Brasília: Briquet de Lemos, 1999.
- MILLER, L. N. Local Citation Analysis of Graduate Biology Theses: Collection Development Implications. **Science and Technology Librarianship**, v. 64, jan. 2011. Disponível em: <http://www.istl.org/11-winter/refereed3.html>. Acesso em: 25 fev. 2022.
- MOREIRA, S. V. Análise documental como método e como técnica. In: DUARTE, J.; BARROS, A. (orgs). **Métodos e técnicas de pesquisa em comunicação**. São Paulo: Atlas, 2006. 2a ed.
- QUADROS, M. B.; VOGEL, M. J. M. Gerenciadores de referências na Ciência da Informação. In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 8., 2022, Maceió. **Anais...**, Maceió: UFAL, 2022. p. 590-596. Disponível em: <https://ebbc.inf.br/ojs/index.php/ebbc/article/view/119>. Acesso em: 20 abr. 2023.
- RODRIGUES, M. P. L. Citações nas dissertações de mestrado em ciência da informação. **Ciência da Informação**, Rio de Janeiro, v. 11, n. 1, 1982. Disponível em: <http://revista.ibict.br/ciinf/article/view/177>. Acesso em: 25 abr. 2021.

SANT'ANNA, J. Atuação profissional na normalização bibliográfica: um campo promissor para o bibliotecário. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 15, n. 2, p. 216-236, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/113900>. Acesso em: 20 abr. 2023.

SILVA, M. R.; HAYASHI, C. R. M.; HAYASHI, M. C. P. I. Análise bibliométrica e cientométrica: desafios para especialistas que atuam no campo. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 2, n. 1, p. 110-129, jan./jun. 2011. Disponível em: <http://www.revistas.usp.br/incid/article/view/42337>. Acesso em: 16 jun. 2017.

SILVEIRA, M. A.; CAREGNATO, S. E. Percurso histórico-epistemológico dos estudos de citação no Brasil. **Transinformação**, Campinas, v. 29, n. 1, p. 39-55, 2017. DOI: <https://doi.org/10.1590/2318-08892017000100005>.

SMIRAGLIA, R. P. **The elements of knowledge organization**. Switzerland: Springer, 2014.

SOUSA, R. J. P. L. (re)aproximações entre bibliometria e organização da informação: um relato de experiência. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 25, n. 3, p. 734-750, 2020. Disponível em: <https://brapci.inf.br/index.php/res/v/150552>. Acesso em 20 de mar. 2023.

SPERANDIO, D. S. Normalização de Trabalhos Acadêmicos: da teoria à prática. **Biblionline**, João Pessoa, v. 15, n.3, p.68-77, 2019. DOI: <https://doi.org/10.22478/ufpb.1809-4775.2019v15n3.48098>.

WALTMAN, L. A review of the literature on citation impact indicators. **Journal of Informetrics**, v. 10, n. 2, p. 365-391, 2016. DOI: <https://doi.org/10.1016/j.joi.2016.02.007>.

WITTER, G. P.; PÉCOR, G. M. M. Temática das dissertações e teses em biblioteconomia e ciência da informação no Brasil (1970-1992). In: WITTER, G. P. (org.) **Produção científica**. Campinas: Editora Átomo, 1997. p. 77-86.

Metodologia para categorização das concepções científicas em Ciência da Informação

Methodology for categorization of scientific concepts in Information Science

Josiana Florêncio Vieira Régis (1), Guilherme Ataíde Dias (2)

(1) Universidade Federal do Rio Grande do Norte, Biblioteca Central Zila Mamede - BCZM. Rua da Biblioteca, S/N. Campus Universitário da UFRN. CEP 59078-900. Natal, RN, josianavieira@gmail.com, (2) Universidade Federal da Paraíba - Campus I, Departamento de Ciência da Informação – DCI. CEP 58051-900 – João Pessoa, PB, guilhermeataide@gmail.com

Resumo

Este artigo configura-se como um recorte dos resultados de uma tese de doutorado para consolidar e viabilizar a prática de uma metodologia para categorização das concepções científicas da Ciência da Informação. A temporalidade e as apresentações das pesquisas científicas demonstram fatos, conflitos, avanços, regressões, complexidades, contradições, centralizações e flexibilizações dos campos científicos. Diante disso, temos como objetivo, encontrar essas características para viabilizar a materialização da memória coletiva de uma comunidade discursiva, como por exemplo, a Ciência da Informação por meio da criação de um modelo de categorização científica. A metodologia adotada foi a teoria fundamentada nos dados de Strauss e Corbin e a análise de domínio de Joseph Tennis, tendo-se como amostragem 221 teses de doutorado distribuídas entre cinco programas de pós-graduação na área. Os resultados apontam que a categorização científica determina as dimensões de um conteúdo por meio da análise, seleção e registro de unidades de conhecimento e as coocorrências de contextos. Deste modo, podemos acompanhar as características temporais do conhecimento científico e suas intercalações epistemológicas e pragmáticas. Essa categorização funciona como uma espécie de “empreendimento científico” que soluciona os problemas de interpretação, ao mesmo tempo em que abre espaço para novas problematizações.

Palavras-chave: Organização do conhecimento. Conhecimento científico. Categorização científica. Bases metodológicas.

Abstract

This article is an excerpt from the results of a doctoral thesis to consolidate and enable the practice of a methodology for categorization of scientific concepts of Information Science. The temporality and presentations of scientific research demonstrate facts, conflicts, advances, regressions, complexities, contradictions, centralizations and flexibility of scientific fields. Therefore, we aim to find these characteristics to enable the materialization of the collective memory of a discursive community, such as Information Science, through the creation of a model of scientific categorization. The methodology adopted was the theory grounded in data by Strauss and Corbin and the domain analysis by Joseph Tennis, with a sample of 221 doctoral theses distributed among five postgraduate programs in the area. The results indicate that the scientific categorization determines the dimensions of a content through the analysis, selection and registration of knowledge units and the co-occurrences of contexts. In this way, we can follow the temporal characteristics of scientific knowledge and its epistemological and pragmatic intercalations. This categorization works as a kind of “scientific enterprise” that solves the problems of interpretation, at the same time that it makes room for new problematizations.

Keywords: Knowledge Organization. Scientific knowledge. Scientific categorization. Methodological foundations.

Resumen

Este artículo es un extracto de los resultados de una tesis doctoral para consolidar y posibilitar la práctica de una metodología de categorización de conceptos científicos de las Ciencias de la Información. La temporalidad y las presentaciones de la investigación científica evidencian hechos, conflictos, avances, retrocesos, complejidades, contradicciones, centralizaciones y flexibilidad de los campos científicos. Por lo tanto, nuestro objetivo es encontrar estas características para permitir la materialización de la memoria colectiva de una comunidad discursiva, como es la Ciencia de la Información, a través de la creación de un modelo de categorización científica. La metodología adoptada fue la teoría fundamentada en datos de Strauss y Corbin y el análisis de dominio de Joseph Tennis, con una muestra de 221 tesis doctorales distribuidas en cinco programas de posgrado del área. Los resultados indican que la categorización científica determina las dimensiones de un contenido a través del análisis, selección y registro de unidades de conocimiento y las co-ocurrencias de contextos. De esta manera, podemos seguir las características temporales del conocimiento científico y sus intercalaciones epistemológicas y pragmáticas. Esta categorización funciona como una especie de “empresa científica” que resuelve los problemas de interpretación, al mismo tiempo que da lugar a nuevas problematizaciones.

Palabras clave: Organización del conocimiento. El conocimiento científico. Categorización científica. Bases metodológicas.

1 Introdução

O cenário tecnológico digital proporciona mudanças no comodismo cultural de uma sociedade cada vez mais desafiada pelo poder das transformações sociais. Muitos questionamentos científicos simbolizam a necessidade de uma materialização de produtos e serviços que visem contribuir com o saber humano alinhado ao senso comum.

As funções comunicativas com o objetivo de mediar os saberes diminuem a distância entre a domesticação de pensamentos e a difusão de mensagem simplificadas. Para isso, temos o fenômeno criativo como um dos principais elementos difusores na elucidação das mensagens.

A fragmentação das informações disponíveis no âmbito tecnológico instiga os profissionais e pesquisadores da ciência a buscar alternativas de desenvolvimento e processamento de tal forma que possibilite uma aproximação teórica e prática entre especialistas e a comunidade em geral. Neste caso, indicamos o desenvolvimento de recursos eletrônicos baseados nas expectativas do público interessado em conceituações abstratas.

A pluralidade de ideias e das fragmentações de saberes podem provocar represálias quanto a utilização, distribuição, reprodução e materialização do conhecimento. Por isso, a adoção de categorizações científicas que englobam a interação entre os pensamentos e a complexidade dos fatos é primordial na organização do conhecimento científico de uma área.

A temporalidade e as apresentações das pesquisas científicas demonstram fatos, conflitos, avanços, regressões, complexidades, contradições, centralizações e flexibilizações dos campos científicos. Diante disso, temos a responsabilidade de identificar essas características para viabilizar a materialização da memória coletiva de uma comunidade discursiva, como a da Ciência da Informação.

A categorização científica determina as dimensões de um conteúdo por meio da análise, seleção e registro de unidades de conhecimento. A codificação acompanhada de uma categorização também identifica as coocorrências de contextos. Deste modo, podemos acompanhar as características temporais do conhecimento científico e suas interações epistemológicas e pragmáticas. Essa categorização funciona como uma espécie de “empreendimento científico” que soluciona os problemas de interpretação, ao mesmo tempo em que abre espaço para novas problematizações.

Este artigo visa consolidar e viabilizar a prática de uma metodologia para categorização das concepções científicas da Ciência da Informação.

Parafrazeando esse contexto, apresentamos uma metodologia que intensifica essas problematizações científicas e atenua a comunicação com outros grupos. É um pressuposto para a constituição de uma base reflexiva sobre a identificação e escolha de paradigmas emergentes.

2 Os conceitos constituintes dos modelos e concepções da Ciência da Informação

O campo da Ciência da Informação oferece oportunidades e instrumentos para codificação, seleção, transmissão e compartilhamento dos três conceitos constitutivos (DIC - Dados, Informação, Conhecimento) diante uma estrutura comunicacional. Nesse aspecto, Capurro (2006) elucida o fenômeno da mensagem como principal elemento da CI para analisar o impacto social que esses conceitos trazem nas transformações e evolução da sociedade e do conhecimento científico.

É preciso compreender os significados das mensagens, as condições materiais e físicas para a sua transmissão, a credibilidade e relevância para então, justificar uma crença verdadeira para uma dada situação (CAPURRO, 2006). Há uma preocupação maior pela forma como o conhecimento está sendo comunicado ou compartilhado do que pelo acesso ao conhecimento.

Por este motivo, acrescentou-se mais um conceito constituinte da Ciência da Informação, a Mensagem (M). A partir disso, temos o fenômeno DICM (Dados- Informação-Conhecimento-Mensagem) concentrados no domínio universal do conhecimento que representa o significado proposicional dos pensamentos científicos (ZINS, 2007a).

O pragmatismo da Organização do Conhecimento (OC), por exemplo, nos mostra que os conteúdos são transmitidos, na maioria das vezes, por mensagens codificadas pelos emissores e entendidas pelos receptores. O entendimento é algo tendencioso devido ao grau de interpretação de um determinado domínio ou conteúdo. Por isso, torna-se interessante obter um conhecimento local acerca dos ambientes estudados para identificar quais são os melhores instrumentos que contribuem para a relevância do ambiente social.

Os teóricos do conhecimento científico com formação em Ciência da Informação apresentam várias definições da área com múltiplos focos de investigação. Existe a possibilidade de expressar o modo de pensar sob várias concepções do saber científico, dentre elas, temos a objetividade diante de conclusões lógicas e a subjetividade das ideias. Dentro dessa perspectiva, podemos considerar o “coletivo de pensamento” como:

[...] a comunidade das pessoas que trocam pensamentos ou se encontram numa situação de influência recíproca de pensamentos, temos em cada uma dessas pessoas, um portador de desenvolvimento histórico de uma área de pensamento, de um determinado estado do saber e da

cultura, ou seja, de um estilo específico de pensamento (FLECK, 2010, p. 82).

A partir desse ponto de vista, ocorrem ligações evolutivas que constroem uma área do conhecimento por meio de engrenagens científicas para determinar o contexto e o impacto do pensamento. Os vínculos históricos e epistemológicos de uma área participam dessa evolução de pensamentos científicos em sua totalidade, mesmo que não sejam explícitos nas teorias e experiências. O percurso de um domínio do saber também está alinhado às tentativas e condições específicas do processo de organização e representação da informação e do conhecimento.

Os estilos de pensamento de uma comunidade científica são decorrentes dos fatores sociais que influenciam as atividades de conhecimento, dentre elas, o aprendizado, as tradições, as realidades, os problemas e objetivos. Nessas condições, temos os esquemas científicos especializados que fragmentam os pilares do conhecimento de um domínio do saber. Muitas vezes, a complexidade dos fenômenos provoca mudanças ou rompimentos de paradigmas com base na percepção dos fatos científicos (FLECK, 2010; MORIN, 2006).

O mundo científico é repleto de fenômenos e observações experienciadas pelos pesquisadores e isso requer reflexões acerca do caráter polêmico do conhecimento. Nesse caso, “é preciso que o fenômeno seja catalogado, filtrado, depurado, vertido no molde dos instrumentos” (BACHELARD, 2008, p. 17). As atividades científicas captam o pensamento contemporâneo postulada pelo atomismo do conhecimento. A temporalidade é outro elemento que envolve essa organização da informação e do conhecimento, tendo em vista a dinamicidade dos arranjos sistemáticos desencadeados pela estruturação. O objetivo é tornar os instrumentos ativos e presentes nas interações científicas.

Em nível teórico, Oliveira (2011, p. 43) destaca que “a CI apresenta dificuldades relativas à delimitação de seu objeto de estudo, a informação, bem como dificuldades em desenvolver teorias em suas diversas subáreas”. Porém, esse assunto vem sendo discutido em âmbito nacional e internacional, sempre buscando melhorias na sua consolidação teórica.

Com o intuito de conhecer, aprimorar e sintetizar os estilos de pensamento dos pesquisadores da área de Ciência da Informação, Zins (2007b) resolveu indagá-los a respeito do conceito de “Ciência da Informação” em suas mais variadas percepções de mundo. O resultado desta investigação obteve 50 definições deste conceito, gerando assim os 6 (seis) modelos da Ciência da Informação descritas adiante. Este estudo visa apenas demonstrar quais são esses modelos e concepções e integrá-los na metodologia científica criada durante a codificação das categorias e subcategorias.

A história da CI transparece uma perspectiva dialética que caminha rumo a uma maturidade científica diante das conjunturas teóricas e pragmáticas da área. As estratégias de concretude e abstração do conhecimento científico caracterizam novos elementos que vão sendo incorporados ao escopo de problematizações (ARAÚJO, 2014).

O autor relata ainda sobre o conceito de Informação que interfere na evolução da CI. Enumera três conceitos: o primeiro se refere a algo objetivo com propriedades definidas; o segundo conceito é cognitivista (algo subjetivo), produto da interação entre os dados e a mente dos indivíduos; o terceiro conceito é pragmático, onde a informação é algo intersubjetivo, a partir de arranjos de elementos e ações específicas (ARAÚJO, 2014).

Este olhar informacional desencadeia a necessidade de uma análise dos fenômenos estudados, tanto dentro do campo da CI como em outras áreas do conhecimento com novos objetos de estudo, advindas de disciplinas periféricas que contribuem para o desenvolvimento evolutivo da área. Nesse contexto, apresentamos os 6 (seis) modelos da CI elaborados por Zins (2007b) baseados em visões diferentes de vários pesquisadores.

Os modelos denominados por Zins (2007b) são: modelo de **alta tecnologia** - são implementados no domínio de alta tecnologia; modelo de **tecnologia** - são implementados no domínio tecnológico em geral, ou seja, todos os tipos de tecnologias; modelo de **cultura e sociedade** - são implementados no domínio cultural; modelo do mundo humano - são implementados no domínio humano; modelo do **mundo vivo** - são implementados no mundo dos seres humanos, humanos e não-humanos; modelo de **mundos vivo e físico** - são implementados em todos os tipos de organismos biológicos, humanos e não humanos, e todos os tipos de objetos físicos. Esses modelos foram descritos e identificados de acordo com as citações de pesquisadores da CI (Ver Quadro 1).

Quadro 1. *Definição dos modelos da Ciência da Informação*

Tipo de Modelo	Definição de Ciência da Informação
Modelo <i>Hi-Tech</i> (Alta Tecnologia)	<p>A Ciência da Informação envolve o estudo de funções, estrutura, transmissão de informações e gerenciamento de sistemas. Trata-se do estudo de dados, informações, conhecimento e mensagem no domínio coletivo, abordando aspectos mediadores e utilizando alta tecnologia, além de incluir estudos com usuários.</p> <p>A Ciência da Informação compreende esforços teóricos e aplicados para definir informações, processamento computacional e tecnologias afins, bem como a interação dessas informações e sistemas com práticas e</p>

<i>Tipo de Modelo</i>	<i>Definição de Ciência da Informação</i>
	estudos humanos específicos, como negócios, cultura, biblioteconomia e filosofia.
Modelo de Tecnologia	A Ciência da Informação é uma disciplina matemática que estuda tecnologias de transmissão de informações.
Modelo de Cultura e Sociedade	<p>A Ciência da Informação trata do design e uso de sistemas de informação para mediação do conhecimento, bem como a Biblioteconomia.</p> <p>Essas disciplinas se preocupam com teorias, metodologias e procedimentos para individualizar, organizar e disseminar o conhecimento contido em livros e documentos, e conectá-lo com a mente humana. A Ciência da Informação e Biblioteconomia fazem parte de uma ciência geral da comunicação que conecta memórias externas com sistemas cognitivos ou conhecimento.</p> <p>Três concepções: não existe uma Ciência da Informação, mas várias visões diferentes da Ciência da Informação. Uma é a “Ciência da Mensagem”, que é um reconhecimento / redescoberta da base histórica primária da CI a partir da década de 1880. Outra é uma Ciência da Informação mais geral que tenta incluir todo o D-I-C-M. Uma terceira é uma visão restrita à TI, ancorada na tecnologia digital.</p> <p>Ciência da Informação é o estudo de todos os aspectos do gerenciamento de informações (por exemplo, pesquisa, criação de sistemas de TI, armazenamento, alteração, exclusão, manipulação de atualização, ferramentas para desenvolvimento, manipulação, administração, informações sobre informações, introdução aos usuários finais etc.)</p> <p>A Ciência da Informação dedica-se à manipulação de informações por meio da coleta, armazenamento, recuperação, classificação e interpretação de dados, visando compreender os mecanismos subjacentes a esse processo.</p> <p>A Ciência da Informação investiga formas de gerenciar dados com o objetivo de criar, gerenciar e compreender informações, a fim de produzir conhecimento.</p>
Modelo do Mundo Humano	A Ciência da Informação é o estudo da interação entre humanos e informação e todos os mecanismos e elementos do contexto que desempenham um papel nessa interação.

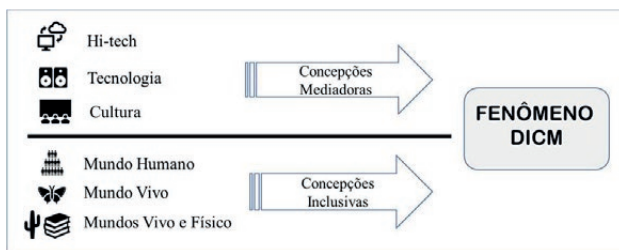
<i>Tipo de Modelo</i>	<i>Definição de Ciência da Informação</i>
	<p>A Ciência da Informação é uma disciplina que abrange o estudo de todas as manifestações da informação. Embora seja comum a atenção se concentrar no armazenamento e recuperação de informações, incluindo a utilização de sistemas de bibliotecas, classificação e indexação, catálogos, além de mecanismos de pesquisa, mapeamento de conceitos, estudos de relevância e recuperação, a disciplina vai além, englobando os comportamentos de pesquisa e recuperação de usuários, necessidades informacionais, comunidades de usuários, design de interfaces homem-computador e visualização de informações.</p> <p>A Ciência da Informação é uma disciplina que abrange a totalidade do processo de comunicação e compreensão, incluindo aspectos intra e interpessoais. Sua abrangência é ampla e inclui desde a Teoria da Informação de Shannon até a semiótica e memética. No entanto, sua vastidão impede que seja definida de forma precisa, a menos que se restrinja e delimite a disciplina, caracterizando-a como algo distinto.</p>
Modelo do Mundo Vivo	A Ciência da Informação é uma área de estudo e prática que busca determinar as leis e os princípios relacionados à análise, design e avaliação de sistemas de dados, informações e conhecimento. Essa disciplina é fundamentada na premissa de que todos os organismos são sistemas que realizam funções cognitivas e afetivas com diferentes graus de eficiência, sendo que cada uma dessas funções pode ser aprimorada por meio da tecnologia criada e aplicada.
Modelo do Mundo Vivo e Físico	A Ciência da Informação é o estudo de sistemas, seus subsistemas e processos de informação e suas inter-relações em diferentes contextos ambientais. Esta definição se aplica aos níveis moleculares e celulares ou a órgãos, organismos, grupos, comunidades ou níveis superiores. A tecnologia da informação preocupa-se com o manuseio e o processamento ideal das informações, geralmente para determinados indivíduos ou organizações, e a bioinformática estendeu recentemente a ciência da informação para os reinos animal e vegetal.

Fonte: adaptado de Zins (2007b, tradução nossa).

Após a definição dos modelos de CI expostos anteriormente, Zins (2007b) estabeleceu dois tipos de concepções existentes na área, são elas: as concepções “mediadoras” do fenômeno DICM. Essa primeira concepção está relacionada aos três primeiros modelos descritos no quadro 1 (modelos de alta tecnologia, tecnologia e cultura). A segunda concepção, denominada “inclusiva”, caracteriza-se por englobar todos os contributos do fenômeno DICM com visões gerais (modelos de mundo humano, mundo vivo e mundos vivo e físico).

Em síntese, destacamos que as concepções mediadoras veem a CI como um campo que busca uma aproximação com outras áreas do conhecimento, enquanto as concepções inclusivas abordam a CI como um campo geral que pertence a todos os campos, subcampos e áreas do conhecimento. Um exemplo de concepções inclusivas são as teorias relacionadas às ciências cognitivas, epistemologia, filosofia da ciência, sociologia do conhecimento, educação, linguística, semiótica, entre outros (ZINS, 2007b). Essas concepções estão ilustradas na Figura 1:

Figura 1. Representação dos modelos de concepções da CI



Fonte: adaptado de Zins (2007b).

Com base na tríade constituinte da CI – Dados, Informação, Conhecimento – entende-se que as perspectivas mediadoras são as mais aplicáveis no domínio universal do conhecimento, incluindo desde os aspectos tecnológicos até as condições sociais e cognitivas. Essas mediações promovem a disseminação do conhecimento humano para o usuário.

Notamos que as concepções trazem alguns questionamentos no que se refere aos fenômenos, domínio e escopo. Zins (2007b) corrobora essa ideia quando questiona: quais são os fenômenos explorados? qual é o domínio do campo? qual é o escopo da exploração? nas pesquisas científicas. As respostas são advindas de uma extensa análise de domínio (TENNIS, 2003) para explorar essas concepções em uma comunidade discursiva.

3 Criação de uma metodologia para mapeamento científico da Ciência da Informação

O profissional da informação que está inserido no contexto das universidades e que atua na produção e divulgação do conhecimento científico deve apresentar recursos que sejam coerentes com os objetivos institucionais das universidades públicas ou privadas. Esses recursos devem possuir estratégias que direcionem as aplicações, teorias, paradoxos, interferências, limites e lacunas das áreas científicas.

A obtenção de recursos informacionais interfere na responsabilidade social do conhecimento científico produzido pelas universidades com a construção, disseminação e, conseqüentemente a reelaboração desse conhecimento gerando novas ações que contribuam para as atividades de ensino, pesquisa e extensão.

A difusão e promoção do conhecimento científico torna-se um dos maiores desafios da universidade, pois tem a missão de converter o conhecimento em algo útil e reutilizável para a sociedade em geral por meio da compreensão e aplicação.

Neste sentido, observamos que a dinâmica informacional das atividades está ligada aos processos do DICM (Dados, Informação, Conhecimento e Mensagem). O relacionamento entre a informação e o conhecimento proporcionam uma otimização dos recursos com o objetivo de atender aos objetivos e necessidades da comunidade científica.

Pensando na obtenção desses recursos para dirimir as necessidades informacionais dos usuários e identificar as lacunas existentes no meio científico de uma área específica, criamos uma metodologia que proporciona uma análise interpretativa dos contextos teóricos e práticos publicados pelas universidades. Utilizamos como base, uma codificação com as 10 (dez) categorias centrais que compõem o mapa do conhecimento de Zins (2007b) para a construção de novas categorias, seguindo as recomendações dos mecanismos de codificação estabelecidos por Strauss e Corbin (2008) – codificação aberta, axial e seletiva - para a construção de uma teoria fundamentada nos dados coletados. Essa codificação corresponde à construção e representação de categorias e subcategorias dos documentos (teses).

Utilizamos a abordagem qualitativa, a pesquisa documental, a análise de domínio de Tennis (2003) e a Teoria Fundamentada nos Dados de Strauss e Corbin (2008). A amostra é constituída por 221 teses de doutorado defendidas no quadriênio compreendido entre 2013 e 2016, distribuídas em 5 (cinco) programas de pós-graduação em Ciência da Informação no Brasil das seguintes universidades: UNESP, USP, UFRJ, UnB e UFMG. A codificação foi executada por meio do software NVIVO.

A codificação e mapeamento científico passou por três tipos principais de codificação: **a codificação aberta, a codificação axial e a seletiva**. A codificação aberta é o processo inicial e analítico que visa identificar, primeiramente em linhas gerais, os conceitos, fenômenos, propriedades e dimensões nos dados coletados. A codificação axial é o agrupamento ou criação de subcategorias de modo associativo e de acordo com as propriedades e dimensões. A codificação seletiva integra e seleciona por meio de um refinamento dos fenômenos, propriedades e dimensões codificadas em cada categoria e subcategoria (STRAUSS; CORBIN, 2008).

Tendo em vista o uso da TFD e da análise de domínio, apresentamos a seguir, as etapas que subsidiaram a criação da metodologia científica para categorização em CI (Ver Quadro 2):

Quadro 2. Etapas da TFD e passos correspondentes

Etapas da TFD	Passos correspondentes
Coleta das teses	Busca online das teses por programa de Pós-graduação em CI no quadriênio Capes 2013 a 2016; Download e arquivamento dos documentos.
Microanálise	Inclusão das categorias predeterminadas de Zins (2007c).
Codificação aberta	Leitura e análise com pré-categorização das teses.
Codificação axial	Análise de domínio dos documentos de acordo com as bases teóricas e metodológicas de Strauss e Corbin (2008), Tennis (2003), Smiraglia (2014; 2015), Hjørland (2002), Hjørland e Albrechtsen (1995), Zins (2007b).
Codificação seletiva	Análise das linhas de conectividade para agrupar as categorias e subcategorias existentes.
Codificação para o processo	Discussão dos resultados; Este procedimento ocorre simultaneamente com a codificação axial e construção de categorias.
Modelo condicional/consequencial e amostragem teórica	Diagramação dos resultados com a elaboração do mapa do conhecimento da Pós-graduação em Ciência da Informação.

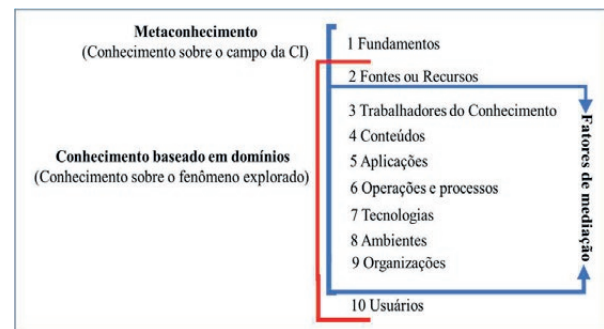
Fonte: Almeida (2021).

A codificação baseada na TFD também utilizou os critérios de Tennis (2003) para delinear os objetos de estudo e as lacunas existentes nas pesquisas, sempre observando as “áreas de Modulação” - define a extensão do domínio e “graus de especialização” - qualifica os graus de especialização e define a intensão de um domínio.

A codificação utilizou 10 (dez) categorias centrais de Zins (2007c) e subcategorias que foram criadas ao longo do mapeamento metodológico. As categorias centrais são denominadas, a saber: Fundamentos, Recursos, Conteúdos, Aplicações, Operações e processos, Tecnologias, Ambientes, Organizações, Usuários e Trabalhadores do conhecimento (Ver Figura 2).

As categorias centrais de Zins (2007c) são provenientes das teorias adjacentes do mapa do conhecimento do autor e das adaptações metodológicas e pragmáticas observadas no campo empírico deste estudo. Demonstramos a seguir, os objetivos de cada categoria, elencados por Almeida (2021):

Figura 2. Categorias do Mapa do conhecimento da CI



Fonte: produzido e adaptado de Zins (2007c, tradução nossa).

A categoria 1 “Fundamentos” engloba os estudos de base filosófica, histórica, epistemológica e metodológica da Ciência da Informação. É composta por teoria, pesquisa, educação e história da Ciência da Informação. Correspondem às disciplinas centrais e princípios fundamentais para o desenvolvimento das bases mencionadas.

Os núcleos que identificam e justificam a existência das bases clássicas e coletivas de conhecimento, buscam um consenso histórico, metodológico, filosófico e epistemológico da área. O movimento discursivo dos traços da Arquivologia, Biblioteconomia, Museologia e Ciência da informação enfatiza as especificidades e instrumentos de cada área com vistas a cooperação e união a partir da potencialização de um olhar informacional (ARAÚJO, 2014).

Esta categoria centrou-se na indexação e codificação das informações presentes nos resumos e nos objetivos gerais e específicos das teses por se tratar das bases referenciais e constituintes da Ciência da Informação.

A categoria 2 “Recursos” enfoca os recursos gerais do conhecimento que são utilizados na produção do conhecimento registrado. Apontam os conhecimentos teóricos e práticos relacionados ao desenvolvimento, manutenção e gerenciamento dos recursos de informação. Esses recursos apoiam os pesquisadores durante o processo de produção e comunicação científica no âmbito do compartilhamento das

informações. São constituídos por fatores moderadores da criação e produção do conhecimento, seguindo o pensamento de Nonaka *et al.* (2000). Em linhas gerais, os Recursos do Conhecimento são definidos como:

Meios de que se pode dispor ou ainda, ferramentas utilizadas na finalidade de apoiar o armazenamento, compartilhamento e reutilização de conhecimento científico, cujo intuito maior é a integração e/ou relacionamento entre autores, pesquisadores, professores e estudantes para o trato de informação de interesse em comum. Servem em sua essência, como um modo de obter a atualização da informação apropriada ao tema pesquisado e estudado (CASSOTTA *et al.*, 2017, p. 17).

Nessa categoria, os recursos foram analisados com base na codificação dos procedimentos metodológicos e do desenvolvimento teórico das temáticas das teses e classificados em: epistemológicos, científicos, gerenciais, educacionais, informacionais, digitais, arquivísticos, bibliográficos e museológicos. Cada recurso possui instrumentos e teorias da CI e áreas afins que contribuem para as investigações científicas e mediam o processo de busca e atualização das informações. A democratização do conhecimento advém da caracterização das pesquisas de modo a integrar as diferentes linhas de pensamento.

A categoria 3 “Trabalhadores do conhecimento” determina o perfil da comunidade científica na área de Ciência da Informação. Cita os autores, orientadores e coorientadores que estruturam e sistematizam o conhecimento científico em suas pesquisas. Apresenta codificações com base na extração de informações referentes aos níveis de graduação, especialização, mestrado, doutorado e pós-doutorado contidas nos currículos lattes de cada pesquisador citado anteriormente.

Drucker (2002) considera o trabalhador do conhecimento como um indivíduo com elevado nível intelectual que possui seus próprios meios de produção e seu conhecimento. É caracterizado como um profissional que se identifica com uma área do conhecimento, ou seja, é especialista na administração do conhecimento específico de uma área. O autor complementa que “a comunidade deles será cada vez mais de pessoas que compartilham os mesmos conhecimentos altamente especializados, não importando onde eles trabalhem ou para quem trabalhem (DRUCKER, 2002, p. 12)”.

Os trabalhadores do conhecimento estão sempre buscando novas competências e habilidades para enfrentar os desafios da sociedade contemporânea. São conscientes das mudanças tecnológicas e culturais, estabelecendo novas identidades sociais e relacionamentos por meio das redes colaborativas no âmbito da sua produtividade.

Nessa categoria, os produtores e mediadores do conhecimento científico (trabalhadores do

conhecimento) são agrupados e categorizados por grau de formação acadêmica: graduação, especialização, mestrado, doutorado e pós-doutorado. Cada grau de formação é dividido em subcategorias com as denominações dos cursos. As informações são extraídas da Plataforma CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) por meio dos currículos Lattes dos pesquisadores.

O nível de formação “Graduação” é classificado de acordo com os parâmetros gerais de classificação para os cursos de educação superior elaborada pelo Ministério da Educação (MEC/INEP).

Essa classificação é denominada “Classificação Internacional Normalizada da Educação Adaptada para Cursos de Graduação e Sequenciais de Formação Específica (Cine Brasil)”. A metodologia de classificação desses cursos obedece ao agrupamento das semelhanças temáticas entre eles.

Apresenta 11 (onze) áreas de conhecimento que são utilizadas como referência para a classificação oficial dos cursos no Censo da Educação Superior: programas básicos; educação; artes e humanidades; ciências sociais, comunicação e informação; negócios, administração e direito; ciências naturais, matemática e estatística; computação e tecnologias da informação e comunicação (TIC); engenharia, produção e construção; agricultura, silvicultura, pesca e veterinária; saúde e bem-estar; e serviços (INEP, 2020).

Para os níveis de formação “Mestrado” e “Doutorado” são consideradas as denominações dos cursos informados nos currículos Lattes dos trabalhadores do conhecimento. Já para o nível de formação “Pós-doutorado” não há especificação e nomeação da área no currículo Lattes, conseqüentemente não possuem subcategorias.

A categoria 4 “Conteúdos” aborda os tipos e problematizações relacionados ao conteúdo do conhecimento mediado em vários tipos de estruturas. Geralmente, as categorias baseadas em conteúdos estão interligadas às questões baseadas em recursos, ou seja, são complementares. Por exemplo, podemos relacionar os recursos aos vasos e os conteúdos ao que existe dentro deles.

Os conteúdos são ordenados e categorizados de acordo com as ementas dos Grupos de Trabalho (GTs) da Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB): Estudos Históricos e Epistemológicos da Ciência da Informação; Organização e Representação do Conhecimento; Mediação, Circulação e Apropriação da Informação; Gestão da Informação e do Conhecimento; Política e Economia da Informação; Informação, Educação e Trabalho; Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação; Informação e Tecnologia; Museu, Patrimônio e Informação;

Informação e Memória; Informação & Saúde (ANCIB, 2020).

A Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB) apresenta objetivos e estruturas que são coerentes com as agendas de pesquisa e com o incentivo à pesquisa científica na área de Ciência da Informação. Os conteúdos apresentados nas ementas dos GTs promovem o desenvolvimento da pesquisa e cooperação entre profissionais, pesquisadores, discentes e docentes da pós-graduação. O fomento à divulgação de estudos avançados na área contribui para o aperfeiçoamento intelectual dessa comunidade científica (ANCIB, 2020).

Os conteúdos das teses foram codificados de acordo com a constituição, desenvolvimento e conceituação teórica e aplicada dos campos científicos, tendo-se por base as ementas dos Grupos de Trabalho (GTs) da ANCIB. A análise e codificação dos conteúdos foram extraídos do Título, Resumo, Palavras-chaves e leitura dinâmica do referencial teórico e metodológico em busca da temática geral que representa cada documento (tese).

A categoria 5 “Aplicações” aborda questões relacionadas ao desenvolvimento de recursos projetados para o encontro das necessidades e interesses do usuário. Refere-se a criação de protótipos e modelos teóricos e aplicados com funções e finalidades direcionadas para a aquisição de um determinado conhecimento.

Burke (2016) enfatiza a necessidade de valorização de um conhecimento alternativo, também denominado local, prático e contextualizado. Em outras palavras, algumas formas de conhecimento exigem um estreitamento de visão por parte dos usuários e dos trabalhadores do conhecimento. É importante fazer a seguinte pergunta: a aplicação do conhecimento é útil para quem ou para quem?

Para facilitar a busca e recuperação das aplicações, os protótipos e modelos conceituais são agrupados por semelhanças de temáticas e finalidades de acordo com a realidade observada nas pesquisas. A tradução do conhecimento refletida em um método de aplicação, protótipo ou propostas de melhorias amplia o nível de compreensão acerca da utilidade da produção do conhecimento científico e possibilita uma eficácia e eficiência nas atividades de visualização e recuperação.

A categoria Aplicações foi codificada com base no desenvolvimento de estruturas metodológicas e de modelos teóricos e aplicados encontrados nas teses com o objetivo de clarificar e atender as necessidades informacionais dos usuários. A categoria foi delimitada por temas (subcategorias) referentes aos modelos encontrados: Estudos métricos; Preservação digital; Repositórios digitais; Mediação da informação; Competência em informação; Gestão da informação e do conhecimento; Arquitetura da informação;

Ontologias; Recuperação da informação; Organização da informação e do conhecimento

A categoria 6 “Operações e processos” inclui as ações (principais processos e operações) envolvidas em mediar a produção do conhecimento, dentre eles: organização, representação, preservação, documentação, busca e recuperação de conhecimento, mensuração, avaliação, disseminação, processamento, armazenamento, publicação, visualização e divulgação.

No sentido amplo e organizacional, Choo (2006, p. 17) argumenta que “o coração da organização do conhecimento está centrado na administração dos processos de informação que constituem a base para criar significado, construir conhecimento e tomar decisões”. Todos os processos acima mencionados nesta categoria, requerem etapas científicas de interpretação (destaque dos aspectos objetivos do ambiente), seleção (redução de ambiguidades) e retenção (armazenamento da criação de significado - produtos - para o futuro) do conhecimento.

Esta categoria consistiu na codificação da principal operação e/ou processo envolvido nas pesquisas. A operação e o processo abrangem um conjunto de atividades interligadas que promovem a produção e transformação do conhecimento científico para otimizar os resultados de pesquisa. A síntese e seleção dessa operação e processo envolveram a leitura do resumo e partes significativas das teses, como a introdução, objetivos e procedimentos metodológicos.

A categoria 7 “Tecnologias” abrange os instrumentos tecnológicos (linguagens, ferramentas e software) utilizados como meios para o tratamento das informações na execução das operações e processos delimitados na categoria 6. Esta categoria foi codificada com base nas teses que apresentaram meios tecnológicos para intermediar os resultados e objetivos das pesquisas.

Essas tecnologias visam facilitar a mediação do conhecimento científico. As linguagens são recursos utilizados por programadores no desenvolvimento e implementação de software nas áreas de aplicações científicas e empresariais (SEBESTA, 2018) enquanto as ferramentas são meios intangíveis e digitais que auxiliam a organização e gerenciamento das informações.

Pressman e Maxim (2016, p. 41) definem um software como:

[...] instruções (programas de computador) que, quando executadas, fornecem características, funções e desempenho desejados; estruturas de dados que possibilitam aos programas manipular informações adequadamente; e informação descritiva, tanto na forma impressa quanto na virtual, descrevendo a operação e o uso dos programas.

O desenvolvimento de um software possui estruturas e atividades complexas compostas por ferramentas e linguagens de programação, adotadas por tecnologias e pessoas em diferentes domínios. Takahashi (2009, p. 83) reforça a necessidade em “distinguir o conhecimento como ferramenta para mediar a interação com a realidade e o conhecimento como modo de busca sistemática da estrutura da realidade”. Nesse contexto, o autor atribui o termo “conhecimento tecnológico do tipo científico” ao corpo de conhecimentos, compatível com a ciência contemporânea e controlável pelo método científico, o qual é empregado para controlar, transformar ou criar coisas ou processos, naturais ou sociais. Além disso, apresenta características sucessivas de reconstruções e formulações teóricas nas comunidades científicas.

O uso de ferramentas tecnológicas otimiza as atividades metodológicas e customiza a produtividade dos trabalhadores do conhecimento diante dos desafios contínuos da sociedade. Podemos chamar o uso das tecnologias como um processo de codificação do conhecimento para cumprir o seu papel de armazenamento e posterior captura. Essa perspectiva também envolve a construção de aparatos tecnológicos para estudar as teorias e aplicações adjacentes do campo do conhecimento.

A categoria 8 “Ambientes” corresponde a uma perspectiva ambiental mais ampla (macroambiente) do conhecimento registrado, ou seja, o ambiente principal de informação que integra a produção do conhecimento dos diversos aspectos da sociedade contemporânea. O ambiente geral é caracterizado por um conjunto de condições comuns:

As condições tecnológicas, econômicas, legais, políticas, culturais, demográficas e ecológicas são fenômenos ambientais que formam um campo dinâmico de forças que interagem entre si. Esse campo de forças tem efeito sistêmico. O ambiente geral é genérico e comum para todas as organizações, mas cada uma delas tem o seu ambiente particular: o ambiente de tarefa (CHIAVENATO, 2004, p. 514).

Trazendo a ideia de Chiavenato (2004) para esta categoria, subentende-se que o ambiente tarefa constitui o ambiente de operações mais próximo da investigação científica em conformidade com o segmento do conhecimento registrado, seja teórico ou aplicado, composto por entradas (objetivos da pesquisa) e saídas (resultados da pesquisa).

Esta categoria foi codificada por meio do ambiente real ou objeto de estudo das pesquisas (teses). Nesse ambiente, o pesquisador necessita coletar dados e informações para observar e esclarecer os fenômenos de sua pesquisa, seja ela teórica ou aplicada. O ambiente real também ajuda a definir a problemática e as questões metodológicas da pesquisa científica.

A categoria 9 “Organizações” refere-se ao ambiente empírico da pesquisa científica. Corresponde ao aspecto organizacional da provisão de informações. São agrupadas tanto no aspecto organizacional (divisão por setores da economia) quanto no aspecto da organização de informações (divisão por funções). Todos esses aspectos estão relacionados ao gerenciamento do conhecimento organizacional constituído por 3 (três) elementos: pessoas, entidades sociais e conhecimento.

As organizações são categorizadas em 2 (dois) grupos com respectivas subcategorias: tipo funcional (organizações de memória) e tipo organizacional (setores da sociedade - primeiro setor, segundo setor, terceiro setor; setor internacional e organizações científicas).

As organizações do tipo funcional correspondem às “organizações de memória” que visam atender as necessidades informacionais dos usuários por meio das principais atividades: identificação, aquisição, processamento e transmissão, caracterizadas pela tangibilidade dos produtos de informação (BORGES, 2008) e são destinadas aos museus e patrimônio cultural, bibliotecas, centros de documentação, arquivos e hemerotecas. As organizações do tipo organizacional são os setores da sociedade que correspondem aos interesses públicos, privados e sociais, respectivamente. Cada um deles apresenta a sua composição e finalidade.

Ainda no tipo organizacional, tem-se as organizações do “setor internacional” que engloba todas as pesquisas centradas no ambiente empírico internacional. E para as “organizações científicas”, a ideia conceitual concentra-se nos princípios da Administração Científica, criada por Taylor com alguns aspectos como ponto de partida:

É o primeiro esforço científico para analisar e padronizar os processos produtivos com o objetivo de aumentar a produtividade e a eficiência; Desenvolve técnicas e métodos que racionalizam a produção, logrando forte aumento da produtividade (CHIAVENATO, 2004, p. 72).

Tendo-se como base as ideias de Taylor e Chiavenato, podemos considerar as organizações científicas como “empresas” ou “associações” que promovem o compartilhamento e a divulgação do conhecimento científico. Apresentam soluções e princípios distribuídos por temáticas para viabilizar o aprimoramento das competências e incentivar as especializações por meio da colaboração entre os trabalhadores do conhecimento com fins científicos. Essa sistemática desenvolve um aumento da eficiência e da eficácia nos índices de produtividade científica.

Esta categoria apresentou as codificações das teses que desenvolveram suas pesquisas por meio de estruturas organizacionais em diversos setores da economia sob as condições políticas, econômicas e sociais. Essa categoria tem conexões teóricas e práticas com a categoria “Ambientes”, uma vez que:

Assim como na natureza, os ambientes organizacionais contêm outras organizações e estas, a princípio, são capazes de influenciar seus ambientes, desempenhando um papel ativo, especialmente quando agem em conjunto (TURETA; ROSA; ÁVILA, 2006, p. 8).

A categoria 10 “Usuários” exemplifica os usuários potenciais dos recursos vinculados e do conhecimento produzido e registrado. Esses usuários podem ser identificados por meio das recomendações de pesquisa (recomendações teóricas e operacionais) e dos grupos e comunidades direcionados aos interesses informacionais.

O fluxo de informações e o acesso aos recursos de conhecimento influenciam os padrões de busca dos usuários. O principal objetivo desta categoria é desenvolver significados, onde o usuário pode ver o que foi desenvolvido nas pesquisas e criar a sua interpretação (ideia). A partir dessa percepção, os usuários interessados nas temáticas apresentadas, podem realizar novas pesquisas, desenvolver melhorias, validar teorias e práticas e realizar ações sistemáticas do conhecimento científico.

Belkin (1980) percebe que o usuário nem sempre sabe expressar a sua necessidade de informação, visto que ainda não sabe exatamente o que está precisando ou faltando para solucionar o seu problema. Por isso, Choo (2006, p. 101) complementa que “satisfazer uma necessidade de informação vai muito além do que encontrar informações que respondam aos questionamentos dos usuários”. Desta forma, a contribuição da representação do conhecimento em níveis categóricos facilita a compreensão dos processos e gera valor aos objetivos e à intensidade da busca informacional.

Essa categoria foi indexada a partir das considerações finais e recomendações de pesquisas futuras registradas nas teses codificadas para que possam auxiliar na busca por interesses informacionais de usuários potenciais em determinadas temáticas da CI.

Enfatizamos que a categorização científica está direcionada à dimensão aplicada da Organização do Conhecimento por considerar a autenticidade da tradução do conhecimento exposta nesta investigação por meio da adoção dos principais termos técnicos e contextos encontrados nas pesquisas de doutorado. O significado e a sinalização do uso potencial do conhecimento categorizado podem proporcionar relações e reflexões entre a realidade observada, o aspecto temporal das pesquisas e as ações científicas.

As características sociais e epistemológicas advêm a partir dessa categorização com uma forma de organizar, investigar e orientar o desenvolvimento científico da Ciência da Informação. A comunicação e mediação desse desenvolvimento requerem estruturas que possibilitem adaptações contínuas.

4 Conclusão

A Ciência da Informação possui um papel crucial no avanço do conhecimento científico: construir pontes entre as áreas do conhecimento por meio de estratégias de organização e recuperação da informação e do conhecimento que visem responder aos questionamentos dos usuários finais e potenciais.

Destacamos que as bases metodológicas criadas neste estudo podem subsidiar a expansão da sua aplicabilidade na evolução do conhecimento científico e identificar em qual concepção teórica, a Ciência da Informação está mais operante em um determinado espaço de tempo.

Os 6 (seis) modelos da CI elaborados por Zins (2007b) baseados em visões diferentes de vários pesquisadores: modelo de hi-tech (alta tecnologia); modelo de tecnologia; modelo de cultura e sociedade; modelo do mundo humano; modelo do mundo vivo e modelo do mundo vivo e físico podem ser identificados em todas as produções científicas da área, sendo que uns têm mais incidência do que outros em cada aspecto temporal das pesquisas.

As transformações sociais, culturais, epistêmicas e tecnológicas da sociedade exigem que o meio científico descubra a dialogicidade das ideias com a finalidade de solucionar os problemas informacionais. Para isso, a CI aborda o campo da Organização e Representação do Conhecimento, como por exemplo, a elaboração e utilização de categorias que possuem uma estruturação simples e objetiva para facilitar a mediação do conhecimento.

A comunicação científica necessita de uma dinamicidade do pensamento sistêmico de cada comunidade discursiva. Nesse sentido, observamos na codificação, um coletivismo metodológico diante da pluralidade de disciplinas e linhas de pesquisa das teses de doutorado, em que reforça a construção de pensamentos inter, multi e transdisciplinares, ao mesmo tempo em que alerta para a identificação das tendências e campos de atuação científica.

Essas tendências têm chances de acompanhar os progressos da quarta revolução industrial (SCHWAB, 2016) e da mais recente quinta revolução industrial (MARTINEZ, 2019) com o uso de tecnologias ubíquas e da inteligência artificial para alavancar as ferramentas tecnológicas executadas no mapeamento e mediação do conhecimento.

Os modelos argumentativos entre as teorias e as práticas são relevantes para a definição dos fenômenos revelados nas pesquisas sob uma perspectiva categórica. A ideia da categorização científica possibilita o fomento dos grupos de pesquisa e pós-graduação da área, assim como a determinação das futuras agendas de pesquisa ou estrutura curricular.

Referências

- ALMEIDA, Josiana Florêncio Vieira Régis de. **Contribuição metodológica para estruturação e mediação do conhecimento científico em Ciência da Informação**. 2021. 280 f. Tese (Doutorado em Gestão e Organização do conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2021.
- ARAÚJO, Carlos Alberto Ávila. **Arquivologia, biblioteconomia, museologia e ciência da informação: o diálogo possível**. Brasília, DF: Briquet de Lemos, 2014.
- ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO (ANCIB). **Estatuto**. Disponível em: <https://ancib.org/wp-content/uploads/2020/10/Estatuto-2018-compactado.pdf>. Acesso em: 15 dez. 2020.
- BACHELARD, Gaston. **O novo espírito científico**. Lisboa: Edições 70, 2008.
- BELKIN, N. J. Anomalous States of Knowledge as a basis for information retrieval. **Canadian Journal of Information and Library Science**, [S.l.], v. 5, n. 1, p. 133-143, 1980.
- BORGES, Mônica Erichsen Nassif. O essencial para a gestão de serviços e produtos de informação. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 5, n. 2, p. 115-128, 2008. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2007>. Acesso em: 14 dez. 2020.
- BURKE, Peter. **O que é história do conhecimento?** São Paulo: Editora UNESP, 2016.
- CAPURRO, Rafael. **Knowledge Map of Information Science. Rafael Capurro's responses to Chaim Zins**. 2006. Disponível em: <http://www.capurro.de/zins.html>. Acesso em: 25 ago. 2019.
- CASSOTTA, Maria Luiza Jurema; LUCAS, Alexandre; BLATTMANN, Ursula; VIERA, Angel Freddy Godoy. Recursos do conhecimento: colaboração, participação e compartilhamento de informação científica e acadêmica. **Informação & Sociedade: Estudos**, João Pessoa, v. 27, n. 1, p. 17-34, jan./abr. 2017.
- CHIAVENATO, Idalberto. **Introdução a teoria geral da administração: uma visão abrangente da moderna administração das organizações**. 3. ed. compacta. Rio de Janeiro: Elsevier, 2004.
- CHOO, Chun Wei. **A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões**. 2. ed. São Paulo: Editora SENAC, 2006.
- DRUCKER, Peter F. **O melhor de Peter Drucker: o homem, a administração, a sociedade**. São Paulo: Nobel, 2002.
- FLECK, Ludwik. **Gênese e desenvolvimento de um fato científico**. Belo Horizonte: Fabrefactum, 2010.
- HJØRLAND, B. Domain analysis in information Science: eleven approaches, tradicional as well as innovative. **Journal of Documentation**, [S.l.], v. 58, n. 4, p. 422-462, 2002.
- HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. **Journal of the American Society for Information Science and Technology**, [S.l.], v. 46, n. 6, p. 400-425, 1995.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Estrutura CineBrasil 24082020. Disponível em: http://download.inep.gov.br/pesquisas_estatisticas_indicadores_educacionais/cinebrasil/documentos/Estrutura-Cine-Brasil-24082020.xlsx. Acesso em: 15 dez. 2020.
- MARTINEZ, Inma. **La quinta revolución industrial: como la comercialización del espacio se convertirá en la mayor expansión industrial del siglo XXI**. Tradução de Veronica Puertollano López. **España**: Deusto, 2019.
- MORIN, Edgar. **A cabeça bem-feita: repensar a reforma, reformar o pensamento**. Tradução de Eloá Jacobina. 7. ed. Rio de Janeiro: Bertrand Brasil, 2006.
- NONAKA, I.; TOYAMA, R.; KONNO, N. SECI. Ba and Leadership: a unified model of dynamic knowledge creation. **Long Range Planning**, [S.l.], v. 33, 2000.
- OLIVEIRA, Alfredo Gontijo de. Complementaridade e transdisciplinaridade. In: PAULA, João Antônio de. **A transdisciplinaridade e os desafios contemporâneos**. Belo Horizonte: Editora UFMG, 2011.
- PRESSMAN, R. S.; MAXIM, B. R. **Engenharia de software: uma abordagem profissional**. 8. ed. Porto Alegre: AMGH, 2016.
- SCHWAB, Klaus. **A quarta revolução industrial**. São Paulo: Edipro, 2016.
- SEBESTA, Robert W. **Conceitos de linguagem de programação**. 11. ed. Porto Alegre, Bookman, 2018.
- SMIRAGLIA, R. P. **The elements of knowledge organization**. Switzerland: Springer International Publishing, 2014.
- SMIRAGLIA, Richard. **Domain analysis for knowledge organization: tools for ontology extraction**. Reino Unido: Elsevier Chandos Publishing, 2015. 110p.
- STRAUSS, Anselm; CORBIN, Juliet. **Pesquisa qualitativa: técnicas e procedimentos para o desenvolvimento de teoria fundamentada**. 2. ed. Porto Alegre: Artmed, 2008. 288p.
- TAKAHASHI, Ricardo H. C. **A estrutura do conhecimento tecnológico do tipo científico**. Belo Horizonte: Editora UFMG, 2009.
- TENNIS, Joseph T. Two Axes of Domains for Domain Analysis. **Knowledge Organization**, [S.l.], v. 30, n. 3-4, p. 191-195, 2003.
- TURETA, César; ROSA, Alexandre Reis; ÁVILA, Silvio César. Da teoria sistêmica ao conceito de redes interorganizacionais: um estudo exploratório da teoria das organizações. **Revista de Administração da UNIMEP**, [S.l.], v. 4, n. 1, p. 1-16, 2006.
- ZINS, C. Conceptual approaches for defining 'data', 'information', and 'knowledge'. **Journal of the American Society for Information Science and Technology**, [S.l.], v. 58, n. 4, p. 479-493, 2007a.
- ZINS, C. Conceptions of Information Science. **Journal of the American Society for Information Science and Technology**, [S.l.], v. 58, n. 3, p. 335-350, 2007b.
- ZINS, C. Knowledge map of Information Science. **Journal of the American Society for Information Science and Technology**, [S.l.], v. 58, n. 4, p. 526-535, 2007c.

Mineral Exploration Knowledge Organization System (MEKOS): um modelo conceitual para o domínio da exploração mineral

*Mineral Exploration Knowledge Organization System (MEKOS):
a conceptual model for the mineral exploration domain*

Gabriela Silva Caetano (1); Carlos Henrique Marcondes de Almeida (2)

(1) Universidade Federal de Minas Gerais, Av. Pres. Antônio Carlos, 6627 - Pampulha, Belo Horizonte - MG, 31270-901; gbrila.caetano@gmail.com (2) Universidade Federal Fluminense, R. Des. Ellis Hermydio Figueira, 783 - Aterrado, Volta Redonda - RJ, 27213-145; ch_marcondes@id.uff.br

Resumo

No contexto corporativo, os SOC's são criados para descrever o domínio de conhecimento do negócio dentro do funcionamento interno dos sistemas. Esta pesquisa teve como objetivo a elaboração de um modelo conceitual para o domínio da exploração mineral. Sua primeira etapa consistiu na elaboração de um esquema básico dos processos de negócio ao qual o modelo de domínio do conhecimento servirá. Realizou-se pesquisa documental para o conhecimento dos principais SOC's da área de Geociência. Após foi aplicado o método dedutivo, como forma de validar o modelo conceitual. O modelo foi denominado *Mineral Exploration Knowledge Organization System (MEKOS)*. Entende-se que ele deva ser um pouco mais do que um mapa estático, como se apresentam os modelos conceituais. Portanto, ele foi modelado no software Multites, especializado na construção de taxonomias e tesouros. Esta pesquisa pretende contribuir para a facilitação do processo de indexação dos documentos e sua recuperação. Com o desenvolvimento de SOC's dentro das corporações podemos avançar na unificação das informações, como resultado teremos um melhor gerenciamento dos ativos de informação, além de tornar a inteligência de negócio mais precisa e confiável. A realização de um estudo sobre a modelagem de domínios do conhecimento para a exploração mineral possui relevância tanto para a aplicação prática quanto teórica. Por um lado, visa à futura construção de um SOC numa empresa de mineração e, por outro, ao debater a Organização do Conhecimento no contexto corporativo, contribui com a solidez desse campo de pesquisa, tendo em vista a economia crescente em volta dos bens minerais.

Palavras-chave: Modelo Conceitual. Ontologias Corporativas. Exploração Mineral.

Abstract

In the corporate context, KOS are created to describe the business knowledge domain within the inner workings of systems. This research aimed to develop a conceptual model for the field of mineral exploration. Its first step consisted in the elaboration of a basic schema of the business processes which the knowledge domain model will serve. Documentary research was carried out to understand the main KOS in the area of Geoscience. Afterwards, the deductive method was applied, as a way to validate the conceptual model. The model was named *Mineral Exploration Knowledge Organization System (MEKOS)*. It is understood that it should be a little more than a static map, as conceptual models are presented. Therefore, it was modeled in Multites software, specialized in building taxonomies and thesauri. This research intends to contribute to the facilitation of the document indexing process and its recovery. With the development of KOS within corporations, we can advance in the unification of information, as a result we will have better management of information assets, in addition to making business intelligence more accurate and reliable. Carrying out a study on the modeling of knowledge domains for mineral exploration is relevant for both practical and theoretical application. On the one hand, it aims at the future construction of a KOS in a mining company and, on the other hand, by debating the Organization of Knowledge in the corporate context, it contributes to the solidity of this field of research, in view of the growing economy around the minerals.

Keywords: Conceptual Model. Corporate Ontologies. Mineral Exploration.

Resumen

En el contexto corporativo, los SOC se crean para describir el dominio de conocimiento empresarial dentro del funcionamiento interno de los sistemas. Esta investigación tuvo como objetivo desarrollar un modelo conceptual para el campo de la exploración mineral. Su primer paso consistió en la elaboración de un esquema básico de los procesos de negocio a los que servirá el modelo de dominio del conocimiento. Se realizó una investigación documental para comprender los principales SOC's en el área de Geociencias. Posteriormente se aplicó el método deductivo para validar el modelo conceptual. El modelo se denominó *Mineral Exploration Knowledge Organization System (MEKOS)*. Se entiende que debe ser más que un mapa estático, como se presenta en los modelos conceptuales. Por lo tanto, se modeló en el software Multites, especializado en la construcción de taxonomías y tesouros. Esta investigación pretende contribuir a la facilitación del proceso de indexación de documentos y su recuperación. Con el desarrollo de los SOC dentro de las corporaciones, podemos avanzar en la unificación de la información, como resultado tendremos una mejor gestión

de los activos de información, además de hacer que la inteligencia de negocios sea más precisa y confiable. Realizar un estudio sobre el modelado de dominios de conocimiento para la exploración mineral es relevante tanto para su aplicación práctica como teórica. Por un lado, apunta a la futura construcción de un SOC en una empresa minera y, por otro lado, al debatir la Organización del Conocimiento en el contexto empresarial, contribuye a la solidez de este campo de investigación, en vista de la creciente economía en torno a los minerales.

Palabras clave: Modelo conceptual. Ontologías corporativas. Exploración de minerales.

1 Introdução

O conhecimento, para ser compartilhado, necessita de formalização lógica de fatos e ideias, aspectos semânticos que representam a realidade tratada. A organização do conhecimento está presente no controle do vocabulário quando este estipula conceitos, evidencia suas relações, como a espiral do conhecimento de Ranganathan (2009) confere movimentos de geração de ideias conectadas, tal como os modelos conceituais, que se constituem como primeira etapa para a elaboração de um Sistema de Organização do Conhecimento (SOC). No contexto corporativo, os SOCs são criados para descrever o domínio de conhecimento do negócio dentro do funcionamento interno dos sistemas. Eles dão acesso a uma gama variada de recursos informacionais, em especial documentos técnicos. A descrição deste domínio é usada nas operações de representação da informação que estão em linguagem natural no documento técnico produzido pela instituição. Assim, os SOCs auxiliam a precisão da recuperação da informação a partir da qualidade da entrada dos metadados que representam os documentos e da saída com resultados relevantes para os usuários.

Esse trabalho compõe uma parte dos resultados de uma dissertação de mestrado [1] que propôs a criação de um modelo conceitual para o domínio da exploração mineral.

2 Aportes teóricos

Os modelos conceituais representam uma categorização possível de um domínio do conhecimento. Essa categorização refletirá em uma estrutura de conceitos que formará uma linguagem controlada. Essa linguagem se baseia no processo cognitivo de apreensão do conhecimento utilizado para representar os objetos de interesse neste domínio ou consultá-los.

Domínio do conhecimento pode ter outros sinônimos como área do conhecimento, universo conceitual ou, simplesmente, domínio. No caso de SOCs em ambientes corporativos esses termos expressam o ambiente, empresa, local que se pretende aplicar os conhecimentos da organização do conhecimento para criação de um SOC para formalizar as estruturas conceituais desse domínio.

Essa representação de conceitos a partir de domínios de conhecimento pode facilitar a detecção da produção de

conhecimento de um dado domínio, já que o mundo é pensado e entendido a partir da linguagem.

A falta de padronização na prática da organização do conhecimento favorece a uma compreensão equivocada dos conceitos advindos dos domínios, dificultando também uma base terminológica apropriada, ou ainda um vocabulário que facilite a recuperação de informações relevantes para usuários e que possa ser compartilhado entre domínios como é o caso das ontologias (NOVO, 2014, p. 32).

Elaborar um SOC é uma atividade que implica na modelagem de domínio do conhecimento. De acordo com Souza (2017),

durante o processo de modelagem que compreende a construção de um SOC, conceitos são selecionados, representados e relacionados. A modelagem de um SOC pode ter como objetivo criar uma aplicação utilitária para uso em sistemas de informação específicos ou a representação de um domínio do conhecimento, que pode ter um âmbito de aplicação mais amplo dentro da comunidade científica. (SOUZA, 2017, p. 37).

A representação de um domínio do conhecimento é o princípio norteador para a organização de informação (CAMPOS; GOMES, 2003). As autoras Campos e Gomes, em outro artigo de 2014, ao abordarem a modelagem de domínios de acordo com Barry Smith e Ingetraut Dahlberg, complementam que:

Um dos fatores fundamentais na elaboração de modelos conceituais de domínios é a determinação da unidade de representação que irá compor a estrutura de classes deste domínio, ou seja, de que natureza são os elementos de representação que serão utilizados na taxonomia de um domínio. (CAMPOS; GOMES, 2014, p. 84).

A modelagem de um domínio do conhecimento está associada à representação das coisas no mundo, tem como objetivo entender, comunicar e raciocinar sobre uma porção da realidade (GUIZZARDI, 2005). Essa abstração constitui como uma importante ferramenta para aquisição de conhecimento. Para tanto, é necessário selecionar os objetos de maior relevância dentro do universo investigado de maneira que “constroem-se, assim, esquemas abstratos da realidade, nos quais as coisas são reduzidas a seus perfis mais convenientes” (SAYÃO, 2001, p.82).

Hjørland (2021, tradução proponente), no verbete *Domain analysis da Encyclopedia of Knowledge Organization*, define que um domínio pode ser uma disciplina, mas não somente. Trata-se de uma divisão do trabalho cognitivo que pode ser – teoricamente,

socialmente ou de forma institucionalizada – dividido. Portanto, os modelos de conhecimento não são divisões prontas do mundo, mas interpretações dinâmicas de recortes da realidade e se desenvolvem a partir de um objetivo.

A “análise de um domínio” é uma expressão também usada na Engenharia de Software para aquilo que é considerado, na Organização do Conhecimento, a classificação facetada (PRIETO-DÍAZ, 1990, tradução proponente). Além disso, o conceito domínio foi utilizado para relacionar o domínio à visão cognitiva para o desenvolvimento de inteligências artificiais.

O objeto de um domínio, ao qual se referem Hjørland e Hartel (2003), é considerado tudo o que o compõe. Numa perspectiva terminológica, são todos os substantivos, verbos e adjetivos utilizados para descrever esse campo de conhecimento. Termos que, externalizados ou comunicados, compõem um vocabulário.

Já as interações sociológicas, isto é, as atividades e vivências de quem faz parte daquele domínio, Saracevic (1975, tradução proponente) conceitua como “a visão do conhecimento do sujeito”. A literatura ressalta que os domínios nunca são inequívocos, eles são caracterizados por essa processualidade. Por isso, os domínios nunca estarão congelados no tempo, pois se alterarão à medida que novos conhecimentos são gerados e processos precisam ser revistos.

No contexto de empresas, Varma (2007) cita a importância dos SOCs tais como a ontologia para melhor gerenciamento dos ativos informacionais.

À medida que as organizações percebem a necessidade de quebrar os silos de informações existentes e questionar os limites estabelecidos no fluxo de informações na organização, há uma tendência emergente de adotar uma visão mais holística da arquitetura da informação. Essa tendência se deve ao fato de muitas empresas não responderem com rapidez suficiente às mudanças nas condições e serem capazes de se adaptar ao estouro de informações. Espera-se que o crescente interesse em ontologias e classificação de informações resulte em um melhor gerenciamento dos ativos de informação em uma organização. (VARMA, 2007, tradução proponente).

A informação enquanto ativo informacional é assim considerada, pois agrega valor à empresa, sendo documentos ou bases de dados que servem de embasamento para a execução das atividades. A organização e representação do conhecimento auxiliará na criação de estruturas categoriais, nas quais o conhecimento é representado por uma classificação materializada em um SOC. Como podem existir diversas maneiras ou critérios para classificar os objetos em um domínio essa classificação deve resultar em um consenso ou acordo entre os usuários deste domínio.

3 Metodologia para construção do Modelo Conceitual

Nesta pesquisa, a autora, que trabalha como profissional de informação, assume o papel de modelizadora (CAMPOS, 2001), sendo o profissional que intermedia os processos de elaboração do domínio do conhecimento e os usuários, promovendo o consenso no que diz respeito às entidades do Modelo Conceitual e os relacionamentos entre estas.

Nas subseções a seguir serão explicitadas as abordagens seguidas para a construção do Modelo Conceitual.

3.1 Cadeia de Valor do negócio

A primeira etapa da pesquisa consistiu na elaboração de um esquema básico, baseado nos processos de negócio ao qual o modelo de domínio do conhecimento servirá, o qual denominamos Cadeia de Valor.

A área da Exploração Mineral é dividida em três grandes categorias: Prospecção Mineral, etapa de estudo de áreas com potencial econômico; Tecnologia Mineral, fase onde são estudadas tecnologias para melhor aproveitamento dos minerais encontrados; e Desenvolvimento de Projetos que ocorre concomitante às diferentes fases de avaliação do depósito mineral.

Foram identificados 17 termos, provenientes dos processos-chave da empresa, nesta pesquisa denominaremos de Cadeia de Valor.

3.1 SOCs da geociências

Realizou-se pesquisa documental para o conhecimento dos principais sistemas de organização do conhecimento da área de Geociência. Inicialmente, verificou-se se o SOC estava disponível para uso, podendo ser pesquisável ou ter capacidade de ser processado, através de OCR por exemplo.

O primeiro SOC analisado foi o AMTG (COORDINATING COMMITTEE FOR GEOSCIENCE PROGRAMMES IN EAST AND SOUTHEAST ASIA, 2006), um importante instrumento para compartilhamento de dados geocientíficos entre os países asiáticos. Possui 633 páginas, com termos em 11 línguas, o que possibilita uma grande abrangência de integração. Observou-se que *Geochemistry* (Geoquímica) e *Mineralogy* (Mineralogia) são termos listados como categorias e termos. Em *Geotechnics* (Geotecnia), é listado *Engineering Geology* (Engenharia Geológica) como categoria e depois como “use” para o termo mesmo termo *Geotechnis*.

Já o FintoGeo (GEOLOGICAL SURVEY OF FINLAND, 2016), é uma ontologia completa para a área de geociências, sobretudo, por possuir outros domínios do conhecimento. Assim como em outros SOCs analisados, requer definição dos termos para auxílio do escopo em que ele deve ser utilizado. Além disso, alguns

termos não possuem indicação de nenhuma das relações utilizadas, tais como o *Broader Concept*, *Narrower Concept*, e *Related Concept*.

O GEODESC (CPRM, 2005) é um importante, se não o principal, tesouro de geociências no Brasil. No entanto, são necessários alguns apontamentos a partir da análise feita do tesouro. Em alguns momentos, os termos são muito abrangentes e, em outros, muito restritos como, por exemplo, em Geoquímica que representado como TR (termo restrito) possui termos como “Biogeoquímica”, “elementos e gases”, e “Base de Dados”.

O *Glossary of Geology* (NEUENDORF; MEHL JUNIOR, 2011) é o SOC analisado que mais possui termos. Apesar de ser um glossário, apresenta diversos tipos de relações possíveis entre os termos. O *American Institute Geosciences* possui também um tesouro, o GEOREF, ao qual não foi possível acessar para esta pesquisa. O glossário e o tesouro são usados para indexar documentos em várias bases bibliográficas de Geociências, como por exemplo o *GeoScienceWorld*, o que evidencia forte influência sobre o SOCs.

A USGS *Geological Survey* (USGS, 2023) é uma importante instituição para a área de geociências. O USGS *Thesaurus* possui poucos termos, mas conceitua cada termo, o que é um diferencial entre os demais. Apesar de não ser exaustivo, e não indexar termos que aparecem na Cadeia de Valor, os termos de ambos são similares. A relação hierárquica do USGS *Thesaurus* possui mais de um nível, como, por exemplo, o termo “análise química”, ao ser pesquisado, recupera um termo geral e vários específicos.

3.2 Análise dos SOCs

A partir disso, dentro dos SOC identificados foram pesquisados os termos da cadeia de valor pelo modelizador, que verifica a adequação do mapeamento. Esse procedimento está de acordo com o que prevê a Norma ISO 25964-2:2013, “a mediação humana deve ser usada sempre que possível para alcançar um desempenho de recuperação superior e evite falsas inferências de equivalências inexatas”. (ISO 25964-2:2013, p. 31). Para cada termo da cadeia de valor, foi elaborada uma ficha de análise contendo os termos gerais; termos específicos; qual definição; qual sua faceta/classe.

A comparação dos termos do mapeamento de processos nos SOCs visou a entender como outras instituições que elaboraram SOCs para a área de geociências estruturaram esse universo de conhecimento.

As relações foram baseadas na norma ISO 25964-2:2013 que descreve a “equivalência”, “hierárquica” e “associativa”. Optou-se por dividir a relação hierárquica em duas para diferenciar a relação explícita do SOC daquela feita a partir do conhecimento do modelizador, conforme previsto também na ISO 25964-2:2013, nos

casos em que nenhum equivalente exato esteja disponível, “o indexador faz uma escolha entre qualquer mapeamento inexato e/ou hierárquico, selecionando aquele que melhor se adequa ao documento em questão” (ISO 25964-2:2013, p. 31)

A opção “Não encontrado” foi colocada para indicar que não houve nenhum tipo de relação quando for o caso.

Com os insumos da ficha de análise dos termos foi elaborado um modelo conceitual para cada um dos SOCs analisados.

3.3 Entrevistas

Após a elaboração dos mapas conceituais dos cinco SOCs analisados, foi aplicado o método dedutivo.

Como forma de validar o modelo conceitual solicitou a funcionário da empresa que possuem conhecimento de todos os processos por terem um escopo de atuação multidisciplinar, que analisassem cada um dos modelos conceituais e indicassem os termos e/ou estruturas poderiam agregar o modelo conceitual. Entrevistou-se quatro integrantes da amostra a partir de um roteiro semiestruturado. No final das entrevistas, foram gerados 20 mapas conceituais comentados.

Nas entrevistas, foram feitos comentários não somente sobre estrutura ali apresentada, mas também indicados outras mudanças que poderiam ser pertinentes. Como exemplo, uma pessoa entrevistada pontuou que na perspectiva dela o termo *Petrology* (Petrologia) no contexto que estava sendo apresentado, deveria ser na verdade *Geomorphology* (Geomorfologia).

A partir disso, todas as indicações feitas pelos entrevistados foram analisadas, e tudo que fosse um consenso da maioria seria agregado no modelo conceitual da Exploração Mineral, quando houve empate a autora como a 5ª pessoa decidiu se estava adequado ao contexto da Exploração Mineral ou não.

3.4 A construção do modelo conceitual

Nessa etapa, se estrutura o conhecimento do domínio estabelecendo as classes e relações, tendo como resultado um modelo conceitual do domínio.

As relações utilizadas foram somente as hierárquicas previstas pela ISO 25964-2:2013: “*Broader*” - o termo que representa um conceito com um significado mais amplo; “*Narrower*” - o termo que se refere a um conceito com significado mais específico.

O modelo conceitual da exploração mineral foi desenvolvido em até seis níveis hierárquicos: o primeiro, o segundo e parte do terceiro nível são compostos por termos da estrutura da Cadeia de Valor. Os novos termos, retirados a partir dos dados da pesquisa, estão no quarto nível em diante. Ele possui 112 termos, sendo 95 adicionados a partir da análise dos SOCs, e os demais 17 provenientes da Cadeia de Valor.

O diagrama completo pode ser consultado no Apêndice A.

Visando iniciar o desenvolvimento de um sistema de organização do conhecimento para a exploração mineral, o modelo conceitual da exploração mineral foi nominado *Mineral Exploration Knowledge Organization System* (MEKOS).

Entende-se que ele deva ser um pouco mais do que um mapa estático, como se apresentam os modelos conceituais. Portanto, ele foi modelado no software Multites (2023), especializado na construção de taxonomias e tesouros.

4 Metodologia para construção do Modelo Conceitual

Com a intenção de expandir o SOC da Exploração Mineral, o MEKOS foi hospedado em um site como um serviço de web, no qual as pessoas podem ler sobre o MEKOS, navegar por sua taxonomia e o modelo conceitual da Exploração Mineral, além de sugerir novos termos. As figuras abaixo representam prints do MEKOS.

Na página da Home, foram adicionadas informações sobre esta pesquisa, o MEKOS e a autora.,

Figura 1. Home - MEKOS, 2021



Na página “Modelo Conceitual”, foram adicionadas informações sobre Modelos Conceituais e o Modelo Conceitual da Exploração Mineral, que foram retiradas desta dissertação.

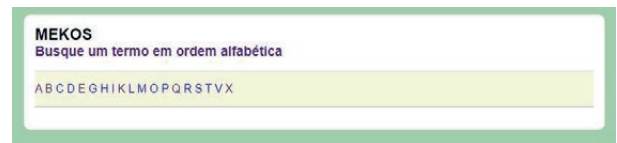
Caetano, Gabriela Silva; Marcondes de Almeida, Carlos Henrique. Mineral Exploration Knowledge Organization System (MEKOS): um modelo conceitual para o domínio da exploração mineral. *Organização e Representação do Conhecimento em diferentes contextos: desafios e perspectivas na era da datificação*, 2023.

Figura 2. Modelo Conceitual - MEKOS, 2021



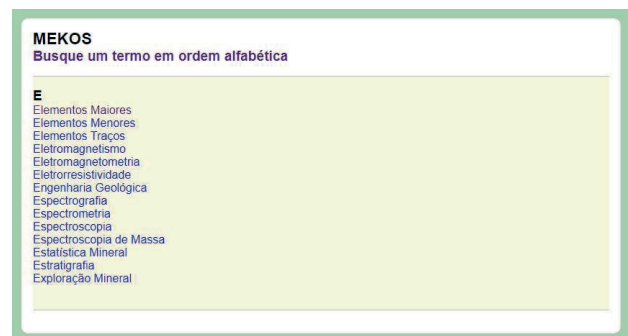
Na página do MEKOS, é possível consultar a taxonomia desenvolvida. Nesta primeira página, pode-se pesquisar e visualizar termos pela ordenação alfabética.

Figura 3. Busque um termo em ordem alfabética - MEKOS, 2021



Após selecionada a letra que corresponde à primeira letra do termo buscado, a tela abrirá todos os termos correspondentes àquela letra.

Figura 4. Termos iniciados com a letra E - MEKOS, 2021



Ao selecionar o termo, sua ficha terminológica será aberta. Nas fichas terminológicas, estão todas as relações e categorias previstas para o termo.

Figura 6. Termo Geofísica - MEKOS, 2021



Na última página, há campos que os usuários podem preencher para sugerirem novos termos a serem agregados ao MEKOS.

Figura 7. Campo para sugestão de termo - MEKOS, 2021

SUGERIR NOVO TERMO

Formulário de sugestão de um novo termo

Nome* Termo

E-mail* Tema

Definição

Prospecção Mineral

Enviar

5 Considerações finais

Esta pesquisa foi motivada pela necessidade de subsídios teóricos para sustentar a construção de um sistema de organização do conhecimento para o domínio da Exploração Mineral. Nesse sentido, a pesquisa buscou desenvolver insumos que refletissem a necessidade do uso de sistemas de organização do conhecimento em corporações, entender como é o desenvolvimento de modelos conceituais, identificar categorias existentes em outros sistemas de organização do conhecimento relacionados com a área de geociências e incluir aquelas pertinentes para construir o modelo conceitual da Exploração Mineral.

Esta pesquisa pretende contribuir para a facilitação do processo de indexação dos documentos e sua recuperação. O objetivo da pesquisa ora relatada, objetivava-se desenvolver um tesouro, no entanto, constatou-se como uma grande tendência na gestão do

conhecimento em organizações, a utilização de ontologias corporativas para sistematizar, e facilitar o acesso e reuso das suas informações e dos conhecimentos gerados.

Com o desenvolvimento de SOCs dentro das corporações podemos avançar na unificação das informações. Como resultado teremos um melhor gerenciamento dos ativos de informação, além de tornar a inteligência de negócio mais precisa e confiável.

A aplicação da organização do conhecimento na área da exploração mineral se mostra com um grande potencial, uma vez que o conhecimento desenvolvido e empregado nesse campo é heterogêneo, abrangendo diferentes domínios científicos. O domínio da exploração mineral possui uma complexa rede de relações entre conceitos desses domínios.

A realização de um estudo sobre a modelagem de domínios do conhecimento para a exploração mineral possui relevância tanto para a aplicação prática quanto teórica. Por um lado, visa à futura construção de um SOC numa empresa de mineração e, por outro, ao debater a Organização do Conhecimento no contexto corporativo, contribui com a solidez desse campo de pesquisa, tendo em vista a economia crescente em volta dos bens minerais. Ações para a padronização do conhecimento da Exploração Mineral auxiliarão na indexação e recuperação da informação que, por sua vez, ajudam a melhorar a produtividade, fomentar a pesquisa e gerar novos conhecimentos. No entanto, ressalta-se que, por ser uma área de grande fator de impacto econômico, a confidencialidade de dados e informações põem em desvantagem estudos de organização do conhecimento sobre a área.

Partindo da premissa de que um modelo conceitual é o primeiro passo para a elaboração de um SOC, foi desenvolvido o site MEKOS objetivando contribuir efetivamente com a organização das informações desse domínio de conhecimento. Acredita-se que a taxonomia resultante, pode ser uma contribuição para outros setores da sociedade que trabalhem com exploração de bens minerais, e que os conhecimentos gerados a partir dos passos metodológicos seguidos possam ser extrapolados para essas outras instituições e outros domínios corporativos semelhantes.

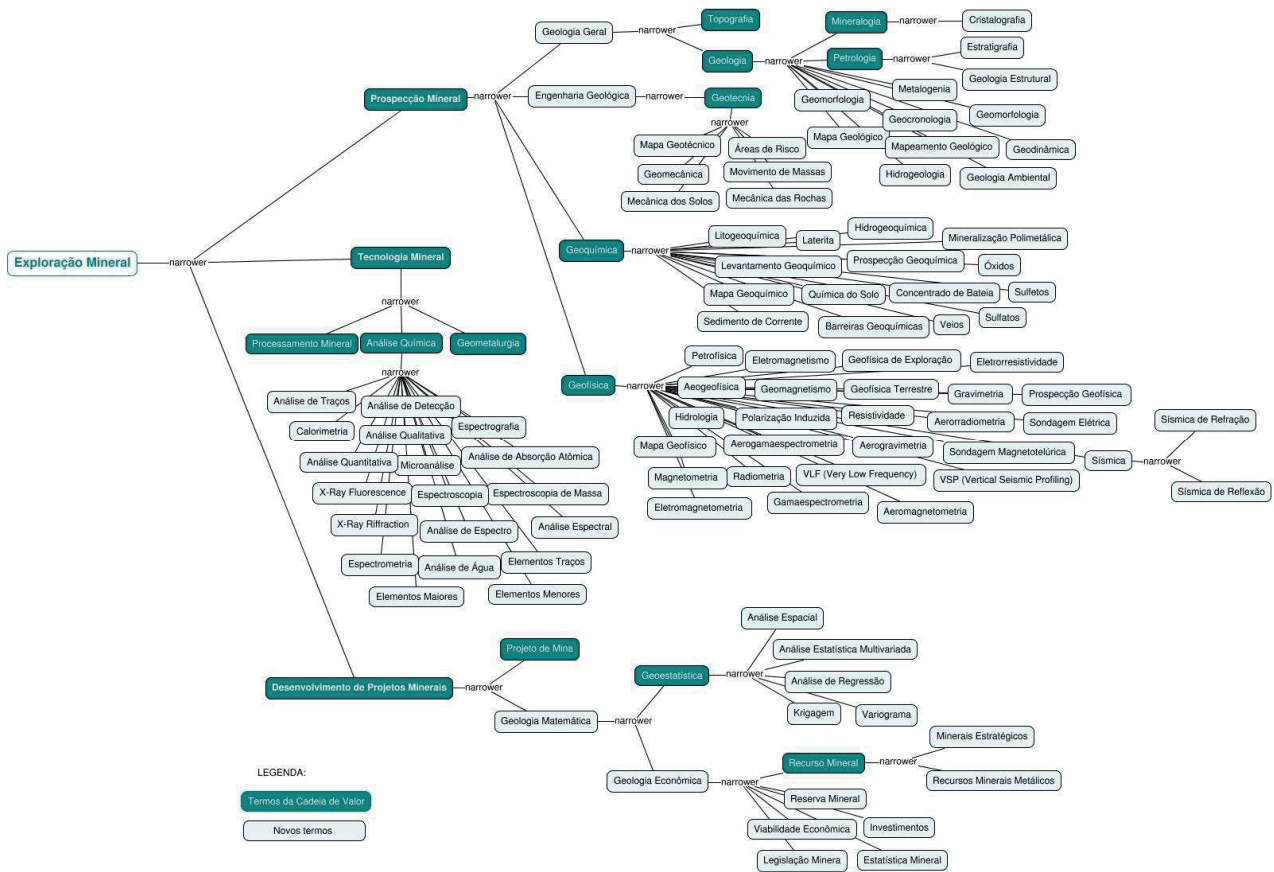
Notas

- [1] CAETANO, Gabriela Silva. Uso e desenvolvimento de sistemas de organização do conhecimento pelas corporações: criação de um modelo conceitual para o domínio da exploração mineral. 2022. 199 f. Dissertação (Mestrado) - Curso de Gestão e Organização do Conhecimento, Universidade Federal de Minas Gerais, Belo Horizonte, 2022. Disponível em: <https://repositorio.ufmg.br/handle/1843/39284>. Acesso em: 15 maio 2023.

Referências

- CAMPOS, M. L. A. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói: Eduff, 2001.
- CAMPOS, M. L. A.; GOMES, H. E. Organização de domínios de conhecimento e os princípios ranganathianos. **Perspect. cienc. inf.**, Belo Horizonte, v. 8, n. 2, p. 150-163, jul./dez. 2003. Disponível em: https://brapci.inf.br/_repositorio/2010/11/pdf_1ec0949a1a_0012907.pdf. Acesso em 15 out. 2020.
- CAMPOS, M. L. A.; GOMES, H. E. Princípios para modelagem de domínio: a posição de Barry Smith e de Ingtraut Dahlberg. **Ciência da Informação**, Brasília, v. 41, n. 1, p. 81-94, jan. 2014. Disponível em: <http://revista.ibict.br/ciin/article/view/1420>. Acesso em: 16 ago. 2020.
- COORDINATING COMMITTEE FOR GEOSCIENCE PROGRAMMES IN EAST AND SOUTHEAST ASIA. **Asian multilingual thesaurus of geosciences**. Thailand: Unesco, 2006. 563 p. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000146998>. Acesso em: 12 maio 2023.
- CPRM. **GEODESC**: vocabulário controlado em geociências. Rio de Janeiro: CPRM, 2005. Disponível em: https://www.cprm.gov.br/publique/media/rede_bibliotecas/geodesc.pdf. Acesso em: 12 maio 2023.
- GEOLOGICAL SURVEY OF FINLAND. **FINTO GEO – Geological ontology**. Helsinque: Geological Survey Of Finland, 2016. Disponível em: <https://dev.finto.fi/geo/en/>. Acesso em: 12 maio 2023.
- GUIZZARDI, G. **Ontological foundations for structural conceptual models**. 2005. 416 f. Thesis. ([Doutorado em Ciência da Computação]) - Universidade de Twente, Enschede, Holanda, 2005.
- HJØRLAND, B. Domain Analysis. In: **ISKO**: Encyclopedia of Knowledge Organization. Disponível em: https://www.isko.org/cyclo/domain_analysis. Acesso em: 12 jun. 2021.
- HJØRLAND, B.; HARTEL, J. Afterword: Ontological, epistemological and sociological dimensions of domains. **Knowledge Organization**, v. 30, n. 3-4, p. 239-245, 2003.
- INTERNATIONAL STANDARD ORGANIZATION. **ISO 25964**: thesauri and interoperability with other vocabularies. Part 2: Interoperability with other vocabularies. Geneve: International Standard Organization, 2013.
- MULTITES. **Thesaurus and Taxonomy Authoring Tools**. Disponível em: <http://multites.net/>. Acesso em: 12 maio 2023.
- NEUENDORF, K. K.e.; MEHL JUNIOR, J. P.; JACKSON, J. A. **Glossary of Geology**. [S.L.]: American Geosciences Institute, 2011.
- NOVO, H. F. **Análise conceitual e cognitiva**: Modac – um modelo dinâmico para auxiliar à construção de Sistemas de Organização do Conhecimento (SOC). 2014. 173 f. Tese (Doutorado em Modelagem da Geração e Difusão do Conhecimento) - Universidade Federal da Bahia, 2014. Disponível em: <https://repositorio.ufba.br/ri/handle/ri/19651>. Acesso em: 28 jun. 2021.
- PRIETO-DÍAZ, R. Domain analysis: An introduction ACM SIGSoft Software. **Engineering Notes**, v. 15, n. 2, p. 47-54, 1990.
- RANGANATHAN, S. R. **As cinco leis da Biblioteconomia**. Brasília: Briquet de Lemos, 2009.
- SARACEVIC, T. Relevance: A review of and a framework for the thinking on the notion in information science. **Journal of the American Society for Information Science**, v. 26, n. 6: p. 321-343, 1975.
- SAYÃO, L. F. Modelos teóricos em Ciência da Informação: abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 82-91, jan./abr. 2001.
- SOUZA, S. T. **Modelagem de domínios em sistemas de organização do conhecimento (SOC)**: uma investigação em tesaurus e ontologias para a informação legislativa. 2017. Tese. 327 f. (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2017.
- USGS. **USGS Thesaurus**. Disponível em: <https://apps.usgs.gov/thesaurus/thesaurus-full.php?thcode=>. Acesso em: 12 maio 2023.
- VARMA, Vasudeva. Use of Ontologies for Organizational Knowledge Management and Knowledge Management Systems. **Integrated Series In Information Systems**, [S.L.], p. 21-47, jan. 2007. Springer US. http://dx.doi.org/10.1007/978-0-387-37022-4_2.

Apêndice A – Modelo conceitual da exploração mineral



Fonte: Elaboração própria, 2021.

Organização do Conhecimento e os desafios tecnológicos da era *Big Data*

Knowledge Organization and the technological challenges of the Big Data era

Fabio Orsi Meschini (1), Marivalde Moacir Francelin (2)

(1) Universidade de São Paulo, Av. Prof. Lúcio Martins Rodrigues, 443 - Butantã, São Paulo - SP, CEP: 05508-020, fabioemeschini@gmail.com, (2) marivalde@usp.br

Resumo

Tendo em vista as complexidades propiciadas pelo constante desenvolvimento tecnológico e seus impactos no ciclo informacional, esta pesquisa objetiva investigar sob quais formas a Organização do Conhecimento compreende e desenvolve as temáticas sobre tecnologias relacionadas ao *Big data*, bem como analisar as perspectivas metateóricas desta produção científica. Para tanto, foram realizados levantamentos bibliográficos de livros e e-books, além da realização de coleta de artigos científicos nas bases LISA, Scopus, Web of Science e BRAPCI para a construção teórica da pesquisa. Na parte empírica, foram investigados os temas retratados nos artigos coletados do contexto tecnológico, bem como as influências teóricas (citações) e perspectivas metateóricas. Os resultados indicam que a Ciência da Informação e a Organização do Conhecimento sofrem notáveis impactos da tecnologia em seu fazer científico. No entanto, seu caráter social propicia pesquisas que abarcam tanto questões tecnológicas quanto sociais, oferecendo subsídios para as questões vivenciadas pelo contexto *Big data*. Conclui-se que os desenvolvimentos tecnológicos e semânticos na gestão deste grande volume de dados existentes na atualidade acarretam uma espécie de garantia democrática que deve ser conduzida pela Ciência da Informação e Organização do Conhecimento e suas interdisciplinaridades com a Ciência da Computação, a Matemática e a Estatística, sob diretrizes humanas e sociais.

Palavras-chave: Organização do Conhecimento. Ciência da Informação. *Big data*. Tecnologia.

Abstract

Bearing in mind the complexities provided by constant technological development and its impacts on the informational cycle, this research aims to investigate in which ways the Knowledge Organization understands and develops themes about technologies related to *Big Data*, as well as to analyze the metatheoretical perspectives of this scientific production. To this end, bibliographic surveys of books and e-books were carried out, in addition to the collection of scientific articles in the LISA, Scopus, Web of Science and BRAPCI databases for the theoretical construction of the research. In the empirical part, the themes portrayed in the above-mentioned articles from the technological context were investigated, as well as the theoretical influences (citations) and metatheoretical perspectives. The results indicate that Information Science and Knowledge Organization suffer notable impacts from technology in their scientific work, however, their social character provides research that encompasses both technological and social issues, offering subsidies for issues experienced by the *Big data* context. It is concluded that the technological and semantic developments in the management of this large volume of data existing today entail a kind of democratic guarantee that must be conducted by these areas and their interdisciplinarity with Computer Science, Mathematics and Statistics.

Keywords: Knowledge Organization. Information Science. *Big data*. Technology.

Resumen

Teniendo en cuenta las complejidades proporcionadas por el constante desarrollo tecnológico y sus impactos en el ciclo informacional, esta investigación tiene como objetivo investigar de qué manera la Organización del Conocimiento comprende y desarrolla temas sobre tecnologías relacionadas con *Big Data*, así como analizar las perspectivas metateóricas de este producción científica. Para ello, se realizaron levantamientos bibliográficos de libros y e-books, además de la recopilación de artículos científicos en las bases de datos LISA, Scopus, Web of Science y BRAPCI para la construcción teórica de la investigación. En la parte empírica, se investigaron los temas abordados en los artículos mencionados desde el contexto tecnológico, así como las influencias teóricas (citas) y las perspectivas metateóricas. Los resultados indican que las Ciencias de la Información y la Organización del Conocimiento sufren impactos notables de la tecnología en su quehacer científico, sin embargo, su carácter social propicia investigaciones que abarcan tanto temas tecnológicos como sociales, ofreciendo subsidios para los problemas que experimenta el contexto *Big data*. Se concluye que los desarrollos tecnológicos y semánticos en la gestión de este gran volumen de datos existentes en la actualidad suponen una especie de garantía democrática que deben llevar a cabo estas áreas y su interdisciplinariedad con la Informática, las Matemáticas y la Estadística.

Palabras clave: Organización del Conocimiento. Ciencias de la Información. Grandes datos. Tecnología.

1 Introdução

As pesquisas que possibilitam abordagens de questões impactantes na sociedade, como as observadas no contexto *Big data* e seus problemas de privacidade, acesso e utilização dos dados, destacam a relevância da Ciência da Informação. Observa-se ainda que os estudos epistemológicos e metodológicos da área representam um dos meios para a consecução de diretrizes organizativas que atuem no contexto *Big data* sob uma perspectiva social. Desde suas origens, a Ciência da Informação considerou os aparatos tecnológicos como soluções para lidar com a organização para a massa informacional que permanece em constante crescimento e transformação. Os impactos tecnológicos são evidentes no fazer científico da área, mas seu caráter social ainda precisa ser destacado nas questões tecnológicas. Nota-se que este perfil social da área tem potencial para subsidiar discussões acerca das questões vivenciadas pelo contexto *Big data*. Pois, como já indicaram Saracenic (1996) e Buckland (2012), um dos papéis da Ciência da Informação é amenizar as tensões existentes entre humanos, tecnologias e informação, priorizando a adaptação das tecnologias e recursos informacionais aos aspectos humanos.

É importante observar que o crescimento exponencial de dados disponíveis em meios virtuais está intrinsecamente relacionado aos constantes avanços tecnológicos, gerando um volume informacional de complexa apreensão e assimilação humana. A criticidade dos elementos trazidos pelo *Big data* fica a cargo dos campos mais voltados as áreas do Direito, Filosofia da Ciência e Ciências Sociais que, por sua vez, buscam identificar questionamentos éticos, sociais e econômicos. Estes dados trazem vestígios de teorias, não estando livres de preconceitos e decisões humanas, fato que influencia as formas sob as quais estes são interpretados. O impacto tecnológico no ciclo informacional propiciou a obtenção de uma diversidade de suportes informacionais que permitiram observar que as ligações entre a informação, os suportes e seus conteúdos informacionais estão cada vez mais amplos. Essa amplitude implica na necessidade de métodos organizativos da massa crescente e diversa de dados. Nesse caso, evidencia-se a necessidade de enfatizar áreas como a Ciência da Informação e a Organização do Conhecimento para lidar com tais complexidades organizacionais.

Quando se analisa o contexto *Big data*, é fundamental considerar a ambivalência existente nesta realidade de dados. Se, por um lado, nota-se a possibilidade de aprimoramento e expansão da representação e organização do conhecimento, por outro, são necessárias análises cuidadosas sobre esse contexto na medida em que as complexidades éticas também crescem exponencialmente. Destaca-se, portanto, que o contexto do *Big data* levanta discussões epistemológicas, metodológicas, éticas e tecnológicas relacionadas ao conheci-

mento. Dessa forma procura-se investigar sob quais formas a Organização do Conhecimento compreende e desenvolve as temáticas sobre tecnologias relacionadas ao *Big data*, bem como analisar as perspectivas meta-teóricas desta produção científica.

2 Organização do Conhecimento, *Big Data* e a era digital

As formas de lidar com a circulação de conhecimentos pela sociedade sempre foram desafiadoras. Desde a criação da imprensa em 1450 por Gutenberg até a validação enciclopédica do século XVIII, que viabilizou a disponibilização de conhecimentos empíricos e teóricos existentes na época, as definições sobre o que era compreendido como conhecimento por determinado momento histórico foram recorrentes. A evolução científica e tecnológica influenciou as maneiras pelas quais as pessoas acessam dados, informação e conhecimento (BURKE, 2003).

Vale ainda observar que esta ampliação de conhecimento pela sociedade estimulou a criação de um mercado editorial para lidar com a produção e distribuição deste conhecimento. Assim, a noção de informação como mercadoria começa a ser estruturada e as ofertas de diferentes suportes possibilita o acesso ao conhecimento sob variados formatos, abrangendo tanto leituras mais técnicas e fracionadas quanto mais intensivas e acadêmicas. Este início de diversidades na apresentação de conteúdos informacionais e de conhecimento já estabelece uma relação com a realidade a ser vivenciada posteriormente com a Internet. De acordo com Burke (2003), o vasto oferecimento de informações e conhecimento impõe desafios organizativos e de confiabilidade nos conteúdos elaborados, a interligação entre os diferentes intelectuais torna-se necessária para compreender mais extensamente os fenômenos, já que individualmente não é possível alcançar a totalidade das informações disponíveis.

Nas observações realizadas por Burke sobre a história social do conhecimento é possível encontrar, considerando, evidentemente, os respectivos períodos e espaços históricos, algumas similaridades com a realidade *Big data* no que diz respeito às dificuldades na intensa produção de dados e suas demandas organizacionais, envolvendo dinamicidades nos formatos e suportes informacionais, monopólios institucionais na gestão de dados e confiabilidade nos conteúdos disponíveis. Esses aspectos permanecem presentes e são objetos de discussões nas áreas ligadas a dados, informação e conhecimento. Sob um olhar holístico, salienta-se que as semelhanças entre o passado e o presente na circulação de dados, informação e conhecimento indicam oportunidades para que a Organização do Conhecimento aprimore seus recursos para contribuir com esta realidade massiva de dados.

Impactos ocasionados pelo *Big data* interferem no fazer científico e, portanto, também são sentidos pela Organização do Conhecimento que, por sua vez, trabalha com volumes informacionais cada vez mais crescentes em decorrência da tecnologia. Hjørland (2016, p.478) ressalta que a área deve ser compreendida “[...] como uma base de conhecimento que pode ser aplicada a todas as plataformas tecnológicas”. Considerar a Organização do Conhecimento em seu sentido mais amplo possibilita apresentar contribuições para diversas esferas da sociedade. Organizar conhecimento tem se tornado cada vez mais urgente, especialmente por meio de serviços mais dinâmicos e alinhados com a realidade tecnológica.

A amplitude na qual a Organização do Conhecimento desempenha suas atividades científicas pode levar a crer que todos os objetos e documentos teriam um potencial informacional, dificultando uma delimitação das fronteiras da área. No entanto, conforme Gnoli (2012, p. 269) observa, o foco deve permanecer no “[...] conteúdo do assunto, o que importa não é o objeto material, mas seu uso para transmitir conhecimento”. O autor salienta que os dados presentes no *Big data* têm possibilidades informacionais que merecem ser exploradas pela Organização do Conhecimento. A flexibilização estrutural da representação da informação e do conhecimento deve ser um objetivo para a área, já que possibilitaria diferentes combinações das dimensões (fenômeno, perspectiva e suporte) de assunto.

É possível observar que uma classificação dos recursos informacionais baseada em dimensões representadas por fenômenos, perspectivas e suportes, expressas pelos metadados destes recursos, pode ser um caminho viável para a realidade massiva de dados. Com isso, novas conexões de conhecimento seriam obtidas, atendendo expectativas do contexto *Big data*, como a própria classificação de dados.

As dimensões classificatórias que abrangem fenômeno, perspectiva e suporte seriam maneiras mais adequadas para uma recuperação informacional mais ampla, já que seria possível combinar estes elementos entre diversos recursos informacionais. Novas conexões estimulariam o alcance de mais conhecimentos que estão dispersos pela classificação disciplinar. O caminho interdisciplinar se apresentaria como uma alternativa fundamental para a área compreender o cenário tecnológico o que, por sua vez, contribuiria para o desenvolvimento de diretrizes sociais, éticas e organizacionais.

San Segundo Manuel e Martínez-Ávila (2014, p. 97) salientam que “[...] o pensamento digital molda nossa realidade e sua forma organizacional”, portanto, novas formas de estruturação e organização mostram-se necessárias na realidade massiva de dados. Os autores observam que os aparatos tecnológicos fazem parte do pensamento humano e as discussões sobre as influências da tecnologia são urgentes, uma vez que, a Orga-

nização do Conhecimento, em seu âmbito digital, relaciona-se, tanto com a materialidade, ou seja, a tecnologia em si, quanto com a cultura que está depositada amplamente nos meios digitais.

A Organização do Conhecimento, de acordo com Guimarães (2017), vem conquistando um protagonismo nas discussões interdisciplinares da Ciência da Informação por meio de subsídios teóricos e metodológicos advindos dos processos de mediação do conhecimento socialmente produzido. Logo, com base em Santos, Neves e Souza (2019), pode-se observar que a Organização do Conhecimento possui uma centralidade nas pesquisas realizadas em Ciência da Informação, uma vez que se relaciona diretamente com processos organizativos e representativos da informação e do conhecimento, tanto no aspecto técnico, prático e teórico. Pode-se afirmar que os estudos realizados por esta área são estratégicos, pois lidam com os desafios impostos sobre os diferentes processos e ciclos informacionais, “considerando a complexidade que envolve a relação entre a informação, os usuários e as condições de mediação da informação” (SANTOS; NEVES; SOUZA, 2019, p. 95).

Os avanços tecnológicos, segundo San Segundo Manuel e Martínez-Ávila (2014, p.98), impactam diretamente a Organização do Conhecimento na medida em que as “[...] conexões de redes digitais podem resultar em uma rede global de inteligências individuais interconectadas, capaz de gerar uma inteligência mais ampla”, quando comparada as individuais. Estabelecer recursos organizativos para a expansão representativa e semântica dos grandes volumes de dados impõe novos modelos, ferramentas e paradigmas para lidar com a realidade digital. As complexidades advindas dos intensos desenvolvimentos tecnológicos também perpassam pela emulação das relações sociais existentes (e suas desigualdades) na realidade e que são transferidas para os meios digitais, tornando o contexto tecnológico um reflexo da sociedade.

Nesse sentido, San Segundo (2013) destaca que as influências tecnológicas nos fluxos informacionais culminaram na existência de um paradigma tecnológico, no qual novas relações sociais são constituídas, acarretando formas inéditas de vivenciar o contexto informacional e tecnológico. Esse paradigma tem por base compilar grandes volumes de informação para o oferecimento de produtos baseados nas preferências dos usuários por meio de algoritmos, o que ocasiona uma produção em série destes produtos, buscando uma personalização que os tornem mais atrativos ao público almejado. Pode-se dizer que a representação do mundo caminha para a virtualidade e as fronteiras entre o real e o virtual estão cada vez mais difusas, fato que amplia os conflitos éticos informacionais, pois a virtualização centrada em poucas empresas pode interferir no exercício da cidadania.

As discussões tecnológicas realizadas atualmente acabaram por elevar a importância do paradigma social no contexto informacional pesquisado na Organização do Conhecimento e Ciência da Informação, trazendo a informação como um fenômeno que abrange as perspectivas social e humana. Esta importância social na organização, representação e circulação dos registros informacionais e de conhecimento acabam por “[...] determinar que a sua distribuição e a sua recuperação sejam consideradas como questões sociais, políticas e econômicas” (SANTOS; NEVES; SOUZA, 2019, p.98). Vale salientar que estes quesitos sociais, econômicos e políticos tornam-se ainda mais relevantes e complexos em uma realidade massiva de dados permeada por conflitos éticos relacionados ao uso e acesso de dados sobre diversos setores da sociedade.

Assim, verifica-se a existência, segundo Martínez-Ávila (2015, p.486), de “[...] uma desconexão entre a teoria e os fundamentos da organização do conhecimento e as aplicações mais tecnológicas de ponta”. Isso talvez tenha influência do cenário que separa estudos teóricos dos práticos, favorecendo o perfil tecnológico mais aplicado. A intensidade tecnológica, de acordo com o autor, tende a relegar para segundo lugar as experiências humanas em Organização do Conhecimento que, por sua vez, são fundamentais para um desempenho tecnológico e de pesquisa alinhados com as demandas sociais.

As teorias e as especialidades em Organização do Conhecimento têm sua importância nas relações com a Tecnologia, tendo em vista a superioridade das vivências e inteligências humanas no que se refere à Inteligência Artificial. Porém, os seguintes aprimoramentos, conforme aponta Martínez-Ávila (2015), precisam ser consolidados na relação da Organização do Conhecimento com a Tecnologia: atribuição de termos (ampliação da precisão na classificação e indexação automática); alinhamentos dos vocabulários (incompatibilidade e ausência de interoperabilidade entre os vocabulários dos domínios, fato que amplia a distância entre máquinas e humanos); avaliação e projetos de sistemas (problemas de comunicação com os desenvolvedores de sistemas para melhoria de seus usos); análise de domínio (predominância de viés individual nas pesquisas nas áreas de Ciências Cognitivas e Inteligência Artificial em detrimento de uma consideração sobre aspectos culturais, históricos e sociais); marcação social (diminuta qualidade na atribuição de metadados realizada por leigos ou especialistas no assunto marcado, mas sem conhecimentos sobre Organização do Conhecimento); e, aspectos sociológicos e compromissos éticos (compromissos com as questões éticas, culturais e sociais para reduzir problemas relacionados aos interesses comerciais).

Quando se analisa o contexto *Big data*, é fundamental considerar a ambivalência existente na realidade de

dados. Se, por um lado, nota-se a possibilidade de aprimoramento e expansão da organização do conhecimento, por outro, são necessárias análises cuidadosas sobre o contexto de dados, na medida em que as complexidades éticas também crescem exponencialmente. O *Big data* levanta discussões epistemológicas, metodológicas, éticas e tecnológicas relacionadas ao conhecimento. Este cenário abrangente e complexo tem implicações sobre privacidade, propriedade intelectual, política e sociedade. Além disso, são suscitadas questões direcionadas à qualidade, procedência, credibilidade e precisão na representação e organização dos dados (HAJIBAYOVA; SALABA, 2018). Em síntese, é importante defender uma visão cuidadosa e crítica dos impactos do *Big data* na Organização do Conhecimento, pois as complexidades e implicações do contexto de dados devem ser consideradas sob o aspecto humano. O vislumbre com as vantagens do *Big data* deve ser equalizado com a realização de pesquisas sobre suas implicações sociais, éticas e científicas, e que embasa a construção adequada de sistemas de organização do conhecimento que valorizem o desenvolvimento humano.

As inserções de processos automatizados na indexação e recuperação da informação dificultam a avaliação da qualidade dos itens recuperados em uma busca informacional porque não apresentam uma precisão adequada no assunto. As discussões em Organização do Conhecimento, de acordo com Hjørland (2000), devem ser amplamente voltadas para a otimização das formas de representação e recuperação da informação e conhecimento para os usuários. Os processos automáticos são aliados às interpretações e avaliações realizadas por humanos que, por sua vez, não podem ser substituídas. Dessa maneira, a Organização do Conhecimento precisa permanecer elaborando metodologias referentes à análise de domínio, inserindo o “[...] estudo dos domínios do conhecimento e das ‘comunidades discursivas’”, representados por coletividades sociais que, por sua vez, “desenvolvem conhecimentos comuns a partir de uma linguagem profissional comum, canais de comunicação comuns, bases de dados, etc” (HJØRLAND, 2000, p.68). Estas pesquisas sobre domínios, segundo Hjørland, aproximam a área de uma atuação mais direcionada a elementos sociológicos, acarretando reflexões relacionadas à quais objetivos, necessidades e interesses a Organização do Conhecimento deve servir. Deve-se elencar na área a presença de princípios mais voltados ao cenário comercial, como os observados em grandes empresas de tecnologia e redes sociais. No entanto, os princípios sociais, democráticos e culturais devem ter protagonismo nas questões informacionais.

Para Khan et al. (2017), o *Big data* exige uma capacidade de agregação e correlação de dados em uma diversidade de fontes que torna as análises complicadas na realidade massiva de dados. É importante a análise de dados para previsões mais eficazes, porém, coletar e

estocar dados não são os principais problemas. O que se coloca como a principal dificuldade no momento é a complexidade de relacionar e analisar os resultados dessas relações de dados. Por outro lado, respeitando-se a necessidade de segurança dos dados e a garantia da privacidade dos usuários, o *Big data* pode representar uma oportunidade de extrair informações relevantes para a sociedade.

Uma dessas garantias pode ser situada na elaboração de vocabulários. Dessa maneira, Barité (2014) destaca a necessidade de articulação dos processos relacionados aos vocabulários controlados com os novos recursos tecnológicos. Esses recursos podem facilitar o acesso aos dados e informações por meio de uma apresentação organizada e racional dos vocabulários controlados, dos itens informacionais e de acordo com as demandas de cada comunidade de usuários. Evidencia-se, também nesse contexto, a relevância da defesa de uma complementaridade entre a linguagem natural e o vocabulário controlado, objetivando um resultado mais eficaz para os usuários. Como os dados são geralmente sobre pessoas, a relação entre linguagem natural e controlada é interessante para o *Big data*, ressaltando que ontologias e folksonomias tem um aspecto social destacado em suas ações. É visível a necessidade de adaptabilidade dos vocabulários controlados na era digital. A complementaridade entre a linguagem natural e a controlada parece um caminho viável, bem como considerar os usuários e suas ações (aspecto social) no processo de recuperação e acesso informacional. A ontologia e a folksonomia são elementos importantes nesses processos.

O enfrentamento das instabilidades trazidas pelos contínuos fluxos do desenvolvimento tecnológico, segundo Szostak, Gnoli e López-Huertas (2016), deve ser concebido como uma oportunidade de inovação, adaptação, transformação e revolução dos processos executados pela Organização do Conhecimento. Conectar-se com as demandas dos avanços da tecnologia, alinhando-se com uma perspectiva interdisciplinar, amplia as possibilidades de estabelecer ligações entre os dados. Este cenário maximiza os potenciais valores informacionais existentes nos dados que estão em meio a um ambiente digital caótico, com sobrecarga de informações e evidentes necessidades organizacionais conectivas de conteúdos. Os sistemas de classificação atuais precisam ser mais adequados às demandas da contemporaneidade, com seus diversos bancos de dados digitais, e ao desenvolvimento necessário da Web Semântica. Portanto, essa proposta classificatória tem por finalidade “[...] organizar entendimentos acadêmicos e, assim, facilitar a comunicação e, esperançosamente, a síntese” (SZOSTAK; GNOLI; LÓPEZ-HUERTAS, 2016, p.210). Aprimoramentos nos processos comunicacionais e de sínteses são fundamentais para lidar com o contexto massivo de dados e suas relações com o *Big data*. Estas contribuições tornam-se ainda mais importantes na medida em que estimulam os processos co-

municacionais tanto em domínios específicos quanto entre domínios diferentes.

O ciclo envolvendo os desenvolvimentos científicos e tecnológicos para Morin (2013, p.63) possui interações que podem ser plenamente observadas e analisadas, dadas as suas ligações intrínsecas, uma vez que a ciência “[...] permite produzir a tecnologia e esta permite o desenvolvimento da ciência, que por sua vez, desenvolve a tecnologia”. É importante discutir as relações entre conhecimento, tecnologia e ciência na medida em que os aparatos tecnológicos têm o potencial de ampliar o campo do conhecimento que “[...] pode ser visto, percebido, observado e concebido” (MORIN, 2013, p.64). Morin ressalta que o excedente informacional causado pela tecnologia pode gerar desconhecimento. De seu ponto de vista, para se assimilar informações e construir conhecimentos é necessária a existência de uma estrutura teórica que aporte sentidos para estas informações, bem como a constituição de estruturas mentais suficientes para alicerçar a consecução de conhecimentos. O excesso de informação pode interferir nesses fluxos de assimilação informacional e construção de conhecimentos na medida em que compromete o estabelecimento das articulações cognitivas e das reflexões, restringindo, de certa forma, o conhecimento. O progresso do conhecimento deve estar alinhado a uma concepção de integração, na qual os conhecimentos especializados se comuniquem entre si, evitando o “[...] desmembramento do conhecimento, a destruição do conhecimento-sabedoria, ou seja, do conhecimento que alimenta nossa vida” (MORIN, 2013, p.99), contribuindo com o aspecto humano do aperfeiçoamento. As discussões sobre elementos tecnológicos e seus impactos na sociedade devem, para Morin (2013, p.110), considerar a existência da lógica das máquinas artificiais, cada vez mais presentes no cotidiano social, partindo do princípio que não são aplicados “[...] esquemas tecnológicos apenas ao trabalho manual ou mesmo à máquina artificial, mas também às nossas próprias concepções de sociedade, vida e homem”.

Assim, é importante refletir, conforme análises de Morin (2013), sobre os impactos tecnológicos na sociedade, sendo que a união dos problemas éticos com as possíveis contribuições da ciência pode estimular a concepção de uma vida social com um perfil mais humanista. Seguir os preceitos tecnológicos sem reflexões, buscando não compreender para quais caminhos e sob quais objetivos estes preceitos preconizam, pode ser uma decisão perigosa para a conquista de uma vida social mais justa. Para Morin, pensar sobre o saber exige uma série de reflexões baseadas em um considerável volume de informação, onde se torna necessário ter consciência do próprio patrimônio informacional, bem como objetivar a interpretação correta dos dados que constituem o ambiente e evitar as orientações advindas da racionalização e sua visão totalizante do mundo a partir de dados parciais. Esta busca por uma

unicidade de princípios racionalizantes, abre espaço para que os indivíduos sejam objetificados quantitativamente em prol de interesses econômicos e afins, sujeitos, portanto, a manipulações sociais.

3 Procedimentos metodológicos

A pesquisa é considerada bibliográfica, na medida em que foram realizados levantamentos bibliográficos de livros e e-books, além da realização de coleta de artigos científicos nas bases LISA, Scopus, Web of Science e BRAPCI por meio de buscas envolvendo os seguintes termos e expressões: “*Big data*”; algorit*; epistemolog*; “Information Science”; “knowledge organization”; “information organization”; “Ciência da Informação”; “organização do conhecimento” e “organização da informação”.

O universo da parte empírica, composta pelas análises temáticas e metateóricas, foi constituído a partir da coleta de artigos nas bases de dados Scopus, Web of Science, LISA e BRAPCI, por meio de buscas realizadas com os seguintes termos: (“*big data*” AND “knowledge organization” OR “organização do conhecimento”) e (“*big data*” AND “information organization” OR “organização da informação”). Destaca-se que as buscas foram realizadas sem delimitação de tempo nas bases de dados, objetivando-se uma recuperação temporal mais abrangente. Ressalta-se ainda que foram utilizados filtros restringindo os resultados a artigos da área de Ciência da Informação. Com o auxílio do Microsoft Excel e Word, foi analisada a existência de sobreposição entre os artigos das diferentes bases.

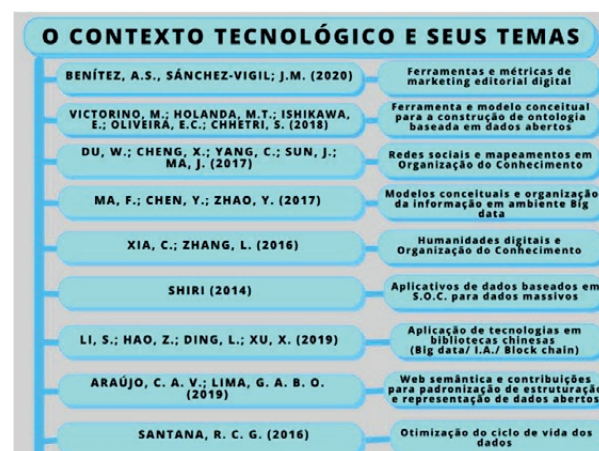
Após a etapa acima, elaborou-se a leitura dos textos completos dos artigos, almejando-se confirmar se eles realmente possuíam correspondência com a busca realizada, reduzindo, assim, possíveis inconsistências. O total de artigos recuperados foi de 47 e, após a verificação da sobreposição, sobraram 30. Ressalta-se que 4 artigos foram excluídos do universo de pesquisa por não apresentarem discussões diretamente relacionadas ao tema *big data*. Portanto, o universo de pesquisa foi composto por 26 artigos, no entanto, 9 artigos foram selecionados deste universo por abrangerem pesquisas que objetivam a consecução de recursos tecnológicos para lidar com o *Big data*, propondo estratégias que propiciem serviços digitais capazes de usufruir dos benefícios existentes neste contexto massivo de dados. Constituindo, assim, o “contexto tecnológico” desta pesquisa, no qual foram investigados os temas retratados neste referido contexto, bem como as influências teóricas (citações) e perspectivas metateóricas.

4 Contexto tecnológico e seus temas

A Figura 1 apresenta de forma sintética os principais assuntos pesquisados e as autorias presentes no contex-

to tecnológico da Organização do Conhecimento. As contribuições advindas das pesquisas integrantes deste contexto são apresentadas nos parágrafos seguintes.

Figura 1. O contexto tecnológico e seus temas.



Fonte: Autores.

As pesquisas alocadas no contexto tecnológico e representadas pelas seguintes autorias de Benítez e Sánchez-Vigil (2020), Victorino, Holanda, Ishikawa, Oliveira e Chhetri (2018), Du, Cheng, Yang, Sun e Ma (2017), Ma, Chen e Zhao (2017), Xia e Zhang (2016), Shiri (2014), Li, Hao, Ding e Xu (2019), Araújo e Lima (2019) e Santana (2016), concentram-se em propiciar recursos que aprimorem as possibilidades informacionais presentes no *Big data*. Perspectiva que pode acarretar a realização de novos estudos informacionais na Organização do Conhecimento. Embora estas pesquisas não aprofundem as discussões éticas e sociais, é interessante observar as considerações que elas apresentam sobre as influências do *Big data* nos comportamentos sociais, bem como a busca por melhores serviços aos cidadãos por meio da extração de informações relevantes para a sociedade. Preocupações sobre a recuperabilidade informacional e suas interligações com a descoberta e construção de novos conhecimentos por meio de aperfeiçoamentos semânticos na realidade de dados também se mostram presentes na condução destas pesquisas. Os desenvolvimentos tecnológicos e semânticos na gestão do grande volume de dados existentes na atualidade acarretam uma espécie de garantia democrática que deve ser conduzida pela Ciência da Informação e Organização do Conhecimento e suas interdisciplinaridades com a Ciência da Computação, Matemática e Estatística.

5 Influências teóricas e as características metateóricas do contexto tecnológico

As influências teóricas e as características metateóricas observadas no contexto tecnológico são explicitadas na figura 2. Quanto às menções de autorias nas referências, nota-se uma dispersão, pois foram encontradas 25

autorias diferentes com apenas duas menções, fato que indica a busca por embasamentos teóricos advindos de diferentes autores, dada a amplitude temática e as complexidades oriundas das pesquisas que abrangem aspectos tecnológicos e seus desafios éticos e sociais. As principais influências teóricas deste contexto são representadas por Berners-Lee (2001a; 2001b; 2006), Bizer (2007; 2011; 2012), Liu (2014; 2015a; 2015b), Salvador Benítez (2010; 2014; 2019), Sun (2013; 2015; 2016), Brasil (2011; 2014; 2016a; 2016b) e Zhdanova (2005a; 2005b; 2006; 2008).

Sobre as perspectivas metateóricas do contexto tecnológico, nota-se que o âmbito “Intelectual externo” é o que possui mais ocorrências, seis (6), abrangendo os seguintes autores: Benítez e Sánchez-Vigil (2020), Victorino, Holanda, Ishikawa, Oliveira e Chhetri (2018), Du, Cheng, Yang, Sun e Ma (2017), Xia e Zhang (2016), Li, Hao, Ding e Xu (2019) e Santana (2016). Observa-se também três (3) ocorrências dos âmbitos “Intelectual interno”, contemplando os seguintes pesquisadores: Ma, Chen e Zhao (2017), Shiri (2014) e Araújo e Lima (2019). O âmbito “Social externo” possui três (3) ocorrências representadas pelos pesquisadores: Du, Cheng, Yang, Sun e Ma (2017), Araújo e Lima (2019) e Santana (2016).

Figura 2. Principais influências teóricas e características metateóricas do contexto tecnológico.

Principais influências teóricas do contexto tecnológico (em menções nas referências)		Principais características metateóricas do contexto tecnológico (em nº de ocorrências)	
<p>BRASIL ZHANOVA, A.V. 4</p> <p>BERNERS-LEE, T. BIZER, C. LIU, W. SALVADOR BENÍTEZ, A. SUN, J. 3</p>		<p>Intelectual interno (paradigmas cognitivos; escolas de pensamento; mudanças paradigmáticas; escolas de pensamento; ferramentas metateóricas; teorias) 3</p>	<p>Intelectual externo (o uso de conceitos emprestados da: Filosofia, Economia; Linguística etc.) 6</p>
<p>BEYER, M. A. LANEY, D. CYGANIAK, R. BORGMAN, C. L. CONEGLIAN, C. S. CRAWFORD, K. DEMCHENKO, Y. MEMBREY, P. LAT, C. D. KLEEDORFER, F. MO, X.H. NOORDEN, R. 2</p> <p>SANTANA, R. C. G. SHIRI, A. SOUZA, R. R. ALMEIDA, M. B. SUN, Y. WANG, S. W. YAN, X. YOURDON, E. ZHANG, Y. PENG, J. HUANG, D. LI, F. 2</p>		<p>Social interno (paradigmas comuns; colégios invisíveis; escolas; redes; backgrounds; individuais) 0</p>	<p>Social externo (Impacto da sociedade; Instituições; Raízes históricas). 3</p>
Contexto Tecnológico			

Fonte: Autores.

Salienta-se que as características do contexto tecnológico possuem os embasamentos teóricos advindos de um núcleo constituído por diversos pesquisadores, dada a abrangência temática e complexidades existentes nas pesquisas envolvendo tecnologia. Os autores mencionados nas referências desse contexto, ao abordarem questões relacionadas à Web Semântica, Dados linkados, Web de dados, Interligações de banco de dados, Machine Learning, Política de dados nas redes sociais, Redes sociais científicas e personalização de conteúdo, Computação social, Dados abertos, Governança digital e Ontologia orientada para a comunidade, ratificam os esforços da Ciência da Informação e Organização do Conhecimento na integração de instrumentos e meto-

dologias tecnológicas para o tratamento adequado da massa de dados existente atualmente, incluindo-se, ainda que timidamente, discussões relacionadas a privacidade, democracia e cidadania.

Quanto às características metateóricas, observa-se a prevalência do “Intelectual externo”, fato que reforça a busca por elementos teóricos e demais recursos tecnológicos advindos de outras áreas para a expansão semântica dos dados e, com isso, viabilizar a ampliação de possibilidades informacionais aos usuários. Este âmbito metateórico abrange colaborações e interdisciplinaridades oriundas da Ciência da Computação, Humanidades digitais, Inteligência Artificial, Block chain e Matemática, sob uma orientação de interoperabilidade informacional entre os diferentes domínios científicos. Nota-se, ainda, a ocorrência do âmbito “Intelectual interno”, no qual se tem o objetivo de desenvolver os recursos já existentes na área, almejando a inovação e adaptação à realidade massiva de dados. Este âmbito traz questões relacionadas à Recuperação da Informação, Vocabulários controlados e Web Semântica, sob um norte de desenvolvimento. Já o âmbito “Social externo” aponta a existência da relevância social na construção das teorias e metodologias discutidas na Organização do Conhecimento. Os temas discutidos neste âmbito abarcam discussões sobre os impactos sociais da intensa utilização de dados no fazer científico, bem como o uso consciente de dados governamentais para a garantia democrática de direitos, respeitando-se a privacidade dos usuários, uma vez que as questões sociais e democráticas devem ser um dos nortes para a realização de pesquisas na Organização do Conhecimento sobre *Big data*.

6 Conclusões

A Organização do Conhecimento, e seus importantes métodos analíticos e representativos informacionais, indica a existência de experiências científicas para lidar com o contexto *Big data* e suas complexidades interpretativas, evitando a confusão presente no fenômeno de dados entre aquisição (facilitada pelas tecnologias) e análise e interpretação (dificultadas pela vastidão de dados) (BORGMAN, 2015). O campo tem possibilidades de agir como um mediador entre as questões relacionadas às facilidades de aquisição de dados e sua complexa interpretação, contribuindo para a expansão semântica do contexto massivo, principalmente no que tange a dados científicos.

A importância da análise dos dados em um contexto *Big data* reside nas possibilidades de conexões entre estes dados, facilitando a definição de quais parâmetros de busca são viáveis para a consecução de respostas a determinadas questões. Ressalta-se que essas conexões e cruzamentos permitem que as tomadas de decisões sejam mais assertivas já que mais fontes e visões dife-

rentes podem ser acessadas. No entanto, aferir o que pode ser útil nesta abundância de dados torna-se um desafio complexo.

Ao discutir a necessidade de inclusão da diversidade e da interdisciplinaridade nas questões informacionais atuais, o foco da Organização do Conhecimento deve ser o usuário, perfil humanizador também preconizado no contexto *Big data*. O impacto social das atividades da Organização do Conhecimento ressignificaram a centralidade do usuário para a relevância da área, já que sem usuário não existe Ciência da Informação e Organização do Conhecimento. A adaptabilidade na era digital e seus complexos elementos são meios para o pleno desenvolvimento das áreas. Nesse sentido, as novas dimensões da informação acarretaram em uma releitura de suas fronteiras e procedimentos organizativos em um fluxo informacional amplo e descontínuo.

A existência de pesquisas em Organização do Conhecimento abordando questões relacionadas, ainda que em fases iniciais, à condução de desenvolvimentos tecnológicos e semânticos na gestão de grandes volumes de dados acarreta uma espécie de garantia democrática que deve ser conduzida pela Ciência da Informação e Organização do Conhecimento no contexto *Big data*.

Referências

- ARAÚJO, C. A. V.; LIMA, G. A. B. de O. Estruturação e representação semântica de *big data* no contexto de base de dados governamentais. **Múltiplos Olhares em Ciência da Informação**, v. 9, n. 2, 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/137107> Acesso em: 20 ago. 2021.
- BARITÉ, M. El control de vocabulario en la era digital: revisión conceptual. **Scire**, v.20, n.1, p.99-108, 2014. Disponível em: <https://www.ibersid.eu/ojs/index.php/scire/article/view/4196/3766> Acesso em: 03 nov. 2021.
- BENÍTEZ A.S., SÁNCHEZ-VIGIL; J.M. Tools and metrics in editorial marketing: from *big data* to artificial intelligence. **Scire**, v.26, n.1, p.35-46, 2020. Disponível em: <https://www.ibersid.eu/ojs/index.php/scire/article/view/4682/4254> Acesso em: 20 ago. 2021.
- BERNERS-LEE, T. **Linked data-design issues**. 2006. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 10 abr. 2023.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, [s. l.], v. 284, n. 5, p. 29-37, 2001.
- BIZER, C.; BONCZ, P. A.; BRODIE, M. L.; ERLING, O. The meaningful use of *Big Data*: four perspectives-four challenges. **Sigmod Record**, [s. l.], v. 40, n. 4, p. 56-60, 2011.
- BIZER, C.; CYGANIAK, R. **D2R server-publishing relational databases on the semantic web**. 2012. Disponível em: d2rq.org/d2r-server. Acesso em: 12 abr. 2023.
- BIZER, C.; CYGANIAK, R.; GAUB, T. **The RDF book mashup: from web APIs to a web of data**. 2007. Disponível em: sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-248/paper4.pdf. Acesso em: 11 abr. 2023.
- BORGMAN, C. *Big data, little data, no data: scholarship in the networked world*. Cambridge: MIT Press, 2015. E-book. Disponível em: <https://doi.org/10.7551/mitpress/9963.001.0001>. Acesso em: 18 jun. 2020.
- BRASIL. **Decreto** nº 8.638, de 15 de janeiro de 2016 Diário Oficial da União, Brasília, DF, 18 jan. 2016a. Seção 1, p. 2. Disponível em: <http://www2.camara.leg.br/legin/fed/decret/2016/decreto-8638-15-janeiro-2016-782270-publicacaooriginal-149245-pe.html>. Acesso em: 23 nov. 2022.
- BRASIL. **Decreto** nº 8.777, de 11 de maio de 2016. Diário Oficial da União, Brasília, DF, 12 mai. 2016b. Seção 1, p. 21. Disponível em: <http://www2.camara.leg.br/legin/fed/decret/2016/decreto-8777-11-maio-2016-783082-norma-pe.html>. Acesso em: 23 out. 2022.
- BRASIL. **Lei** nº 12.527, de 18 de novembro de 2011. Diário Oficial da União, Brasília, DF, 18 nov. 2011. Seção 1, p. 1. Edição Extra. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.html. Acesso em: 14 jun. 2022.
- BRASIL. **Lei** nº 12.965, de 23 de abril de 2014. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Diário Oficial da União, Brasília, DF, 24 abr. 2014. Seção 1, p. 1. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm#art32. Acesso em: 14 jun. 2022.
- BURKE, P. **Uma história social do conhecimento**: de Gutenberg a Diderot. Rio de Janeiro: Zahar, 2003.
- BUCKLAND, M. What kind of Science can Information Science be? **Journal of the American Society for Information Science and Technology (ASIS & T)**, [s.l.], v.63, n.1, p.1-7, 2012. Disponível em: <https://onlinelibrary-wiley.ez67.periodicos.capes.gov.br/doi/epdf/10.1002/asi.21656>. Acesso em 24 out. 2020.
- DU, W.; CHENG, X.; YANG, C.; SUN, J.; MA, J. Establishing interoperability among knowledge organization systems for research management: a social network approach. **Scientometrics**, v.112, n.3, p.1489-1506, 2017. Disponível em: [https://link.springer-com.ez67.periodicos.capes.gov.br/content/pdf/10.1007/s11192-017-2457-0.pdf](https://link.springer.com.ez67.periodicos.capes.gov.br/content/pdf/10.1007/s11192-017-2457-0.pdf) Acesso em: 20 ago. 2021.
- GNOLI, C. Metadata about what? Distinguishing between ontic, epistemic, and documental dimensions in Knowledge Organization. **Knowledge Organization**, v.39, n.4, p.268-275. 2012. Disponível em: https://www.nomos-elibrary.de/10.5771/0943-7444-2012-4-268.pdf?download_full_pdf=1 Acesso em: 28 out. 2021.

- GUIMARÃES, J. A. C. et al. A dimensão teórica da Análise de Domínio na produção científica brasileira de Ciência da Informação. In: PINHO, F. A.; GUIMARÃES, J. A. C. (orgs.). **Memória, tecnologia e cultura na organização do conhecimento**. Recife: Ed. UFPE, 2017. p. 34-47. E-Book (Série: Estudos Avançados em Organização do Conhecimento, v. 4). Disponível em: <https://isko.org.br/wp-content/uploads/2021/05/Proceedings-ISKO-Brasil-2017.pdf> Acesso em: 06 ago. 2021.
- HAJIBAYOVA, L.; SALABA, A. Critical questions for *big data* approach in knowledge representation and organization. In: RIBEIRO, F.; CERVEIRA, M. E. (ed.). **Challenges and opportunities for Knowledge Organization in the Digital Age**. Baden-Baden: Ergon, 2018 p. 144-151. Disponível em: https://books.google.com.br/books?hl=pt-BR&lr=&id=MyuDDwAAQBAJ&oi=fnd&pg=PA144&ots=MUMbrgeHHL&sig=t4svqy_GIQixvov5XPYbrzt71A4#v=onepage&q&f=false. Acesso em: 23 abr. 2021.
- HJØRLAND, B. Knowledge organization. Should librarians organize all information on the internet? **Human IT**, v.4, n.4, p. 48-70, 2000. Disponível em: https://www.researchgate.net/publication/294469755_Knowledge_organization_-_Should_librarians_organize_all_information_on_the_internet Acesso em: 04 nov. 2021.
- HJØRLAND, B. Knowledge organization (KO). **Knowledge Organization**, v.43, n.6, p. 475-484, 2016. Disponível em: https://www.nomos-elibrary.de/10.5771/0943-7444-2016-6-475.pdf?download_full_pdf=1 Acesso em: 04 nov. 2021.
- KHAN, S. et al. A survey on scholarly data: from *big data* perspective. **Information Processing & Management**, v. 53, n. 4, p. 923-944, 2017. Disponível em: <https://www-sciencedirect.ez67.periodicos.capes.gov.br/science/article/pii/S0306457316305994?via%3Dihub> Acesso em: 27 set. 2021.
- LI, S.; HAO, Z.; DING, L.; XU, X. Research on the application of information technology of *Big Data* in Chinese digital library. **Library Management**, v.40, n.8-9, p.518-553, 2019. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/LM-04-2019-0021/full/pdf?title=research-on-the-application-of-information-technology-of-big-data-in-chinese-digital-library> Acesso em: 20 ago. 2021.
- LIU, W. **Digital humanities and linked data**. 2014. Disponível em: <http://sociedade.library.sh.cn/IT4L2014>. Acesso em: 10 abr. 2023.
- LIU, W.; ZHANG, C.; XIA, C. Controle normativo na era da world wide web. **Journal of library science in China**, [s. l.], v. 3, 2015.
- LIU, W.; ZHOU, D. M. Forward-looking of the next decade library technology trends. **Library Journal**, [s. l.], v. 34, n. 1, p. 4-12, 2015.
- MA, F.; CHEN, Y.; ZHAO, Y. Research on the organization of user needs information in the *big data* environment. **Electronic Library**, v.35, n.1, p.36-49, 2017. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/EL> Acesso em: 20 ago. 2021.
- MARTÍNEZ-ÁVILA, D. Knowledge Organization in the intersection with information technologies. **Knowledge Organization**, v.42, n.7, p.486-498, 2015. Disponível em: https://www.nomos-elibrary.de/10.5771/0943-7444-2015-7-486.pdf?download_full_pdf=1 Acesso em: 03 nov. 2021.
- MORIN, E. **Ciência com consciência**. 13. ed. Rio de Janeiro: Bertrand Brasil, 2013.
- SALVADOR BENÍTEZ, A. Entre la privacidad y el negocio: condiciones de uso y política de datos en redes sociales. **Derecom**, [s. l.], p. 139-161, 2019.
- SALVADOR BENÍTEZ, A.; GUTIÉRREZ DAVID, E. Redes sociales y medios de comunicación: desafíos legales. **El Profesional de La Información**, [s. l.], v. 19, n. 6, p. 667-674, 2010.
- SALVADOR BENÍTEZ, A.; OLIVERA ZALDUA, M.; SÁNCHEZ VIGIL, J. M. Redes sociales y medios de comunicación: políticas de uso de contenidos audiovisuales. In: SÁNCHEZ, J. S.; GARCÍA GARCÍA, F. **Tecnología y narrativa audiovisual**. Madrid: Fragua, 2014. p. 717-735.
- SAN SEGUNDO, R.; MARTÍNEZ-ÁVILA, D. Digital as a hegemonic medium for epistemology and knowledge organization. **Advances in Knowledge Organization**, v.14, p. 96-100, 2014. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/aiko_vol_14_2014_14.pdf Acesso em: 03 nov. 2021.
- SAN SEGUNDO, R. Panorama de investigación en Organización del conocimiento en su dimensión epistemológica. In: DODEBEL, Vera; GUIMARÃES, José Augusto Chaves (org.). **Complexidade e organização do conhecimento, desafios de nosso século**. Rio de Janeiro: ISKO-Brasil; Marília: FUNDEPE, 2013. p. 26-33. E-Book (Série: Estudos Avançados em Organização do Conhecimento, v. 2). Disponível em: <https://isko.org.br/wp-content/uploads/2021/05/Proceedings-ISKO-Brasil-2013.pdf> Acesso em: 28 jul. 2021.
- SANTANA, R. C. G. Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. **Informação & Informação**, v. 21, n. 2, p. 116-142, 2016. Disponível em: <https://brapci.inf.br/index.php/res/v/35252> Acesso em: 20 ago. 2021.
- SANTOS; R. F.; NEVES, D. A. de B.; SOUZA, E. D. de A. Organização do Conhecimento como domínio de estudo da Ciência da Informação: uma reflexão a partir dos aspectos epistemológicos. In: BARROS, T. H. B.; TOGNOLI, N. B. (orgs). **Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas**. Belém: ED. da UFPA, 2019. p. 95-104. E-Book (Série: Estudos Avançados em Organização do Conhecimento, v. 5). Disponível em: <https://isko.org.br/wp-content/uploads/2021/02/LIVRO-ISKO-BRASIL-EDICAO-BELEM.pdf> Acesso em: 28 jul. 2021.
- SARACEVIC, T. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan. /jun. 1996.
- SHIRI, A. Linked data meets *Big data*: A Knowledge Organization Systems Perspective. **Advances in Classification Research Online**, [s.l.], v.24, p.16-20, 2014. Disponível em: https://era.library.ualberta.ca/items/2edaa5e0-3034-448e-bb5a-1b931f84ff78/view/643d8d9b-6280-4e28-aa0e-6fedc2a23a9e/ACRO_24_1_16.pdf. Acesso em: 8 jun. 2020.

SUN, J.; MA, J.; LIU, Z.; MIAO, Y. Leveraging content and connections for scientific article recommendation in social computing contexts. **The Computer Journal**, [s. l], v. 57, n. 9, p. 1331-1342, 2013.

SUN, J.; XU, W.; MA, J.; SUN, J. Leverage RAF to find domain experts on research social network services: a *big data* analytics methodology with map reduce framework. **International Journal of Production Economics**, [s. l], v. 165, p. 185-193, 2015.

SUN, J.H.; WEI, J.; SUN, J.M. Research and practice on the construction of makerspace in Chinese university libraries. **Library Theory And Practice**, [s. l], v. 38, n. 7, p. 80-84, 2016.

SZOSTAK, R.; GNOLI, C.; LÓPEZ-HUERTAS, M. **Interdisciplinary Knowledge Organization**. Cham, Switzerland: Springer, 2016. 227 p. E-book.

VICTORINO, M.; HOLANDA, M.T.; ISHIKAWA, E.; OLIVEIRA, E.C.; CHHETRI, S. Transforming open data to linked open data using ontologies for information organization in *big data* environments of the brazilian government: the brazilian database government open linked data DBgoldbr. **Knowledge Organization**, v.45, n.6, p.443-466, 2018. Disponível em: <https://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=>

1&sid=875022b2-29c1-4708-baed-035b445f73d8%40pdc-v-sessmgr03 Acesso em: 27 set. 2021.

XIA, C.; ZHANG, L. The application of linked data in Shanghai library's service of genealogy digital humanities. **Library Journal**, v.35, n.10, p.26-34, 2016. Disponível em: <http://www.libraryjournal.com.cn/EN/Y2016/V35/I10/26> Acesso em: 20 ago. 2021.

ZHDANOVA, A. V. Towards a community-driven ontology matching. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE CAPTURE: ACM, 3., 2005, **Proceedings [...]**. 2005. p. 221-222.

ZHDANOVA, A. V. Community-driven ontology construction in social networking portals. **Web Intelligence And Agent Systems: An International Journal**, [s. l], v. 6, n. 1, p. 93-121, 2008.

ZHDANOVA, A. V.; KRUMMENACHER, R.; HENKE, J.; FENSEL, D. Community-driven ontology management: Deri case study. In: INTERNATIONAL CONFERENCE ON WEB INTELLIGENCE, 2005. **Proceedings [...]**. 2005. p. 73-79.

ZHDANOVA, A. V.; SHVAIKO, P. **Community-driven ontology matching**. Berlin: Springer, 2006.

Percepção bibliotecária sobre política de indexação: um estudo em bibliotecas especializadas

Librarian perception of indexing policy: a study in specialized libraries

Emilly Leticia Vieira de Souza (1), Lais Pereira de Oliveira (2)

(1) Universidade Estadual Paulista Júlio de Mesquita Filho, Avenida Hygino Muzzi Filho, 737, Marília, São Paulo, elv.souza@unesp.br. (2) Universidade Federal de Goiás, Avenida Esperança s/n, Câmpus Samambaia, Goiânia, Goiás, laispereira2@ufg.br

Resumo

Aborda a percepção do profissional bibliotecário sobre política de indexação. Objetiva identificar o entendimento bibliotecário sobre política de indexação na atuação dentro de bibliotecas especializadas, mediante análise de considerações dos profissionais a respeito do que é, dos elementos e aspectos que compõem o instrumento diretriz do processo de indexação, bem como das circunstâncias que envolvem a política de indexação e as unidades especializadas. Metodologicamente, constitui pesquisa descritiva de natureza qualitativa, desenvolvida a partir da aplicação de questionário a bibliotecários que atuam em bibliotecas especializadas. Os resultados indicam que os profissionais têm entendimento conceitual e concepção consciente e comum do aspecto geral de uma política de indexação, enquanto instrumento formal, normativo, agregador e essencial ao serviço bibliotecário na atividade de indexação em biblioteca especializada. Conclui-se que os profissionais identificam, definem, caracterizam e reconhecem a importância do uso de uma política de indexação no cotidiano das unidades de informação especializadas, porém não a têm formalizada e não a utilizam, o que possibilita a aplicação de estudos a respeito das motivações do não uso de uma política de indexação.

Palavras-chave: Tratamento temático da informação. Indexação de assunto. Atuação bibliotecária. Biblioteca especializada.

Abstract

It approaches the perception of the librarian about the indexing policy. It aims to identify the librarian's understanding about the indexing policy in the performance of specialized libraries, by analyzing the professionals' considerations about what it is, the elements and aspects that make up the guideline instrument of the indexing process, as well as the circumstances involving the indexing policy and the specialized units. Methodologically, it constitutes descriptive research of a qualitative nature, developed from the application of a questionnaire to librarians who working in specialized libraries. The results indicate that the professionals have a conceptual understanding and conscious and common conception of the general aspect of an indexing policy, as a formal instrument, normative, aggregating and essential to the librarian service in the indexing activity in a specialized library. It is concluded that professionals identify, define, characterize and recognize the importance of using an indexing policy in the routine of specialized units, but it is not formalized or used, which enables the application of studies on the motivations for not using an indexing policy.

Keywords: Subject representation. Indexing. Librarian performance. Specialized library.

Resumen

Aborda la percepción de los bibliotecarios sobre la política de indización. Tiene como objetivo identificar la percepción del bibliotecario sobre la política de indización en el desempeño dentro de las bibliotecas especializadas, mediante el análisis de las consideraciones de los profesionales sobre lo que es, los elementos y aspectos que componen el instrumento directriz del proceso de indización, así como las circunstancias que envuelven la política de indización y las unidades especializadas. Metodológicamente constituye una investigación descriptiva de carácter cualitativo, desarrollada a partir de la aplicación de un cuestionario a bibliotecarios que laboran en bibliotecas especializadas. Los resultados indican que los profesionales tienen una comprensión conceptual y una concepción consciente y común del aspecto general de una política de indización, como instrumento formal, normativo, agregador y esencial al servicio bibliotecario en la actividad de indización en una biblioteca especializada. Se concluye que los profesionales identifican, definen, caracterizan y reconocen la importancia de utilizar una política de indización en el cotidiano de las unidades especializadas, pero no la tienen formalizada y no la utilizan, permite la aplicación de estudios sobre las razones para no utilizar una política de indización.

Palabras clave: Tratamiento temático de la información. Indización. Trabajo de bibliotecario. Biblioteca especializada.

1 Introdução

No bojo da organização do conhecimento (OC) há processos, produtos e instrumentos específicos que viabilizam o tratamento e representação da informação. No que se refere a estes últimos, tendo a estrutura do conhecimento contida nos documentos como objeto, a OC se vale de instrumentos na forma de sistemas de classificação, tesouros, entre outras linguagens (GUIMARÃES *et al.*, 2015), para que esse conhecimento registrado possa ser organizado.

Linguagens documentárias representam uma ferramenta essencial para fins de descrição do conteúdo documental (GUIM; FUJITA, 2015), com aporte significativo na atribuição de termos a um documento. Todavia, a sistematicidade do processo de indexação de assunto também é substancialmente assistida pela política de indexação (PI). Entendida como documento formal, que contém diretrizes instrutivo-reguladoras da atribuição de assunto, a política de indexação direciona a ação bibliotecária nesse âmbito, refletindo um alinhamento político, normativo, estratégico e funcional em prol da evidência temática de um documento.

A política de indexação constitui-se a partir dos objetivos e características de uma dada unidade de informação (KOCHANI; BOCCATO; RUBI, 2012), alinhando-se à cultura organizacional e ao contexto informacional. Destarte, tendo em vista que a qualidade da indexação condiciona a recuperação de informações úteis (FUJITA, 2016), o conjunto normativo-orientador da política de indexação cumpre com função das mais relevantes no sistema de recuperação da informação, uma vez que proporciona uma ação indexadora mais assertiva e expressiva na determinação do conteúdo documental.

A práxis, porém, não é a única via observável nessa conjuntura, considerando que a investigação em PI se faz, também, extremamente necessária. Mesmo porque a prática bibliotecária é beneficiada com as pesquisas realizadas no domínio do tratamento temático da informação (GOMES; LIMA, 2021).

Ademais, assumindo posição decisiva na agregação de valor para disponibilização do conhecimento registrado, o tratamento temático da informação eleva, também, a política de indexação, tornando-a cada vez mais necessária (GUIMARÃES, 2004). Além do que há, ainda, dificuldades consideráveis na consolidação de políticas de indexação (RODRIGUES; FUJITA; DAL'EVEDOVE, 2017), sobretudo no que diz respeito à formalização desse instrumento em bibliotecas.

Sumariamente, ao considerar a biblioteca especializada como um ambiente destinado a organizar e armazenar materiais necessários ao seu público, em uma área específica de conhecimento e atuação, destaca-se a importância de organizar de modo orientado o acervo do

qual dispõe, para melhor atendimento das demandas informacionais institucionais. Cerne no qual a política de indexação se mostra fundamental.

Nesse sentido, a presente pesquisa busca identificar a percepção bibliotecária sobre PI na atuação dentro de bibliotecas especializadas. Pode, então, contribuir para o referencial teórico em organização do conhecimento, especificamente o que se dedica aos instrumentos utilizados no desenvolvimento dos processos de tratamento temático da informação, como a política de indexação.

2 Política de indexação

Política de indexação é um instrumento de tratamento temático da informação que delinea a indexação e, nesse sentido, auxilia o processo de mediação da informação entre unidades informacionais. Neste trabalho são representadas pelas bibliotecas, e por seus respectivos usuários, tendo como principal objetivo auxiliar e facilitar o processo de busca, recuperação e uso da informação.

Fujita (2012) define a política de indexação como um conjunto de decisões que esclareçam os interesses e objetivos de um sistema de informação e do sistema de recuperação da informação, não sendo esta apenas uma lista de procedimentos a serem seguidos, pois considera, também, os objetivos e a subjetividade das instituições e de seus usuários.

Conforme afirma Rubi (2012), todos os setores de uma biblioteca contêm políticas gerais e específicas que direcionam questões administrativas e tomadas de decisões. A PI faz parte desse conjunto de políticas, sendo uma aliada importante para que o bibliotecário realize um trabalho racional e objetivo, com base nas orientações por ela pontuadas.

Fundamentalmente, a indexação se constitui enquanto atividade de evidência temática, corporificada em termos que refletem o conteúdo documental. Sendo assim, a política atua na construção de diretrizes com base em critérios pré-estabelecidos que auxiliam nas tomadas de decisões, com vistas a propiciar uma indexação que represente o documento para que esse possa ser recuperado pelo usuário de modo assertivo.

Desse modo, ao estruturar uma política de indexação o bibliotecário indexador encontra suporte à sua atividade. Mais do que o amparo ao seu trabalho técnico-intelectual, a política viabiliza uma indexação mais eficiente, qualificando a recuperação da informação (GOMES; LIMA, 2021).

Termos que refletem o assunto documental são, pois, extraídos e referenciados de forma condensada durante a indexação (RODRIGUES; FUJITA; DAL'EVEDOVE, 2017). Razão pela qual diretrizes que orientem essa consecução são primordiais, como

forma de dirimir erros e inconsistências na representação de assunto.

Além disso, deve-se ressaltar que a atribuição de pontos de acesso temáticos é via para posterior localização do documento (GOMES; LIMA, 2021). Contudo, essa ação reveste-se de uma subjetividade considerável que é minimizada ao ser amparada por diretrizes base.

2.1 A importância da política de indexação

O acesso à informação nem sempre acontece com a facilidade e a rapidez necessárias em meio ao excesso de conteúdos cotidianamente produzidos e registrados (BENTES PINTO, *et al.*, 2015). Nesse sentido, o tratamento temático da informação reveste-se de caráter estratégico, sobretudo na ação de indexação de assunto.

Segundo Fujita (2012), a política de indexação é norteadora de princípios e critérios que guiarão a tomada de decisões para facilitação, otimização e racionalização dos processos, tornando-a indispensável por conta da organicidade, objetividade e direcionamento por ela proporcionados, colaborando com um andamento da ação de indexação pautada em diretrizes e caminhos norteados, com vistas a garantir seu aprimoramento e a obtenção de resultados eficazes conforme as necessidades dos usuários.

A PI pode ser explicada a partir de dois eixos: o eixo horizontal, envolvendo questões administrativas, recursos humanos, financeiros, clientela; e o eixo vertical, abrangendo as questões processuais da atividade de organização e representação da informação (FUJITA, 2012). Esses elementos estão diretamente relacionados com as tomadas de decisões impactantes na recuperação da informação, seleção, aquisição, cobertura, entrada, saída, indexação, avaliação. Pois, consequentemente, todos influenciam nas decisões e nos resultados do processo de indexação.

De acordo com Gil Leiva (2012), a indexação envolve processos cognitivos, dos quais, por meio desses é realizada a organização da informação, utilizando o discurso textual como forma de comunicação mediante livro, artigo, tese, legislação, ou seja, uso da comunicação textual. Ao fazer relação entre organização da informação e o discurso textual é perceptível que há um estreito relacionamento entre a comunicação e a indexação, uma vez que a estruturação e a compreensão textual, juntamente com os seus critérios comunicacionais são elementos que abrangem o antes, durante e pós ato de indexar.

Portanto, ao tratar da importância da PI é de suma relevância considerar que a indexação envolve processos cognitivos de organização da informação e da comunicação para que seja concretizado (GIL LEIVA, 2012). Diante disso, a qualidade da indexação advém da análise e tradução de documentos com base em aspectos lógicos e linguísticos, que envolvem a interpretação e

representação do conteúdo informacional, para haver comunicação entre tal conteúdo e o usuário.

Logo, é o indexador que atribui sentido à interlocução, ou seja, faz a mediação entre o leitor e o alcance da resposta para sua necessidade, representados pela busca a informação, e o sistema, com suas possibilidades de recuperação. Desse modo, a política de indexação tem grande importância, por envolver orientações para o processo de tomada de decisões impactantes na recuperação da informação.

2.2 A política de indexação nas unidades de informação: biblioteca especializada

O processo de indexar envolve uma ampla gama de tipos de materiais e suportes. Conforme Gil Leiva (2012) qualquer objeto pode ser indexado, facilitando seu armazenamento e recuperação em bases de dados, independentemente de suas características. Ainda segundo o autor, o tamanho do objeto indexado não tem relação direta com o número de palavras-chave, assuntos ou descritores empregados em sua representação, ou seja, as características da indexação ficam a critério do profissional e das necessidades da unidade, porém esses devem ser normatizados e justificados para que haja uma padronização da ação.

Ademais, há que se atentar que permeiam a indexação: parâmetros de análise de assunto, a relação do conteúdo informativo e o contexto da unidade de informação, além de diretrizes registradas na política de indexação da instituição (SANTOS *et al.*, 2019). Conjuntura esta, desafiadora e dinâmica. Tendo esse fato como base, a biblioteca especializada tende a atender tanto as suas características essenciais ligadas à área com que atende, quanto aos seus usuários.

Como definição desse tipo de unidade, tem-se que “a biblioteca especializada é uma unidade que se dedica a organização e disseminação de informações sobre um assunto ou grupo de assuntos em particular” (CARIBÉ, 2017, p. 193). Ainda, considera-se que são ambientes existentes com o objetivo de auxiliar no realizar das atividades profissionais da área que abrangem, exigindo um ambiente informacional para a atribuição de conhecimento e atualizações do ramo.

Assim, são unidades que precisam do estabelecimento de uma política de indexação para o funcionamento adequado do local, possibilitando ao seu público a facilitação da busca e obtenção de resultados, tendo em vista a especificidade de seus acervos e das solicitações dos usuários atendidos.

2.3 Diretrizes para a elaboração de uma política de indexação

No que se refere à construção da política de indexação, há que se observar o tipo de unidade de informação, bem como o público-alvo e o acervo disponível, para

determinação dos níveis de exaustividade e especificidade da indexação, entre outros aspectos. Desse modo:

A política de indexação pode ser definida por diversos fatores em uma unidade de informação, desde a seleção de tipos de documentos a serem indexados, procedimentos de análise e representação de assuntos, aspectos qualitativos da indexação como precisão, especificidade, exaustividade e revocação, instrumentos de controle de vocabulário tais como linguagem documentária ou opção por trabalhar com linguagem natural, uso de programas de indexação automática, além da avaliação da indexação pela consistência e pela recuperação (FUJITA, 2016, p. 4-5).

Na mesma direção, Kochani, Boccato e Rubi (2012, p. 264) destacam que “a política de indexação deve ser desenvolvida a partir dos objetivos e especificidades de características de cada organização”. Isso ocorre para que as unidades tenham suas demandas atendidas de forma direcionada e eficaz, ou seja, precisam obter respostas assertivas às buscas realizadas e auxiliar no desenvolvimento das atividades fim dos usuários de uma biblioteca especializada.

Além do desenvolvimento estrutural, há que se atentar para as percepções existentes em torno da política de indexação, por parte dos agentes que atuam tanto em sua concepção quanto em sua aplicação: os bibliotecários. Portanto, a seção seguinte descreve os aspectos metodológicos utilizados na busca por compreender a percepção e prática dos bibliotecários em unidades especializadas, com a subsequente exposição dos resultados obtidos com a pesquisa.

3 Metodologia

A pesquisa é descritiva e de abordagem qualitativa. Volta-se à investigação de questões práticas da temática diretamente aos bibliotecários pesquisados, aos quais os comportamentos são postos em observação, com dados coletados via questionário, sendo este instrumento constituído por perguntas abertas e fechadas a respeito do tema proposto (GIL, 2008).

A pesquisa teve como universo bibliotecas especializadas localizadas em Goiânia – Goiás, a partir de levantamento das unidades de informação dessa tipologia, com atuação em áreas distintas. A amostra estabeleceu-se por acessibilidade e conveniência (GIL, 2008). Nessa perspectiva, foram convidadas 8 bibliotecas, das quais metade responderam ao formulário, que ficou disponível por 4 meses, período estabelecido para a coleta dos dados.

O corpus envolveu, portanto, quatro bibliotecas especializadas identificadas e caracterizadas conforme o Quadro 1.

Mediante as bibliotecas especializadas (Quadro 1), apresenta-se, também, a elaboração e organização do questionário, formulado a partir de fundamentação teó-

rica na literatura da área, para que fossem obtidos dados em relação a conceituação e características observadas pelos bibliotecários de unidades especializadas em relação a política de indexação, com base no que se identificou como aspectos essenciais de observação e discussão. Para isso, distribuiu-se em três seções principais, nas quais foram destrinchadas mediante questionamentos que esclarecessem cada tópico principal.

Quadro 1. Caracterização das bibliotecas investigadas.

Unidades	Representação	Área de especialização
Unidade 1	U1	Patrimônio cultural brasileiro material e imaterial, arquitetura, engenharia, arqueologia, história e afins
Unidade 2	U2	Estatística e geociências
Unidade 3	U3	Direito
Unidade 4	U4	Geociências

Fonte: Dados da pesquisa (2023).

Estruturalmente, o formulário continha inicialmente perguntas voltadas a caracterização do perfil profissional e da unidade em relação a atuação. Em seguida, para observar o contato dos respondentes com a PI, inseriram-se questões a respeito do contato com o instrumento, a existência ou não dele na unidade e os fatores e elementos presentes no mesmo, somados à questão de qual seria a importância do estabelecimento de um material desse tipo para a rotina local.

Por fim, a última subdivisão é de grande ênfase na discussão do tema aqui proposto, pois se volta para a compreensão da percepção dos respondentes sobre política de indexação, momento em que foram postas indagações a respeito do entendimento, utilidade, questões de elaboração, elementos envolvidos pelo processo, vantagens de implementação, importância para tomada de decisão e adaptação, como implementar uma PI e para finalizar foram questionados sobre a relevância e utilização de um guia ou manual de indexação.

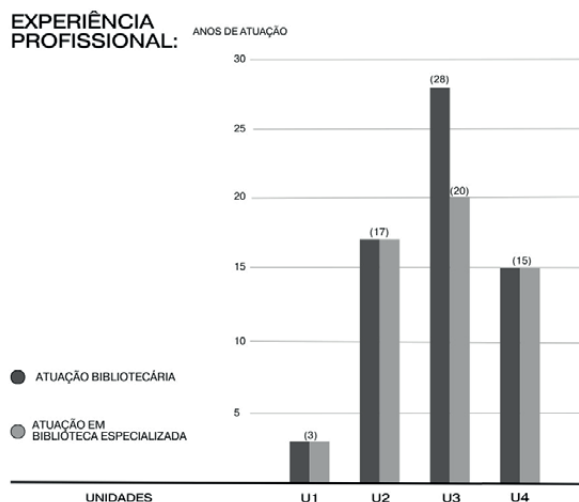
Explicitadas as abordagens presentes no questionário, coloca-se na seção seguinte a análise e interpretação das informações prestadas pelas unidades.

4 Análise e interpretação dos dados

A coleta de dados via questionário foi organizada em frentes de discussão em torno da temática política de indexação. Portanto, mediante a organização e disposição das questões foi possível obter informações sobre caracterização; contato; percepção; aplicabilidade e usabilidade de uma política de indexação.

Com base nos dados de caracterização coletados, demonstra-se na Figura 1 a apresentação do tempo de atuação bibliotecária dos representantes das bibliotecas especializadas pesquisadas.

Figura 1. *Atuação bibliotecária*



Fonte: *Dados da pesquisa (2023)*.

Nota-se, pela Figura 1, que três das quatro bibliotecas especializadas investigadas (U2, U3 e U4) contam com bibliotecários com considerável tempo de atuação profissional, acima dos 15 anos. Além do que possuem vasta experiência em unidade de informação especializada.

Tendo isto como ponto de partida, em relação ao contato com a temática “política de indexação” verificou-se pelas respostas que é unânime na atuação profissional como bibliotecário; em leitura da literatura da área de Biblioteconomia e na preparação para concurso público. Também, é importante afirmar que apenas dois dos respondentes afirmaram ter tido contato em capacitações/eventos/formação continuada na área de Biblioteconomia e em estágio obrigatório ou não-obrigatório feito durante a graduação.

Esse contato e proximidade com o tema “política de indexação”, seja durante a formação ou atuação, são importantes para que o profissional adquira familiaridade e conhecimento das vantagens que uma PI bem elaborada e estabelecida atribui ao contexto especializado, principalmente em meio à característica pontuada por Caribé (2017) sobre cada biblioteca especializada ter unicidade ao se dedicar a organização e disseminação de assuntos específicos. Ademais, o contato com o assunto representa uma forma de aproximar o bibliotecário desse conjunto de decisões que é a política de indexação (FUJITA, 2012). Até porque a PI têm sido progressivamente mais necessárias do ponto de vista do tratamento temático da informação (GUIMARÃES, 2004), uma vez que ampara a organização do conheci-

mento registrado e a alocação desse nos ambientes informacionais.

Após a apresentação contextual dos bibliotecários à frente das bibliotecas especializadas, destacam-se informações sobre qual o vínculo de cada instituição com a política de indexação. Sobre esse aspecto, observou-se que nenhuma das unidades investigadas conta com uma PI formalmente estabelecida. Reforça-se aqui, como expõe a literatura, que essa é mais uma entre as políticas gerais e específicas em bibliotecas (RUBI, 2012) sem existência formal e, portanto, dificilmente consolidada (RODRIGUES; FUJITA; DAL'EVEDOVE, 2017).

Por conseguinte, indagadas a respeito de quais os fatores presentes na política de indexação estabelecida ou no instrumento de orientação que norteia a prática indexadora, as bibliotecas especializadas demonstram certa sistematização nessa atividade. A U2 indica haver uma orientação para indexação, para padronizar o processo realizado por parte dos bibliotecários até que a política de indexação seja elaborada. Isso porque há poucos anos biblioteca passou por uma mudança do sistema de gestão de acervo e no momento trabalha para a implementação de um repositório. A U3 diz que há orientações técnicas para a rede de bibliotecas da instituição.

O detalhamento de apenas duas das quatro bibliotecas especializadas (U2 e U3) em relação ao que utilizam para nortear o processo de indexação instiga saber a forma com que realizam a atividade, já que não possuem diretrizes formais do processo; e quanto às que mencionam a presença de orientações é importante observar duas questões, uma que não consideram o material orientador existente como uma política de indexação e outra a forma pela qual são direcionadas pelo documento mencionado a partir das decisões e considerações ao longo do processo. Posto isto, traz-se destaque para a importância de a política de indexação ser desenvolvida com base nos anseios e especificidades locais (KOCHANI; BOCCATO; RUBI, 2012), ou seja, não é suficiente existir orientações, essas devem ser elaboradas, formalizadas e utilizadas conscientemente com vistas a auxiliar e aprimorar os serviços prestados à instituição e o atendimento ao próprio público.

Nessa perspectiva, as bibliotecas especializadas investigadas destacaram, respectivamente, a adoção de alguns elementos em detrimento de outros considerados importantes, mas, dos quais não necessariamente fazem uso, no processo de indexação dos materiais da unidade, expostos no Quadro 2.

Como se observa no Quadro 2, nível de especificidade, capacidade de revocação e capacidade de precisão são elementos do processo de indexação considerados importantes por todas as bibliotecas especializadas

investigadas. Todavia, são colocados em prática por apenas metade das unidades, no caso dos dois primeiros e, por uma, no caso do último elemento.

Quadro 2. Elementos do processo de indexação em bibliotecas especializadas

Elementos	Adotados				Considerados importantes			
	U1	U2	U3	U4	U1	U2	U3	U4
Nível de exaustividade		X				X	X	X
Nível de especificidade	X	X			X	X	X	X
Capacidade de revocação	X	X			X	X	X	X
Capacidade de precisão		X			X	X	X	X
Linguagem controlada		X	X	X		X	X	X
Linguagem natural	X		X	X	X			X

Fonte: Dados da pesquisa (2023).

Outra questão a ser destacada é o quanto nível de exaustividade é expressamente visto em sua relevância (75% das unidades), embora somente uma biblioteca procure adotá-lo na prática de indexação, o que pode estar relacionado a dificuldades de execução frente a outros procedimentos, ou mesmo, de compreensão das particularidades que envolvem esse elemento do processo de indexação. Quase que na mesma proporção, linguagem natural é adotada por 75% das bibliotecas, mas, considerada relevante por apenas duas.

No quadro geral, tem-se que a U1 considera precisão importante, porém não adota; a U2 utiliza e traz como relevante todos os elementos citados menos a linguagem natural; a U3 aplica somente linguagem controlada e natural, mesmo definindo todos como importantes, com excessão da linguagem natural, apesar de utilizá-la; a U4 também aplica somente linguagem controlada e natural, porém considera todos os elementos relevantes, sem exceções.

Diante dessas colocações, vale evidenciar a significância positiva de todas as unidades reconhecerem o dever de utilizar praticamente todos os elementos, uma vez que sua combinação permite uma personalização da representação de assunto via indexação que pode impactar na recuperação dos materiais do acervo. Percebe-se, então, que há discernimento a respeito dos elementos expostos (Quadro 2) estruturados em uma PI, por mais que não haja o estabelecimento e uso integral. O que evidencia

a essencialidade do estabelecimento de um documento norteador de princípios e critérios facilitadores das tomadas de decisões (FUJITA, 2012), que represente a abordagem do material (RODRIGUES; FUJITA; DAL'EVEDOVE, 2017) e a facilite localização dele (GOMES; LIMA, 2021) em concordância com as devidas prioridades da unidade.

Em relação à linguagem natural, elemento não sinalizado pela U2 e U3, mas utilizado pelas duas, reflete que a falta de diretrizes aprimoradoras do uso desse tipo de abordagem influencia no reconhecimento de sua devida utilidade e, conseqüentemente, da importância que carrega. Assim, observa-se que é utilizado por elas, pois no seu princípio prático tende a refletir a linguagem comumente empregada na unidade e por seus usuários. Além do que não exige esforços relativos à adoção ou elaboração de linguagem documentária na forma de tesouro para a indexação e recuperação por assunto da informação via sistema.

Deve-se ressaltar que, no espaço destinado a menção de outros elementos foi mencionado pela U3 a adoção de orientações técnicas e a experiência do indexador como elemento considerado importante. Pontuações relevantes, ao considerar o fato de as duas influenciarem na qualidade da indexação, porque tanto orientações quanto as experiências do profissional impactam nas tomadas de decisões e na consciência das escolhas e dos resultados a serem obtidos. Sob esse aspecto, tem-se que na política de indexação o profissional encontra amparo para a realização das atividades e para a indexação eficiente (GOMES; LIMA, 2021), e define as características do processo com base no que a unidade necessita (GIL LEIVA, 2012; KOCHANI; BOCCATO; RUBI, 2012).

A respeito dos aspectos da indexação adotados e considerados importantes na realização da indexação, apresenta-se o Quadro 3.

Como se nota no Quadro 3, cobertura de assunto, seleção e aquisição de documentos e processo de indexação são os aspectos do processo de indexação considerados importantes por todas as bibliotecas especializadas investigadas. Da mesma forma, são colocados em prática pela maioria das unidades.

Destacadamente, tempo de resposta do sistema é o aspecto que menos observado pelas bibliotecas especializadas, embora visto em sua relevância pela maior parte delas. O mesmo se vê em relação a avaliação do sistema, que não é adotado por nenhuma unidade. Esse tipo de constatação demonstra a necessidade de maior associação da prática indexadora com a recuperação da informação, da mesma forma que a percepção ampliada sobre o sistema de informação.

No quadro geral, a U1 utiliza apenas seleção e aquisição de documentos e estratégia de busca, e considera

importante, além desses, somente cobertura de assunto e o processo de indexação. Já a U2 adota todos com exceção da avaliação do sistema. A U3 não usa tempo de resposta e avaliação do sistema. A U4 faz uso exclusivamente da cobertura de assunto, desconsiderando todos os outros aspectos, apesar de mencionar a importância de todos, assim como U2 e U3. A U1 traz como importante apenas a cobertura de assunto, a seleção e aquisição de documentos e o processo de indexação.

Quadro 3. Aspectos da indexação em bibliotecas especializadas

Aspectos	Adotados				Considerados importantes			
	U1	U2	U3	U4	U1	U2	U3	U4
Cobertura de assunto		X	X	X	X	X	X	X
Seleção e aquisição de documentos	X	X	X		X	X	X	X
Processo de indexação		X	X		X	X	X	X
Estratégia de busca	X	X	X			X	X	X
Forma de saída		X	X			X	X	X
Tempo de resposta do sistema		X				X	X	X
Avaliação do sistema						X	X	X

Fonte: Dados da pesquisa (2023).

Como em relação aos elementos, portanto, também os aspectos do processo de indexação são vistos em sua importância. Ainda assim, a práxis precisa se ampliar em boa parte deles, tendo em consideração a qualidade da indexação para a recuperação de informações úteis (FUJITA, 2016).

Sobre os questionamentos do nível de importância, o grau de influência e o que as unidades compreendem como conceito de PI dispõem-se as respostas no Quadro 4.

Como destacado no Quadro 4, as quatro unidades reconhecem ser importante (U1) e muito importante (U2, U3, U4) a implementação de um PI, seja em biblioteca especializada no geral ou na própria unidade em que atuam. Como, também, destacam influência relativa (U1) e extrema (U2, U3, U4) da política de indexação nas decisões. Além disso, todas atribuem ao conceito de indexação o caráter de documento normativo e instrutivo do processo de indexar, com reflexo nas tomadas de decisões da biblioteca. Assim, possuem concep-

ção consciente e comum do aspecto geral de uma PI, tanto quanto um entendimento conceitual semelhante, atrelado ao caráter de instrumento formal para indexar.

Quadro 4. Percepções sobre política de indexação

Unidade	Nível de importância de implementação: Especializada em geral e Unidade respondente	Influência nas decisões	Concepção de política de indexação
U1	Importante	Relativamente	“Instrumento com diretrizes para tomadas de decisões.”
U2	Muito importante	Extremamente	“Determina as diretrizes do processo de indexação realizado na unidade, que impacta diversos outros processos e serviços da biblioteca.”
U3	Muito importante	Extremamente	“Diretrizes básicas para melhorar o desempenho da atividade de tratamento da informação que reflita na recuperação do documento.”
U4	Muito importante	Extremamente	“Diretrizes de indexação estabelecidas em um documento formal adotado pela instituição.”

Fonte: Dados da pesquisa (2023).

Acerca da visão de para que serve uma política de indexação no contexto especializado em que atuam, coletou-se: U1 “no contexto de rede de bibliotecas que pertencemos, uma política de indexação traria orientação e uniformidade aos trabalhos nacionalmente oferecidos”; U2 “Serve para determinar diante do contexto de cada unidade de informação o caminho que será seguido

pelas bibliotecárias e bibliotecários para realizar o processo de indexação. Tendo em vista a recuperação do objeto indexado pelo usuário”; U3 “Servir de guia para a tomada de decisões no momento de tratar a informação especializada”; e U4 “Para nortear o processo de indexação”. Além das observações sobre a serventia da política de indexação, todas demonstraram que a utilizariam, fato que traz questionamentos do porquê ainda não instituíram o documento já que possuem visão positiva a respeito desta, de seus elementos e aspectos.

Consonante a isso, a(s) principal(is) vantagem(ens) de construir e implementar uma política de indexação para as instituições, mencionadas nas respostas, estão ligadas a consistência das operações (U1,U2); otimização do serviço (U3, U4), sendo mencionadas também garantia da recuperação do item indexado pelo usuário (U2); racionalização dos processos, consistência das operações e segurança na tomada de decisões (U3); e otimização do serviço, racionalização dos processos e consistência das operações (U4). Essas atribuições positivas são advindas da organização e disponibilização de relações lógicas e consistentes do processo de racionalização das diretrizes de indexação conforme as necessidades e objetivos da unidade ao qual é direcionada. Processos esses discutidos e denominados por Gil Leiva (2012) como cognitivos, nos quais são meios de realização da organização da informação.

Ao serem indagados sobre a importância da política de indexação para um novo profissional integrado à instituição, as respostas são demonstradas no Quadro 5.

Conforme o Quadro 5, reconhecida a utilidade da política para orientação e auxílio das ações de indexação tomadas na unidade, pontua-se as colocações dos profissionais das unidades em relação a importância que uma PI teria para um novo bibliotecário ou estagiário ao adentrar na instituição. Dessa forma, a U1 diz ser importante e a U2, U3 e U4 trazem como muito importante, resposta condizente com a indicação de quais ações devem ser tomadas por um recém-contratado, em que todas citam a precisão de conhecer os parâmetros, processos e instrumentos de trabalho. Reforça-se, pois, que a constituição da política de indexação atrelada às características da unidade, como apreçoam Kochani, Boccato e Rubi (2012), representa um meio para que esse contexto seja devidamente assimilado pelos novos profissionais.

Outro ponto de discussão colocado foi a respeito da forma com que os bibliotecários utilizariam uma PI no dia a dia. Nesse contexto, a U1 diz que “tendo em vista a rotina de trabalho, acredito que especialmente no momento de análise do material antes de realizar a catalogação”; a U2 “sempre que necessário, para realizar as atividades relacionadas ao processo de indexação, da mesma forma que utilizo os demais manuais e políticas da Biblioteca.”; a U3 “consulta frequente como instrumento de trabalho. Também é norteador na percepção

de melhorias”; e a U4 “para consultas na rotina de trabalho”. Logo, ficam evidentes as pontuações em relação a frequência e a utilização do material na rotina do local, aspecto característico do objetivo primordial da política de indexação que é orientar os processos e as decisões que o envolve na prática profissional, em especial na crescente demanda informacional que dificultam o a facilidade de acesso às informações (BENTES PINTO, *et al.*, 2015).

Quadro 5. *Integralização de profissional*

<i>Unidade</i>	<i>Nível de importância para adaptação e conhecimento de decisões</i>	<i>Principais ações a serem tomadas por um recém-contratado</i>
U1	Importante	“Seguir os parâmetros já estabelecidos pela casa. E havendo necessidade de inovação, sugerir melhorias aos processos.”
U2	Muito importante	“Conhecer a política e os processos que ela envolve na teoria e na prática.”
U3	Muito importante	“Primeiramente conhecer a missão, a visão e os valores da Instituição; Estudar/Ler as normas internas e os instrumentos de trabalho; Conhecer as rotinas e as necessidades informacionais dos usuários; Estabelecer contato com outras instituições/profissionais que trabalham com a mesma especialidade; Se possível, solicitar treinamento institucional para melhor desempenho das atividades.”
U4	Muito importante	“Ter conhecimento da política. Receber treinamento adequado para uso da política.”

Fonte: *Dados da pesquisa (2023).*

Indagados em relação a uma unidade que não possui uma PI implantada e de quais medidas devem ser tomadas para a possível implementação, os respondentes trouxeram as seguintes afirmações: U1 “analisar a viabilidade e enviar proposta à chefia do centro de documentação da rede nacional de bibliotecas do instituto”; U2 “fazer um diagnóstico de como o processo de indexação é realizado, verificar como o usuário recupera as informações no sistema (fazer testes), anotar as tomadas de decisões quando ao processo de indexação, estudar o tema, verificar políticas de outras instituições que tenham características semelhantes”; U3 “consci-

entizar, apresentar propostas para a administração superior”; e U4 “formação de grupo de trabalho interdisciplinar para estudo e elaboração da política”. Sobre as medidas citadas anteriormente, somente a U1 não as tomaria, comentário justificado por afirmação dizendo que “toda mudança e melhoria que envolve um contexto de rede nacional é muito complexa e normalmente onerosa, pois se fizer algo que atenda a uma unidade federativa, é necessário fazer igualmente para todas (equipamento, treinamento, etc). Portanto, deixo essa tarefa para a pessoa competente no Centro de Documentação da sede”.

Acerca das considerações feitas, destaca-se a significância do fato das unidades pensarem em questões de observações, estudos e testes para aplicação de uma política e de levá-la aos seus superiores para aprovação. São ponderações realizáveis e úteis à qualidade do processo, porém o foco deve estar na realização da indexação e nos seus resultados, pois apesar da nomenclatura e complexidade da política de indexação, é fato que ela está ligada a organização registrada de práticas que já são realizadas ou conhecidas nas unidades. Ou seja, é um material que possibilita reflexões sobre aquilo que é posto em prática ou que ainda não é feito, e os impactos que traz para a utilização do acervo e o atendimento aos usuários. Vale ressaltar, também, que a PI, ao ser explicada por Fujita (2012) a partir de dois eixos, evidencia que reflete nos aspectos de organização e representação da informação, como também em questões de cunho administrativo, financeiro, recursos humanos e clientela.

5 Conclusão

A pesquisa buscou identificar a percepção bibliotecária sobre política de indexação, a partir de um estudo com profissionais responsáveis por bibliotecas especializadas. Diante disso, compreendeu-se que os bibliotecários investigados, além de possuírem vasta experiência na profissão, em sua maioria também trabalham há anos em unidade dessa tipologia. Logo, mesmo diante de diferentes nichos de atuação dessas bibliotecas de caráter especializado, alcançaram-se resultados muito semelhantes a partir de respostas emitidas por U1, U2, U3 e U4.

Desse modo, as considerações feitas refletem conhecimento e entendimento do que vêm a ser, da composição e da importância de uma política de indexação. Afinal, por mais que não a tenham estabelecida formalmente, os bibliotecários atuantes em biblioteca especializada demonstram conhecimento acerca de alguns dos seus elementos e aspectos, relevantes no processo de indexação. Além desse fato, a pesquisa permitiu perceber que as bibliotecas especializadas que integraram o corpus investigativo estão abertas para implementação e uso de uma PI, na condição de autoriza-

ção e apoio de seus superiores na instituição. Tal aspecto reforça a condição de que a futura implantação deva ocorrer em consonância com a cultura organizacional existente e, observando as especificidades de cada unidade.

Conclui-se, assim, que o amparo institucional pode incentivar as unidades a implantarem um documento como a política de indexação, que auxiliaria e contribuiria positivamente para a rotina de atendimento das demandas dos usuários. Razão pela qual pesquisas futuras podem se dedicar à compreensão das reais motivações do não estabelecimento, até então, de PI no âmbito estudado, uma vez que têm pensamentos alinhados quanto as atribuições e qualidades de uma política de indexação para bibliotecas especializadas.

Referências

- BENTES PINTO, Virginia; RABELO, Camila Regina de Oliveira; GIRÃO, Igor Peixoto Torres; RAULINO NETO, Francisco Hilton Rodrigues. A percepção dos profissionais e dos estudantes de Biblioteconomia acerca da práxis do arquivato da informação. **Folha de Rosto**, v. 1, n. 2, p. 17-28, jul./dez. 2015.
- CARIBÉ, Rita de Cássia do Vale. A biblioteca especializada e o seu papel na comunicação científica para o público leigo. **RICI**, Brasília, v. 10, n. 1, p. 185 -203, jan. /jul. 2017.
- FUJITA, Mariângela Spotti Lopes. A linguagem documental na negociação de uma política de indexação para bibliotecas universitárias: procedimentos e estratégias da pesquisa-integral. **Revista Conhecimento em Ação**, Rio de Janeiro, v. 1, n. 1, p. 2-17, jan./jun. 2016.
- FUJITA, Mariângela Spotti Lopes. A política de indexação para representação e recuperação da informação. In: GIL-LEIVA, Isidoro; FUJITA, Mariângela Spotti Lopes (Eds.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 17-30.
- GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 5.ed. São Paulo: Atlas, 1999. 200 p.
- GIL LEIVA, Isidoro. Aspectos conceituais da indexação. In: GIL-LEIVA, Isidoro; FUJITA, Mariângela Spotti Lopes (ed.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 31-106.
- GOMES, Rainer Finelli; LIMA, Gercina Ângela de. A política de indexação nos estudos publicados na área de Biblioteconomia e Ciência da Informação: uma revisão sistemática de literatura. **Em Questão**, Porto Alegre, v. 27, n. 1, p. 210-236, jan./abr. 2021.
- GUIM, Vera Lúcia Ribeiro; FUJITA, Mariângela Spotti Lopes. As linguagens de indexação e a análise de domínio. In: GUIMARÃES, José Augusto Chaves; DODEBEL, Vera (org.). **Organização do conhecimento e diversidade cultural**. Marília: ISKO-Brasil; FUNDEPE, 2015. p. 125-134.
- GUIMARÃES, José Augusto Chaves. As políticas de indexação como elemento para a gestão do conhecimento nas organizações. In: VIDOTTI, Silvana Aparecida Borsetti Gregorio

(coord.). **Tecnologia e conteúdos informacionais**: abordagens teóricas e práticas. São Paulo: Polis, 2004. p. 43-52.

GUIMARÃES, José Augusto Chaves; SANTOS, André Ynada dos; SALES, Rodrigo de; MATOS, Daniela Fernanda de Oliveira. A dimensão conceitual da organização do conhecimento nos congressos da NASKO: análise de conteúdo Bardiana. In: GUIMARÃES, José Augusto Chaves; DODEBEI, Vera (org.). **Organização do conhecimento e diversidade cultural**. Marília: ISKO-Brasil; FUNDEPE, 2015. p. 55-72.

KOCHANI, Ardala Ponce; BOCCATO, Vera Regina Casari; RUBI, Milena Polsinelli. Política de indexação para clippings: otimização do tratamento temático em coordenadorias de comunicação social brasileiras. **Revista Interamericana de Bibliotecología**, Medellín, v. 35, n. 3, p. 257-273, 2012.

RODRIGUES, Talita Andrade; FUJITA, Mariângela Spotti Lopes; DAL'EVEDOVE, Paula Regina. Política de indexa-

ção em bibliotecas destinadas às necessidades dos usuários deficientes visuais. **Revista Conhecimento em Ação**, Rio de Janeiro, v. 2, n. 2, p. 56-72, jul./dez. 2017.

RUBI, Milena Polsinelli. Política de indexação. In: GIL LEIVA, Isidoro; FUJITA, Mariângela Spotti Lopes (ed.). **Política de indexação**. São Paulo: Cultura acadêmica; Marília: Oficina Universitária, p. 107-120, 2012.

SANTOS, Raimunda Fernanda dos; SILVA, Jefferson Higinda; ALBUQUERQUE, Maria Elizabeth Baltar Carneiro de; OLIVEIRA, Henry Poncio Cruz de. Implicações da Pós-verdade na indexação de recursos informacionais. In: BARROS, Thiago Henrique Bragato; TOGNOLI, Natalia Bolfarini (org.). **Organização do conhecimento responsável**: promovendo sociedades democráticas e inclusivas. Belém: Ed. da UFPA, 2019. p 85-94.

Perspectivas do conhecimento arquivístico no cenário de acesso à informação e proteção de dados pessoais

Perspectives of archival knowledge in the information access and personal data protection scenario

Denize Laureano Rocha (1), Clarissa M. dos Santos Schmidt (2)

(1) Universidade Federal Fluminense, Niterói (RJ), denizerocha@id.uff.br. (2) clarissaschmidt@id.uff.br

Resumo

Diferentes possibilidades informacionais contidas nos documentos de arquivo, quais sejam, aquelas vinculadas ao contexto de produção e ao conteúdo, serviram de prisma para explorar as perspectivas do conhecimento arquivístico no cenário de acesso à informação e proteção de dados pessoais em ambientes digitais. A partir do desenvolvimento das questões propostas, objetiva-se refletir sobre a representação de elementos constitutivos dos documentos de arquivo e o uso de suas informações de contexto e de conteúdo aplicadas à realidade contemporânea. O procedimento metodológico utilizado foi o teórico-bibliográfico, de modo a explorar conceitos que permitam pensar acerca de perspectivas do conhecimento arquivístico no cenário de acesso à informação e proteção de dados pessoais. A investigação encontrou na própria arquivística elementos para ampliar a representação, controle, recuperação e acesso às informações contidas nos documentos, empregando coordenadamente as informações de contexto e de conteúdo como subsídios à *práxis* nos arquivos, independente de qual fase ou suporte estejam. Além disso, recursos de outras áreas do saber, como os esquemas de metadados e Sistemas de Organização do Conhecimento, podem ser utilizados para complementar os métodos e técnicas arquivísticos.

Palavras-chave: Acesso. Documento digital. Contexto de produção documental. Conhecimento arquivístico.

Abstract

The different informational possibilities contained in archival documents, that is, those linked to the context of production and to the content, served as a prism to explore the perspectives of archival knowledge in the scenario of access to information and protection of personal data in digital environments. From the development of the proposed questions, the objective is to reflect on the representation of constituent elements of archival documents and the use of their context and content information applied to contemporary reality. The methodological procedure used was the theoretical-bibliographical one, in order to explore concepts that allow thinking about perspectives of archival knowledge in access to information and protection of personal data scenario. The investigation found in archival science itself elements to expand the representation, control, retrieval and access to the information contained in the documents, using coordinately contextual and content information as subsidies to the *praxis* in the archives, regardless of what phase or support they are. Furthermore, resources from other areas of knowledge, such as metadata schemas and Knowledge Organization Systems, can be used to complement archival methods and techniques.

Keywords: Access. Electronic records. Contexto. Archival knowledge.

Resumen

Las diferentes posibilidades informacionales contenidas en los documentos de archivo, es decir, aquellas vinculadas al contexto de producción y contenido, sirvieron de prisma para explorar las perspectivas del conocimiento archivístico en el escenario del acceso a la información y protección de datos personales en ambientes digitales. A partir del desarrollo de las preguntas propuestas, el objetivo es reflexionar sobre la representación de los elementos constitutivos de los documentos de archivo y el uso de su información contextual y de contenido aplicada a la realidad contemporánea. El procedimiento metodológico utilizado fue el teórico-bibliográfico, con meta de explorar conceptos que permitan pensar perspectivas de conocimiento archivístico en el escenario del acceso a la información y protección de datos personales. La investigación encontró en la archivística misma elementos para ampliar la representación, control, recuperación y acceso a la información contenida en los documentos, utilizando información contextual y de contenido de forma coordinada como subsidios a la *praxis* en los archivos, independientemente de su fase o soporte. Además, los recursos de otras áreas del conocimiento, como los esquemas de metadatos y los Sistemas de Organización del Conocimiento, se pueden utilizar para complementar los métodos y técnicas de archivo.

Palabras clave: Acceso. Documento de archivo digital. Contexto. Conocimiento Archivístico.

1 Introdução

As questões abordadas no presente trabalho fazem parte de uma pesquisa de Doutorado em andamento no Programa de Pós-Graduação em Ciência da Universidade Federal Fluminense. Nossa investigação dialoga com elementos da realidade da sociedade contemporânea e ganha novos contornos durante seu desenvolvimento por influência de mudanças correntes no contexto social no qual está inserida.

Busca-se desenvolver uma reflexão sobre a representação de elementos constitutivos dos documentos de arquivo e o uso de suas informações de contexto e de conteúdo, aplicada à realidade contemporânea. Para tanto, o procedimento metodológico escolhido foi o teórico-bibliográfico, de modo a explorar conceitos que permitam refletir acerca de perspectivas do conhecimento arquivístico no cenário de acesso à informação e proteção de dados.

À vista disso é importante destacar que o uso de tecnologias da informação molda a organização e o funcionamento da sociedade ao longo do tempo. Em ambientes cada vez mais digitais, pessoas e organizações registram suas ações e atividades em plataformas, sistemas, e softwares que produzem e acumulam dados, informações e documentos em ambientes digitais. A utilização destes ambientes para a produção, uso, armazenamento e acesso às informações e documentos arquivísticos impõe a constante necessidade de aprimorar os conceitos e técnicas utilizados para acompanhar o ritmo acelerado da transformação digital que ocorre no mundo.

Nesse cenário, a tecnologia pode favorecer a democratização do acesso à informação, bem como facilitar a produção e troca de informações entre pessoas e organizações, impulsionando a acumulação de documentos e colaborando com o exercício da cidadania ao ultrapassar barreiras que são próprias da realidade física tangível.

Em contraponto, ao mesmo tempo desperta preocupações em torno de como controlar, representar, e manter as características dos documentos que são produzidos, utilizados e custodiados em ambientes digitais ao longo do tempo e no espaço.

Assim, cabe esclarecer que as características de imparcialidade, autenticidade, naturalidade, inter-relacionamento e unicidade dos documentos de arquivo debatidas por Duranti (1994) continuam sendo elementos fundamentais para que estes registros sirvam às suas finalidades e constituam provas confiáveis das ações que atestam, independente do meio ou suporte no qual são produzidos.

Nessa perspectiva, a partir da conjuntura delineada até aqui, nos aproximamos do elemento que garante ao documento as características inscritas por Duranti e a

qualidade de “arquivístico”: seu contexto de produção. Este elemento compreende “tudo o que está fora do documento e tem significado para o seu sentido” (DURANTI; EASTWOOD; MACNEIL, 2002, p. 18).

Schmidt e Smit (2015, p. 4) atribuem às informações de contexto, também chamadas de extrínsecas, caráter prioritário quando se trata dos documentos de arquivo, haja vista que nelas reside a manifestação dos princípios da proveniência e da ordem original dos documentos.

Ademais, para melhor entendimento do que abrange o contexto de produção do documento de arquivo, é possível compreendê-lo a partir de cinco perspectivas: 1) Contexto de Procedimentos; 2) Contexto de Proveniência; 3) Contexto Documental; 4) Contexto Jurídico-Administrativo; 5) Contexto Tecnológico (CONARQ, 2020, p. 19).

O contexto de procedimentos compreende todo o “conjunto de normas internas que regulam a produção, tramitação, uso e arquivamento dos documentos da instituição”; o de proveniência “organogramas, regimentos e regulamentos internos que identificam a instituição produtora de documentos”; o contexto documental diz respeito ao “código de classificação, guias, índices e outros instrumentos que situam o documento dentro do conjunto a que pertence, ou seja, do fundo”; o jurídico-administrativo refere-se ao “conjunto de leis e normas externas à instituição produtora de documentos as quais controlam a condução das atividades dessa mesma instituição”; e, por fim (e igualmente relevante), o contexto tecnológico, que é o “ambiente tecnológico (hardware, software e padrões) que envolve o documento” (CONARQ, 2020, p. 19).

Consequentemente, o contexto envolve toda a estrutura e o ambiente transversais à ação registrada no documento, inclusive as plataformas, sistemas, hardwares e softwares utilizados para a produção de documentos e informações. Diante disso, os ambientes digitais devem obrigatoriamente preconizar a manutenção das características contempladas nestes cinco elementos de contexto objetivando conferir qualidade arquivística aos documentos.

Nessa mesma linha de pensamento, Tognoli, Rodrigues e Guimarães (2019) reforçam que o contexto também é fator fundamental para pensar a organização do conhecimento arquivístico, haja vista que é o pilar de sustentação do valor probatório e histórico dos documentos.

Cabe esclarecer que entende-se por conhecimento arquivístico aquele gerado ao longo do processo de análise do órgão produtor, seus documentos, funções e atividades, constituindo um capital essencial para alicerçar o desempenho das funções arquivísticas (TOGNOLI; RODRIGUES; GUIMARÃES, 2019, p. 71).

Por sua vez, o direito de acesso à informação, a transparência, a preocupação com sigilo, proteção de dados pessoais e dados sensíveis, são pontos eloquentes nos debates relacionados ao dever das instituições e direitos dos cidadãos. Por isso, o arquivo, e consequentemente o conhecimento arquivístico, assumem papel fundamental nesse debate.

Deste modo, lançamos mão da concepção das possibilidades informacionais, apresentadas por Schmidt e Smit (2015), para fundamentar as reflexões propostas na intenção de reforçar o papel dos arquivos no cenário de acesso à informação, sigilo, transparência e proteção de dados pessoais.

[...] o documento de arquivo possui duas possibilidades informacionais, isto é, as informações de contexto, de “fora”, e as de conteúdo, de “dentro”. Nesse sentido, no campo dos arquivos a função classificação preocupa-se em representar as informações de “fora” de modo a garantir os vínculos, o caráter orgânico e consequentemente o valor primário e probatório do documento de arquivo. Trata-se de elemento caro à área, pois é o que atribui identidade e autonomia ao conhecimento arquivístico (SCHMIDT; SMIT, 2015, p. 4).

Assim, chegamos ao entendimento de que tanto as informações de contexto quanto às informações de conteúdo constituem o insumo necessário aos processos de análise que geram o conhecimento arquivístico.

2 Possibilidades informacionais

Partindo do ponto de vista de Schmidt e Smit (2015, p. 4), consideramos “duas possibilidades informacionais” para potencializar o conhecimento arquivístico, isto é, as informações de contexto, que estão ‘fora’, e as de conteúdo, que estão ‘dentro’ do documento de arquivo. Vale ressaltar que no contexto (fora) reside o vínculo arquivístico e o caráter orgânico dos documentos, o que outorga-lhes o valor probatório.

A separação didática entre estas duas categorias de elementos possibilita o exame minucioso da constituição dos documentos de arquivo para propor novas perspectivas que harmonizem a necessidade de manter as características fundamentais dos documentos com as novas demandas apresentadas pela sociedade nos dias de hoje.

O contexto, elemento extrínseco ao documento, envolve toda a estrutura na qual é produzido e situa-o como parte de um todo, permitindo reconhecê-lo em um conjunto orgânico e não como unidade isolada. Além disso, ao reconhecer e controlar esse contexto, torna-se possível estabilizar o sentido do documento em seu conjunto, ao longo do tempo e no espaço.

Logo, as atividades inerentes ao campo dos arquivos traçam a execução de suas funções, à priori, em torno das informações que estão “fora” do documento de

arquivo, isto é, do seu contexto. Representar esse contexto é o que garante a manutenção do documento de arquivo enquanto prova do funcionamento de um determinado órgão produtor.

Por seu turno, as informações de conteúdo também devem ser consideradas em arquivos, tendo em vista sua importância na conformação do conhecimento arquivístico, bem como na compreensão dos significados dos documentos pelos usuários. Porém, conforme destacam Schmidt e Smit,

O acesso à informação nos arquivos pressupõe, portanto, a solidariedade entre dois tipos de informação (contexto e conteúdo) e somente o efetivo entrelaçamento das duas categorias informacionais gera as condições para que os arquivos preencham seu papel social, ou seja, a organização do acesso a um determinado tipo de informação. (SCHMIDT; SMIT, 2013, p. 579).

Tradicionalmente na arquivística as informações de contexto são o centro das atenções nas ações que compreendem os documentos durante a produção e tramitação; as informações de conteúdo, por sua vez, só ganham os holofotes na fase permanente, quando os documentos já concluíram sua tramitação (SCHMIDT; SMIT, 2013, p. 585).

Essa perspectiva é condicionada pela herança do pensamento sobre as atividades nos arquivos considerando documentos em suportes convencionais, e traça uma separação entre contexto e conteúdo que limita as possibilidades de acesso e uso das informações e não condiz mais com o panorama tecnológico que abriga as circunstâncias de produção dos documentos nas organizações contemporâneas.

[...] em tempo presente, o documento arquivístico eletrônico é o majoritariamente produzido, além dos procedimentos de pesquisa presencial nas instituições terem sido abreviados, sendo cada vez mais ampliada a necessidade de consultas virtuais, com respostas rápidas e objetivas. Além disso, muitas organizações possuem situações de recuperação de conteúdo em documentos de arquivo nos momentos de produção e trâmite documental, em serviços de protocolo, nos bancos de dados, nos sistemas eletrônicos, e na gestão de seus acervos e informações de maneira geral (SCHMIDT; SMIT, 2013, p. 585).

Observa-se, então, que planos de classificação, tabelas de temporalidade de documentos, inventários e guias são instrumentos de gestão e de pesquisa que materializam a realização das funções arquivísticas e o conhecimento arquivístico produzido no processo de análise dos documentos. Estes instrumentos são responsáveis por representar a complexa teia relacional composta pelos documentos, o órgão produtor, suas funções e atividades. Além disso, também viabilizam o controle, gestão, recuperação e acesso às informações contidas nos documentos de arquivo.

Outrossim, ferramentas metodológicas de outras áreas do saber podem ser utilizadas como importantes recursos para complementar a representação dos elementos constitutivos dos documentos de arquivo, isto é, as informações de contexto e as informações de conteúdo.

Nessa linha de pensamento, consideramos que o uso de metadados, recurso advindo da tecnologia da informação, pode contribuir para o registro de elementos intrínsecos e extrínsecos dos documentos quando utilizados para descrever e viabilizar a recuperação, gerenciamento, compreensão e preservação de documentos arquivísticos ao longo do tempo (CONARQ, 2020, p. 36).

A utilização desse recurso para representar os elementos constitutivos dos documentos de arquivo deve ser planejada e padronizada de acordo com as necessidades identificadas em um estudo prévio que contemple a produção documental e seus contextos. Os parâmetros adotados podem ser registrados em um “Esquema de Metadados”, que é o instrumento que materializa o plano lógico adotado, demonstra as relações entre os elementos elencados, estabelece regras e orienta a utilização e gestão dos metadados (CONARQ, 2020, p. 28).

Portanto, observamos que o “Esquema de Metadados” pode ser associado aos instrumentos de gestão e de pesquisa arquivísticos, para compor o conjunto de ferramentas necessárias para representar, controlar e manter os documentos de arquivo, seus contextos e significados.

Do mesmo modo, a perspectiva dos Sistemas de Organização de Conhecimento (SOCs), própria da área de Organização do Conhecimento, também pode ser utilizada como recurso associado para esquematizar, organizar e representar as informações de contexto e de conteúdo, bem como o conhecimento arquivístico produzido ao longo das análises inerentes às funções arquivísticas.

SOCs são sistemas conceituais semanticamente estruturados que contemplam termos, definições, relacionamentos e propriedades dos conceitos. Na organização e recuperação da informação, os SOC cumprem o objetivo de padronização terminológica para facilitar e orientar a indexação e os usuários. Quanto à estrutura variam de um esquema simples até o multidimensional, enquanto que suas funções incluem a eliminação da ambiguidade, controle de sinônimos ou equivalentes e estabelecimento de relacionamentos semânticos entre conceitos (CARLAN; BRÄSCHER, 2011, p. 54).

É importante destacar que o cenário atual condensa a complexidade de gerenciar o alto volume de documentos e informações produzidos em meio digital, concomitantemente à crescente demanda da sociedade por transparência e acesso à informação e, por

consequência, a necessidade de assegurar o direito à privacidade, sigilo e proteção de dados pessoais.

Diante disso, consideramos que a utilização de metadados e de Sistemas de Organização do Conhecimento podem ser uma via de acesso para representar, controlar e manter o sentido dos documentos arquivísticos, seu contexto e seus elementos constitutivos.

Ademais, Schmidt e Smit (2013, p. 585) sugerem que a utilização de diferentes formas de organização e recuperação das duas possibilidades informacionais favorecem a disponibilidade e o acesso às informações contidas nos documentos de arquivo.

Nesse panorama vale esclarecer que não há oposição entre as duas possibilidades informacionais. As informações de contexto e de conteúdo são complementares entre si, de forma que o uso coordenado destes dois elementos constitutivos do documento de arquivo viabiliza a recuperação e acesso às informações sem prejuízo da manutenção do estatuto arquivístico do documento.

É importante reforçarmos que o estatuto probatório do documento de arquivo se dá pela sua natureza contextual e não pela natureza do assunto. Portanto, na teoria arquivística, as informações de contexto são as mais significativas para a classificação documental. Porém, como vimos, na sociedade da informação e em rede, de contínuo desenvolvimento das tecnologias digitais e acentuada troca de informação e conhecimento, o acesso ao conteúdo do documento, isto é, conteúdo substantivo – natureza do assunto, é algo que deve e passa a ser contemplado na atividade do profissional da informação (SCHMIDT; SMIT, 2013, p. 585).

Desta feita, é possível observar que a associação de recursos inerentes à arquivística também pode constituir um caminho para aprimorar a representação, controle e manutenção dos documentos e informações arquivísticas em seu contexto de produção.

Significa dizer que informações de contexto e informações de conteúdo podem ser empregadas simultaneamente nas funções arquivísticas para ampliar a representação, o entendimento e as possibilidades de acesso aos documentos e informações contidas nos documentos de arquivo.

Nessa lógica, ao romper com a fronteira contida na ideia de que informações de contexto são insumos direcionados aos instrumentos de gestão, ao passo que informações de conteúdo são objetos pertencentes somente aos instrumentos de pesquisa, é possível alcançar a elaboração de instrumentos amplos que sejam capazes de representar, organizar, dar sentido, controlar e articular todos os elementos e atores que coexistem no funcionamento das organizações e na produção de documentos de arquivo.

Portanto, é possível observar que encontramos na própria arquivística recursos que já nos são conhecidos e que podem ser utilizados sob outro prisma para ampliar o acesso à informação.

Por fim, a proposta apresentada pela presente reflexão é aplicar a concepção dos SOC à arquivística e utilizar os metadados para promover a representação, controle e manutenção do significado das informações de contexto e de conteúdo dos documentos de arquivo no desenvolvimento de instrumentos, modelos e sistemas que viabilizem o acesso, uso e proteção dos dados e informações.

3 Acesso à informação e proteção de dados pessoais: direitos fundamentais ligados aos arquivos

O acesso à informação é um direito intrinsecamente ligado ao exercício da cidadania e aos direitos humanos fundamentais. A discussão em torno do direito de acesso à informação emerge no cenário político nacional e internacional no contexto pós-guerra, junto às questões políticas ligadas à ciência e tecnologia (GOMEZ, 2002).

Entretanto, essa questão já era uma preocupação no Século XVIII, quando sob forte influência do iluminismo no período Pós-Revolução Francesa, foi publicada a Declaração dos Direitos do Homem do Cidadão, que defende a livre comunicação do pensamento e das opiniões, estabelecendo um elo entre acesso à informação e exercício da democracia (CAMPOS; 2020).

Cabe aqui esclarecer que ao falar em direito de acesso à informação compreende-se essencialmente as sociedades organizadas em torno de regimes democráticos, haja vista que essa discussão não tem espaço nos preceitos de sociedades ditatoriais.

Dito isto, podemos observar que o direito de informação está refletido em diversos documentos que versam sobre os direitos humanos ao redor do mundo: Declaração Universal dos Direitos Humanos (1948 - Organização das Nações Unidas - ONU), Pacto Internacional dos Direitos Cívicos e Políticos (1966 - ONU), Declaração Interamericana de Princípios de Liberdade de Expressão (2000 - Comissão Interamericana de Direitos Humanos), Convenção da Assembléia-Geral das Nações Unidas (2005 - ONU), além da Parceria para Governo Aberto ou Open Government Partnership - OGP, que foi criada em 2011 na intenção de fomentar o acesso à informação, transparência e participação social, que teve adesão de 75 países, incluindo o Brasil (CAMPOS, 2020, p. 29).

Além disso, entre as décadas de 1980 e 1990 entidades como o Banco Mundial (BM) e o Fundo Monetário Internacional (FMI) também trataram sobre a necessidade de empenhar esforços em torno de atos normativos e legislativos que regulamentassem o direito de acesso à informação nos países-membros como

mecanismo de combate à corrupção (MARTINS; PRESSER, 2015).

Essas iniciativas produziram reflexos na Constituição da República Federativa do Brasil, aprovada em 5 de outubro de 1988, que preconiza em seu Art. 5º, inciso XXXIII, o direito de acesso à informação, posteriormente regulamentado pela Lei nº 12.527, de 18 de novembro de 2011.

XXXIII - todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado (BRASIL, 1988).

Todavia, vale destacar que:

Os levantamentos das experiências internacionais mostram que, antes do Brasil, cerca de 90 países já haviam regulamentado o direito de acesso à informação sob o controle de órgãos públicos, também chamado de direito ao saber, ao conhecimento ou à verdade. Presente na Suécia há mais de 200 anos, nos últimos 10 anos, este reconhecimento se generalizou por todas as regiões do mundo, sendo tema central de discussão nas principais democracias (CAMPOS; 2020, p. 30).

A Constituição Federal de 1988 também determina no Art. 216, § 2º que “cabem à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem” (BRASIL, 1988).

Outrossim, a previsão de direitos e responsabilidades pela Constituição não enseja obrigatoriamente a sua implementação. Três anos após a promulgação da Carta Magna, já havia sido aprovada a Lei de Arquivos, Lei nº 8.159, de 8 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados, e em seu Art. 1º determina que:

Art. 1º É dever do Poder Público a gestão documental e a proteção especial a documentos de arquivos, como instrumento de apoio à administração, à cultura, ao desenvolvimento científico e como elementos de prova e informação” (BRASIL, 1991).

Todavia, somente vinte e três anos depois após a promulgação da Constituição foi sancionada a Lei de Acesso à Informação (LAI), Lei nº 12.527, de 18 de novembro de 2011, que regula o acesso à informação e que passou a vigorar em 2012, ano em que também entrou em vigor o Decreto nº 7.724, de 16 de maio de 2012, que regulamenta a LAI.

A Lei de Acesso à Informação tem como norte a transparência e baseia-se no acesso como regra, mas, como toda regra, estabelece também suas exceções. Regras e exceções também são estabelecidas pela Lei nº 13.709, de 14 de agosto de 2018, que faz referência à LAI, mas é centrada nos requisitos para o tratamento e proteção de dados pessoais, e tem como objetivo

proteger os direitos fundamentais de liberdade, de privacidade e o livre desenvolvimento.

As informações e os dados, objetos centrais destas leis, são os mesmos encontrados nos documentos de arquivo, matéria-prima do trabalho dos arquivistas.

Nesse cenário, restringir as análises e atividades desenvolvidas pelos arquivistas somente aos elementos de contexto pode dificultar a recuperação e acesso, bem como subtrair o valor real dos documentos de arquivo enquanto ativos essenciais para a garantia do acesso à informação e proteção de dados pessoais.

Por isso, ao expandir a perspectiva de análise e tratamento dos documentos ao considerar tanto os elementos de contexto quanto os elementos de conteúdo dos documentos, possibilita o emprego de métodos e técnicas que proporcionem maior controle e ampliação na representação das informações. Essa escolha deve refletir na elaboração de instrumentos técnicos que contemplem uma maior gama de caracteres sobre os documentos, o que favorece a recuperação e acesso à informação, promovendo a transparência.

4 A perspectiva dos Sistemas de Organização do Conhecimento (SOC)

Sistema de Organização do Conhecimento (SOC) é um termo natural da área da Organização do Conhecimento (OC) que contempla diversos instrumentos elaborados em diferentes estruturas e funções, que objetivam a organização, gestão e acesso à informação e ao conhecimento produzido e acumulado em um dado contexto. Por ter concepção flexível pode ser utilizado em diversas áreas do saber, e o desenvolvimento de um SOC incorpora as características específicas de cada contexto de produção e acumulação das informações e do conhecimento a que se referem com a finalidade de representá-los, torná-los compreensíveis e acessíveis (MAZZOCHI, 2018).

Nessa perspectiva, Barros, Bastos e Santos (2022, p. 3) afirmam que o objetivo de um SOC é o mesmo princípio norteador da organização do conhecimento, ou seja, promover a recuperação e o acesso às informações, e, ainda que o alicerce da arquivística esteja no contexto, o desenvolvimento de um SOC se dá em torno da representação do conteúdo. Desta maneira, ainda para os autores:

[...] o relacionamento entre organização do conhecimento e arquivologia hoje se encontra plenamente em consolidação, tal é o fato da comunidade arquivística na organização do conhecimento ter discutido no passado recente, e continuar discutindo, premissas arquivísticas nos ambientes típicos da OC tradicional, citando aqui nominalmente a International Society for Knowledge Organization (BARROS; BASTOS; SANTOS, 2022, p. 4)

Já Hodge (2000, p. 1) afirma que “por serem mecanismos de organização da informação, os sistemas de organização do conhecimento estão no coração de cada biblioteca, museu e arquivo”. Por sua vez, Zeng (2008) categorizou os tipos de SOC com base na complexidade das suas estruturas e alcance das funções conforme ilustração do Apêndice A.

A representação gráfica elaborada por Zeng relaciona os SOC mais simples aos mais complexos, o que demonstra a ampla gama de ferramentas e instrumentos utilizados e que podem ser perspectivados como Sistemas de Organização do Conhecimento.

Nesse sentido, Barros, Bastos e Santos (2022, p. 5), afirmam que, na arquivística, as ferramentas utilizadas como SOC estão situadas na quarta categoria do esquema estabelecido por Zeng, ou seja, voltadas para a eliminação da ambiguidade, controle de sinônimos e estabelecimento de relações hierárquicas.

Os autores constataam, ainda, uma limitação em relação ao desenvolvimento dos instrumentos arquivísticos, que não avançaram em direção à quinta e sexta categorias de SOC, onde é possível estabelecer relações associativas e apresentar as propriedades do conhecimento representado. Exemplificam com os instrumentos provenientes das funções classificação e descrição, que transportaram essa perspectiva enrijecida dos documentos em papel para os ambientes digitais (BARROS; BASTOS; SANTOS, 2022, p. 5).

Vejam, compreendemos que a hierarquia do ponto de vista classificatório no contexto arquivístico é fundamental, mas não deve ser a única perspectiva. A administração tornou-se cada vez mais descentralizada, e o grande acúmulo documental deixou de ser em mesas nas repartições, passando a acontecer em servidores e nuvens públicas ou empresariais.

Quando analisamos a complexidade da descrição arquivística, percebemos que seu maior problema se dá devido à subjetividade dos resumos-síntese que são construídos nos mais variados instrumentos, e que os pontos de acesso são muitas vezes puramente descritivos e não temáticos. Fazemos aqui duas perguntas retóricas: eles são, afinal, construídos para quem? Para os arquivistas ou para quem irá utilizá-los? (BARROS; BASTOS; SANTOS, 2022, p. 5).

Na esteira dessa discussão, trazemos novamente as questões de acesso à informação e de proteção de dados pessoais, que não são contempladas pelos instrumentos de gestão e de pesquisa desenvolvidos a partir da perspectiva arquivística tradicional.

Nesse sentido, é possível vislumbrar a aproximação das funções e instrumentos arquivísticos, das informações de conteúdo dos documentos de arquivo, utilizando essa possibilidade informacional e a perspectiva dos Sistemas de Organização do Conhecimento para nos aproximar de modelos que sejam capazes de estabelecer

relações associativas e apresentar propriedades nas representações do conhecimento contidas nesses instrumentos; é claro, sem prejuízo do contexto enquanto elemento fundamental definidor do documento de arquivo.

Ainda considerando o esquema elaborado por Zeng, outra perspectiva a ser explorada nos SOC é do eixo “Estrutura”, que sai de modelos planos para modelos multidimensionais. Nesse sentido, ao conciliar elementos de contexto e conteúdo nos SOC para arquivos, torna-se possível alcançar diversas dimensões na representação das informações e do conhecimento arquivístico.

Ademais, é possível vislumbrar ao menos quatro dimensões principais igualmente importantes que podem ser representadas ao aplicar a perspectiva dos SOC aos instrumentos arquivísticos: dimensão teórico-conceitual, dimensão legal, dimensão tecnológica, dimensão metodológica.

A dimensão teórico-conceitual refere-se às funções e princípios arquivísticos como base de sustentação imprescindível à elaboração de qualquer instrumento arquivístico, seja de gestão ou de pesquisa. A dimensão legal abarca os marcos legislativos que devem ser observados na elaboração do SOC. Já a dimensão tecnológica abrange os requisitos, softwares e outros recursos tecnológicos que sejam utilizados no contexto de produção, uso ou gerenciamento de documentos, informações e conhecimentos que servem de matéria-prima para o SOC. Por fim, e igualmente importante, a dimensão metodológica contempla as decisões sobre quais requisitos e modelos nortearão a elaboração do Sistema de Organização do Conhecimento materializado em instrumentos arquivísticos.

Assim, utilizando a perspectiva dos SOC, propomos a ampliação dos caracteres representados pelos instrumentos arquivísticos com vias a contemplar tanto os elementos de contexto quanto os de conteúdo dos documentos de arquivo, aumentando o alcance dessas ferramentas com relação ao controle e recuperação das informações para viabilizar o acesso, sem prejuízo da manutenção do significado destes documentos em seu contexto de produção.

5 Considerações finais

A sociedade atual impõe aos arquivos demandas por acesso às informações contidas nos documentos, provocando nos profissionais da área a necessidade em pensar nos métodos e técnica empregados a fim de garantir a recuperação e acesso às informações, sem perder de vista a imprescindibilidade da manutenção do caráter arquivístico dado pelo contexto de produção dos documentos.

Na mesma medida em que cresce a urgência por acesso à informação também emerge a necessidade de controle e segurança dos dados pessoais contidos nos documentos de arquivo.

Soma-se a este panorama a acelerada produção, troca e uso de informações, possibilitadas pelos ambientes, plataformas e sistemas digitais.

Diante disso, encontramos como caminho a utilização dos metadados como recurso e dos Sistemas de Organização do Conhecimento (SOCs) como perspectiva a serem associados às possibilidades informacionais (informações de contexto e informações de conteúdo) para viabilizar o adequado registro e controle dos documentos em seu contexto, ensejando assim a representação do conhecimento arquivístico e assegurando a recuperação e o acesso à informação de maneira eficaz e autêntica.

Referências

BARROS, T. H. B.; BASTOS, C. M. C.; SANTOS, A. C. R. D. **Sistemas de organização do conhecimento no contexto da arquivologia**: aportes metodológicos para seu desenvolvimento. *Acervo - Revista do Arquivo Nacional*, v. 35, n. 2, p. 1-20, 2022. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/194695>. Acesso em: 28 fev. 2023.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em 28 fev. 2023.

BRASIL. **Lei nº 8.159, de 8 de janeiro de 1991**. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/L8159.htm. Acesso em 28 fev. 2023..

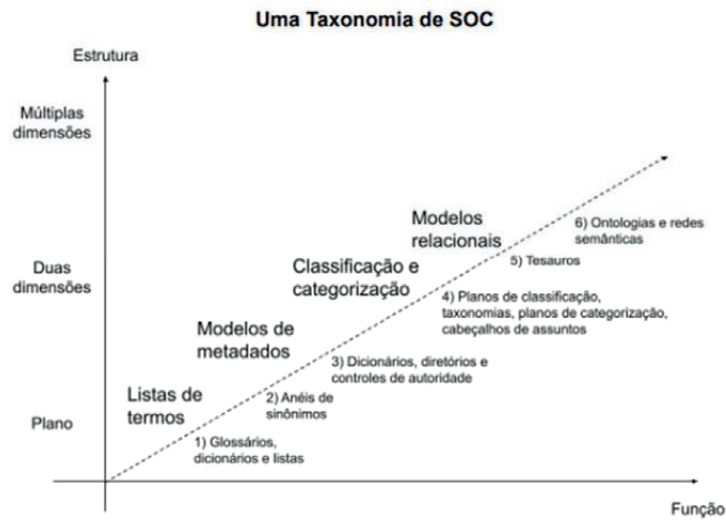
BRASIL. **Lei nº 12.527, de 18 de novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/12527.htm. Acesso em 28 fev. 2023.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/13709.htm. Acesso em 28 fev. 2023..

BRASIL. **Decreto nº 7.724, de 16 de maio de 2012**. Regulamenta a Lei nº 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/decreto/d7724.htm. Acesso em 28 fev. 2023.

- CAMPOS, J. M. S. Lei de Acesso à Informação: Direito Comparado entre Brasil e Estados Unidos. **Revista Acadêmica da Faculdade de Direito do Recife**, v.91, n.2, p.25-48 Set. 2020.
- CARLAN, E.; BRÄSCHER, M. Sistemas de organização do conhecimento na visão da ciência da informação. **Revista Ibero-Americana de Ciência da Informação**, v. 4 No 2, n. 2, p. 53-73, 2011. DOI: 10.26512/rici.v4.n2.2011.1675 Acesso em: 26 out. 2022.
- CONSELHO NACIONAL DE ARQUIVOS. **Glossário dos Documentos Arquivísticos Digitais**. Conselho Nacional de Arquivos (CONARQ), 2020. Disponível em: https://www.gov.br/conarq/pt-br/assuntos/camaras-tecnicas-setoriais-inativas/camara-tecnica-de-documentos-eletronicos-ctde/glosctde_2020_08_07.pdf. Acesso em 05 nov. 2021.
- DURANTI, Luciana. Registros documentais contemporâneos como provas de ação. **Revista Estudos Históricos**, Rio de Janeiro, v. 7, n. 13, p. 49-64, jul. 1994.
- DURANTI, Luciana; EASTWOOD, Terry; MACNEIL, Heather. **Preservation of the Integrity of Electronic Records**. Springer Science+Business Media Dordrecht, 2002.
- GOMEZ, M. N. G. de. (2002). Novos cenários políticos para a informação. **Ciência Da Informação**, 31(1). <https://doi.org/10.18225/ci.inf.v31i1.975>
- HODGE, G. **Systems of Knowledge Organization for Digital Libraries: beyond traditional authorities files**. Washington, DC, the Council on Library and Information Resources. 2000. Disponível em: <https://www.clir.org/pubs/reports/pub91/>. Acesso em: 26 out. 2022.
- MARTINS, C. J. B. N.; PRESSER, N. H. A promoção da cidadania por meio do acesso à informação. **Pesquisa Brasileira em Ciência da Informação e Biblioteconomia**, v. 10, n. 1, 2015. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/27734>. Acesso em: 28 fev. 2023.
- MAZZOCCHI, Fulvio. “Knowledge organization system (KOS)”. **Knowledge Organization** 45, 2018, no.1: 54-78. Also available in ISKO Encyclopedia of Knowledge Organization, eds. Birger Hjørland and Claudio Gnoli, <http://www.isko.org/cyclo/kos>. Acesso em 26 out. 2022.
- SCHMIDT, Clarissa; SMIT, Johanna. Organização e representação da informação em arquivos: uma análise a partir da função classificação. En **XII Congreso ISKO España y II Congreso ISKO España-Portugal**, 19-20 de noviembre, 2015, Organización del conocimiento para sistemas de información abiertos. Murcia: Universidad de Murcia.
- SCHMIDT, Clarissa; SMIT, Johanna. **Organização da Informação e arquivos: diferentes perspectivas informacionais em torno do documento de arquivo. Informação e/ou Conhecimento: as duas faces de Jano**. Porto. Faculdade de Letras da Universidade do Porto - CETAC.MEDIA, p. 579-588. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=4608091>. Acesso em 28 fev 2023.
- TOGNOLI, N. B.; RODRIGUES, A. C.; GUIMARÃES, J. A. C. Definindo o conhecimento arquivístico: estruturas conceituais. **Informação & Informação**, v. 24, n. 2, p. 58-75, 2019. DOI: 10.5433/1981-8920.2019v24n2p58 Acesso em: 26 out. 2022
- ZENG, M. L. Knowledge organization systems (KOS). **Knowledge Organization**, Frankfurt, v. 35, n. 2-3, p. 160-182, 2008.

Apêndice A – Síntese das estruturas e funções da SOC



	1	2	3	4	5	6
Eliminando ambiguidades	XXX		XXX	XX	XXXX	XX
Controlando sinônimos		XXXX	XXX	XX	XXXX	XX
Estabelecendo relações hierárquicas			X	XXX	XXX	XXX
Estabelecendo relações associativas					XXXX	XXXXXX
Apresentando propriedades						XXXXXX

Fonte: Os autores.

Possibilidades para a indexação em museus: identificação e análise de orientações para leitura técnica de objetos museológicos

Possibilities for indexing in museums: identification and analysis of guidelines for technical reading of museum objects

Mateus da Silva Reis (1), Mariângela Spotti Lopes Fujita (2), Roberta Cristina Dal'Evedove Tartarotti (3)

(1) Universidade Estadual Paulista (UNESP), Av. Hygino Muzzi Filho, 737 – Mirante – Marília/SP – CEP 17.525-900, mateusreis42@hotmail.com, (2) mariangela.fujita@unesp.br, (3) Universidade Estadual de Campinas (UNICAMP), Cidade Universitária Zeferino Vaz – Barão Geraldo – Campinas/SP – CEP 13.083-970, roberta_tartarotti@yahoo.com.br.

Resumo

A pesquisa objetiva identificar e analisar orientações para a leitura técnica de objetos museológicos presentes nas principais obras do eixo de Documentação Museológica. Trata-se de uma pesquisa bibliográfica e utiliza como abordagem metodológica a revisão de livros/manuais, como Camargo-Moro (1989), Ferrez (1994), Cândido (2006), Botallo (2010), Padilha (2014) e Roberts (2015). Como resultados, notou-se que as bases para fundamentação da leitura técnica dentro da Documentação Museológica são breves. Observou-se que os textos analisados são de extrema importância para o corpus prático da gestão de acervos, no entanto, as orientações cessam neste ponto. Temáticas como a Leitura Documental e Indexação devem ser melhor desenvolvidas no campo museológico, buscando relacionar teorias de interpretação e simbolismo do objeto junto às variáveis de leitor x texto x contexto. É provável que as variáveis propostas estejam presentes em alguma parte do bojo museológico, não necessariamente retidas na Documentação Museológica, aguardando um estreitamento interdisciplinar.

Palavras-chave: Indexação; Leitura Documental; Documentação Museológica; Objeto Museológico.

Abstract

The research aims to identify and analyze guidelines for the technical reading of museological objects present in the main works of the axis of Museological Documentation. This is a bibliographic research and uses as methodological approach the review of books/manuals, such as Camargo-Moro (1989), Ferrez (1994), Cândido (2006), Botallo (2010), Padilha (2014) and Roberts (2015). As results, it was noted that the bases for substantiation of technical reading within the Museological Documentation are brief. It was observed that the texts analyzed are extremely important for the practical corpus of collection management, however, the guidelines cease at this point. Themes such as Documentary Reading and Indexing should be better developed in the museological field, and should relate theories of interpretation and symbolism of the object to the variables of reader x text x context. It is likely that the proposed variables are present in some part of the museological bulge, not necessarily retained in the Museological Documentation, waiting for an interdisciplinary narrowing.

Keywords: Indexing; Documentary Reading; Museum Documentation; Museum Object

Resumen

La investigación tiene como objetivo identificar y analizar orientaciones para la lectura técnica de objetos museológicos presentes en las principales obras del eje de Documentación Museológica. Es una investigación bibliográfica y utiliza como abordaje metodológico la revisión de libros/manuales, como Camargo-Moro (1989), Ferrez (1994), Cândido (2006), Botallo (2010), Padilha (2014) y Roberts (2015). Como resultados, se constató que las bases de fundamentación de la lectura técnica dentro de la Documentación Museológica son escuetas. Se observó que los textos analizados son de extrema importancia para el corpus práctico de la gestión de colecciones, sin embargo, las directrices cesan en este punto. Temas como Lectura Documental e Indización deben ser mejor desarrollados en el campo museológico, y deben relacionar teorías de interpretación y simbolismo del objeto junto a las variables lector x texto x contexto. Es probable que las variables propuestas estén presentes en alguna parte del bulto museológico, no necesariamente retenidas en la Documentación Museológica, a la espera de un estrechamiento interdisciplinar.

Palabras clave: Indización; Lectura documental; Documentación museológica; Objeto museológico.

1 Introdução

A Leitura Documental (LD) é entendida como etapa inicial no processo de Indexação e busca identificar os conceitos de um recurso informacional tratado, por intermédio de um indexador. A leitura técnica constrói a interação entre o leitor/indexador e o conteúdo em questão, partindo de estratégias cognitivas e metacognitivas, e com a influência de fatores socioculturais que irão resultar na construção de representações de recursos informacionais (FUJITA; NARDI; SANTOS, 1998). Neste contexto, a Leitura Documental abarca três variáveis, sendo estas: o texto, referente à estrutura textual do documento; o leitor, ligado à imagem do indexador; e o contexto, referente à Indexação do conteúdo em sistemas de recuperação (FUJITA, 2004). A Indexação e a Leitura Documental configuram-se como elementos necessários para a representação e recuperação da informação, partindo da premissa que a análise de conteúdo realizada pela leitura soluciona a problemática acerca da tradução de termos que serão utilizados no processo de Indexação, possibilitando uma relação estável entre usuários em diferentes unidades informacionais, tal como os museus (FUJITA, 2004).

Entretanto, a escassez de investigações destes conceitos e suas variáveis nos espaços museológicos resulta em um desconhecimento de possíveis modelos de leitura para análise de assunto que poderiam aprimorar o tratamento temático do objeto museológico. Esta escassez é justificada pelo tratamento da prática museológica (destaca-se a Documentação Museológica e a Conservação Preventiva) enquanto processo repetitivo e imponderável, ocultando possíveis fatores ligados ao significado e contexto sociocultural do objeto e sua relação junto à cognição, metacognição e interpretação do museólogo responsável pelo tratamento informacional. Considerando a insuficiência de discussões voltadas à Indexação e modelos de leitura presentes no campo da Museologia, o questionamento deste estudo surge a partir da inquietação em investigar a existência de orientações para leitura do objeto em literaturas especializadas que abordam o tratamento informacional em espaços museológicos.

O trabalho objetiva identificar e analisar orientações para a leitura técnica de objetos museológicos presentes nas principais obras de Documentação Museológica, sendo estas caracterizadas por livros/manuais com maior índice de citação no eixo proposto. Especificamente, o estudo procura: a) delinear os conceitos de Indexação e Leitura Documental e suas variáveis; b) abordar a questão da Documentação Museológica enquanto norteadora para o tratamento de objetos; c) formular e discutir o compilado de diretrizes de leitura técnica de objetos a partir do referencial teórico. O trabalho proposto

também procura desenhar um percurso teórico voltado à definição geral de Indexação, reconhecendo a Leitura Documental e suas variáveis como elemento constituinte para a compreensão e representação de um recurso informacional.

2 Indexação e Leitura Documental

A Indexação é definida a partir das funções de controle, registro e representação das informações de um determinado assunto, com intuito de otimizar as estratégias de busca em bibliotecas especializadas e centros de documentação (GUIMARÃES, 2009). Gil Leiva (2012, p. 31) afirma que a Indexação proporciona a construção de palavras-chave, índices e cabeçalhos de assunto, originados de uma sucessão interativa de processos mentais, “que têm a ver com a percepção da informação, da memória e da compreensão”. Para o autor, os processos mentais executados por seres racionais para a recepção seletiva da informação, sua codificação simbólica, seu armazenamento e recuperação, equivalem ao conceito de processo cognitivo.

Mai (2005) aponta que o objetivo do processo de Indexação é determinar o assunto temático do recurso informacional e expressá-lo em instrumentos para a representação, tal como os códigos de classificação, os cabeçalhos de assunto e os descritores. Neste propósito, o indexador busca indexar a partir de estratégias bem estruturadas e lógicas para uma eficaz análise de assunto.

No documento “Princípios de Indexação”, criado pela World Information System For Science And Technology, sistema ligado à Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), a operacionalização do processo de Indexação é dividida em dois momentos: o estabelecimento dos conceitos tratados em um documento; e a tradução dos conceitos nos termos de linguagem de Indexação. Especificamente, o primeiro momento (estabelecimento de conceitos do documento) se subdivide em três etapas: a) compreensão do conteúdo como um todo, equivalente à observação de dados obrigatórios para a formação do documento. Como exemplo, destaca-se a estrutura formal de um texto científico, no qual contém título, introdução, seções, subseções, tabelas, gráficos, etc.; b) a identificação dos conceitos que representam este conteúdo, onde o indexador deve seguir uma lógica de seleção de termos ligados ao conteúdo do objeto, priorizando fenômenos, processos e operações; c) seleção de conceitos válidos, determinada pelo objetivo ao qual o documento está sendo indexado, a fim de delinear quais assuntos serão utilizados para recuperar a informação representada. (WORLD INFORMATION..., 1981; SOUSA; FUJITA, 2014).

Na NBR 12676 (ISO 5864), a indexação é dividida em três estágios, sendo estes: o exame do documento e estabelecimento do assunto; a identificação de conceitos; e a tradução dos conceitos nos termos de linguagem de indexação, tais como os tesouros e vocabulários controlados. Conforme a Norma, o processo de identificação de conceitos é indicado através de questionamentos que o indexador deve fazer ao analisar o texto, destacando características como o assunto do documento, as relações semânticas e a interdisciplinaridade do assunto (ABNT, 1992, p. 2).

Para Guimarães, Ferreira e Freitas (2013), a Indexação tem por objetivo descrever o conteúdo informacional do documento, conforme seu assunto, retirando elementos sintéticos e representativos, traduzidos em termos padronizados. O “conteúdo semântico”, conforme o referido autor, configura-se como uma forma de acesso à informação, denominando-se como representação temática.

Silva, Sousa e Bandeira (2012) definem a representação temática, de forma ampla, como a representação do conteúdo do documento, a partir de símbolos, permitindo a identificação do assunto.

No caso de documentos ligados às esferas da Arquivologia e Biblioteconomia, a representação temática visa o conteúdo escrito, a que o documento se refere. Entretanto, a representação de documentos arquivísticos também busca analisar o contexto, a fim de representar o assunto de forma completa. Na Museologia, a representação temática foca na investigação do contexto do documento museológico, buscando compreender a relação indivíduo x objeto, com intuito de representar este documento a partir de termos ligados às dimensões históricas, estilísticas e iconográficas.

Em complemento, Gil Leiva (2012) aborda que as diferentes formas que os profissionais lidam com a Indexação de recursos informacionais são agrupadas em três correntes: a centrada no documento, onde os indexadores “realizam uma descrição precisa e fiel do documento, sem considerar o contexto ou as necessidades de informação dos usuários que se utilizam dela” (SOERGEL, 1985; LANCASTER, 1991 apud GIL LEIVA, 2012, p. 70); a centrada nos usuários, onde são selecionados e traduzidos conceitos em conformidade à linguagem de Indexação baseada nas necessidades de usuários; e a centrada no domínio, na qual gerencia elementos para além do documento, e se baseia no “conhecimento profundo da organização (histórico, objetivos, metas, pessoas e relacionamentos, fluxos de informação, etc.)” (MAI, 2005).

Em suma, Gil Leiva (2008) determina a Indexação enquanto processo realizado com objetos inclinados à representação por conceitos, “a fim de atender as solicitações dos usuários, finalizando com o

atendimento e satisfação da necessidade de informação” (GIL LEIVA, 2008).

A Indexação surge com o aumento das publicações periódicas, no qual eram demandados, por parte das unidades e centros de informação, processos que realizassem a seleção e representação de conceitos dos documentos, objetivando a recuperação da informação (FUJITA, 2003). Neste contexto, as políticas de Indexação buscam formular técnicas e normativas que auxiliem o indexador no processo de representação do assunto, além de delimitar a “[...] cobertura temática em níveis qualitativos e quantitativos tendo em vista os domínios de assuntos e as demandas dos usuários” (FUJITA, 2012, p. 17). Fujita e Santos (2016, p. 62) complementam que “o estabelecimento da política de Indexação é, portanto, imprescindível em sistemas de recuperação da informação, posto que estabelece seus principais atributos e orienta a tomada de decisões para seu planejamento”.

2.1. *Leitura Documental*

Neste cenário de operações de tratamento informacional de conteúdo e de processos de organização do conhecimento, destaca-se a Leitura Documental como a etapa primária que se responsabiliza pela identificação de conceitos, a partir de estratégias e elementos sociocognitivos, possibilitando o processo de análise documental e a construção de produtos resultantes da Indexação.

Segundo Fujita (2004), a leitura é um ato social existente sob um processo comunicacional entre autor e leitor, partindo de objetivos estabelecidos de ambos em um determinado contexto. Entretanto, a autora aponta que diferente da leitura convencional, ou leitura para formação, o modelo Leitura Documental se caracteriza como uma prática profissional que busca “atingir objetivos profissionais dentro de um contexto de trabalho”.

Assim, Fujita (2006) afirma que a Leitura Documental surge na dificuldade do indexador diante do processo da análise de assunto dos documentos. Partindo de elementos e estratégias metacognitivas, a leitura técnica de documentos é iniciada quando o indexador explora a estrutura textual, buscando encontrar o conteúdo temático que compõe o documento (GUIMARÃES; MORAES; GUARIDO, 2007).

Para Fujita, Nardi e Santos (1998), a busca na estrutura do texto é definida como uma etapa não linear, realizada por um indexador com conhecimento prévio e é consciente dos pontos necessários para a leitura e identificação dos conceitos do documento. Neste contexto, o indexador, ou documentalista, é caracterizado como o leitor proficiente que capta as ideias principais, atribui significados e compreende o tema (FUJITA; NARDI; SANTOS, 1998). À vista disso, a figura do indexador se transporta para um

“leitor no ato de análise de um documento com a finalidade de realizar a Indexação para representação do conteúdo por meio de termos que serão, posteriormente, recuperados por um usuário do sistema de informação” (FUJITA, 2004).

Conforme Giasson (1993 apud FUJITA; NARDI; SANTOS, 1998), a compreensão da estrutura textual envolve três variáveis: o leitor, o texto e o contexto. Segundo o autor, o leitor carrega elementos que analisam o sujeito, com base em conhecimentos prévios e habilidades aplicadas no processo de leitura. Para Fujita (1999), durante a leitura são ativados elementos como o conhecimento de vocabulário, assunto e de mundo. Todavia, também interagem neste processo “restrições do contexto do leitor (seu conhecimento prévio, valores, crenças), restrições do texto (intenções do autor refletidas no contexto lingüístico) e restrições do contexto da realização da tarefa de leitura (interesse e objetivo do leitor, estado psicológico...)” (CAVALCANTI, 1989 apud FUJITA, 1999, p. 109). No que se refere ao texto, Giasson define como o material a ser lido, ressaltando a estrutura, forma e conteúdo.

O contexto é definido pelo autor como elementos externos que envolvem o texto, classificados como “contexto psicológico (intenção de leitura, interesse pelo texto...), o contexto social (por exemplo, as intervenções de professores e dos colegas...) e o contexto físico (o tempo disponível, o barulho...)” (GIASSON, 1993 apud FUJITA; NARDI; SANTOS, 1998, p. 16). No contexto sociocultural, Fujita (2006) aponta a necessidade da inclusão de profissionais que trabalham com estratégias de leitura e Indexação no cenário cultural de determinado grupo. Esta necessidade corresponde à importância da reflexão sobre a produção e disseminação de conhecimento resultante de fontes informacionais e documentos registrados em diversos suportes e ligados a grupos e espaços culturais, como os documentos museológicos investigados no campo da Museologia.

No universo de bens culturais materiais, o estabelecimento da leitura técnica recorre à esfera semiótica para o processo de Indexação, na qual oferece vantagens à análise de imagens para interpretação de conteúdo. Para Gatto e Almeida (2020, p. 442), os três passos para esta análise voltam-se à “identificação dos elementos, em seguida, sua existência para, finalmente, gerar seu significado no que tange os assuntos”.

Conforme Golub et al. (2022), realizar as etapas de indexação no vasto campo do patrimônio cultural requer que os profissionais de museus tenham a habilidade de destacar características especiais dos itens analisados. Nesse sentido, o referido autor lista as interfaces a serem observadas, sendo: a heterogeneidade dos objetos, detalhada pela variação

na forma e no conteúdo em um mesmo item; a tematicidade, com variação em cada contexto – social, cultural e histórico – ao qual o item está inserido; a tradução e adequação em linguagens especializadas, delimitando a restrição de representação dos bens culturais por conta de descritores não abrangentes; a representação visual, voltada à qualidade da e aprofundamento visual de imagens digitalizadas; e, por fim, a “recuperação de imagens baseada em conteúdos e a visão computacional na indexação de assuntos” (GOLUB et al., 2022, p. 215).

Na pesquisa, Golub et al. reitera os cuidados para o tratamento informacional com o apoio dos metadados, indicando possíveis restrições nas entidades de obra física e de conteúdo subjetivo. Com isso, os modelos conceituais tornam-se opções eficazes para o desenvolvimento e expansão da tematicidade da informação, baseados na leitura e análise dinâmica dos itens culturais a serem indexados.

3 Metodologia

A presente pesquisa é de natureza bibliográfica e utiliza como abordagem metodológica a revisão de principais obras no eixo de tratamento informacional em museus. Para alcançar o objetivo proposto, foram investigados seis textos voltados à Documentação Museológica e descrição de acervo, sendo estes os de Camargo-Moro (1989), Ferrez (1994), Cândido (2006), Botallo (2010), Padilha (2014) e Roberts (2015). O critério utilizado para delimitar e justificar a relevância das obras foi o nível de citação presente na plataforma Google Acadêmico. Para além, as obras selecionadas justificam-se como manuais base que guiam o tratamento informacional e o processo de Documentação Museológica em espaços museológicos.

Os procedimentos metodológicos se caracterizaram da seguinte maneira: a) a utilização dos termos “Documentação Museológica”, “Documentação em Museus”, “Museu e Aquisição” e “Tratamento Informacional em Museus” na busca da plataforma Google Acadêmico; b) aplicação do filtro para resultados voltados aos livros/manuais que abordem o eixo temático; c) seleção das leituras mais citadas, no período de 2007 a 2021; d) identificação de passagem voltadas à leitura técnica de acervos; e) destaque das principais orientações da temática; f) tabulação e análise crítica dos dados obtidos.

Além disso, a escolha dos termos nesta etapa inicial justifica-se pela intenção de isolar as bibliografias do eixo da Documentação Museológica, não adotando descritores como “leitura” e “leitura documental”, visto que estes poderiam abrir margem para uma densa massa de resultados que se desviariam da esfera prática e teórica da Museologia.

A partir da estratégia de busca apresentada, foram recuperadas e selecionadas as principais obras voltadas à Documentação Museológica e seus respectivos índices de citação, conforme a Tabela 1:

Tabela 1. *Obras selecionadas e seus respectivos índices de citação*

<i>Obra</i>	<i>Autor (a)</i>	<i>Ano de publicação</i>	<i>Citações Google Acadêmico (2007 – 2021)</i>
Documentação museológica: teoria para uma boa prática	Helena Dodd Ferrez	1994	189
Documentação Museológica	Maria Inês Cândido	2006	109
Documentação Museológica e Gestão de Acervo	Renata Cardozo Padilha	2014	78
Museu: Aquisição- Documentação	Fernanda Camargo-Moro	1986	50
Diretrizes em Documentação Museológica	Marilúcia Botallo	2010	30
Inventário e Documentação	Andrew Roberts	2015	19

Fonte: Elaborado pelo autor.

Pós recuperação dos primeiros dados, foram investigadas no corpus das obras possíveis estratégias para a leitura técnica de objetos museológicos, acompanhadas no tópico a seguir.

4 Resultados

Ao analisarmos os textos propostos, foram primeiramente notadas questões inerentes à presença das palavras “leitura técnica” e “Leitura Documental”. As percepções de análise e leitura, encontradas nas obras, voltadas ao objeto museológico são resumidas sob o termo generalizado de “observação/identificação”, estreitando a margem para discussão de conceitos como a própria leitura e compreensão de seus elementos. Mesmo apresentando sugestões de leitura dos elementos iconográficos dos objetos e os fatores que os rodeiam (contexto e o próprio indexador), a interpretação sujeita-se à uma confusão e variedades de termos, optando por identificar o assunto do objeto sob o comando de “visualização geral do objeto”, e não “leitura técnica do objeto”. Ou seja, é constatado que a leitura técnica está presente na prática museológica, mas com outros nomes.

Outro ponto identificado foi a presença destas diretrizes em tópicos não condizentes com a análise de

conteúdo de objetos museológicos. Em grande parte, as passagens de leitura técnica existentes na investigação ressaltaram a não formação objetiva das etapas de descrição e representação documental, direcionando o museólogo especialista para a inclusão de dados do objeto em um instrumento de pesquisa, mas sem orientar de que forma Ler, Identificar, Selecionar e Traduzir as informações associadas.

Um fator de relevância na análise proposta foi o reconhecimento da ficha catalográfica/de catalogação/inventário, abordada em todas as obras, como chave para a leitura técnica. Por ser composto de entidades e propriedades relacionados entre si, o instrumento delineia um roteiro cognitivo para a compreensão do objeto, através de metadados que direcionam para as perguntas: “O que é o objeto?” e “Onde buscar os assuntos referentes a ele?”. Nesse sentido, a ficha direciona, de forma breve, o profissional aos principais elementos do recurso informacional.

A partir da identificação e seleção das obras, foram tabulados os dados que sinalizaram quais foram as obras analisadas, se havia a presença de orientações de leitura técnica e de que forma estas se apresentavam ao profissional museólogo, sendo conferidos na Tabela 2.

Tabela 2. *Obras e orientações de Leitura Documental para acervos museológicos*

<i>Obra/Autoria</i>	<i>Há orientações de Leitura Técnica?</i>	<i>Quais?</i>
Museu- Aquisição/ Camargo-Moro (1986)	Não	-
Documentação Museológica: Teoria para uma Boa Prática/Ferrez (1994)	Não	-
Documentação Museológica/ Cândido (2006)	Sim	“A leitura formal do quadro deve ser feita, preferencialmente, observando o sentido horizontal, dividido por planos (níveis de profundidade, de baixo para cima e da esquerda para a direita) e apontando representações em destaque, contrastes de traços e de policromia, etc”. (CÂNDIDO, 2006, p. 59). “Na descrição formal da peça, deve-se partir do geral para o particular, informando, inicialmente, o material e formato do objeto. Preferencialmente, deve-se observar o sentido horizontal, de baixo para cima, destacando-se elementos ornamentais, recursos pictóricos, etc”. (CÂNDIDO, 2006, p. 60).

<i>Obra/Autoria</i>	<i>Há orientações de Leitura Técnica?</i>	<i>Quais?</i>
Diretrizes em Documentação Museológica/ Botallo (2010)	Não	-
Documentação Museológica e Gestão de Acervo/ Padilha (2014)	Sim	“Descrição intrínseca: refere-se à descrição física do objeto, como, por exemplo, a identificação de marcas ou de algumas assinaturas legíveis, entre outros”. (PADILHA, 2014, p. 52)
Inventário e Documentação/ Roberts (2015)	Sim	“[...] anote qualquer característica própria, inscrições ou marcas e o estado de conservação do objeto”. (ROBERTS, 2015, p. 42).

Fonte: Elaborado pelo autor.

Na obra *Documentação Museológica*, de Cândido (2006), são abordados direcionamentos para a compreensão do recurso informacional, onde roteiriza para que lado observar e iniciar a leitura. Nota-se que o foco desta “sugestão” é a relação leitor (museólogo) x texto (objeto), sem a especificação de refletir sobre o contexto, condicionando a identificação de pontos materiais do objeto, como características “ornamentais, recursos pictóricos, etc” (CÂNDIDO, 2006, p. 60).

Em Padilha (2014), a leitura técnica surge de forma tímida na obra, sendo sugerida como forma de identificação de marcas e assinatura legíveis em objetos. A seleção desta passagem enquanto mecanismo para compreensão do objeto é justificada pelo reconhecimento de marcas, escrituras, ranhuras, etc. como informações complementares para a temática do objeto. Entretanto, a inexistência de roteiros para as estratégias cognitivas de busca do conteúdo torna a sugestão inviável para a Indexação. A mesma situação se repete em Roberts (2015), orientando o profissional ao recorte de inscrições, marcas e estado de conservação do objeto, mas inviabilizando o exercício de metacognição direcionado ao “O que é o objeto?” e “Qual é o assunto do objeto?”.

Como já explicado, as obras tratadas aqui majoritariamente abordam, mesmo que inconscientemente, a leitura técnica através do preenchimento da ficha catalográfica. Entretanto, o dispositivo em questão não pode ser utilizado como base para entender a relação leitor (museólogo) x texto (objeto) x contexto, podendo ser explorada através de modelos estruturados, o que não é apresentado nos resultados obtidos.

A mais visível orientação de leitura técnica é a observada em Cândido (2006), definindo formalmente os passos a serem tomados na compreensão do objeto. Pós a identificação, notou-se que a diretriz ainda atinge níveis abaixo da média, por não ativar elementos como

conhecimento de mundo, vocabulários e assunto, previsto em Fujita (1999).

As obras de Camargo-Moro (1986), Ferrez (1994) e Botallo (2010) integram um campo de desconhecimento quanto à leitura técnica. Mesmo sendo importantes obras para o tratamento informacional de objetos e para a documentação administrativa de espaços museológicos, os textos não citam em nenhum momento um roteiro ou orientação de como ler o objeto e reconhecer sua estrutura, salvo o informado para a construção da ficha catalográfica, o que não se enquadra nesta pesquisa.

As variáveis leitor x texto x contexto no processo de leitura, apresentadas por Giasson (1993), não são encontradas nos textos abordados. Entretanto, na esfera museológica, a reflexão da relação entre indivíduo x objeto x realidade tornam o campo fértil para a investigação de leituras em diferentes acervos. É provável que as variáveis propostas estejam presentes em alguma parte do bojo museológico, não necessariamente retidas na Documentação Museológica, aguardando um estreitamento interdisciplinar. A problemática observada nos resultados volta-se à ausência de relações consistentes entre a Museologia, a Organização e Representação do Conhecimento e a própria Ciência da Informação, impossibilitando a formulação de uma rede teórico-prática integrada por elementos, instrumentos e estratégias que auxiliariam na representação informacional de ponta.

5 Considerações Finais

A Leitura Documental constitui o processo de Indexação como elemento para a compreensão estrutural de um documento/objeto. Formada por etapas bem definidas, a Leitura Documental objetiva estabelecer estratégias que possam colaborar na representação de assunto de diferentes domínios. Na Museologia, notou-se que o caminho para entender a Leitura Documental e suas variáveis ainda é longo. Entretanto, sua esfera teórica oferece elementos para o desenvolvimento conciso de modelos de leitura e para o entendimento da Indexação no campo.

Notou-se na pesquisa que as bases para fundamentação da leitura técnica dentro da Documentação Museológica são breves. Os textos analisados são de importância para o corpus prático da gestão de acervos, mas as orientações cessam neste ponto. Fatores como a Leitura Documental e Indexação devem ser melhor elaborados no campo museológico, devendo relacionar teorias de interpretação e simbolismo do objeto junto às variáveis de leitor x texto x contexto.

A indicação de possibilidades para a leitura técnica de documentos de museu em diversos suportes pode

potencializar o tratamento informacional destes, atenuando a questão do contexto como principal variável para a representação temática do conteúdo. Nesse sentido, questionamentos sobre como as interfaces do tema, significado, uso e histórico do documento museológico se relacionam às variáveis leitor x texto x contexto, bem como quais os desafios e possibilidades para a construção de políticas de leitura e indexação em acervos de museu, surgem para balizar as próximas investigações do tema.

Referências

- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676: Métodos para análise de documentos: Determinação de seus assuntos e seleção de termos de indexação**. Rio de Janeiro, p. 2. 1992.
- BOTTALLO, M. Diretrizes em documentação museológica. In: ASSOCIAÇÃO CULTURAL DE AMIGOS DO MUSEU CASA DE PORTINARI. **Documentação e conservação de acervos museológicos: diretrizes**. Brodowski: Associação Cultural de Amigos do Museu Casa de Portinari; São Paulo: Secretaria de Estado da Cultura de São Paulo, 2010, p. 48-79.
- CAMARGO-MORO, F. Museu: **Aquisição e Documentação**. Livraria Eça Educadora. Rio de Janeiro, p. 82-83, 1986.
- CÂNDIDO, Maria Inez. Documentação Museológica. **Caderno de Diretrizes Museológicas**. 2ª edição. Rio de Janeiro, IPHAN, 2006, p. 33-92.
- FERREZ, Helena Dodd. Documentação museológica: teoria para uma boa prática. **Cadernos de Ensaio**, v. 2, p. 64-74, 1994.
- FUJITA, M. S. L. Abordagem cognitiva e sócio-cognitiva da leitura documentária na formação inicial do indexador: análise da perspectiva individual em contexto sócio cultural. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 7, 2006, Marília. Anais... Marília, UNESP, 2006
- FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para Indexação. **Revista digital de biblioteconomia e ciência da informação**, v. 1, n. 1, 2003. Disponível em: <http://eprints.rclis.org/6266/>. Acesso em: 7 ago. 2020.
- FUJITA, M. S. L. A leitura do indexador: estudo de observação. **Perspectivas em ciência da informação**, v. 4, n. 1, 1999. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/23267>. Acesso em: 14 set. 2020.
- FUJITA, M. S. L. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. **DataGramaZero: Revista de Ciência da Informação**, v. 5, n. 4, 2004.
- FUJITA, M. S. L. A política de Indexação para representação e recuperação da informação. In: GIL LEIVA, I.; FUJITA, M. S. L. (Org.). **Política de Indexação**. São Paulo: Cultura Acadêmica, 2012. Disponível em: https://www.marilia.unesp.br/Home/Publicacoes/politica-de-indexacao_ebook.pdf. Acesso em: 14 set. 2020.
- FUJITA, M. S. L.; NARDI, M. I. A.; SANTOS, S. A leitura em análise documentária. **Transinformação**, v. 10, n. 3, 1998.
- FUJITA, M. S. L.; SANTOS, L. B. P. Política de Indexação em bibliotecas universitárias: estudo diagnóstico e analítico com pesquisa participante. **Transinformação**, v. 28, n. 1, p. 59-76, 2016. Disponível em: https://www.scielo.br/scielo.php?pid=S0103-37862016000100059&script=sci_arttext. Acesso em: 14 set. 2020.
- GATTO, A. C.; ALMEIDA, C. C. Modelo semiótico de leitura documentária para Indexação de fotografias. In: FUJITA, M. S. L.; ALVES, R. C. V.; ALMEIDA, C. C. **Modelos de leitura documentária para Indexação: abordagens teóricas interdisciplinares e aplicações em diferentes tipos de documentos**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2020, pp. 425-444
- GIL LEIVA, I. Manual de indización: teoría y práctica. Gijón: Trea, 2008.
- GIL LEIVA, I. Aspectos conceituais da Indexação. In: GIL LEIVA, I.; FUJITA, M. S. L. (Eds.). **Política de Indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012. p. 31-106.
- GUIMARÃES, José Augusto Chaves. Abordagens teóricas de tratamento temático da informação (TTI): catalogação de assunto, Indexação e análise documental. Ibersid: revista de sistemas de información y documentación, v. 3, p. 105-117, 2009.
- GUIMARÃES, José Augusto; FERREIRA, Gustavo M.; FREITAS, Maria Fernanda M. Correntes teóricas do tratamento temático da informação: uma análise de domínio da presença da catalogação de assunto e da Indexação nos congressos de ISKO-Espanha. 2013.
- GOLUB, K.; ZIOLKOWSKI, P. M.; ZLODI, G. Organizing subject access to cultural heritage in Swedish online museums. **Journal of Documentation**, v. 78, n. 7, pp. 211-247. 2022.
- MAI, J-E. Analysis in indexing: document and domain centered approaches. **Information Processing & Management**, v. 41, n. 3, p. 599- 611. 2005.
- PADILHA, Renata Cardozo. Documentação museológica e gestão de acervo. **Florianópolis: FCC**, 2014.
- RIBEIRO, Célia. Metacognição: um apoio ao processo de aprendizagem. **Psicologia: reflexão e crítica**, v. 16, n. 1, p. 109-116, 2003.
- ROBERTS, A. Inventário e Documentação. In: PATRICK, J. **Como Gerir um Museu: Manual Prático**. Paris: Franly Sa, p. 1-252, 2015.
- SILVA, M.; SOUSA, D. E.; BANDEIRA, P. A representação temática em documentos arquivísticos: o caso da Indexação documental realizada pelos alunos de Arquivologia da UFPB. InCID: Revista de Ciência da Informação e Documentação, v. 3, n. 2, p. 124-141, 13 dez. 2012. Disponível em: <https://www.revistas.usp.br/incid/article/view/48657>. Acesso em: 14 set. 2020.

SOUSA, B. P. de; FUJITA, M. S. L. A análise de assunto no processo de Indexação: um percurso entre teoria e norma. **Informação & Sociedade: Estudos**, [S. l.], v. 24, n. 1, 2014. Disponível em: <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/16281>. Acesso em: 28 nov. 2021.

WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY. Princípios de Indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 10, n. 1, mar. 1981.

Potencial da aplicação de técnicas de Processamento de Linguagem Natural (PLN) e de *Machine Learning* (ML) no processo de categorização para a Análise de Domínio (AD)

Natural Language Processing (NLP) and Machine Learning (ML) techniques potential of applying in the Categorization Process for Domain Analysis (DA)

Maria Lígia Triques (1), Ananda Fernanda de Jesus (2), Ana Cristina de Albuquerque (3), José Eduardo Santarem Segundo (4)

(1) Universidade Estadual de Londrina (UEL), Rodovia Celso Garcia Cid, PR-445, Km 380 - Campus Universitário, Londrina - PR, CEP 86057-970, mlgia.triques@uel.br, (3) albuana@uel.br, (2) Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP-Marília), Av. Hygino Muzzi Filho, 737 - Bairro Mirante, Marília/SP - CEP 17.525-900, af.jesus@unesp.br, (4) santarem@usp.br.

Resumo

Objetivo: Tendo em vista a aproximação com as abordagens teórico-metodológicas da Análise de Domínio (AD), objetiva-se verificar a potencialidade da aplicação de técnicas de Processamento de Linguagem Natural (PLN) e de *Machine Learning* (ML) no processo de categorização temática de artigos científicos, em especial no que concerne a aplicação de algoritmos não supervisionados, com o intuito de simular a análise de um *corpus* teórico inexplorado. **Metodologia:** A pesquisa caracteriza-se como teórica e aplicada, com resultados qualitativos e quantitativos baseados na literatura dos temas em questão e na análise de um estudo realizado em caráter ilustrativo acerca dos resultados de um levantamento bibliográfico, tendo como recorte artigos científicos sobre o patrimônio cultural. **Resultados:** Verifica-se que a contribuição do PLN e ML se encaixa na Abordagem 11 - Cognição científica, conhecimento especializado e inteligência artificial, pois permitir criar modelos mentais de um domínio ou métodos para obter conhecimento para a produção de sistemas especializados a partir do mapeamento dos termos recorrentes, podendo ser combinado a demais abordagens previstas pela AD para ampliar seu potencial. **Conclusão:** De modo preliminar, conclui-se que ao aplicar as técnicas de PLN e ML antes mesmo de uma análise manual prévia do *corpus*, é possível pré-estabelecer padrões que podem facilitar a sua compreensão e discussão. Contudo, o potencial das técnicas de PLN e ML para categorização se mostra ampliado ao aliar demais abordagens que possam contribuir para reconhecer um domínio específico, pois este possui características próprias que determinarão a escolha das categorias, assim como das relações conceituais, impactando as formas de análise.

Palavras-chave: Análise de Domínio. Processamento de Linguagem Natural. Aprendizado de máquina. Categorização.

Abstract

Objective: In view of the approximation with the theoretical-methodological approaches of Domain Analysis (DA), the objective is to verify and discuss the potential of applying Natural Language Processing (NLP) and Machine Learning (ML) techniques in the categorization process thematic of scientific articles, in particular with regard to the application of unsupervised algorithms, with the aim of simulating the analysis of an unexplored theoretical corpus. **Methodology:** The research is characterized as theoretical and applied, with qualitative and quantitative results based on the literature of the themes in question and on the analysis of a study carried out in an illustrative way about the results of a bibliographical survey, with a selection of scientific articles on the heritage cultural. **Results:** It is verified that the contribution of the NLP e ML fits in Approach 11 - Scientific cognition, specialized knowledge, and artificial intelligence, as it allows the creation of mental models of a domain or methods to obtain knowledge to produce specialized systems from the mapping of terms recurrent problems and can be combined with other approaches provided by DA to expand its potential. **Conclusion:** Preliminary, it is concluded that by applying the techniques of PLN and ML even before a previous manual analysis of the corpus, it is possible to pre-establish standards that can facilitate its understanding and discussion. However, the potential of NLP and ML techniques for categorization is increased when combining other approaches that can contribute to recognizing a specific domain, as this has its own characteristics that will determine the choice of categories, as well as conceptual relationships, impacting the ways of analysis.

Keywords: Domain Analysis. Natural Language Processing. Machine Learning. Categorization.

Resumen

Objetivo: En vista de la aproximación con los enfoques teórico-metodológicas del Análisis de Dominio (AD), el objetivo es verificar y discutir el potencial de aplicar técnicas de Procesamiento del Lenguaje Natural (PLN) y Aprendizaje Automático (ML) en el proceso de categorización temática de artículos científicos, en particular en lo que se refiere a la aplicación de algoritmos no supervisados, con

el objetivo de simular el análisis de un corpus teórico inexplorado. **Metodología:** La investigación se caracteriza por ser teórica y aplicada, con resultados cualitativos y cuantitativos basados en la literatura de los temas en cuestión y en el análisis de un estudio realizado de manera esclarecedora sobre los resultados de un levantamiento bibliográfico, con una selección de artículos científicos sobre el patrimonio cultural. **Resultados:** Se verifica que el aporte del PLN e ML se encuadra en el Enfoque 11 - Cognición científica, conocimiento especializado e inteligencia artificial, ya que permite la creación de modelos mentales de un dominio o métodos para obtener conocimiento para la producción de sistemas especializados a partir del mapeo de términos problemas recurrentes y se puede combinar con otros enfoques proporcionados por la AD para expandir su potencial. **Conclusión:** Preliminarmente, se concluye que aplicando las técnicas de PLN y ML incluso antes de un análisis manual previo del corpus, es posible preestablecer estándares que pueden facilitar su comprensión y discusión. Sin embargo, el potencial de las técnicas de PNL y ML para la categorización se incrementa al combinar otros enfoques que pueden contribuir a reconocer un dominio específico, ya que este tiene características propias que determinarán la elección de categorías, así como las relaciones conceptuales, impactando en las formas de análisis.

Palabras clave: Análisis de dominio. Procesamiento natural del lenguaje. Aprendizaje automático. Categorización.

1 Introdução

Criar categorias é um ato presente em diversas atividades humanas, desde as ações cotidianas, até os processos profissionais mais complexos. Pode ser entendido como o ato de expor ideias em uma configuração lógica e organizada mediante distribuição que tem por base a forma que melhor compreende-se o mundo e seus elementos (MARTINS, 2016). É, portanto, uma atividade que parte de uma intenção e, por conseguinte, é contextual.

A noção de categoria, em si, possui uma natureza abstrata, pois, de acordo com Barité (1999), quando parte-se de uma orientação filosófica, se caracteriza por expressões ou propriedades que podem ser percebidas em qualquer objeto ou coisa, por serem generalizantes e instrumentais, no sentido da observância quanto às propriedades para análise do mundo material (BARITÉ, 1999).

Para o autor, as categorias são abstrações que se materializam de forma simples, mas que, com as ferramentas intelectuais proporcionam representações do mundo físico, através de sistemas de conceitos especificamente utilizados para a organização do conhecimento, de forma geral e para as classificações, de maneira mais específica (BARITÉ, 1999). A função ou as características de uma categoria permitem que conceitos sejam agrupados e relações conceituais se façam com maior ou menor grau a partir dessas relações (MOREIRA, 2018).

Assim, o processo de categorização, de acordo com Jacob (2004), se dá em um determinado contexto, sendo que esse pode variar e com ele a composição da categoria é a própria base tanto para a flexibilidade quanto para o poder da categorização. Consequentemente, a forma como algo é categorizado cria um contexto ou estrutura conceitual que não apenas fornece informações sobre a entidade, mas também molda a interação do indivíduo com ela (JACOB, 2004).

Desta forma, no processo de categorização “[...] se distribuyen en categorias los elementos destinados a una clasificación, considerando sus semejanzas, afinidades y diferencias.” (BARITÉ, 2015, p. 45).

Assim sendo, a categorização é um processo inerente ao percurso metodológico das práticas científicas, uma vez que está intimamente vinculada à capacidade de sistematizar os conhecimentos já produzidos em função dos objetivos da pesquisa para obter familiaridade com o domínio pesquisado e identificar potenciais lacunas a serem pesquisadas.

Com apresentam Freitas, Bufrem e Breda, (2016, p. 6, tradução nossa)

A categorização como forma de sistematizar e organizar informação e conhecimento é uma diretriz, não apenas para padronizar sistemas de informação, bancos de dados e bibliotecas, mas para facilitar o desenvolvimento da produção científica em um campo com análise de domínio, abrangendo estudos métricos, bem como a recuperação de informações do usuário [1].

Nesse sentido, a Análise de Domínio, caracterizada pela possibilidade de visualizar os aspectos de um determinado domínio do conhecimento, apresenta-se como uma abordagem e um método que permite delinear características de acordo com categorias e critérios definidos. Guimarães, Pinho e Ferreira (2012, p. 34) reconhecem que:

[...] ao buscar caracterizar um domínio científico a partir do conhecimento por ele produzido, a análise de domínio manifesta-se como um processo organizativo por excelência, que alia teoria e prática para propiciar uma visão mais abrangente dos principais conceitos da área.

Unindo tanto teoria quanto prática, a Análise de Domínio reúne em seus pressupostos metodológicos, subdisciplinas como a bibliometria, organização do conhecimento, recuperação da informação e letramento informacional (HJØRLAND, 2004), distribuídas em onze abordagens propostas por Hjørland (2002). Estas buscam identificar os parâmetros do universo em que um domínio opera (SMIRAGLIA, 2015) e mapear os diferentes atores, instituições e processos de comunicação em um dado domínio, tendo como uma das principais bases, as relações contextuais estabelecidas.

Estudos como os de Smiraglia (2015) e Guimarães e Tognoli (2015), trazem demais possibilidades de abordagens que contribuem e ampliam os estudos da

Análise de Domínio, reafirmando que esse é um campo de diálogo frutífero.

Com o potencial de contribuir com as práticas de categorização, e, portanto, com os processos relacionados à Análise de Domínio, as

[...] técnicas de processamento de linguagem natural (PLN) estão buscando aproximar a linguagem computacional da linguagem humana, buscando oferecer formas de transformar o que uma pessoa escreve e fala em algo compreensível para as máquinas” (CONEGLIAN, 2020, p. 18).

O PLN pode ser aplicado em conjunto com técnicas de *Machine Learning (ML)*, cuja base opera em torno dos agentes computacionais ou algoritmos, os tornam capazes de aprender com a experiência e identificar padrões.

A escolha do algoritmo ou conjunto de algoritmos a ser utilizado no processo de predição é contextual, dependendo das características do *corpus* e dos objetivos da atividade a serem realizadas, destacando-se dois grandes conjuntos de algoritmos: aprendizado supervisionado, que utiliza dados para treinamento, cujo resultado é conhecido e explicitado para o algoritmo. Assim, o algoritmo conhece a solução e a partir dos dados definirá quais são os aspectos que devem ser considerados para classificar algo em uma categoria. No segundo tipo, aprendizado não-supervisionado, não há um resultado ou a solução desejada previamente, sendo o treino realizado então, com padrões estatísticos nos conjuntos de dados (CONEGLIAN, 2020, p. 127).

Técnicas de PLN já foram citadas como relevantes a exploração da linguagem de um determinado domínio. Hjørland (2002) reconhece que é natural recorrer a aspectos linguísticos, uma vez que a tarefa de recuperação envolve expressões e objetos linguísticos. Logo, o autor destaca o fato das técnicas linguísticas computacionais serem úteis na análise estatística da distribuição de termos de um texto, contribuindo sobretudo para ampliar metodologias de construção de vocabulários específicos de domínio, como os tesouros.

Nesse sentido, de modo a incitar uma aproximação com as abordagens metodológicas da Análise de Domínio, o objetivo foi verificar a potencialidade da aplicação de técnicas de Processamento de Linguagem Natural (PLN) e de *Machine Learning (ML)* no processo de categorização temática de artigos científicos, em especial no que concerne a aplicação de algoritmos não supervisionados, com o intuito de simular a análise de um *corpus* teórico inexplorado, isto é, em que categorias não podem ser definidas *a priori*.

2 Metodologia

A pesquisa proposta caracteriza-se como teórica e aplicada, com resultados qualitativos e quantitativos baseados tanto na literatura dos temas em questão,

quanto na análise de um estudo realizado em caráter ilustrativo, com a finalidade de verificar o potencial das técnicas de PLN e ML na categorização de resultados de um levantamento bibliográfico, tendo como recorte artigos científicos.

O referido estudo foi elaborado com o propósito de simular o processo de familiarização com um domínio de interesse, processo esse no qual um dado *corpus* teórico é categorizado pelo pesquisador, agrupando os documentos por semelhança e identificando categorias de análise com base em padrões identificados.

Para atuar como *corpus* foram obtidos artigos científicos sobre a temática “patrimônio cultural”, limitando-se a artigos publicados em inglês. Para compor a amostra de teste foi considerada a estratégia (TS= (“cultural heritage")), sendo selecionada uma amostra aleatória de 100 documentos, sendo considerados apenas os anos pares.

O *corpus* passou por um processo de pré-processamento dos dados, com aplicação de filtros e da elaboração de uma lista de *stopwords* em inglês. As *stopwords* são palavras que se repetem ao longo do texto, mas que não refletem o seu conteúdo, são exemplos: artigos, conectivos, números, tais como páginas e anos. Em seguida, foi aplicado um algoritmo de clusterização hierárquica (*hierarchical clustering algorithm*), utilizando a métrica de distância *cosine*.

Com base nos resultados desse estudo algumas possibilidades foram aventadas sobre o potencial das técnicas de Processamento de Linguagem Natural (PLN) e de *Machine Learning (ML)* para mapeamento de um domínio.

3 A Análise de Domínio

Tendo como base às práticas, interesses e limites de especialização de um determinado grupo de indivíduos (BARITÉ, 2013), um domínio pode ser entendido como um corpo de conhecimento que é comum a uma comunidade discursiva (HJØRLAND, 2017).

Thellefsen e Thellefsen (2004, p. 179) entendem os domínios do conhecimento como “[...] uma demarcação de um determinado conhecimento, seja ele fixado num contexto profissional ou não”, e que tem seus próprios “[...] instrumentos, conceitos, significados, estruturas de informação, necessidades informacionais e critérios relevantes [...]” (HJØRLAND, 2002, p.258) refletidos nas comunidades discursivas e que podem ser mapeados mediante análise.

Enquanto metodologia para a Ciência da Informação, a Análise de Domínio foi introduzida pelos trabalhos de Hjørland e Albrechtsen (1995; 1998), e abordagens específicas foram estabelecidas por Hjørland (2002). O autor propõe:

C11 - dados abertos, instituições, informação, web e modelo	10
C12 - modelo, edificação, dado, digital e cultural	8
C13 - universidade, material, análise, uso e histórico	5
C15 - conhecimento, comunidade, gerenciamento, social e estudo	5

Fonte: Autores (2023).

As categorias C8 e C14 correspondem aos artigos considerados *outliers*, não sendo estes incorporados a análise e, dessa forma, sendo 13 o número de categorias consideradas significativas.

A compreensão dos motivos que levam esses documentos a não se encaixarem em nenhuma dessas categorias irá depender da análise posterior do pesquisador dos documentos completos, em que será possível identificar se esses documentos realmente são pertinentes para o domínio estudado e o que os torna únicos em relação aos demais, investigando a sua potencial inclusão em alguma das categorias existentes.

Para além da análise dos destaques na nuvem de palavras, a compreensão aprofundada de cada categoria gerada, visando compreender a sua relevância e pertinência em relação ao domínio estudado, também será impactada pela análise do pesquisador, a ser realizada por meio de um estudo aprofundado dos textos completos. Com a realização dessa análise, pode existir a necessidade de excluir ou agrupar categorias.

Isso dependerá, sobretudo, do propósito do pesquisador com relação ao domínio estudado, ficando a cargo dele criar critérios que garantam o padrão de escolha e, por conseguinte, o rigor científico.

Como os autores já possuíam certa familiaridade com o domínio estudado, foi possível entender o porquê da existência de certas categorias, confirmar hipóteses pré-existentes. Nesse sentido, ressalta-se o impacto da familiaridade do pesquisador com o domínio estudado, já que essa familiaridade promove maior facilidade em relação a esse processo posterior de análise das categorias, viabilizado a criação de inferências a respeito dos documentos completos nelas agrupados apenas com a visualização dos destaques nas nuvens de palavras de cada documento.

Com esse conhecimento apriorístico, foi possível confirmar que grande parte dos artigos se encaixam nas categorias estabelecidas pelo algoritmo, como no exemplo do Quadro 2, que apresenta o título e as palavras-chave de cada um dos 16 artigos incluídos na categoria C10.

Quadro 2. Artigos categorizados na C10

ID	TÍTULO	PALAVRAS-CHAVE
53	Mixed reality experiences in museums: Exploring the impact of functional elements of the devices on visitors' immersive experiences and post-experience behaviours	Mixed reality; Cultural heritage museum; Immersive devices; Functional elements; immersive experiences; Post-experience behaviours
32	(Re)presenting heritage: Laser scanning and 3D visualisations for cultural resilience and community engagement	Built heritage visualisation; community engagement; laser scanning; local heritage; resilience
94	Leveraging collective intelligence: from univocal to multivocal representation of cultural heritage	Crowdsourcing; Collective intelligence; User-generated content; Cultural heritage crowdsourcing; Foucault's power-knowledge construct; Representation of cultural expressions
95	Moving from experiment to reality: choices for cultural heritage institutions and their government	Não consta
23	Oral Memory Construction for Ancient Villages in Digital Age: A Case Study of Shiquan Miao Village in Chongqing	Oral memory; Ancient villages; Memory construction; Shiquan Miao Village
78	Intangible Cultural Heritage Presentation and Preservation - Challenges and Opportunities for Museum Specialists	Intangible Cultural Heritage (ICH); ICH Digitisation; Safeguarding; Museum Specialists' Competencies
13	Towards unified retrieval system for GLAM institutions in India: Designing a prototype for biblio-cultural information space	Library discovery; Single-window search interface; Cultural heritage resources; Digital archive; ArchivesSpace; Koha; OAI/PMH; VuFind
28	Stakeholders' Attitudes Towards the Management and Preservation of Digital Cultural Heritage Resources in Ghana	Não consta
45	Integration of Digital Cultural Heritage Resources in China: Understanding Public Expectations	Digital culture heritage integration; Cultural institutions; Public expectation; Digital culture service; China

ID	TÍTULO	PALAVRAS-CHAVE
54	HerCulB: content-based information extraction and retrieval for cultural heritage of the Balkans	Information extraction; Content-based search; Natural language processing; Intangible cultural heritage
26	An Ontology for Architectural Heritage: Historical Figures and Organizations	Ontology; Semantic Modelling; Bulgarian Revival Architecture; Historical Figures; Cultural Heritage.
44	CRMCR - a CIDOC-CRM extension for supporting semantic interoperability in the conservation and restoration domain	Semantic Web; Cultural Heritage; Ontology
12	Integration of models for linked data in cultural heritage and contributions to the FAIR principles	Cultural Heritage; Linked Open Data; Data Integration; Semantic Web; FAIR Principles
37	A machine learning framework for enhancing digital experiences in cultural heritage	Cultural heritage; Digital heritage; Deep learning; Machine learning
9	3D Data Acquisition and Modelling of Complex Heritage Buildings	3D data acquisition; Semantic modelling; Digital heritage Optimized data capturing; Heritage documentation
33	A Further Step Toward Elaborating Guidelines for Three-dimensional Scanning of Cultural Heritage in Korea: Ongoing Efforts and Challenges	Cultural heritage; Guidelines; 3D scanning, Architecture; Archaeology; Cultural heritage policy; Korean case studies

Fonte: Autores (2023).

Verifica-se que a C10 reúne de fato artigos que tratam principalmente de temáticas no âmbito digital, incluindo a análise e a discussão de serviços e produtos, bem como ferramentas relacionadas ao patrimônio cultural, como ontologias, modelagem de dados e informações, web e *linked data*, princípios de qualidade e interoperabilidade, entre outros.

Além da confirmação de hipóteses, os resultados ainda auxiliam na descoberta de novos padrões de assuntos, ajudando o pesquisador a elaborar novas hipóteses e inferências por meio da visualização das tendências e conexões que se revelaram a partir da recorrência dos termos.

Dessa forma, os resultados auxiliam na percepção das palavras ou termos mais usados para representar determinados assuntos. O PLN justamente permite esse mapeamento contribuindo com mais de uma abordagem de AD já proposta.

Para a construção de vocabulários específicos de domínio como os tesouros, técnicas de PLN e ML tem maior visibilidade e aplicação (HJØRLAND, 2002). Contudo, considera-se que a contribuição do PLN é mais ampla e se encaixa na *Abordagem 11 - Cognição científica, conhecimento especializado e inteligência artificial*, proposta por Hjørland (2002), pois permitir criar modelos mentais de um domínio ou métodos para obter conhecimento para a produção de sistemas especializados, o que inclui mapear um domínio por meio da produção científica de um dado assunto. Destaca-se aqui a perspectiva do autor em relação ao potencial das formas de análise de domínio com origem na Ciência da Computação.

Nesse sentido, verifica-se que as potencialidade da aplicação de técnicas de PLN e ML no processo de categorização temática de artigos científicos, podem ser combinadas principalmente com a *Abordagem 9 - Estudos terminológicos, linguagens para propósitos específicos (LSP), semântica de bases de dados e estudos de discurso*. Uma vez que seu enfoque é na forma como se “[...] organizam palavras, textos e enunciados em um domínio segundo critérios semânticos e pragmáticos.” (HJØRLAND, 2002, p. 451), a combinação desses estudos validaria a categorização de forma contextual.

É possível citar como exemplo da importância da combinação de abordagens, o resultado da categoria C1 gerada no estudo em questão. A palavra com maior peso nessa categoria foi “*paper*”, seguida de “*area*”, “*study*”, “*science*”. “*material*”. Apenas com os resultados obtidos pela aplicação de PLN, o contexto podem ser inconclusivo, mesmo considerando os demais termos de destaque. Ao verificar o título, as palavras-chave e o resumo dos artigos dessa categoria, foi possível determinar que a palavra “*paper*” não se trata da sua tradução para “artigo científico”, e sim da tradução literal para “papel”, na medida que os textos abordam a temática de conservação e restauração de artefatos de papel.

Com isso, ainda pode-se citar a combinação do PLN aplicado a artigos científicos com a *Abordagem - 5 Estudos bibliométricos* cuja proposta é organizar padrões sociológicos de reconhecimento explícito entre documentos individuais permitindo a visualização estatística desses resultados. Ou ainda com a *Abordagem 10 - Estruturas e instituições da comunicação científica* e a *Abordagem - 6 Estudos históricos* que proporcionam aos resultados da categorização automatizada a validação com demais padrões do contexto de análise.

Portanto, reafirma-se que é a triangulação de métodos e resultados que garante a validação do estudo, bem como contribui com o próprio PLN na replicação das análises, permitindo, desse modo o mapeamento do domínio.

5 Conclusão

Enquanto forma de perceber o mundo ao redor, a categorização tem função no modo de identificar a dinâmica em que algo opera.

Ao aplicar as técnicas de PLN e ML antes mesmo de uma análise manual prévia do *corpus*, é possível pré-estabelecer padrões que podem facilitar a sua compreensão e discussão ao fornecer uma visão geral dos principais aspectos que o compõem.

Por outro lado, uma maior compreensão tanto do domínio como do funcionamento do algoritmo permite refinar os resultados por meio da parametrização, levando a criação de categorias mais significativas. Isso significa que a familiaridade com o tema, isto é, com suas relações conceituais, e com a técnica permite que a análise seja feita de maneira mais acurada, já que o pesquisador é capaz de entender os motivos de determinados agrupamentos e reconhecer esses padrões, nomear e descrever cada uma das categorias para, assim, realizar inferências a respeito delas.

Uma vez que a Análise de Domínio é um procedimento que permite e pressupõe a triangulação, isto é, o uso de mais métodos como forma de abordagem, unindo teoria e prática, o uso das técnicas de PLN e ML se tornam mais viável e com resultados mais significativos ao ser complementado por demais modos de análise.

Os resultados obtidos no presente estudo, de caráter inicial, exploratório e com resultados ilustrativos, permitem concluir que existe um potencial de contribuição das técnicas de ML e PLN para instrumentalizar e operacionalizar pesquisas baseadas em Análise de Domínio. Para que os procedimentos e limites dessa contribuição possam ser delimitados, tornam-se pertinentes e necessários não só estudos práticos mais aprofundados sobre as técnicas, mas sobretudo acerca das comunidades discursivas e do contexto empregado que influenciarão as técnicas, pois um domínio específico possui suas características que determinarão o resultado da escolha das categorias, assim como das relações conceituais que delas podem se originar.

Diante disso, propõe-se para trabalhos futuros a realização de um estudo mais detalhado, que forneça maiores subsídios para delinear o potencial do PLN e ML como uma das abordagens próprias de AD.

Notas

[1] Categorization as a way to systematize and organize information and knowledge is a guideline not only to standardize information systems, databases and libraries, but to facilitate the development of scientific production in a field with domain analysis, covering metric studies, as well as user's information retrieval (FREITAS; BUFREM; BREDA, 2016, p. 6).

Referências

- ARANGO, G.; PRIETO-DIAZ, R. Domain Analysis: Acquisition of reusable information for software construction. **IEEE Comp. Society Press**, 1989.
- BARITÉ, M. G. **Guía metodológica para el acceso, el análisis y la organización temática de documentos jurídicos**. Universidad de la República, 1999.
- BARITÉ, M. **Diccionario de Organización del Conocimiento**: clasificación, indización, terminología. 5. ed. Montevideo: PRODIC, 2013.
- CONEGLIAN, C. S. **Recuperação da Informação com abordagem semântica utilizando Linguagem Natural: a Inteligência Artificial na Ciência da Informação**. 2020. 194 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, Marília, 2020. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/193051/oneglian_cs_dr_mar.pdf?sequence=3&isAllowed=y. Acesso em: 23 mar. 2023.
- DIAS, C. da C. A análise de domínio, as comunidades discursivas, a garantia de literatura e outras garantias. **Informação & Sociedade**: Estudos, João Pessoa, v. 25, n. 2, p. 7-17, maio/ago. 2015. Disponível em: <https://brapci.inf.br/index.php/res/v/92174>. Acesso em: 23 mar. 2023.
- FREITAS, J. L.; BUFREM, L. S.; BREDA, S. M. Methodological choices for research in Information Science: Contributions to domain analysis. **Transinformação**, Campinas, v. 28, n. 1, p. 05-13, jan./abr. 2016. Disponível em: <https://www.scielo.br/j/tinf/a/PBGJwdSD5phK5xywng3jQq/abstract/?lang=en>. Acesso em: 23 mar. 2023.
- GUIMARÃES, J. A. C.; TOGNOLI, N. B. Provenance as a domain analysis approach in archival knowledge organization. **KO KNOWLEDGE ORGANIZATION**, [S.l.], v. 42, n. 8, p. 562-569, 2015. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2015-8-562.pdf>. Acesso em: 23 mar. 2023.
- GUIMARÃES, J. A.; PINHO, F. A.; FERREIRA, G. M. Relações teóricas da organização do conhecimento com as abordagens de catalogação de assunto, indexação e análise documental: uma análise de domínio da revista Scire (1995-2010). // **Scire**, [S.l.], v. 18, n. 2, p. 31-41, jul./dic. 2012. Disponível em: <https://repositorio.unesp.br/bitstream/handle/11449/114698/ISSN11353716-2012-18-02-33-43.pdf?sequence=1&isAllowed=y>. Acesso em: 11 jun. 2022.
- HERNANDEZ, B. F. **Análise dos estudos retóricos de gênero como abordagem para a análise de domínio**. 2020. Dissertação (Mestrado em Ciência da Informação) - Centro de Ciências da Educação, Universidade Federal de Santa Catarina, Florianópolis, 2020. Disponível em: <https://repositorio.ufsc.br/handle/123456789/216643>. Acesso em: 11 jun. 2022.
- HJØRLAND, B. Domain analysis in information science: eleven approaches-traditional as well as innovative. **Journal of Documentation**, [S.l.], v. 58, n. 4, p. 422-462, 2002.
- HJØRLAND, B. Domain analysis: A socio-cognitive orientation for information science research. **Bulletin of the**

American Society for Information Science and Technology, [S.l.], v. 30, n. 3, p. 17-21, 2004.

HJØRLAND, B. The socio-cognitive theory of users situated in specific contexts and domains. *In*: FISHER, K.; ERDELEZ, S.; MCKECHNIE, L.E.F. (ed.) **Theories of information behavior**. Medford, N.J: Information Today, 2005, p. 339-343.

HJØRLAND, B. Domain analysis. **KO Knowledge Organization**, v. 44, n. 6, p. 436-464, 2017. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2017-6-436.pdf>. Acesso em: 25 maio 2023.

HJØRLAND, B.; ALBRECHTSEN, H. An Analysis of Some Trends in Classification Research. **Knowledge Organization**, [S.l.], v. 26, p. 131-39, 1998.

HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. **Journal of the American society for information science**, [S.l.], v. 46, n. 6, p. 400-425, 1995. Disponível em: [https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/\(SICI\)1097-4571\(199507\)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y](https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1097-4571(199507)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y). Acesso em: 23 mar. 2023.

JACOB, E. K. Classification and Categorization: A Difference that Makes a Difference. **LIBRARY TRENDS**, [S.l.], v. 52, n. 3, p. 515-540, 2004. Disponível em: <https://www.ideals.illinois.edu/items/1778/bitstreams/6309/object?dl=1>. Acesso em: 23 mar. 2023.

MARTINS, G. K. A representação do conhecimento em uma perspectiva fenomenológica. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Bahia. **Anais [...]**. Bahia: 2016.

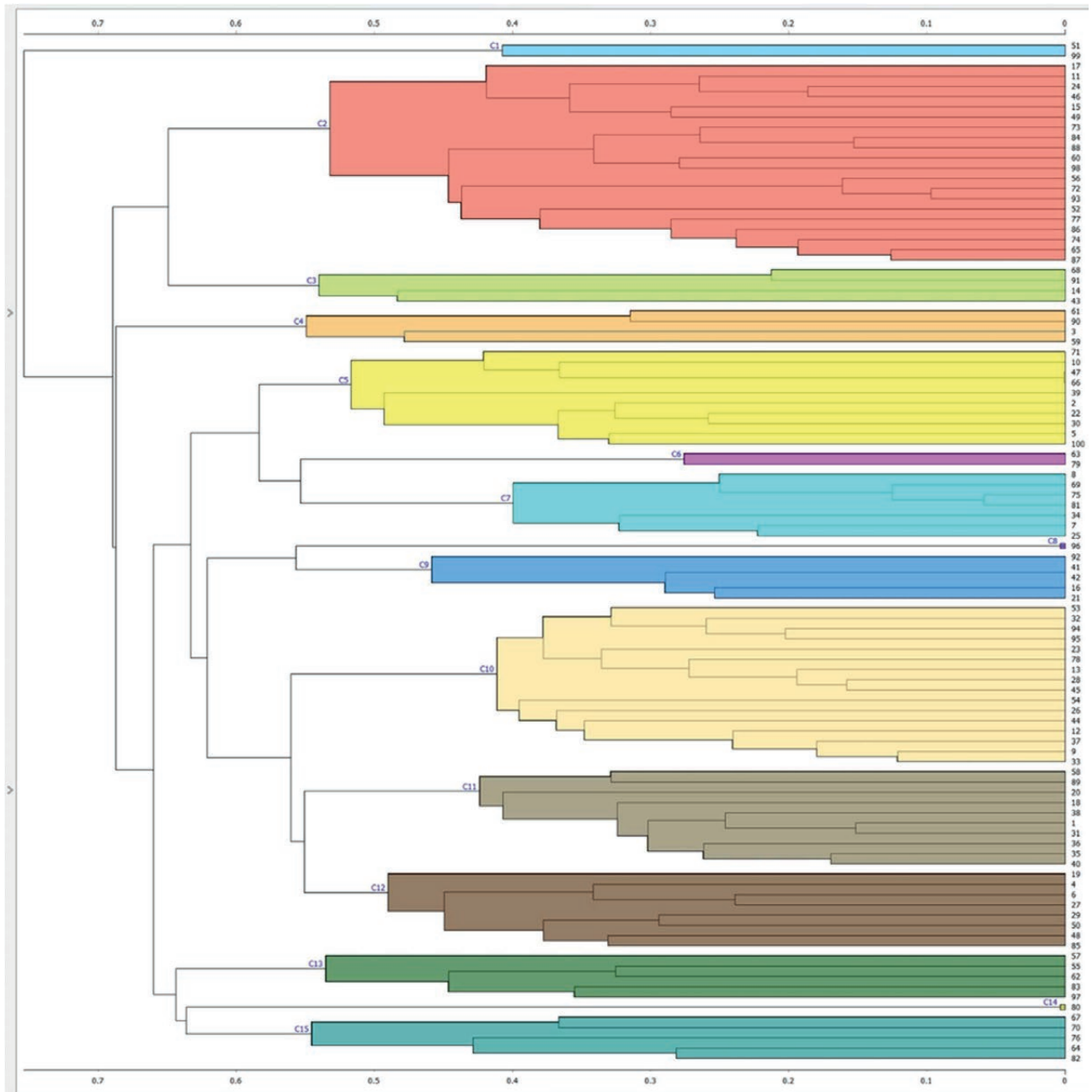
MOREIRA, W. **Sistemas de organização do conhecimento: aspectos teóricos, conceituais e metodológicos**. 2018. 164 f. Tese (livre-docência) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2018. Disponível em: <http://hdl.handle.net/11449/190878>. Acesso em: 23 mar. 2023.

NEIGHBORS, J. M. **Software construction using components**. University of California, Irvine, 1980.

SMIRAGLIA, R. P. Domain analysis of domain analysis for knowledge organization: Observations on an emergent methodological cluster. **KO Knowledge Organization**, [S.l.], v. 42, n. 8, p. 602-614, 2015. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2015-8-602.pdf>. Acesso em: 23 mar. 2023.

THELLEFSEN, T. L.; THELLEFSEN, M. M. Pragmatic semiotics and knowledge organization. **KO Knowledge Organization**, [S.l.], v. 31, n. 3, p. 177-187, 2004. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2004-3-177.pdf>. Acesso em: 23 mar. 2023.

Apêndice A – Lista de categorias geradas



Fonte: Autores (2023).

A presença *online* dos autores mais produtivos da ISKO Brasil: indicadores de atividade, visibilidade e impacto nas plataformas sociais

The online presence of ISKO Brasil's most productive authors: indicators of activity, visibility and impact on social platforms

Francielle Franco dos Santos (1), Maria Cláudia Cabrini Grácio (2), Caliel Cardoso de Oliveira (3)

(1) Universidade Estadual Júlio Mesquita Filho, francielle.franco@unesp.br, (2) cabrini.gracio@unesp.br, (3) Universidade Federal do Rio Grande do Sul, caliel.co@gmail.com

Resumo

Objetivo: caracteriza a presença *online* dos pesquisadores mais produtivos da ISKO-Brasil através de indicadores de atividade, visibilidade e impacto nas plataformas sociais. **Metodologia:** estudo quali-quantitativo que utiliza procedimentos altmétricos e análises estatísticas para caracterizar a presença *online* dos pesquisadores mais produtivos da ISKO-Brasil. **Resultados:** identifica 13 pesquisadores como os que mais publicaram em todas as edições do evento, destes, apenas 7 tinham perfil simultaneamente em todas as plataformas. Através da análise dos perfis nas plataformas sociais, por meio de 19 indicadores, identificou-se que as dimensões de atividade e impacto apresentaram maior número de indicadores, ampliando assim as possibilidades de análise. **Conclusão:** a presença *online* dos pesquisadores mais produtivos da ISKO Brasil é relevante, porque além de manter perfis ativos, fornecem indicadores capazes de caracterizar as dimensões propostas. Todos os pesquisadores possuem perfis ativos em pelo menos 4 das 7 plataformas analisadas. Além disso, 7 pesquisadores têm perfis em todas as plataformas, portanto denota-se motivação para que os pesquisadores mantenham perfis ativos em diferentes plataformas. Sugere-se adicionar redes de uso não acadêmico, como o Twitter, além de análises qualitativas para compreender as motivações dos pesquisadores para a manutenção de perfis em plataformas sociais.

Palavras-chave: Presença *online*. Plataformas sociais. Altmétria. ISKO Brasil.

Abstract

Objective: characterizes the online presence of ISKO-Brasil's most productive researchers through indicators of activity, visibility and impact on social platforms. **Methodology:** qualitative and quantitative study that uses altmetric procedures and statistical analysis to characterize the online presence of the most productive researchers at ISKO-Brasil. **Results:** identifies 13 researchers as those who published the most in all editions of the event, of which only 7 had a profile simultaneously on all platforms. Through the analysis of profiles on social platforms, using 19 indicators, it was identified that the dimensions of activity and impact had a greater number of indicators, thus expanding the possibilities of analysis. **Conclusion:** the online presence of ISKO Brasil's most productive researchers is relevant, because in addition to maintaining active profiles, they provide indicators capable of characterizing the proposed dimensions. All researchers have active profiles on at least 4 of the 7 platforms analyzed. In addition, 7 researchers have profiles on all platforms, so there is motivation for researchers to maintain active profiles on different platforms. It is suggested to add networks for non-academic use, such as Twitter, in addition to qualitative analyzes to understand researchers' motivations for maintaining profiles on social platforms.

Keywords: Online presence. Social platforms. Altmetric. ISKO Brazil.

Resumen

Objetivo: caracterizar la presencia en línea de los investigadores más productivos de ISKO-Brasil a través de indicadores de actividad, visibilidad e impacto en las plataformas sociales. **Metodología:** estudio cualitativo y cuantitativo que utiliza procedimientos altmétricos y análisis estadístico para caracterizar la presencia en línea de los investigadores más productivos de ISKO-Brasil. **Resultados:** identifica a 13 investigadores como los que más publicaron en todas las ediciones del evento, de los cuales solo 7 tenían perfil simultáneamente en todas las plataformas. A través del análisis de perfiles en plataformas sociales, utilizando 19 indicadores, se identificó que las dimensiones de actividad e impacto tenían un mayor número de indicadores, ampliando así las posibilidades de análisis. **Conclusión:** la presencia en línea de los investigadores más productivos de ISKO Brasil es relevante, porque además de mantener perfiles activos, proporcionan indicadores capaces de caracterizar las dimensiones propuestas. Todos los investigadores tienen perfiles activos en al menos 4 de las 7 plataformas analizadas. Además, 7 investigadores tienen perfiles en todas las plataformas, por lo que existe motivación para que los investigadores mantengan perfiles activos en diferentes plataformas. Se sugiere agregar redes para uso no académico, como Twitter, además de análisis cualitativos para comprender las motivaciones de los investigadores para mantener perfiles en las plataformas sociales.

Palabras clave: Presencia en línea. Plataformas sociales. Altmétria. ISKO Brasil.

1 Introdução

A Organização do Conhecimento (OC), enquanto disciplina ou subárea da CI, nasceu e desenvolveu-se ao longo da história da Ciência da Informação, tendo seu papel consolidado após a Segunda Guerra Mundial com o aumento de informações que necessitavam de organização. Nesse sentido, surge e desenvolve-se uma comunidade científica para abrigar os pesquisadores da área, a ISKO.

A revolução informacional do pós-guerra não trouxe apenas mais volume de informação, junto com ela vieram as constantes transformações tecnológicas, sobretudo após o surgimento e popularização da internet, na década de 1990. Desse modo a ciência é atravessada por esse contexto e as atividades de pesquisa, que tem como premissa básica a comunicação e a visibilidade acadêmica, passam a considerar os ambientes virtuais, sobretudo, as redes sociais acadêmicas.

Portanto, partindo da premissa de que a presença *online* de pesquisadores em diferentes plataformas é cada vez mais relevante e necessária, indaga-se como se caracteriza a presença *online* dos pesquisadores mais produtivos da ISKO Brasil através dos indicadores de atividade, visibilidade e impacto nas plataformas da web social?

2 Referencial Teórico

2.1 A área de organização do conhecimento e a ISKO Brasil

A área conhecida como Organização do Conhecimento (OC) ou Organização e Representação do Conhecimento (ORC) é uma disciplina correlata à Ciência da Informação (CI) sendo vista ora como subárea da CI, ora como disciplina própria, mas ainda próxima ao arcabouço oferecido pelos acadêmicos da Informação como um todo. Os conceitos fundamentais da OC foram propostos pela teórica alemã Ingetraut Dahlberg que ainda na década de 1970 formalizou importantes fundamentos da área, especialmente dentro das proposições da chamada Teoria do Conceito. É Dahlberg, por exemplo, que define o Conceito como sendo a “compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico” (DAHLBERG, 1978; MELO; BRASCHER, 2014). É também Dahlberg que esquematiza de maneira formal as características que podem existir nos conceitos, bem como as espécies de relacionamentos que podem se formar entre eles. Outros teóricos, como Birger Hjørland (2008), oferecem considerações significativas sobre a relação entre a OC e a CI, apontando que muito do arcabouço teórico desenvolvido ao longo do tempo na CI são criação de diferentes perspectivas dentro da OC,

contribuindo para uma especialização da área em diferentes pontos de vista e perspectivas de estudo e.

Cabem ainda algumas considerações sobre a história da *International Society for Knowledge Organization* (ISKO). Dahlberg (2013) narra que a fundação da ISKO, ocorrida em 1989, se deu a partir de divergências de alguns teóricos que faziam parte da *Society for Classification*, mas desejavam dedicar-se ao estudo do Conceito e do Conhecimento em detrimento das pesquisas matemáticas e estatísticas realizadas naquela sociedade científica. Dessa forma, a ISKO é fundada com o intuito de explorar as contribuições teóricas oferecidas por Dahlberg e seus colegas: o Conceito como uma unidade de conhecimento e a estruturação dos sistemas de Conhecimento. Dessa forma, a ISKO se consolidou ao longo das décadas como a principal sociedade científica para o estudo e avanço da Organização do Conhecimento enquanto disciplina por si só, relacionando-se e integrando diferentes áreas do conhecimento. Dahlberg conclui que, enquanto organização científica especializada, a ISKO tem um papel fundamental em “decidir” quanto ao status da OC, quer seja enquanto ciência por si só, quer seja como subdisciplina ou conjunto de teorias e práticas em outras áreas (Ibid, 2013).

Nesse sentido, partindo da definição de que a ISKO é uma comunidade científica, composta por pesquisadores reunidos ao redor de interesses comuns, presume-se que, além de interesses, estes sujeitos possam compartilhar ainda comportamentos similares. Portanto, cabem estudos de diferentes tipologias a fim de caracterizar esta comunidade, sobretudo em meios de socialização de produtos científicos mais recentes, como as plataformas da web social.

2.2 Reputação e Visibilidade Acadêmica

No contexto acadêmico, visibilidade é a capacidade de influenciar determinado público, ou ainda, de responder às demandas informacionais de uma comunidade científica. Desse modo, pode-se dizer que a visibilidade é intrínseca e inerente ao processo da comunicação científica. Os caminhos da ciência podem ser visualizados, objetivamente, por meio de indicadores de visibilidade dos diferentes atores do processo científico (BARCELOS; MARICATO, 2021; OLIVEIRA; GRÁCIO, 2013).

Nesse sentido, a visibilidade pode ser alcançada pela simples presença de um pesquisador em determinada rede. Envolve três elementos: “o ‘eu’ e o ‘outro’ e a relação entre ambos” (RECUERO, 2009, p. 109). É matéria-prima para o desenvolvimento de valores como a reputação. Assim, a reputação é construída a partir da percepção que o outro têm de determinadas ações ou comportamentos (RECUERO, 2009). Desse modo, visibilidade e reputação são relacionados na medida em que uma na medida em que uma precede a outra.

As informações sobre quem somos e o que pensamos são engendradas em diferentes sistemas de classificação que vão definir os diferentes tipos de reputação. Desse modo, pode-se inferir que existem diferentes sistemas de reputação que têm por objetivo mensurar e classificar pessoas, grupos, objetos ou ideias. A principal função desses mecanismos é alimentar sistemas de recomendação, como os serviços do Google ou da Amazon, a partir da avaliação de usuários (MASUM; ZHANG, 2004). Nesse sentido, nem sempre as indicações do que é melhor é declarado de maneira explícita, muitas vezes o uso individual de cada pessoa é fator determinante para a reputação.

No ambiente acadêmico, a citação pode ser considerada um ativo na construção de reputação, contudo alguns fatores podem interferir na probabilidade de um autor ser citado, afetando a percepção da comunidade. O primeiro deles é o “efeito Mateus”, em que os pesquisadores altamente citados seguirão sendo cada vez mais mencionados, isso porque a contribuição de trabalhos desenvolvidos por pesquisadores de renome é tida como certa (WANG; BARABÁSI, 2021). Outra questão é o vínculo entre pesquisadores e instituições: aqueles com visibilidade têm mais chances de integrarem equipes em instituições de impacto, bem como as instituições que têm pesquisadores renomados têm vantagens na disputa de financiamentos e recursos públicos. Portanto, estes são sistemas que se retroalimentam, assim, reputação e visibilidade não são apenas questões individuais; o reconhecimento científico tem impacto coletivo (WANG; BARABÁSI, 2021; HERMAN; NICHOLAS, 2019).

Entretanto, o fazer científico vai além da publicação de artigos em periódicos, compreendendo um conjunto numeroso de tarefas até que as descobertas sejam comunicadas (ORDUÑA-MALEA; MARTÍN-MARTÍN; LÓPEZ-CÓZAR, 2016). Dado o exposto, entende-se então, que no sistema de reputação acadêmica, a publicação não é garantia de impacto, visto que quem o estabelece é a comunidade acadêmica ao adotar a publicação na fundamentação das suas pesquisas (WANG; BARABÁSI, 2021). Portanto, se faz necessário compreender as práticas científicas para além da publicação de artigos, visto que todas as atividades de pesquisa têm em si incorporadas a busca pela reputação (HERMAN; NICHOLAS, 2019).

2.3 Atividades de pesquisa, suas dimensões e mensurações

Durante a trajetória acadêmica de um pesquisador as diferentes funções e tarefas desempenhadas podem ser analisadas em distintas dimensões, sobretudo no contexto digital no qual o monitoramento é premissa básica. Para Orduña-Malea, Martín-Martín e López-Cózar (2016), essas dimensões das funções desempenhadas por um acadêmico podem ser definidas em: a) Atividade: capacidade de construir

conhecimento; b) Visibilidade: manifestação evidente da atividade de um autor; c) Impacto: repercussão que a produção de autor obtém, influenciando em âmbito acadêmico, profissional ou social.

A literatura apresenta ainda outras classificações para o fazer da ciência, como a apresentada por Herman e Nicholas (2019). A partir da lista de atividades acadêmicas proposta por Boyer (1990), Herman e Nicholas (2019) enumeram mais de 30 atividades desenvolvidas por pesquisadores ao longo das suas carreiras e como cada uma delas pode contribuir na construção da reputação. A lista de atividades foi distribuída nas seguintes categorias: a) investigação; b) integração; c) aplicação; d) ensino; e) cocriação. (HERMAN; NICHOLAS, 2019).

Para este estudo, interessa-nos particularmente a categoria de cocriação, sobretudo no contexto da web social. Herman e Nicholas (2019) adicionaram a cocriação na lista desenvolvida na década de 1990, porque naquela época não era possível vislumbrar o impacto que as tecnologias de comunicação e informação, principalmente as plataformas sociais, teriam na difusão do conhecimento. Adicionalmente, interessa aqui a dimensão de impacto proposta por Orduña-Malea, Martín-Martín e López-Cózar (2016). Entretanto, entende-se que todas as dimensões são complementares visto que para causar impacto um pesquisador precisa desenvolver visibilidade.

Muitas, senão todas, as práticas científicas têm por finalidade a manutenção da reputação. O papel da avaliação é fundamental neste processo porque ela é o instrumento através do qual se pode refutar ou confirmar reputações (MASUM; ZHANG, 2004). Herman e Nicholas (2019) defendem que a reputação acadêmica está mais ligada às práticas de divulgação científica do que necessariamente às atividades de pesquisa. Assim, compreende-se que existe um mecanismo capaz de quantificar o impacto da atividade para cada prática de difusão do conhecimento (MASUM; ZHANG, 2004).

As métricas tradicionais do impacto acadêmico, como a quantidade de artigos publicados em periódicos e a contagem das suas citações, ainda são as bases do sistema de reconhecimento acadêmico (HERMAN; NICHOLAS, 2019). Assim, as medidas de impacto tradicionais, baseadas em citação, podem funcionar como uma espécie de bússola, que pode ajudar os pesquisadores a navegarem na próxima fase histórica da ciência: a do acesso aberto (WILLINSKY, 2010). Apesar das evidências de fragilidades apontadas na literatura, as métricas tradicionais ainda são os indicadores mais eficazes para mensurar os impactos da ciência (HERMAN; NICHOLAS, 2019). Por outro lado, não se pode ignorar o atual contexto informacional e as possibilidades descortinadas pelas plataformas da web social.

2.4 Comunicação e divulgação científica na web social

Para além de gerar e aumentar o conhecimento em benefício da sociedade, é de interesse dos cientistas trabalhar em prol do aumento de sua reputação (HERMAN; NICHOLAS, 2019). Notadamente, a divulgação de resultados de pesquisa na web, tem grande potencial na construção de reputação acadêmica e o acesso aberto é um dos fatores que permite o alcance global. Tornar as pesquisas acessíveis faz com que elas possam ser encontradas e visualizadas por meio de mecanismos de buscas, como o Google Scholar (WILLINSKY, 2010).

No fluxo tradicional da comunicação científica a divulgação dos resultados só era possível no fim da investigação. O uso massivo de plataformas da web social por parte dos pesquisadores permitiu que diferentes etapas da pesquisa pudessem ser compartilhadas de maneira aberta. Desse modo, estes novos canais de comunicação e divulgação científica podem ser considerados complementares, quando colocados em comparação com os meios tradicionais (HERMAN; NICHOLAS, 2019). Contudo, quem controla o mercado de publicações de alto impacto são as grandes editoras científicas, causando interferências no sistema de reputação acadêmica e apreensão nos autores que desejam compartilhar etapas da pesquisa que ainda não passaram pelo crivo da revisão por pares (LÓPEZ-CÓZAR; MARTÍN-MARTÍN, 2020).

A explosão das redes sociais e o uso para a promoção de informações são os grandes fenômenos infocomunicacionais deste século. O uso destes canais pelos mais variados setores da sociedade interfere diretamente em suas reputações (MASUM; ZHANG, 2004). Questiona-se, então, o quanto a reputação acadêmica poderia ser ampliada se não ficasse restrita apenas ao ambiente acadêmico. As redes sociais ampliam o alcance, atingindo comunidades de interesse não especializadas que podem ser de grande valia para os pesquisadores e para a ciência (HERMAN; NICHOLAS, 2019).

2.5 Presença online de pesquisadores em plataformas sociais e os principais indicadores

O uso de sites de redes sociais por pesquisadores tem sido apontado como fonte de estresse ou como canal de baixa credibilidade (SUGIMOTO *et al.*, 2017). Entretanto, criar e manter um perfil em diferentes plataformas sociais, acadêmicas ou não, pode gerar visibilidade e, por consequência, fortalecer sua reputação (HERMAN; NICHOLAS, 2019; ARAÚJO, 2014). Nesse sentido, devem ser realizadas escolhas estratégicas sobre quais plataformas utilizar de acordo com os objetivos e públicos que deseja atingir.

A percepção e o uso das plataformas por parte dos pesquisadores variam de acordo com as áreas de pesquisa (MARTÍN-MARTÍN; ORDUÑA-MALEA;

LÓPEZ-CÓZAR, 2018a). Todavia, é mais comum que se faça uso delas no sentido de manter uma espécie de cartão de visitas e acumular capital simbólico (SUGIMOTO *et al.*, 2017). Não há consenso na literatura sobre a presença on-line de pesquisadores nas plataformas sociais, o que abre caminho para diferentes estudos que possam investigar esse tema.

Para Sugimoto *et al.* (2017), os pesquisadores usam as plataformas sociais para disseminação, consumo e comunicação de suas publicações. As principais diferenças entre as plataformas sociais são o público e os objetivos de compartilhamento do conteúdo. Nesse sentido, elas são independentes entre si, porém complementares (RECUERO, 2009). O *upload* de um artigo científico em seu perfil no ResearchGate configura um exemplo do uso complementar das redes sociais, ampliando as chances de recuperação por mecanismos de busca. Caso este pesquisador publique um *tweet* indicando o doi deste artigo, ele pode iniciar uma conversa em torno do tema com um público que eventualmente não usa o ResearchGate. Ademais, publicar um doi no Twitter viabiliza o rastreamento pelos agregadores altmétricos. Desse modo, o uso de diferentes plataformas pelos pesquisadores ajuda na construção da reputação acadêmica (RECUERO, 2009).

Dado o exposto, destaca-se que as diferentes plataformas sociais apresentam diversas possibilidades de uso, sempre tendo como premissa o compartilhamento de informações e a interatividade entre seus usuários. Orduña-Malea, Martín-Martín e López-Cózar (2016) reúnem 93 indicadores coletados em diferentes plataformas da web social, agrupando-os de acordo com a sua natureza, sendo elas: publicação; citação; uso; difusão, comentário ou discussão; avaliação; conectividade social; e indicadores combinados. Ainda em um exercício de redução, esses indicadores foram distribuídos de acordo com as dimensões de atividade, visibilidade e impacto. O Quadro 1, apresenta a lista de indicadores explorados neste estudo, de acordo com as plataformas selecionadas para coleta dos dados.

Investigações que têm indivíduos como objeto de estudo sempre enfrentam desafios, inclusive com as métricas mais tradicionais (TORRES-SALINAS; MILANÉS-GUISADO, 2014). A busca por nomes enfrenta obstáculos como a ambiguidade, erros ortográficos ou alteração por questões matrimoniais (WANG; BARABÁSI, 2021). Uma das formas de minimizar este problema foi o desenvolvimento de plataformas de identificadores únicos (IDs). Essas plataformas fornecem IDs através dos quais é possível acessar o percurso histórico de um pesquisador pelo agrupamento de sua produção de forma sistematizada. Destacam-se: ResearcherID e Scopus Author Profile.

As plataformas sociais de uso acadêmico, além de apresentar as produções, também estimulam a comunicação científica. Além disso, destaca-se que o compartilhamento de informações nessas redes também propicia a conectividade social entre os pares, aumentando a colaboração científica. Outro fator relevante são as oportunidades de criação de grupos de discussão com base em interesses comuns, dado o rompimento de barreiras, como a geográfica. Entre as plataformas mais relevantes destacam-se: Google Scholar Citations, ResearchGate e Academia.edu (UNESCO, 2015; ORDUÑA-MALEA; MARTÍN-MARTÍN; LÓPEZ- CÓZAR; 2016).

As plataformas mencionadas permitem a criação de perfis através dos quais é possível adicionar diferentes formatos de produção acadêmica: artigos, capítulos de livros, conjuntos de dados, apresentações, vídeos, entre outros. Esses conteúdos, ou a interação por eles estimulada, geram indicadores que podem ser considerados complementares às métricas tradicionais (HERMAN; NICHOLAS, 2019), porque capturam mais do que apenas aspectos do produtivismo acadêmico.

Os agregadores altmétricos foram criados para que todos os indicadores fornecidos pelas plataformas da web social pudessem ser organizados e para que as menções on-line das produções acadêmicas pudessem ser visualizadas. O Altmetric.com (ALT) e o Impactstory (IMP) são mencionados no Manifesto Altmétrico como um dos maiores provedores de pontuação altmétrica.

Algumas plataformas apresentam funcionalidades muito próximas e seus indicadores também apresentam similaridades (como o índice h, no GSC, RG, ACA, RID e SAID). Entretanto, pode-se justificar a criação e a manutenção de perfis em plataforma similares pelo fato de que cada um tem suas especificidades na mensuração de diferentes aspectos do impacto científico (MARTÍN-MARTÍN *et al.*, 2016), assim, faz sentido a presença multiplataforma. Além das limitações de cada rede, destaca-se a dependência gerada a partir de indicadores de uma única plataforma, considerando a volatilidade das fontes de dados (ORDUÑA-MALEA; LÓPEZ-CÓZAR, 2017). Portanto, para o desenvolvimento de investigações como a presente pesquisa, é necessário considerar múltiplas fontes para coleta de dados.

3 Metodologia

Esta investigação tem natureza quali-quantitativa, com caráter descritivo e utiliza-se de indicadores altmétricos e análises estatísticas para caracterizar a presença *online* de um conjunto de pesquisadores. O

corpus partiu da análise dos 260 trabalhos publicados nas 5 edições do Congresso Brasileiro em Organização do Conhecimento (ISKO-Brasil), entre os anos de 2011 e 2019. A base de dados Brapci foi utilizada para a recuperação dos artigos publicados. A extração das autorias de cada documento foi realizada de forma manual e resultou em 334 pesquisadores. A aplicação da lei do elitismo de Price nesse universo de pesquisadores resultou em 18 autores. Todavia, o *corpus* de estudo foi ajustado para os 13 autores, dada a grande dispersão de autores com a mesma quantidade de artigos a partir do 14º autor. A Tabela 1, no Apêndice A, apresenta a lista dos 13 autores mais produtivos em todas as edições da ISKO Brasil.

A lista de pesquisadores possibilitou o levantamento dos perfis e indicadores nas plataformas sociais (SANTOS, 2023). Foram selecionadas plataformas que permitissem a coleta dos indicadores de acordo com os grupos apresentados no quadro 1: ResearcherID (RID), Scopus Author Profile (SAID), Google Scholar Citations (GSC), ResearchGate (RG), Academia.edu (ACA), Altmetric Explorer (ALT) e ImpactStory (IMP).

A coleta ocorreu no dia 23 de março de 2023, de forma manual a partir da url dos perfis em cada plataforma. Todas as plataformas, com exceção do ResearchGate e do Academia.edu, possibilitam a busca por meio do identificador Orcid de cada pesquisador. Desse modo, foi realizado o levantamento dos identificadores na Orcid com os nomes de cada pesquisador para posterior busca dos perfis em cada plataforma.

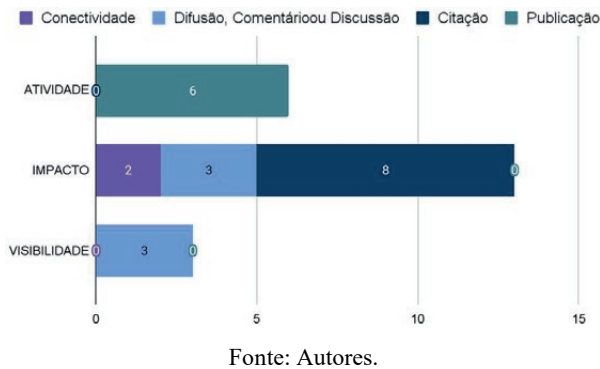
A localização e extração dos dados no Google Scholar Citations foi realizada pelo software Publish or Perish. No ResearchGate e Academia.edu, utilizando os nomes completos de cada pesquisador, ou variações como nome e sobrenome do meio, ou nome e apenas o último sobrenome, foram realizadas buscas por meio dos sistemas de busca disponibilizados nas plataformas.

4 Resultados

Foram coletados 19 indicadores nas 7 plataformas com o objetivo de caracterizar as dimensões de atividade, impacto e visibilidade (Figura 1).

As dimensões de atividade e impacto apresentam maior número de indicadores (6 e 13), portanto elas dispõem de maiores possibilidades de análise. A dimensão de atividade é mensurada através dos indicadores de publicação, coletados na GSC, RID, SAID, RG e ACA. A dimensão de impacto tem a maior diversidade de indicadores sendo eles conectividade, difusão, comentário ou discussão; e citação, coletados em todas as 11 plataformas.

Figura 1. Número de indicadores coletados por dimensões da pesquisa científica



Por fim, a visibilidade é a dimensão que oferece a menor variedade de indicadores, sendo representada pelo indicador de difusão, comentário ou discussão, coletado na ALT e na IMP.

Por meio dos indicadores (Tabela 1, no apêndice A, em anexo), observa-se que os autores mais produtivos da ISKO Brasil estão presentes de forma consistente nas plataformas (7 dos 13 pesquisadores estão presentes em todas as plataformas), ou seja, além de manterem perfis ativos, apresentam indicadores com potencial de análise. Os indicadores mais expressivos podem ser visualizados no GSC e os menos expressivos são encontrados nos agregadores altmétricos, ALT e IMP.

De acordo com as dimensões aqui propostas, a atividade de um pesquisador pode ser mensurada através de indicadores de publicação, citação e índice h. A tabela 2 apresenta a estatística descritiva destes indicadores.

Tabela 2. Estatísticas descritivas do indicador de índice h

Indicador	N	Média	Erro/Desvio	Mínimo	Máximo
h_RID	11	4,27	2,102	1	7
h_SAID	11	3,36	2,618	0	7
h_GSC	11	15,27	8,510	1	28
h_RG	11	7,18	3,945	0	13

Fonte: Autores.

Nota-se que 11 pesquisadores têm perfil simultaneamente nas plataformas RID, SAID, GSC e RG. Para esse conjunto de 11 pesquisadores, há diferença estatística significativa quanto ao índice h nessas 4 plataformas. Observa-se que a média do índice h no GSC é bastante superior ao calculado pelas outras redes, especialmente em relação ao RID e SAID. Além disso, RID e SAID apresentam valores próximos e RG um valor intermediário entre as duas extremidades.

O GSC tem foco em organizar e apresentar as publicações de um pesquisador e não fomenta a conectividade social. Destaca-se que o índice h fornecido pelo GSC é aceito por diversas instituições,

incluindo a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) e a Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp), como um parâmetro importante na avaliação de pesquisadores – o que reforça o foco da plataforma mais na produção e citação do que em conectividade e uso. Ademais, contribui para explicar o fato de ser uma plataforma em que todos os pesquisadores analisados mantêm um perfil ativo.

O conjunto de funcionalidades disponíveis aliado à ampla cobertura faz dessa plataforma a melhor ferramenta para a análise bibliométrica de algumas disciplinas, especialmente aquelas ligadas às áreas de humanidades e ciências sociais (MARTÍN-MARTÍN *et al.*, 2016). Entretanto, é preciso cautela na utilização do Google Scholar (fonte de informação do GSC) como fonte de análise para avaliação da pesquisa, conforme mencionado por Caregnato (2012).

Por outro lado, o Google Scholar tem grande capacidade de recuperação, devido ao seu poder de indexação. Por meio dessa plataforma, é possível encontrar documentos que ainda não estão relacionados em bases que precisam de atualização manual, como o Currículo Lattes (CAREGNATO, 2012). Martín-Martín, Orduña-Malea e López-Cózar (2018b) corroboram ao afirmar que o GSC tem capacidade de fornecer uma imagem bastante precisa de uma comunidade científica. Portanto, pode-se concluir que, neste recorte de pesquisa, a dimensão de atividade pode ser coberta e caracterizada através do índice h nas plataformas. Além disso, existem evidências estatísticas de correlação entre esses indicadores, ou seja, as tendências de citação e índice h são muito parecidas nas plataformas.

Referente a dimensão de Impacto, mensurada através de indicadores de publicação e citação, foram coletados dados nas plataformas GSC, RID, SAID, RG e ACA. As tabelas 3 e 4 apresentam os valores estatísticos destes indicadores.

Tabela 3. Estatísticas descritivas do indicador de publicação

Indicador	N	Média	Erro/Desvio	Mínimo	Máximo
Pub_RID	9	30,67	25,710	10	84
Pub_SAID	9	20,00	18,330	6	50
Pub_GSC	9	199,89	164,058	9	499
Pub_ACA	9	38,56	41,192	0	102

Fonte: Autores.

Quanto ao indicador de publicação, para esse conjunto de pesquisadores, há diferença estatística significativa entre as 4 plataformas. Observa-se que a média do indicador de publicação no GSC também é bastante superior ao calculado pelas outras 3 plataformas, especialmente em relação ao RID e SAID.

Tabela 4. *Estatísticas descritivas do indicador de citação*

Indicador	N	Média	Erro Desvio	Mínimo	Máximo
Pub_RID	9	30,67	25,710	10	84
Pub_SAID	9	20,00	18,330	6	50
Pub_GSC	9	199,89	164,058	9	499
Pub_ACA	9	38,56	41,192	0	102

Fonte: Autores.

O teste estatístico não-paramétrico de Friedman apontou ainda que há diferença estatística significativa entre os indicadores de citação disponibilizados por essas 4 plataformas, com as tendências de diferenças entre elas semelhantes à observada para no indicador de índice h disponibilizados pelas mesmas plataformas.

Quanto à dimensão de impacto, destaca-se que a análise aqui proposta ateu-se ao impacto acadêmico (número de citações), porque todas essas plataformas são de natureza acadêmica, destinadas a disseminar, otimizar e ampliar a comunicação científica (entre pares). Entretanto, os agregadores alométricos (IMP e ALT) capturam as menções *online* em outros canais, nos quais o público acadêmico não é maioria.

Conforme os dados apresentados no conjunto de indicadores coletados (SANTOS, 2023), 8 pesquisadores têm presença no ALT e em IMP simultaneamente. O teste estatístico de Wilcoxon demonstrou que há diferença estatística significativa entre os indicadores de menção dessas duas plataformas, uma vez que todos eles têm valor diferente de zero para menções na plataforma ALT e valor igual a zero para o mesmo indicador na IMP.

Analisando especificamente os agregadores alométricos, destaca-se que o Altmetric.com é o provedor alométrico que mais captura menções em redes sociais, blogs e portais de notícias (ORTEGA, 2020) e por essa razão é o mais adotado no desenvolvimento de análises alométricas. A comparação entre os dois agregadores evidencia a maior cobertura no Altmetric.com. Segundo Ortega (2020), o foco do IMP está na criação dos perfis de pesquisadores. Assim, mesmo que a IMP apresente presença dos pesquisadores, a ALT ainda apresenta dados mais relevantes. O fato de a manutenção dos documentos na IMP ser realizada de forma manual (por meio do upload ou da vinculação do Orcid) pode afetar a cobertura da plataforma.

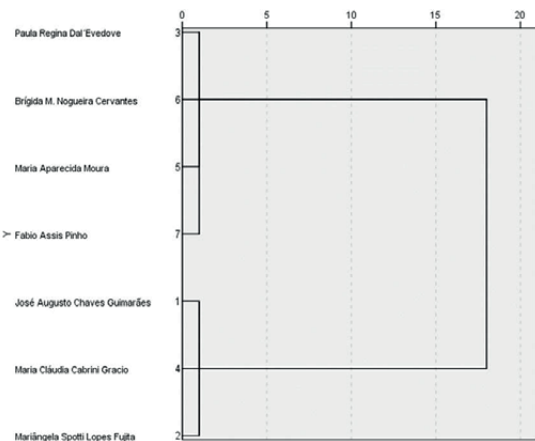
Sobre a conectividade social, dos 13 pesquisadores, 11 têm perfil na RG e na ACA. O teste estatístico aplicado evidenciou que não há diferença estatística significativa entre a quantidade de seguidores em ambas as plataformas. Para caracterizar coletar indicadores de conectividade social, dentro da dimensão de impacto, destaca-se a importância de analisar redes sociais não

acadêmicas, como o Twitter. Esta é a rede de uso pessoal em que mais é possível observar e rastrear o impacto acadêmico para além do público acadêmico (ORDUÑA-MALEA; LÓPEZ-CÓZAR, 2017), portanto, pode ser definida como a principal fonte de análises da disseminação e impacto social das publicações acadêmicas (ORTEGA, 2020).

Este cenário apresenta uma imagem que demonstra que, para além dos recursos de contagem de citação (foco do GSC), os pesquisadores têm interesse em dar visibilidade para suas publicações em plataformas com outros recursos (seguir, seguidores, adição de outros formatos de publicação). Através dos perfis na RID e SAID denota-se a penetração desses autores em bases de dados relevantes mundialmente, isso porque que ambas coletam publicações da Web of Science e da Scopus. A produção deste conjunto de pesquisadores está indexada por grandes bases, o que nos leva a entender que a presença desse grupo de autores nessas plataformas demonstra o alcance da sua produção para além das fronteiras da ciência brasileira.

Além das dimensões analisadas, destacam-se as similaridades entre os pesquisadores do universo analisado. Para que as comparações pudessem ser realizadas de maneira equivalente foram considerados aqueles pesquisadores que tivessem presença em todas as plataformas mencionadas. Desse modo, dos 13 pesquisadores relacionados na tabela 1, foram identificados 7 com presença simultânea em todas as plataformas. A Figura 2 demonstra visualmente as similaridades entre os estes 7 pesquisadores.

Figura 2. *Combinação de cluster de distância redimensionado*



Fonte: Autores.

O dendrograma apresenta o conjunto de pesquisadores quanto a presença nas plataformas, ou seja, os 4 pesquisadores no agrupamento superior têm presenças similares nas redes, bem como os 3 pesquisadores no agrupamento inferior apresentam similaridades entre si. Para além dos indicadores numéricos apresentados no estudo, outras hipóteses podem ser levantadas. Um dos

indícios para a presença massiva nas plataformas é a percepção, por parte dos pesquisadores, da importância da socialização e disseminação dos produtos científicos em diferentes canais. Essas práticas têm impacto positivo na avaliação dos programas de pós-graduação e nas instituições de filiação dos pesquisadores.

Nesse sentido, o agrupamento inferior tem similaridade mais coesa porque os três pesquisadores estão vinculados à mesma instituição, a UNESP. O Programa de Pós-Graduação em Ciência da Informação da Unesp foi avaliado com nota 6 na última avaliação da Capes (2017-2020), sendo considerado de excelência nacional e com forte incentivo à internacionalização. Além disso, estes 3 pesquisadores são bolsistas de produtividade da Capes, ou seja, apresentam grau de produtividade elevado e por consequência teriam maior motivação para socializar as publicações em diferentes canais, possivelmente como forma de uma prestação de contas pelo investimento público em suas carreiras.

Além da UNESP (nota 6), outras instituições com nota 5 (UFMG) ou 4 (UEL, UFRJ, UFPE, USP), na última avaliação da Capes (2017-2020), fazem parte da lista de instituição de afiliação dos pesquisadores, o que reforça a ideia de que estar vinculado à programas de pós-graduação bem avaliados pode ser motivação para a manutenção da atenção on-line dos pesquisadores, porque é através do rastro que suas publicações deixam na web social que a instituição ganha reconhecimento.

5 Considerações Finais

A presença *online* dos pesquisadores mais produtivos da ISKO-Brasil pode ser considerada relevante, ou seja, além de perfis ativos foram coletados indicadores que possibilitam caracterização das dimensões de atividade, impacto e atividade. Mesmo que a dimensão de visibilidade tenha sido caracterizada com número e variedade menor de indicadores coletados, considerou-se satisfatório para os objetivos propostos.

As dimensões de atividade e impacto apresentaram maior número de indicadores e, dessa maneira, foi possível analisar a presença de forma ampliada. Entretanto, a dimensão de atividade foi medida apenas pelo indicador de publicação, o que demonstra o potencial de caracterização no âmbito acadêmico e menos social. Nesse sentido, a fim de complementar análises como esta, sugere-se para pesquisas futuras, a coleta de indicadores em outras redes, como o Twitter por exemplo, para que a dimensão de impacto possa considerar também a divulgação científica.

Portanto, a presença on-line deste conjunto de pode ser definida como massiva e relevante, considerando que 7 pesquisadores (de uma amostra de 13) possuem perfil ativo em todas as plataformas e todos os pesquisadores mantêm perfis em pelo menos 4 das 7 plataformas.

Desse modo, denota-se que os pesquisadores têm interesse em manter perfis em diferentes plataformas sociais. Sugere-se um aprofundamento qualitativo a fim de compreender as percepções e motivações dos pesquisadores quanto ao uso destas plataformas, além da inclusão de outras redes e indicadores, sobretudo para cobrir aqueles que não puderam ser aprofundados aqui como conectividade social.

Referências

- ARAÚJO, Ronaldo Ferreira de. Ciência 2.0 e a Presença *Online* de Pesquisadores: visibilidade e impacto. **Ciência da Informação em Revista**, v. 1, n. 3, p. 32-40, 2014.
- BARCELOS, Janine; MARICATO, João de Melo. Visibilidade e engajamento público na web 2.0: um estudo altmétrico a partir dos artigos publicados na Scientific Data. **Em Questão**, v. 27, n. 1, p. 263-285, 2021.
- BOYD, Danah M.; ELLISON, Nicole B. Social network sites: Definition, history, and scholarship. **Journal of Computer-Mediated Communication**, v. 13, n. 1, p. 210-230, 2007.
- CAREGNATO, Sônia Elisa. E. Google Acadêmico como ferramenta para os Estudos de Citações: Avaliação da Precisão das Buscas por Autor. **PontodeAcesso**, [S. l.], v. 5, n. 3, p. 72-86, 2012. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/5682>. Acesso em: 11 abr. 2023.
- DAHLBERG, Ingetraut. Brief Communication: What is Knowledge Organization? **Knowledge Organization**, v. 41, n.(1), p. 85-91, 2013. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2014-1-85.pdf>. Acesso em: 18 abr. 2023.
- DAHLBERG, Ingetraut. Knowledge Organization: A New Science? **Knowledge Organization**, v. 33, n.1, 2006. Disponível em: https://www.ergon-verlag.de/isko_ko/downloads/ko3320061c.pdf. Acesso em: 18 abr. 2023.
- DAHLBERG, Ingetraut. Teoria do Conceito. **Ciência da Informação**, Brasília, v. 7, n. 2, p. 101-107, 1978. Disponível em: <https://revista.ibict.br/ciinf/article/view/115>. Acesso em: 18 abr. 2023.
- HERMAN, Eti; NICHOLAS, David. Scholarly reputation building in the digital age: An activity-specific approach. Review article. **Profesional de la Información**, v. 28, n. 1, 2019.
- HJØRLAND, Birger. What is Knowledge Organization (KO)? **Knowledge Organization**, v. 35, n. 3/2, 86-101. Disponível em: <https://doi.org/10.5771/0943-7444-2008-2-3-86>. Acesso em: 18 abr. 2023.
- MARTÍN-MARTÍN, Alberto; ORDUÑA-MALEA, Enrique; AYLLÓN, Juan M.; LÓPEZ-CÓZAR, Emilio Delgado. The counting house: Measuring those who count. Presence of Bibliometrics, Scientometrics, Informetrics, Webometrics and Altmetrics in the Google Scholar Citations, ResearcherID, ResearchGate, Mendeley & Twitter. **Arxiv**, 2016.

- MARTÍN-MARTÍN, Alberto; ORDUÑA-MALEA, Enrique; LÓPEZ-CÓZAR, Emilio Delgado. Author-level metrics in the new academic profile platforms: The online behaviour of the Bibliometrics community. **Journal of Informetrics**, v. 12, n. 2, p. 494-509, 2018a.
- MARTÍN-MARTÍN, Alberto; ORDUÑA-MALEA, Enrique; LÓPEZ-CÓZAR, Emilio Delgado. A novel method for depicting academic disciplines through Google Scholar Citations: The case of Bibliometrics. **Scientometrics**, v. 114, p. 1251–1273, 2018b.
- MASUM, Hassan; ZHANG, Yi-Scheng. Manifesto for the Reputation Society. **First Monday**, v. 27, n. 7, 2004.
- OLIVEIRA, Ely F. Tannuri de; GRACIO, Maria C. Cabrini. Visibilidade dos pesquisadores no periódico *Scientometrics* a partir da perspectiva brasileira: um estudo de cocitação. **Em Questão**, v. 18, n. 3, p. 99-113, 2013.
- ORDUÑA-MALEA, Enrique; LÓPEZ-CÓZAR, Emilio Delgado. Performance Behavior Patterns in Author-Level Metrics: A Disciplinary Comparison of Google Scholar Citations, ResearchGate, and Impactstory. **Frontiers in Research Metrics and Analytics**, v. 2, 2017.
- ORDUÑA-MALEA, Enrique; MARTÍN-MARTÍN, Alberto; LÓPEZ-CÓZAR, Emilio Delgado. The next bibliometrics: Almetrics (Author Level Metrics) and the multiple faces of author impact. **Profesional de la Información**, v. 25, n. 3, p. 485-496, 2016. Disponível em: <http://profesionaldelainformacion.com/contenidos/2016/may/18.pdf>. Acesso em: 23 mar. 2023.
- ORTEGA, José Luis. Almetrics data providers: A meta-analysis review of the coverage of metrics and publication. **Profesional de la información**, [S. l.], v. 29, n. 1, 2020. Doi: 10.3145/epi.2020.ene.07. Disponível em: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2020.ene.07>. Acesso em: 31 mar. 2023.
- RECUERO, Raquel. **Redes sociais na Internet**. Porto Alegre: Sulina, 2009.
- SANTOS, Francielle Franco. Presença on-line dos pesquisadores mais produtivos da ISKO Brasil e os principais indicadores coletados. 2023 **Figshare**. (*Dataset*). Disponível em: <https://doi.org/10.6084/m9.figshare.22814495.v1>
- SUGIMOTO, Cassidy R.; WORK, Sam; LARIVIÈRE, Vincent; HAUSTEIN, Stefanie. Scholarly use of social media and altmetrics: A review of the literature. **JASIST – Journal of the Association for Information Science and Technology**, v. 68, p. 2.037-2.062, 2017.
- TORRES-SALINAS, Daniel; CABEZAS-CLAVIJO, Álvaro; JIMÉNEZ-CONTRERAS, Evaristo. Almetrics: nuevos indicadores para la comunicación científica en la Web 2.0. **Comunicar**, v. 21, n. 41, p. 53-60, 2013.
- UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION (UNESCO). **Research Evaluation Metrics**. 2015. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000232210>. Acesso em: 23 jul. 2022.
- WANG, Dashun; BARABÁSI, Albert-László. **The Science of Science**. Cambridge: Cambridge University Press, 2021.
- WILLINSKY, John. Open access and academic reputation. **Annals of Library and Information Studies**, v. 57, p. 296-302, 2010.

Apêndice A – Tabela 1

Tabela 1. Pesquisadores mais produtivos da ISKO Brasil (2011-2019)

Pesquisadores	IES de afiliação	Nº de trabalhos	Edições nas quais publicou				
			2011	2013	2015	2017	2019
José Augusto Chaves Guimarães	UNESP	12	[1], [2], [3]	[4]	[5], [6];	[7], [8], [9], [10], [11]	[12]
Daniel Martínez-Ávila	Univ. de León/ UNESP	11	[1]	[2]	[3], [4], [5], [6], [7]	[8], [9]	[10], [11]
Mariângela Spotti Lopes Fujita	UNESP	10	[1]	[2]	[3], [4], [5]	[6], [7],	[8], [9], [10]
Paula Regina Dal'Evedove	UFSCAR	8	–	[1]	[2], [3]	[4]	[5], [6], [7], [8]
Mario Barité	Univ. de la Republica	8	[1], [2]	[3]	[4], [5], [6]	[7]	[8]
Maria Cláudia Cabrini Grácio	UNESP	7	[1],	–	[2], [3], [4]	[5], [6], [7]	
João Batista Ernesto Moraes	UNESP	7	[1]	–	[2], [3], [4], [5], [6], [7]	–	–
Maria Aparecida Moura	UFMG	7	[1], [2]	[3], [4]	[5], [6]	[7]	–
Brígida M. Nogueira Cervantes	UEL	6	[1]	[2]	[3], [4]		[5], [6]
Marilda Lara Lopes Ginez de Lara	USP	6	[1]	[2]	[3], [4], [5]	[6]	
Marcos Miranda	UFRJ	6	[1], [2],	[3]		[4]	[5], [6]
Fabio Assis Pinho	UFPB	6	[1]	[2]	[3]	[4], [5]	[6]
Roberta Cristina Dal'Evedove Tartarotti	UNICAMP	6	–	[1]	[2], [3]	[4]	[5], [6]

Fonte: Autores.

Apêndice B – Quadro 1

Quadro 1. Plataformas sociais classificadas quanto ao uso, indicadores principais e a natureza da atividade de pesquisa.

Uso/Tipo	Plataforma	Tipos de Indicadores	Natureza dos Indicadores
Identificadores Únicos	ResearcherID (RID)	Publicações, citações, índice h.	Publicação; citação.
	Scopus Autor Profile (SAID)	Publicações, citações, índice h, seguidores.	Publicação; citação.
	Google Scholar Citations (GSC)	Publicações, citações, índice h.	Publicação; citação; uso.
Acadêmico	Research Gate (RG)	Publicações, citações, índice h, leituras, habilidades	Publicação; citação; conectividade social.
	Academia.edu (ACA)	Publicações, citações, índice h, visualizações, seguidores.	Publicação; citação; conectividade social.
Agregadores Alométricos	Altmetric Explorer (ALT)	Menções em diferentes canais on-line.	Difusão, comentário ou discussão; conectividade social.
	Impact Story(IMP)	Menções em diferentes canais on-line.	Difusão, comentário ou discussão; conectividade social.

Fonte: Adaptado de Orduña-Malea; Martín-Martín; López- Cózar (2016).

Santos, Francielle Franco dos; Grácio, Maria Cláudia Cabrini; Oliveira, Caliel Cardoso de. A presença online dos autores mais produtivos da ISKO Brasil: indicadores de atividade, visibilidade e impacto nas plataformas sociais. *Organização e Representação do Conhecimento em diferentes contextos: desafios e perspectivas na era da datificação*, 2023.

Preservação de documentos digitais: estratégias de preservação

Preservation of digital documents: preservation strategies

Izângela Maria Sansoni Tonello de Oliveira (1), João Cairo Ferreira (2)

(1) Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina, Rodovia Celso Garcia Cid, Pr 445 Km 380, Campus Universitário, Londrina - PR, izangela@uel.br, (2) Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina, joao.cairo.ferreira@uel.br

Resumo

Objetivo: Este artigo tem como objetivo realizar uma coleta de dados de referências bibliográficas com o intuito de conhecer as abordagens teóricas e metodológicas publicadas sobre a preservação de documentos digitais em arquivos. **Metodologia:** Foi realizada uma coleta de dados de referências bibliográficas sobre o tema "preservação de documentos digitais em arquivos", priorizando especialmente textos recentes relacionados à arquivologia e biblioteconomia. **Resultado:** O resultado obtido demonstra que há diversas propostas para a preservação digital, porém ainda não há uma proposta totalmente viável e abrangente. Portanto, torna-se necessário adotar alguns métodos para reduzir o risco de perda dos documentos digitais. **Conclusão:** Conclui-se que a preservação digital em arquivos é um tema complexo e desafiador, e que as estratégias de preservação digital devem abranger todos os aspectos relacionados à problemática, incluindo custos, legislação, gestão, acesso, políticas e critérios.

Palavras-chave: Ciência da Informação. Preservação Documental. Preservação Digital. Preservação Digital.

Abstract

Objective: This article aims to collect bibliographic references to understand the theoretical and methodological approaches published on the preservation of digital documents in archives. **Methodology:** A collection of bibliographic references was conducted on the topic of "preservation of digital documents in archives," with a focus on recent texts related to archival studies and library science. **Result:** The findings reveal that there are several proposals for digital preservation; however, there is still no fully feasible and comprehensive proposal. Therefore, it is necessary to adopt certain methods to mitigate the risk of digital document loss. **Conclusion:** In conclusion, digital preservation in archives is a complex and challenging subject, and digital preservation strategies must encompass all aspects related to the issue, including costs, legislation, management, access, policies, and criteria.

Keywords: Information Science. Document Preservation. Digital Preservation. Digital Preservation.

Resumen

Objetivo: Este artículo tiene como objetivo recopilar datos de referencias bibliográficas con el fin de conocer los enfoques teóricos y metodológicos publicados sobre la preservación de documentos digitales en archivos. **Metodología:** Se llevó a cabo una recopilación de datos de referencias bibliográficas sobre el tema de "preservación de documentos digitales en archivos", priorizando especialmente textos recientes relacionados con la archivología y la biblioteconomía. **Resultado:** Los resultados obtenidos demuestran que existen diversas propuestas para la preservación digital, pero aún no hay una propuesta totalmente viable y completa. Por lo tanto, es necesario adoptar algunos métodos para reducir el riesgo de pérdida de los documentos digitales. **Conclusión:** Se concluye que la preservación digital en archivos es un tema complejo y desafiante, y que las estrategias de preservación digital deben abarcar todos los aspectos relacionados con la problemática, incluyendo costos, legislación, gestión, acceso, políticas y criterios..

Palabras clave: Ciencias de la Información. Conservación de documentos. Preservación Digital. Preservación Digital.

1 Introdução

A organização digital e sua preservação é uma área de estudo fundamental para a arquivologia e a biblioteconomia, que buscam estabelecer padrões, normas, políticas, critérios e procedimentos para garantir a perpetuidade de documentos e informações em meio digital. Com o avanço tecnológico e a crescente produção de documentos digitais, torna-se cada vez

mais importante compreender as abordagens teóricas e metodológicas que norteiam a preservação digital. Para tanto, é necessário realizar uma coleta de dados de referências bibliográficas que permita conhecer as diferentes estratégias adotadas no contexto da arquivologia e da biblioteconomia.

Embora a preservação digital seja um desafio recorrente, torna-se ainda mais complexo devido à fragilidade dos documentos digitais com o passar do tempo. A rápida

degradação física, a obsolescência tecnológica, a complexidade e os custos são alguns dos obstáculos enfrentados nesse contexto. Nesse sentido, surgem diferentes estratégias de preservação digital que buscam incorporar todos os aspectos relacionados a essa problemática, como os custos, a legislação, a gestão, o acesso, as políticas e os critérios.

Embora haja diversas propostas de preservação digital disponíveis, ainda não existe uma solução totalmente viável e abrangente para garantir a preservação de documentos digitais a longo prazo. Dessa forma, torna-se necessário adotar métodos que reduzam o perigo da perda desses documentos. Para tanto, é preciso buscar alternativas que possam minimizar os efeitos da rápida degradação física, da obsolescência tecnológica e da complexidade inerentes aos documentos digitais, com o intuito de assegurar a preservação de informações relevantes para a sociedade e a cultura.

2 Levantamento Bibliográfico

A preservação digital surgiu na segunda metade do século passado e intensificou-se neste início do século atual. As preocupações acerca da preservação digital no mundo tiveram sua primeira expressão no contexto mundial a partir de um trabalho realizado no International Council on Archives (ICA), em 1970 (INTERNATIONAL COUNCIL ON ARCHIVES, 1970). No Brasil, apesar de contar com representantes no Committee on Automation do ICA na área de arquivologia, os eventos da área, na década de 1970, abordaram pouco a temática. Somente a partir da década de 90, começaram a surgir publicações que abordavam o tema da tecnologia da informação e dos documentos eletrônicos no Brasil. A Câmara Técnica de Documentos Eletrônicos (CTDE) do Conselho Nacional de Arquivos (CONARQ), órgão encarregado de preparar diretrizes para a gestão e a preservação de documentos arquivísticos no Brasil, é a instituição que tem pesquisado mais profundamente a gestão e a preservação, por longo tempo, de documentos digitais, com abordagem arquivística. O Grupo de Documentos Eletrônicos (GDAE) participa com dois membros nessa Câmara, tendo auxiliado na produção da Carta de Preservação Digital, Glossário de Termos, duas resoluções e do Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos (e-ARQ) (CONARQ, 2009).

O e-ARQ Brasil é uma especificação de requisitos que estabelece um conjunto de condições a serem cumpridas pela organização produtora/recebedora de documentos, pelo sistema de gestão arquivística e pelos próprios documentos a fim de garantir a sua confiabilidade, autenticidade e acesso. Além disso, o e-ARQ Brasil pode ser usado para orientar a identificação de documentos arquivísticos digitais (CONARQ, 2009). Especifica todas as atividades e operações técnicas da

gestão arquivística de documentos, desde a produção, tramitação, utilização e arquivamento até a sua destinação final. Todas essas atividades podem ser desempenhadas pelos Sistemas Informatizados de Gestão Arquivística de Documentos (SIGAD), os quais, se desenvolvidos em conformidade com os requisitos, conferirão credibilidade à produção e à manutenção de documentos arquivísticos. A Carta para a Preservação do Patrimônio Arquivístico Digital é um instrumento oficial sobre a importância e a necessidade de ações para a preservação de documentos digitais (CONARQ, 2009).

A preservação de documentos eletrônicos no Brasil vem chamando a atenção de diversas entidades, entre as quais o Arquivo Nacional, que está conduzindo estudos sobre criação, manutenção e recuperação de arquivos em longo prazo, com base no modelo Open Archive Information System (OAIS). O Modelo Referencial Open Archival Information System (OAIS), desenvolvido pelo Consultative Committee for Space Data Systems (CCSDS), foi aprovado para publicação como padrão ISO 14721:2003 em fevereiro de 2003 e descreve um enquadramento conceitual para um repositório digital genérico, aberto a todas as comunidades, com garantias de confiabilidade. O modelo OAIS consiste em uma estrutura conceitual que disciplina e orienta um sistema de arquivo dedicado a preservar e manter o acesso à informação digital a longo prazo.

Segundo Innarelli (2006), a aplicação do OAIS em arquivos consiste na organização de pessoas e sistemas, tendo como responsabilidade a preservação e o acesso da informação à comunidade interessada, com foco principal na informação digital, nas formas primárias de armazenamento e no suporte da informação para os materiais de arquivos digitais e físicos. A definição de uma política de preservação envolve, geralmente, todas as facetas de um arquivo. Isso implica a criação de políticas de avaliação e seleção de materiais, a identificação de esquemas de metainformação apropriados (metainformação descritiva, técnica, de disseminação, estrutural e de preservação), a definição de estratégias de preservação adequadas a cada classe de objetos digitais, a criação de planos de sucessão para a eventualidade da organização detentora da informação interromper suas atividades, a utilização de modelos sustentáveis de financiamento, entre outros (FERREIRA, 2006).

Cada vez mais, os recursos culturais e educacionais mundiais estão sendo distribuídos e acessados em formato digital, em substituição ao papel. No entanto, a informação digital está sujeita à obsolescência técnica e à degradação física, correndo o risco de perdas. Para enfrentar os novos desafios criados pelo mundo digital, algumas estratégias têm sido propostas para a preservação digital. Os principais métodos recomendados para a preservação podem ser agrupados em dois tipos: estruturais e operacionais. Os métodos

estruturais tratam dos investimentos necessários, enquanto os métodos operacionais tratam das medidas concretas relacionadas à preservação digital. Entre as principais práticas operacionais estão a migração de suporte e o refrescamento do meio (preservação física), a conversão de formatos, a emulação (preservação lógica) e a preservação do conteúdo (intelectual) (MÁRDERO ARELLANO, 2008).

O Arquivo Nacional finalizou alguns manuais de preservação de documentos digitais, como as versões impressas em português dos manuais "Diretrizes do Produtor" e "Diretrizes do Preservador", em colaboração com a Câmara dos Deputados no âmbito do convênio com o Projeto InterPARES. Os livretos impressos serão distribuídos para os órgãos da Administração Pública Federal e para as instituições parceiras do Arquivo Nacional. As "Diretrizes do Produtor" apresentam um conjunto de recomendações e boas práticas voltadas para organizações que elaboram e mantêm documentos arquivísticos digitais. As "Diretrizes do Preservador" têm como objetivo orientar as instituições sobre os melhores procedimentos para a preservação de documentos arquivísticos digitais. O trabalho de tradução foi realizado a partir dos originais em inglês desenvolvidos pelos pesquisadores da Universidade de British Columbia, no Canadá. O conteúdo integral dos manuais está disponível para o público em geral no site: <http://www.conarq.arquivonacional.gov.br/cgi/cgilua.exe/sys/start.htm>.

A preservação digital é um desafio constante e exige a atuação de diversas entidades. No Brasil, o Arquivo Nacional tem se destacado nesse campo, conduzindo estudos sobre a criação, manutenção e recuperação de arquivos a longo prazo, com base no modelo Open Archival Information System (OAIS). Esse modelo fornece uma estrutura conceitual para garantir a preservação e o acesso à informação digital de forma confiável.

A importância da preservação digital é destacada pela Carta para a Preservação do Patrimônio Arquivístico Digital, um instrumento oficial que ressalta a necessidade de ações nesse sentido (CONARQ, 2009). A preservação dos documentos digitais é fundamental devido à crescente distribuição e acesso a recursos culturais e educacionais nesse formato. No entanto, é preciso estar atento à obsolescência tecnológica e à degradação física, que podem colocar em risco a preservação desses materiais.

Para garantir a preservação digital, é necessário adotar estratégias adequadas. Os métodos estruturais, que envolvem investimentos, e os métodos operacionais, que se referem às medidas concretas de preservação, desempenham papéis fundamentais nesse processo (MÁRDERO ARELLANO, 2008). Entre as práticas operacionais recomendadas estão a migração de suporte e o refrescamento do meio físico, a conversão de

formatos, a emulação e a preservação do conteúdo intelectual.

Nesse contexto, o e-ARQ Brasil desempenha um papel crucial como uma especificação de requisitos que estabelece condições a serem cumpridas pelas organizações, sistemas de gestão arquivística e pelos próprios documentos, a fim de garantir sua confiabilidade, autenticidade e acesso (CONARQ, 2009). O e-ARQ Brasil orienta a identificação e a gestão adequada de documentos arquivísticos digitais, abrangendo todas as etapas da sua vida, desde a produção até a destinação final.

Em suma, a preservação digital é uma preocupação crescente em todo o mundo e no Brasil, envolvendo diversas entidades, como o Arquivo Nacional e o CONARQ. A adoção de estratégias de preservação, o uso de modelos como o OAIS e o e-ARQ Brasil, e a criação de diretrizes e boas práticas são passos essenciais para garantir a preservação a longo prazo dos documentos arquivísticos digitais, assegurando seu acesso e perpetuação do patrimônio histórico e cultural.

3 Critérios e Tipos para Preservação Digital

3.1 Critérios para preservação

Para a preservação digital de documentos em arquivos, são necessárias diversas ações, tais como preservação tecnológica, migração, emulação, encapsulamento, adoção de padrões e protocolos, implementação de política de gestão documental e tecnológica, controle público de legitimidade, além de uma política pública que inclua pesquisa científica e ações de arquivos em todos os níveis. Essas medidas dependem da existência de uma legislação adequada.

A maioria das propostas de critérios existentes menciona a influência do modelo de referência OAIS/SAAI, que significa Sistema de Arquivamento Aberto de Informação, e apresenta um conjunto de critérios para o planejamento de repositórios digitais. Essas propostas, em sua maioria provenientes da área arquivística, já contemplam o conceito de arquivos ou repositórios digitais confiáveis. A confiabilidade é um princípio arquivístico fundamental que está intrinsecamente ligado à responsabilidade na custódia dos documentos.

3.2 Migração

A migração, juntamente com a emulação, são atualmente as estratégias mais utilizadas pelas instituições que buscam preservar objetos digitais (FERREIRA, 2006). No entanto, é importante destacar que, apesar do número elevado de alternativas de preservação digital apresentadas, ainda não existem provas conclusivas quanto à eficácia de cada uma delas.

A migração, que é a estratégia mais comum, permite a disponibilidade permanente das informações. Ao

contrário de outras estratégias de preservação de documentos digitais, a migração se preocupa não apenas com o suporte ou o objeto digital em seu formato digital, mas também com o conteúdo intelectual do documento, ou seja, a informação nele contida.

A importância da migração reside na transferência da informação para novos formatos, mantendo a integridade dos dados. Conforme mencionado por Sayão (2005), a migração envolve a transferência da informação digital de uma mídia obsoleta, deteriorada ou instável para um suporte mais novo ou tecnologicamente atualizado. Isso pode incluir a migração de CD-ROM para um CD-ROM mais recente, de disquete para DVD, de um formato ultrapassado para um formato mais atual ou padronizado, como a migração de MSWord para XML ou ASCII, ou ainda a migração de uma plataforma computacional em desuso para uma mais moderna.

Márdero Arellano (2008) acrescenta que o propósito da migração é preservar a integridade dos objetos digitais e garantir a capacidade dos usuários de recuperá-los, exibí-los e utilizá-los de maneiras diferentes diante das constantes mudanças tecnológicas.

Portanto, a importância da migração está em transferir os objetos digitais para novos formatos, assegurando a preservação da integridade da informação. As instituições que possuem extensos acervos utilizam a migração da informação digital de um hardware ou software para outro, garantindo a recuperação dos dados e a preservação de sua integridade.

3.3 Emulação ou Encapsulamento

As soluções baseadas em encapsulamento têm como objetivo preservar os objetos digitais sem modificá-los até o momento em que se tornam efetivamente necessários. A estratégia de encapsulamento consiste em preservar, juntamente com o objeto digital, todas as informações necessárias e suficientes para permitir o futuro desenvolvimento de conversores, visualizadores ou emuladores. Essas informações podem incluir, por exemplo, uma descrição formal e detalhada do formato do objeto preservado.

É importante ressaltar que, ao longo do tempo, o próprio emulador também se tornará obsoleto, sendo necessário convertê-lo para uma nova plataforma ou desenvolver um novo emulador capaz de emular o anterior (FERREIRA, 2006).

Segundo Márdero Arellano (2008), as técnicas de emulação propõem a preservação dos dados em seu formato original por meio de programas emuladores, que podem imitar o comportamento de uma plataforma de hardware obsoleta e emular o sistema operacional correspondente. O processo consiste em preparar um sistema que funcione da mesma maneira que outro de um tipo diferente para processar os programas.

A emulação é a única técnica que pode preservar os objetos digitais originais e sua capacidade de serem executados conforme foram originalmente. Ela deve ser usada quando o recurso digital não pode ser convertido em formatos independentes de software e migrados no futuro. A grande vantagem dessa abordagem é garantir a manutenção de todos os componentes necessários para garantir o acesso. No entanto, ela não resolve necessariamente o problema do acesso e, por essa razão, deve ser utilizada como uma ação complementar a outras estratégias, como a emulação. A principal desvantagem do encapsulamento está no fato de que o software encapsulado continua sujeito a uma obsolescência rápida.

3.4 Reprografia (Xerox)

A tecnologia da digitalização é uma ferramenta altamente flexível para o arquivamento, preservação e acesso a documentos, por meio do armazenamento de suas imagens em formato digital. Para gerenciar esses serviços, foi criado o GED (Gerenciamento Eletrônico de Documentos), que converte informações, como voz, texto e imagens, em formato digital. Esse sistema funciona por meio de softwares e hardwares específicos e faz uso, em geral, de mídias ópticas para o armazenamento.

Um sistema de GED utiliza a tecnologia da informática para capturar, armazenar, localizar e gerenciar versões digitais das informações. As imagens digitais estão se tornando cada vez mais comuns em bibliotecas e arquivos, e a qualidade dos produtos de imagem digital pode ser impressionante. É inquestionável que a qualidade continuará a melhorar à medida que a tecnologia amadurece. No entanto, sem um esforço sério para garantir o acesso aos arquivos digitais de imagens ao longo do tempo, há um grande risco de perdas.

O processo de digitalização envolve a conversão de documentos em papel, fotos, microfilmes, microfichas, jaquetas ou cartões-janela em imagens digitais, por meio do uso de um scanner. Essa tecnologia permite o armazenamento de grandes volumes de documentos em formato digital.

O Conselho Nacional de Arquivos disponibiliza em seu site a Resolução nº 31, que trata das Recomendações para Digitalização de Documentos Arquivísticos Permanentes. Essas recomendações têm como objetivo auxiliar as instituições detentoras de acervos arquivísticos de valor permanente na concepção e execução de projetos e programas de digitalização. A resolução pode ser encontrada no site: <http://www.conarq.arquivonacional.gov.br/cgi/cgilua.exe/sys/start.htm?frominfoindex=21&infoid=508&sid=46>.

Segundo Sant'Anna (2002), a reprodução em microfilme garante, desde que seja gerada com qualidade e devidamente preservada, o acesso às informações por

cerca de meio milênio. No entanto, é sabido que a reprodução a partir de um microfilme tem qualidade inferior à reprodução a partir do original, e o problema dos documentos multimídia ainda persiste.

As técnicas digitais, embora ainda pouco utilizadas para que se possa ter uma opinião solidamente fundamentada, apresentam-se como algo de extraordinário potencial, sem que se vislumbre qualquer obstáculo que inviabilize sua utilização generalizada.

3.5 *Refrescamento*

A ação de refrescamento tem como objetivo transferir a informação digital de um suporte físico de armazenamento para outro mais atual, antes que o primeiro se deteriore (FERREIRA, 2006). O refrescamento apresenta uma grande vantagem, pois não requer um grande investimento em equipamentos nem um elevado nível de conhecimentos técnicos, a menos que a informação contida no suporte esteja protegida de alguma forma.

No entanto, a principal desvantagem apontada é que o refrescamento apenas resolve os problemas de degradação e obsolescência dos suportes utilizados.

3.6 *Atualização de Versões*

Segundo Ferreira (2006), é bastante comum encontrar aplicações de software capazes de abrir ou importar objetos digitais produzidos por versões anteriores da mesma aplicação.

Essas aplicações geralmente permitem gravar os objetos importados no formato mais atual produzido por ela. Essa operação é conhecida como atualização da versão do formato.

A atualização de versões refere-se à atualização de materiais digitais produzidos por um determinado software, por meio de sua regravação em uma versão mais recente do mesmo (FERREIRA, 2006).

3.7 *Conversão para Formatos Concorrentes*

O processo de atualização de versões geralmente é controlado pela organização responsável pelo desenvolvimento da respectiva aplicação de software.

A qualidade da migração depende, portanto, da capacidade dos importadores fornecidos pelo fabricante e do grau de retro compatibilidade oferecido pelo novo formato. Idealmente, o fabricante deve garantir que todos os atributos presentes em uma determinada versão do formato estejam disponíveis na nova versão que o substitui.

No entanto, independentemente do sucesso econômico de um fabricante ou produto de software, os formatos estão constantemente sujeitos à descontinuidade.

Uma maneira de garantir que os objetos digitais sobrevivam a esse tipo de ruptura tecnológica é

convertê-los para formatos de uma linha de produtos concorrente (FERREIRA, 2006).

Conforme também mencionado por Ferreira (2006), a conversão de um objeto digital para um formato que não tenha sido necessariamente desenvolvido pela mesma empresa que elaborou o software proprietário no qual ele foi produzido está restrita a alguns tipos de objetos. Isso visa proteger o conteúdo da descontinuidade dos softwares, ou seja, quando o software não recebe atualizações. No entanto, existem formatos que não dependem de nenhuma aplicação de software específica. Isso ocorre com a maioria dos formatos de imagem (JPEG, TIFF, PNG). Isso permite que os objetos sejam convertidos entre formatos semelhantes, independentemente da aplicação utilizada na sua criação..

3.8 *Normalização*

Quando há um número controlado de formatos, uma mesma estratégia de preservação pode ser aplicada a um maior número de objetos digitais, o que pode levar a uma redução generalizada nos custos de preservação.

A normalização de formatos pode ser implementada de diversas maneiras. Alguns repositórios realizam a conversão automática dos objetos recebidos para um formato único de preservação.

Outros estabelecem políticas de arquivo que limitam os formatos nos quais aceitam informações, o que significa que cabe aos produtores de informação converter seus objetos digitais para os formatos estipulados (FERREIRA, 2006).

O argumento que sustenta essa abordagem é baseado na premissa de que os produtores de informação são as entidades mais adequadas para avaliar a qualidade da conversão realizada.

3.9 *Tradução de Versões e Verbetes (Pedra de Rosetta Digital)*

Em 1822, o paleógrafo francês Jean-François Champollion decodificou a versão egípcia do texto utilizando seus conhecimentos de grego clássico, um idioma amplamente conhecido pelos historiadores daquela época.

Essa descoberta levou à decodificação de muitos outros textos egípcios encontrados em diversos lugares e em diferentes suportes, como monumentos, rochas e papiros.

Um exemplo de aplicação dessa estratégia é imprimir em papel um conjunto representativo de documentos de texto junto com sua representação binária. No futuro, as regras necessárias para interpretar e migrar os objetos para um novo formato poderiam ser inferidas ao comparar os documentos impressos com sua representação binária. Essa estratégia deve ser considerada apenas em situações em que todos os esforços de preservação falharam. É principalmente uma

ferramenta de arqueologia digital e não uma estratégia básica para a preservação de objetos digitais (FERREIRA, 2006).

Além disso, de acordo com Ferreira (2006), a Pedra de Roseta Digital baseia-se em três momentos distintos: o processo de preservação do conhecimento, o registro da codificação do formato de arquivo e do conteúdo em binários, e a recuperação dos dados e reconstrução dos documentos com base nas especificações construídas na primeira etapa.

Essa técnica é denominada Pedra de Roseta Digital devido ao seu objetivo de traduzir para novos softwares os arquivos digitais provenientes de tecnologias já obsoletas, utilizando parâmetros que permitam essa tradução.

3.10 Arqueologia Digital

A arqueologia digital é considerada uma solução parcial de preservação. Não há garantia de que seja possível recuperar a totalidade dos dados nem garantir sua legibilidade. Além disso, existe outra desvantagem significativa: a falta de documentação adequada dificulta a interpretação dos dados recuperados, muitas vezes sendo necessário recorrer a "adivinhação" para estabelecer a identidade, integridade e contexto do documento digital (FERREIRA, 2006).

A arqueologia digital consiste em resgatar recursos digitais que se tornaram inacessíveis devido à obsolescência tecnológica e/ou degradação dos meios de armazenamento. No entanto, essa abordagem não é uma estratégia em si mesma, mas uma alternativa quando os materiais digitais estão fora de um programa de preservação sistemática (FERREIRA, 2006).

Devido ao alto custo envolvido, a arqueologia digital é recomendada apenas para a recuperação e restauração de dados contidos em suportes danificados ou formatos obsoletos cujo valor ou importância justifiquem esse investimento. Em alguns casos, o valor da informação não justifica o custo da operação.

4 Considerações Finais

As estratégias apresentadas neste estudo são resultado da análise das principais propostas de requisitos, padrões e normas encontradas na literatura das áreas da ciência da informação e da arquivologia. Conclui-se que há diversas ações nesse sentido, porém, até o momento, nenhuma delas foi universalmente aceita.

A preservação de documentos em formato digital representa um desafio para os arquivistas, que precisam criar ambientes organizacionais adequados para essa finalidade. Não existe uma única estratégia de preservação até o momento, pois diversas estratégias têm sido propostas, mas cada uma delas é apropriada apenas para determinados tipos de objetos, situações e

instituições, ou seja, são soluções específicas para casos específicos.

Assegurar que as informações digitais, que são geradas continuamente, estejam acessíveis não apenas no presente, mas também para as futuras gerações, é um dos desafios mais urgentes da chamada Era da Informação. Para lidar com esse desafio, são necessárias soluções não apenas técnicas, mas principalmente econômicas e sociais. A utilização de padrões abertos e metadados tem se mostrado um fator fundamental, pois facilitam a gestão, recuperação e utilização da informação digital, além de viabilizar estratégias como migração e emulação.

Essa é uma jornada que deve ser percorrida pelas organizações responsáveis pela preservação e disponibilização de documentos digitais. Tanto na área da arquivística quanto na ciência da informação, há o desafio de estabelecer padrões, políticas e estratégias para as melhores práticas de arquivamento digital.

A preservação de hardware se mostra inviável, especialmente no que diz respeito ao acesso restrito à informação que só pode ser consultada no local físico onde os hardwares preservados estão localizados. Estratégias como o encapsulamento são consideradas dispendiosas e complexas de manter, pois exigem um maior investimento tecnológico, além de especialistas comprometidos com o processo de representação estrutural e interpretação dos documentos pelos softwares e hardwares.

Entre as estratégias atuais apresentadas neste estudo, a migração e a emulação parecem ser os mais viáveis e confiáveis para serem implementadas, pois são técnicas amplamente utilizadas por organizações com grandes acervos, focando no conteúdo dos documentos e não apenas no suporte físico. A preservação é uma medida de segurança que permite a transferência da informação para as gerações futuras.

No entanto, a preservação de longo prazo de documentos digitais ainda é um tema complexo, com questões em aberto. Apesar da existência de várias estratégias, algumas das quais podem ser mais viáveis do que outras, ainda não há evidências conclusivas sobre a eficácia de nenhuma delas a longo prazo.

Os profissionais das áreas documentais podem e devem trabalhar em conjunto com especialistas em tecnologia para avançar com as ações necessárias, adotando uma postura proativa diante da necessidade de preservação de documentos arquivísticos digitais e buscando encontrar e propor soluções adequadas aos desafios impostos pela sociedade do conhecimento moderno. Estudar a preservação de documentos digitais implica analisar como eles sofrem alterações ao longo do tempo ou como poderão sofrer no futuro.

Nesse sentido, é fundamental realizar pesquisas contínuas e colaborativas, envolvendo profissionais de

diferentes áreas, como arquivistas, cientistas da informação e especialistas em tecnologia. Essas parcerias possibilitam a troca de conhecimentos e experiências, contribuindo para o desenvolvimento de estratégias mais eficientes e adequadas à preservação de documentos digitais.

Além disso, é essencial estabelecer diretrizes, políticas e normas que orientem as melhores práticas de preservação, considerando as especificidades de cada contexto e tipo de documento. A criação de padrões de metadados e a adoção de formatos abertos são exemplos de medidas que favorecem a interoperabilidade e facilitam a gestão e o acesso aos documentos digitais.

No entanto, é importante destacar que a preservação de documentos digitais é um desafio em constante evolução. Novas tecnologias surgem, formatos se tornam obsoletos e a quantidade de dados gerados cresce exponencialmente. Portanto, é necessário manter-se atualizado e adaptar as estratégias de preservação de acordo com as mudanças do ambiente tecnológico.

Em suma, a preservação de documentos digitais demanda esforços conjuntos, envolvendo conhecimentos multidisciplinares, ações proativas e a busca por soluções inovadoras. Somente por meio de uma abordagem abrangente e colaborativa será possível garantir a acessibilidade e a preservação desses documentos para as futuras gerações.

Referências

- CONARQ, CONSELHO NACIONAL DE ARQUIVOS. **Carta para a Preservação do Patrimônio Arquivístico Digital**: preservar para garantir o acesso. [S.l.]: UNESCO, 2009. Disponível em: <http://www.documentoseletronicos.arquivonacional.gov.br/cgi/cgilua.exe/sys/start.htm>, Acesso em: 17 de novembro de 2022.
- FERREIRA, Miguel. **Introdução à preservação digital**: conceitos, estratégias e atuais consensos. Portugal, 2006.
- INNARELLI, Humberto Celeste. **Preservação de documentos digitais**: confiabilidade de mídias CD-ROM e CD-R. Dissertação de Mestrado - UNICAMP, 2006.
- MÁRDERO ARELLANO, Miguel Ángel. **Crítérios para a preservação digital da informação científica**. Brasília, 2008.
- SANT'ANNA, Marcelo Leone. **A digitalização de documentos de arquivo**: o caso das plantas de parcelamento do solo de Belo Horizonte. 2002.
- SAYÃO, Luis Fernando. **Preservação digital no contexto das bibliotecas digitais**: uma breve introdução. Salvador: EDUFBA, 2005.

Procedimentos para leitura documentária de objetos museológicos de coleções zoológicas: contribuições para análise de assunto em museus

Procedures for documentary reading of museum objects from zoological collections: contributions to subject analysis in museums

Raul de Azevedo Carvalho (1), Franciele Marques Redigolo (2), Mariângela Spotti Lopes Fujita (3)

(1) Universidade Estadual Paulista (UNESP), Av. Hygino Muzzi Filho, 737 - Mirante - Marília/SP, raulcarvalho.trab@gmail.com (2) Universidade Federal do Pará, R. Augusto Corrêa, 01 - Guamá, Belém - PA, franciele@ufpa.br (3) Universidade Estadual Paulista (UNESP), mariangela.fujita@unesp.br

Resumo

Este artigo tem como objetivo apresentar observações sobre procedimentos para leitura documentária de objetos museológicos de coleções zoológicas com a finalidade de contribuir para a elaboração de uma metodologia para compreensão do conteúdo e análise de assunto. Para tanto foi realizado um levantamento da literatura disponível e coleta introspectiva de dados por meio da aplicação do Protocolo Verbal Individual com cinco profissionais que realizam a catalogação de coleções zoológicas. Os resultados permitiram identificar quais as estratégias utilizadas pelos profissionais para realizar a leitura documentária e determinar o assunto dos seus objetos museológicos e os subsídios para a criação de um modelo de leitura documentária de objetos museológicos para as coleções de zoologia. Conclui-se, portanto, que a análise de assunto no contexto museológico, enquanto primeira etapa para a indexação, precisa ser bem delineada para auxiliar os profissionais no processo de análise de assunto no contexto museológico.

Palavras-chave: Leitura documentária. Análise de assunto. Coleções zoológicas. Museologia.

Abstract

This article aims to present observations on procedures for documental reading of museological objects from zoological collections in order to contribute to the development of a methodology for understanding the content and subject analysis. For this purpose, a survey of the available literature and an introspective collection of data were carried out through the application of the Individual Verbal Protocol with five professionals who perform the cataloging of zoological collections. The results made it possible to identify the strategies used by professionals to carry out documentary reading and determine the subject of their museological objects and the subsidies for the creation of a model of documentary reading of museological objects for zoology collections. It is concluded, therefore, that subject analysis in the museological context, as the first step towards indexing, needs to be well delineated to assist professionals in the subject analysis process in the museological context.

Keywords: Documentary reading. Subject analysis. Zoological collections. Museology.

Resumen

Este artículo tiene como objetivo presentar observaciones sobre procedimientos de lectura documental de objetos museológicos de colecciones zoológicas para contribuir al desarrollo de una metodología para la comprensión del contenido y el análisis del tema. Para ello, se realizó un levantamiento de la literatura disponible y una recolección introspectiva de datos mediante la aplicación del Protocolo Verbal Individual con cinco profesionales que realizan la catalogación de colecciones zoológicas. Los resultados permitieron identificar las estrategias utilizadas por los profesionales para realizar lectura documental y determinar la temática de sus objetos museológicos y los subsidios para la creación de un modelo de lectura documental de objetos museológicos para colecciones de zoología. Se concluye, por lo tanto, que el análisis temático en el contexto museológico, como primer paso hacia la indexación, debe estar bien delineado para ayudar a los profesionales en el proceso de análisis temático en el contexto museológico.

Palabras clave: Lectura documental. Análisis del tema. Colecciones zoológicas. Museología.

1 Introdução

A leitura documentária se destaca como uma fase importante para o tratamento da informação de um documento, pois é a forma como os profissionais iniciam vários processos como a indexação. Esse processo se constitui basicamente no exame e

conhecimento do conteúdo do documento por meio de uma leitura técnica que permite a identificação e seleção dos seus principais conceitos e a representação desses conceitos selecionados através de uma linguagem documentária (DIAS; NAVES, 2013; FUJITA *et al.*, 2017).

Nesse sentido, a leitura documentária combina estratégias de exploração do documento para determinar seu assunto, sendo o primeiro passo para a análise de assunto na indexação, tornando-a fundamental para que os profissionais determinem o conteúdo temático do documento para sua posterior recuperação pelo usuário. Assim, muitos procedimentos utilizados para identificação do conteúdo temático estão estreitamente ligados à estrutura textual composta por elementos pré-textuais, textuais e pós-textuais. No entanto, sua aplicabilidade em diferentes tipologias documentais, como por exemplo os objetos museológicos, ainda carece de maior aprofundamento, pois estes esbarram em aspectos como a tridimensionalidade, o contexto e a subjetividade, sendo ainda um processo pouco sistematizado (CARVALHO, 2022).

Dito isso, em pesquisa sobre o tratamento temático da informação museológica Carvalho (2022) observou por meio da aplicação de Protocolo Verbal Individual (PVI) as estratégias utilizadas para determinar os assuntos durante a leitura documentária de 09 profissionais de museus durante a catalogação de objetos museológicos, 05 da área de zoologia. Os resultados obtidos revelaram como funciona o processo de indexação para organização do conhecimento nos museus com coleções zoológicas. A partir desses resultados, surgiu a necessidade de aprofundamento dos estudos sobre o tratamento temático da informação museológica de coleções zoológicas.

Nosso objetivo, portanto, é apresentar observações sobre procedimentos para leitura documentária de objetos museológicos de coleções zoológicas com a finalidade de contribuir para a elaboração de uma metodologia para compreensão do conteúdo e análise de assunto e, conseqüente, a identificação e seleção de conceitos, que permitam uma melhor recuperação da informação dessa tipologia de objeto informacional.

Para atingir o objetivo proposto foi aplicado o Protocolo Verbal Individual (PVI) buscando identificar as estratégias utilizadas pelos profissionais de museu para determinar o assunto dos objetos museológicos durante o processo de catalogação museológica.

2 A Análise de assunto na catalogação museológica

Museus são espaços fundamentais para a construção do conhecimento científico, podendo conter acervos das mais diversas áreas do conhecimento, se constituindo em espaços de construção e divulgação científica. Souza (2009) destaca os museus como agentes para o desenvolvimento do conhecimento científico, através da pesquisa e da documentação de seus acervos.

Nesse sentido, podemos entender os museus como unidades de informação que tem por objetivo tornar a informação acessível, de maneira fácil, rápida e

confiável (CARVALHO, 2022) e, portanto, se caracterizam não apenas como um espaço de guarda de objetos, mas também como uma “[...] estrutura organizacional de referência [...]” (MENSCH, 1994, p. 15) que se preocupa com a representação e recuperação da informação.

Desse modo, é por meio da documentação museológica, processo esse que, segundo Ferrez (1994), Bottallo (2010) e Padilha (2014), permite transmitir a informação por meio da: aquisição, identificação e registro, registro fotográfico, catalogação, classificação e pesquisa. Ou seja, é o, procedimento que reúne todos os processos que envolvem o tratamento da informação museológica, permitindo representar a informação para disponibilizá-las para seu usuário, especialmente por meio da etapa da catalogação museológica (FERREZ, 1994; CURY, 2008; BOTTALLO, 2010; DESVALLÉES; MAIRESSE, 2013; PADILHA, 2014; MENDONÇA, 2016; ZALEWSKI *et al.*, 2020).

Nesse sentido, a catalogação museológica se torna a etapa mais importante para a representação da informação, exercendo um papel fundamental na recuperação da informação nos museus, pois permite reunir informações que, de outra forma, estariam dispersas (BOTTALLO, 2010), atuando por meio de práticas direcionadas à extração de informações do objeto museológico para compreensão de um contexto (histórico, social, natural, entre outros), para que este possa ser recuperado pelo usuário posteriormente.

É na etapa da catalogação também que o profissional realiza a operação de síntese das informações por meio da relação entre as informações sobre determinado contexto do objeto museológico e seus aspectos físicos, contribuindo solidariamente para seu processo de análise de assunto para a indexação (CARVALHO; FONSECA; REDIGOLO, 2021). Portanto, é nesse momento que o profissional estabelece o

[...] resgate de informações do objeto em si, e sim, buscar através da pesquisa o contexto de produção do bem cultural, com um método capaz de permitir a construção e a comunicação do conhecimento acerca do bem cultural produzido historicamente [...] (NASCIMENTO, 1994, p. 36).

Dito isso, entende-se aqui que a análise de assunto é a primeira e mais importante etapa do processo de indexação e, segundo Fujita (2003), se divide em três fases: leitura documentária, identificação de conceitos e seleção de conceitos, tendo como finalidade a organização da informação e do conhecimento, objetivando a recuperação da informação.

Observa-se que esses elementos da análise de assunto estão presentes na catalogação museológica, a partir da representação das informações que se inicia na leitura documentária realizada com o objetivo de sintetizar as “[...] informações históricas, simbólicas e de usos do

objeto no museu [...]” (PADILHA, 2014, p. 52) e, posteriormente, na identificação e seleção de conceitos que melhor representem o conteúdo informacional do objeto museológico na ficha de catalogação.

Nesse sentido, observa-se que para se ter uma eficácia na catalogação do objeto museológico é preciso tomar decisões em relação à representação do assunto para um determinado grupo de usuários. Essa representação tem início na leitura documentária realizada com o objetivo de atribuição de termos que caracterizam o assunto do documento, não existindo a atribuição correta de termos, uma vez que um mesmo documento pode ser indexado de forma diferente dependendo do interesse de cada grupo de usuários (LANCASTER, 2004).

Ou seja, da mesma forma que os documentos textuais em uma biblioteca, as coleções museológicas constituem um documento informativo (HERNÁNDEZ HERNÁNDEZ, 2016) que continua gerando conhecimento e que, apesar de possuir aspectos complexos como a tridimensionalidade, precisam ser representados para que a informação possa ser recuperada de forma adequada pelos usuários (CERAVOLO; TÁLAMO, 2007; DAVIS, 2020).

Assim, entende-se que a análise de assunto portanto, se insere no processo de documentação museológica a partir da catalogação, visto que esta também se utiliza da leitura documentária para realizar a identificação e a seleção de conceitos para representação da informação dos objetos museológicos. Nessa perspectiva, a análise de assunto nos museus é um importante processo metodológico, pois permite estabelecer estratégias de busca e usabilidade da informação pelo usuário (CARVALHO, 2022).

No entanto, no que diz respeito às diretrizes que deem auxílio ao profissional onde buscar informações sobre o assunto de um objeto museológico, para assim determinar seus pontos de acesso nos catálogos, observa-se uma lacuna teórico-metodológica no campo da Museologia e da própria Ciência da Informação.

Dessa forma, estudos que busquem sistematizar os procedimentos metodológicos de análise de assunto no contexto museológico se tornam ainda mais relevantes para subsidiar a indexação como uma ferramenta de gestão aplicável aos sistemas de recuperação da informação museológica, sendo necessário aprofundar sobre questões relativas aos estudos teóricos-metodológicos da análise de assunto na indexação dentro do contexto museológico.

3 Procedimentos Metodológicos

Trata-se de uma pesquisa qualitativa de caráter exploratório, dividida em duas partes: a) levantamento bibliográfico sobre a análise de assunto e leitura no contexto museológico; e b) aplicação do Protocolo

Verbal Individual (PVI), para identificar as estratégias utilizadas durante a leitura documentária dos objetos museológicos.

Primeiramente, foi realizado um levantamento bibliográfico em fontes de informação da área em bases nacionais e internacionais, onde buscamos identificar as bibliografias relevantes sobre análise de assunto e leitura documentária no âmbito da Museologia e Ciência da Informação. Para a averiguação de dissertações e teses utilizamos a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) e para artigos científicos nacionais e internacionais utilizamos as seguintes bases: Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), *Scielo* e a *Web of Science*, no período de tempo de 2009 a 2022.

Observou-se, no entanto, uma dificuldade em encontrar documentos sobre a temática no que tange a área da Museologia, infere-se, portanto, que essa ainda é uma temática escassa e pouco explorada nesta área do conhecimento. Os documentos recuperados nas bases de dados foram utilizados para a composição teórica deste estudo, quando possuíam caráter pertinente para temática desenvolvida e estavam inseridos no âmbito de discussões da Ciência da Informação e Museologia que compõem os pressupostos teóricos desta pesquisa.

Posteriormente, para a observação da leitura documentária dos profissionais, foi utilizada a coleta introspectiva de dados que consiste na observação da introspecção do profissional durante a realização da tarefa. Segundo Redigolo (2010), esse método preconiza que o sujeito explicita todos os procedimentos realizados em tempo real e em voz alta sendo, portanto, uma técnica exaustiva e que proporciona uma visão ampla sobre a análise dos assuntos.

Como instrumento para coleta introspectiva de dados foi escolhido o Protocolo Verbal Individual (PVI) que permite “[...] observar os processos cognitivos, de forma a conseguir explorar experiências conscientes do sujeito [...]” (REDIGOLO, 2014, p. 126). Dito isso, para alcançar o objetivo geral desta pesquisa foi preciso adaptar esta técnica para a realidade dos museus, uma vez que se trata de uma tipologia de acervo diferente daquela encontrada em bibliotecas.

Assim, foram escolhidas a Coleção de Zoologia da Universidade Federal do Pará (UFPA), ligada ao Museu de Anatomia da UFPA, e as Coleções de Ictiologia, de Herpetologia, de Aracnídeos e de Paleontologia ligadas ao Museu Paraense Emílio Goeldi (MPEG), delimitando o local de aplicação do instrumento de coleta introspectiva de dados. Cada coleção possui um profissional responsável pela catalogação de seu acervo, dessa forma o PVI foi aplicado com cinco profissionais que possuem mais de quatro anos de experiência na atividade.

Para a aplicação do PVI foram utilizados os seguintes procedimentos metodológicos adaptados por Carvalho (2022) com base nos trabalhos de Fujita, Nardi e Fagundes (2003) e Redigolo (2014):

Tabela 1. *Procedimentos para aplicação do PVI*

<i>Etapas</i>	<i>Procedimentos</i>	<i>Objetivos</i>
Procedimentos anteriores à coleta de dados	a) Definição do universo de pesquisa	Definir onde, como e com quem será aplicado o PVI
	b) Seleção do objeto-base (adaptado)	Escolher os objetos (objetos museológicos) para realização da tarefa
	c) Seleção dos participantes	Selecionar os sujeitos que realizarão a atividade
	d) Definição da tarefa	Definir a atividade para a observação dos sujeitos participantes.
	e) Conversa Informal com os sujeitos	Repasar as instruções sobre o PVI antes de sua aplicação
Procedimentos durante a coleta de dados	a) Gravação do “pensar alto”	Gravar os sujeitos realizando sua tarefa individualmente, sem interferência do analista
	b) Entrevista retrospectiva (opcional)	Esclarecer pontos e dúvidas que podem surgir durante a aplicação do PVI
Procedimentos posteriores à coleta de dados	a) Transcrição literal das falas dos participantes	Transcrever as falas dos sujeitos participantes na íntegra
	b) Leitura detalhada dos dados coletados para a construção das categorias de análise	Leitura detalhada das transcrições com intuito de identificar aspectos relevantes para a elaboração das categorias de análise
	c) Construção das categorias de análise	Elaborar categorias de análise para facilitar a compreensão da atividade praticada pelos sujeitos participantes
	d) Retorno aos dados para retirar trechos da discussão que exemplifiquem cada categoria de análise	Reler as transcrições para a retirada de trechos importantes para análise

Fonte: Elaborado pelos autores (2023).

Observa-se que o procedimento “Seleção do texto-base” foi transformado em “Seleção do objeto-base”, visto que, em vez de um documento textual, foram utilizados

objetos museológicos para a realização da tarefa pelos profissionais com o objetivo de identificar suas estratégias e comportamento, durante o processo de leitura documentária. Esta adaptação não interferiu na coleta introspectiva de dados.

Os dados coletados a partir da leitura da transcrição dos PVI, foram organizados e agrupados em dois eixos temáticos e categorias de análise, elaborados a partir das falas de cada sujeito participante, que correspondem às estratégias utilizadas para determinar o assunto durante a catalogação dos objetos museológicos pelos profissionais, sistematizadas na tabela 2:

Tabela 2. *Estratégias para determinação do assunto do objeto museológico*

<i>Eixos</i>	<i>Categorias</i>	<i>Síntese</i>
Exploração do Objeto Museológico	Características Físicas	Identificar os elementos que fazem parte da morfologia do objeto, por meio da análise dos padrões das estruturas do espécime para categorização taxonômica
	Agente	Identificar o agente que praticou a ação
Informações Contextuais	Classificação	Identificar aspectos relacionados ao grupo taxonômico que pertence o objeto analisado
	Geografia	Identificar a procedência, lugares geográficos e localizações do objeto analisado
	Método	Descrever instrumentos, técnicas, métodos, materiais e equipamentos utilizados para coleta do objeto
	Tempo	Identificar a unidade de tempo do acontecimento

Fonte: Elaborado pelos autores (2023).

A partir da observação dos PVI foi possível identificar como os profissionais determinam o assunto de seus objetos. Nas próximas seções são apresentadas as estratégias utilizadas pelos profissionais para realizar a leitura documentária e determinar o assunto dos seus objetos museológicos e os subsídios para a criação de um modelo de leitura documentária de objetos museológicos para as coleções de zoologia.

4 Estratégias para determinação do assunto do objeto museológico

Diferente de um documento textual que possui partes de sua estrutura onde o assunto pode ser identificado explicitamente (título, subtítulo, resumo, introdução, entre outros), o objeto museológico não possui uma estrutura tão explícita, variando de tipologia de objeto.

No entanto, quando se trata de um acervo zoológico, é possível determinar do que se trata o objeto por meio da análise de sua estrutura física com auxílio da literatura da área, permitindo classificá-lo em categorias taxonômicas.

A partir dessa classificação, informações sobre o contexto são importantes, pois permitem que o profissional realize questionamentos como um recurso de inferência dos seus aspectos temáticos. Estes questionamentos são fundamentais e podem ser adotados como parâmetros para o auxílio ao profissional na leitura e representação documentária para contemplar o conteúdo temático do objeto museológico.

Nesse sentido, foi possível perceber que para os objetos museológicos de coleções zoológicas os questionamentos seguiram uma estratégia comum para determinar o assunto entre todos os sujeitos da pesquisa. Primeiramente por meio da análise das características físicas do objeto museológico (morfologia) e comparação com a literatura pertinente, para determinar a que grupo taxonômico pertence. Em seguida, por informações referentes ao contexto como a localização geográfica, grupo taxonômico, instrumentos de coleta, entre outros. Juntos esses dois aspectos se tornam fontes para a descrição temática do objeto museológico das coleções zoológicas.

Verificou-se, portanto, que os sujeitos possuem uma abordagem sistemática para identificação do conteúdo temático, essencial para recuperação da informação pelo usuário, realizando dois tipos de análise: a física e a contextual, para determinar o assunto para fins de recuperação. Assim, para além da análise física, a informação museológica exige um processo flexível de interpretação do contexto (seja ele histórico, natural, de uso ou simbólico) em que o objeto se encontra.

Considerando que a identificação de conceitos é o objetivo da leitura documentária e a operação mais importante da indexação, os resultados obtidos nos levam a concluir que a análise de assunto depende do uso combinado da exploração da estrutura do objeto e da interpretação das informações contextuais. No entanto, cabe destacar a falta de sistematização da leitura documentária no contexto museológico (CARVALHO, 2022), sendo necessário pensar na organização desses procedimentos para compreensão do conteúdo temático do objeto museológico.

Destaca-se também que as estratégias utilizadas pelos profissionais para identificar os conceitos possuem qualidades metodológicas que podem ser aplicadas na instrumentalização da leitura documentária para essa tipologia de objeto, permitindo que o profissional reconheça quais os pontos de acesso que facilitam a identificação dos temas para análise de assunto com fins de indexação, considerando as suas especificidades e peculiaridades.

Portanto, consta, a seguir, os procedimentos para a leitura documentária dos objetos museológicos de coleções zoológicas, como subsídios para criação de um modelo de leitura documentária.

5 Procedimentos para leitura documentária dos objetos museológicos de coleções zoológicas

Os resultados obtidos a partir da investigação das estratégias para determinar o assunto do objeto museológico de coleções zoológicas indicaram subsídios para a sistematização dos procedimentos para a leitura documentária.

A leitura documentária, realizada pelo profissional de museu na análise de assunto visa propiciar a identificação e seleção de conceitos para posterior representação da informação museológica nos catálogos da instituição, de forma que satisfaçam a demanda do usuário.

Verificou-se que o processo para determinar o assunto é fundamentado por procedimentos que consideram as características físicas, as estratégias metacognitivas e a abordagem sistemática, por meio de questionamentos, com objetivo de localizar o objeto em um determinado contexto.

Foi possível, portanto, observar que, embora o trabalho do profissional de museu não obedeça a uma sistemática para realizar a leitura documentária para determinar o assunto do objeto, contemplada por um manual por exemplo, os profissionais analisados possuem aspectos comuns para realizar a representação temática e organizar os seus acervos.

Apresentam-se os procedimentos para leitura documentária com base nas etapas da análise de assunto para indexação propostas por Fujita (2003) e nas observações do PVI:

Tabela 3. *Procedimentos para análise de assunto de objetos museológicos de coleções zoológicas*

<i>Procedimentos</i>	<i>Atividade realizada durante o processo de catalogação</i>
Compreensão do objeto museológico	Análise das características físicas e comparação com a literatura para determinar o grupo taxonômico a que pertence o objeto museológico analisado
Identificação dos conceitos	Identificação dos conceitos por meio de questionamentos sobre o contexto do objeto (histórico, natural, geográfico, de uso ou simbólico).
Seleção dos conceitos para recuperação	Determinação dos conceitos que representarão os conteúdos temáticos na ficha de catalogação do objeto museológico.

Fonte: Elaborado pelos autores (2023).

Os procedimentos propostos acima refletem uma abordagem sistemática de análise de assunto do objeto museológico, de acordo com sua tipologia, que levam em conta as estratégias cognitivas utilizadas para a leitura documentária, pois tem como base os resultados das observações com PVI realizadas por profissionais proficientes.

Assim, os procedimentos de leitura propõem uma sequência de análise que começa pela compreensão do objeto museológico, por meio da análise de suas características físicas com o objetivo de classificá-lo em um grupo taxonômico. A segunda etapa, após essa classificação, passa a abordá-lo de uma forma mais lógica a fim de expressar os termos que melhor o representam, identificando o contexto que o objeto museológico está inserido, de forma a representar e propiciar a seleção dos conceitos válidos para recuperação. Por fim é realizada a seleção dos conceitos julgados relevantes para o acesso à informação, tendo em vista os objetivos da instituição. Cabe ao profissional da informação definir os elementos de descrição que melhor atenderão ao contexto de representação.

Foi possível identificar, portanto, o percurso da leitura do objeto museológico realizada pelo profissional para a representação da informação durante o processo de catalogação. Observa-se que para determinar o assunto dessa tipologia de documentos a leitura documentária feita pelo profissional abrange estratégias de exploração da estrutura do objeto e de elementos contextuais na tentativa de localizar os conceitos pertinentes, sendo uma tarefa exaustiva.

Portanto, pode-se concluir que, apesar da complexidade dos suportes que uma coleção zoológica pode possuir, entende-se que existe a necessidade de representação temática para uma recuperação eficiente pelo usuário e, portanto, subsídios para criação de um modelo de leitura documentária que abranja essa tipologia de acervo.

Dessa maneira, reitera-se sobre a responsabilidade que o profissional de museu deve ter para determinar o assunto dos seus objetos museológicos. Acredita-se, dessa forma, que a adoção de procedimentos metodológicos padronizados para a realização de uma leitura documentária sistematizada, devidamente registrados em um manual, possibilitará aos profissionais de museu a execução dessa atividade de forma mais eficiente e rápida.

Por fim, podemos dizer que existem subsídios para sistematizar esses procedimentos por meio de um modelo de leitura documentária que permita nortear as estratégias cognitivas de exploração do objeto e a identificação e seleção de conceitos para a recuperação da informação pelo usuário.

Cabe destacar que o foco da pesquisa não foi a etapa de tradução, visto que essa ocorre após a leitura documentária e não nos detemos neste aspecto.

Contudo, consideramos importante destacar que essa pode ser uma temática para um desdobramento futuro para essa pesquisa.

Diante do exposto, apresentam-se na seção seguinte as considerações finais a respeito da pesquisa.

6 Conclusão

Observa-se que ainda existem lacunas que precisam ser preenchidas no que diz respeito ao tema da representação temática quando se trata do contexto museológico visto que boa parte da literatura da área da Museologia se volta mais para descrição física dos objetos museológicos. Nesse sentido, a temática da análise de assunto no ambiente museológico ainda carece de literatura e necessita de maiores investigações para tratar de suas especificidades.

No que diz respeito a leitura documentária específica para o objeto museológico de coleções zoológicas, foi possível observar que, para além das diferentes características físicas, essa tipologia de objeto contém características ligadas ao contexto que delineiam seu significado, gerando questionamentos e estratégias que permitem a identificação e seleção dos conceitos para a sua representação nos catálogos.

Nessa perspectiva, portanto, identificar as estratégias e procedimentos para leitura documentária dos objetos museológicos de coleções zoológicas contribui para a ampliação dos estudos sobre representação da informação museológica, com o intuito de acrescentar novos olhares à temática. Assim, esperamos contribuir para o avanço da temática no campo da Museologia e Ciência da Informação.

Referências

- BOTTALLO, M. Diretrizes em Documentação Museológica. *In: Documentação e Conservação de Acervos Museológicos*: Diretrizes. São Paulo: Governo de São Paulo; ACAM Portinari, 2010. p. 48-79.
- CARVALHO, R. A. **A análise de assunto no contexto museológico**: aplicação do protocolo verbal individual nos museus da UFPA e MPEG. 2022. 134 f. Dissertação (Mestrado em Ciência da Informação) - Instituto de Ciências Sociais Aplicadas, Universidade Federal do Pará, Belém, 2022.
- CARVALHO, R.; FONSECA, L. D. P. A.; REDIGOLO, F. M. Diálogo sobre indexação na Biblioteconomia e Museologia. **Informação em Pauta**, Fortaleza, v. 6, 2021. p. 1-24.
- CERAVOLO, S. M.; TÁLAMO, M. F. G. M. Os museus e a representação do conhecimento: uma retrospectiva sobre a documentação em museus e o processamento da informação. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 8., 2007, Salvador. **Anais** [...], Salvador: UFBA, 2007.

- CURY, M. X. **Exposição: concepção, montagem e avaliação.** 2 ed. São Paulo: Annablume, 2008.
- DAVIS, A. Defining Museum. **ICOFOM Study Series**, n. 48-2, 2020. p. 85-94.
- DESVALLÉES, A.; MAIRESSE, F. Musealização. *In*: DESVALLÉES, A.; MAIRESSE, F. (ed). **Conceitos-Chave de Museologia.** São Paulo: Comitê Brasileiro do Conselho Internacional de Museus; Pinacoteca do Estado de São Paulo; Secretaria de Estado da Cultura de São Paulo, 2013.
- DIAS, E. W.; NAVES, M. M. **Análise de assunto: teoria e prática.** 2. ed. Brasília: Briquet de Lemos, 2013.
- FERREZ, H. D. **Documentação museológica: teoria para uma boa prática.** Rio de Janeiro: Minc/Iphan, 1994.
- FUJITA, M. S. L. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**, [s. l.], v. 1, n. 1, 2003. p. 60-90.
- FUJITA, M. S. L.; NARDI, M.I.A.; FAGUNDES, S. A. Observing documentary reading by verbal protocol. **Information Research-an International Electronic Journal**, Sheffield, Univ Sheffield Dept Information Studies, v. 8, n. 4, 31 p., 2003.
- FUJITA, M. S. L.; GIL LEIVA, I; BOCCATO, V. R. C.; INÁCIO, M de O; GUIM, V. L. R; PIOVEZAN, L. B. Procedimentos de Indexação. *In*: FUJITA, M. S. L. (Coord.). **Manual de política de indexação para as bibliotecas universitárias da Unesp.** São Paulo: Unesp, 2017.
- HERNÁNDEZ HERNÁNDEZ, F. Documentary sources of museology: reflections and perspectives. **ICOFOM Study Series**, [s. l.], n. 44, 2016. p. 81-93.
- LANCASTER, F. W. **Indexação e resumos: teoria e prática.** 2. ed. Brasília: Briquet Lemos, 2004.
- MENDONÇA, E. C. Processos de patrimonialização e musealização no âmbito do Programa Nacional de Patrimônio Imaterial: Desafios e potencialidades para a salvaguarda de bens registrados (Brasil). *In*: SEMINÁRIO PATRIMÔNIO, MUSEUS E DESENVOLVIMENTO. Porto, 2016. **Anais** [...], Porto: Universidade do Porto, n. 6, 2016. p. 50-66.
- MENSCH, P. V. **O objeto de estudo da museologia.** Rio de Janeiro: UNIRIO/UGF, 1994.
- NASCIMENTO, R. A. Documentação museológica e comunicação. **Cadernos de Museologia**, [s. l.], n. 3, 1994. p. 31-39.
- PADILHA, R. C. **Documentação Museológica e Gestão de Acervo.** Florianópolis: FCC, 2014.
- REDIGOLO, F. M. **O processo de análise de assunto na catalogação de documentos: a perspectiva sociocognitiva do catalogador em contexto de biblioteca universitária.** 2010. 176 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.
- REDIGOLO, F. M. **O processo de análise de assunto na catalogação de livros em bibliotecas universitárias: aplicação do protocolo verbal.** 2014. 262 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.
- SOUZA, D. M. V. Museus de Ciência, divulgação científica e informação: reflexões acerca de ideologia e memória. **Perspectivas em Ciência da Informação**, [s. l.], v. 14, n. 2, 2009. p. 155-168.
- ZAWLESKI, M. P.; NASCIMENTO, M. M.; SILVEIRA, C. M. S.; LEAL, N. M. P. M. Documentação do museu de ciências naturais carlos ritter: o início de um trabalho. *In*: SEMANA DOS MUSEUS DA UFPEL, 2020, Pelotas. **Anais** [...]. Pelotas: Ed. UFPel, 2020. p. 225-23.

Proposta de classificação para promoção da competência em dados de pesquisa

A classification proposal to promote a research data literacy

Marcelle Costal (1), Luana Sales (2) Marianna Zattar (3)

(1) Fundação Getúlio Vargas (FGV), Rio de Janeiro, RJ, Brasil; Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Brasília, DF, Brasil, marcelle.costal@fgv.br. (2) Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brasil, luanasales@ibict.br. (3) Faculdade de Administração e Ciências Contábeis (FACC), Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, RJ, Brasil, mzattar@facc.ufrj.br.

Resumo

Objetivo: Apresenta as ações de promoção de competência em dados oferecidas por bibliotecas universitárias dos Países Baixos e Reino Unido, visando sistematizar o seu conteúdo e contribuir às ações de competência em dados promovidas por bibliotecários. **Metodologia:** Possui uma abordagem qualitativa a partir de um método exploratório e a criação de categorias. Utiliza para a análise os critérios de Bardin (1977) por meio de um levantamento das ações de competência em dados em ambiente web a partir dos materiais, workshops e treinamentos disponíveis nos portais das bibliotecas no período de outubro de 2020 e março de 2021. **Resultados:** Propõe uma categorização dos assuntos dos conteúdos abordados nas ações de competência em dados de acordo com o quadro de Risdale (2015), uma classificação que expresse os fundamentos da competência em dados e outra para as ações de competência em dados como um processo da Gestão de dados de Pesquisa tendo como princípio os critérios encontrados nos cânones do método analítico sintético de Ranganathan (1967). O cânon que representa esta sistematização é Cânon para Renques, especificamente o Cânon de Sequência Útil conectado aos Princípio Posterior no Tempo. **Conclusão:** Conclui que o escopo da competência em dados está para além da prática e do uso das ferramentas, mas inclui uma dimensão histórica do processo, do seu papel e uma outra, a dimensão crítica dos processos, dos propósitos, das ferramentas, do uso de dados de pesquisas e a sua relação com a sociedade.

Palavras-chave: Competência em informação. Competência em dados. Gestão de dados de pesquisa. Biblioteca universitária. Ciência da Informação.

Abstract

Objective: This study presents actions to promote data literacy offered by university libraries in the Netherlands and the United Kingdom, aiming to systematize their content and contribute to data literacy actions promoted by librarians. **Method:** It has a qualitative approach based on an exploratory method and the creation of categories. Bardin's criteria (1977) are used for the analysis, through a survey of data literacy actions in a web environment based on the materials, workshops and training available on library portals in the period from October 2020 to March 2021. **Result:** It proposes a categorization of the subjects of the content addressed in the data literacy actions according to Risdale's framework (2015), a classification that expresses the fundamentals of data literacy and another for data literacy actions as a data management process. Research data based on the criteria found in the canons of the synthetic analytical method of Ranganathan (1967). The canon that represents this systematization is the Canon for Ranks, specifically the Canon of Useful Sequence connected to the Late Principle in Time. **Conclusions:** It ends that the scope of data literacy goes beyond the practice and use of tools, but includes a historical dimension of the process, its role and another, the critical dimension of processes, purposes, tools, use of data research and its relationship with society.

Keywords: Information literacy. Data literacy. Research data management. University library. Information Science.

Resumen

Objetivo: Presenta acciones para promover la competencia en datos que ofrecen las bibliotecas universitarias de los Países Bajos y el Reino Unido, con el objetivo de sistematizar su contenido y contribuir a las acciones de competencia en datos promovidas por los bibliotecarios. **Metodología:** Tiene un enfoque cualitativo basado en un método exploratorio y la creación de categorías. Para el análisis se utilizan los criterios de Bardin (1977), a través de un relevamiento de acciones competenciales en datos en ambiente web a partir de los materiales, talleres y capacitaciones disponibles en los portales bibliotecarios en el período de octubre de 2020 a marzo de 2021. **Resultados:** Propone una categorización de los sujetos de los contenidos abordados en las acciones de competencia de datos según el marco de Risdale (2015), una clasificación que expresa los fundamentos de la competencia de datos y otra para las acciones de competencia de datos como proceso de gestión de datos de investigación basado en los criterios encontrados en los cânones del método

analítico sintético de Ranganathan (1967). El canon que representa esta sistematización es el Canon de Rangos, específicamente el Canon de Secuencia Útil conectado al Principio Tardío en el Tiempo. **Conclusión:** Se concluye que el alcance de la competencia en datos va más allá de la práctica y uso de herramientas, sino que incluye una dimensión histórica del proceso, su rol y otra, la dimensión crítica de los procesos, propósitos, herramientas, uso de los datos de investigación y su relación con la sociedad.

Palabras clave: Competencia en información. Competencia en datos. Gestión de datos de investigación. Biblioteca universitaria. Ciencia de la Información.

1 Introdução

A produção científica desenvolvida e difundida em suporte digital trouxe à luz questionamentos com particularidades do próprio campo. Este contexto anunciou a necessidade de desenvolver não só instrumentos e ferramentas de gestão que orientem a geração de dados de pesquisa e coleções de dados a partir do seu uso, armazenamento, processamento, análise, compartilhamento e reuso, mas, também, o desenvolvimento de habilidades em informação no uso dessas tecnologias, para a gestão de dados nos mais variados contextos disciplinares, bem como na reflexão crítica de sua natureza.

De acordo com Vilar e Zabukovec (2019) pela crescente necessidade na Europa de prestar contas de seus dados aos financiadores, os pesquisadores seguem avançando na questão e acabam por refletir práticas de Gestão de Dados de Pesquisa (GDP) muito pessoais que trazem dificuldades de interpretação dos dados, dificultando o reconhecimento e o reuso por parte da comunidade científica.

Os pesquisadores precisam gerenciar os seus dados e não sabem como, uma vez que dificilmente recebem orientações, ajuda, ou os serviços prestados não atendem às suas demandas e expectativas fazendo com que não se aprofundem sobre os princípios da GDP (SCHUMACHER; VANDECREEK, 2015).

Borgman (2010), aponta que a biblioteca se tornou o lugar mais provável para a responsabilidade sobre os dados de pesquisa e a sua hospedagem e o desenvolvimento de um conjunto de habilidades e práticas voltadas para a competência de seu uso, compartilhamento, preservação e reuso.

Na mesma perspectiva, Carlson *et al.* (2011) relatam que os profissionais da informação, em especial os bibliotecários, começaram a entender a necessidade de desenvolver meios que auxiliem as pessoas a superar os novos desafios trazidos pela tecnologia referentes à competência em dados.

A gestão de dados e informação é competência do bibliotecário, à medida que as bibliotecas por já desenvolverem a competência em informação, estão bem-posicionadas também para desenvolverem a competência em dados (CALZADA PRADO; MARZAL, 2013). No entanto, para Buys e Shaw (2015, p. 20-21) um dos maiores desafios enfrentados pelos bibliotecários é “conscientizar pesquisadores

universitários de que a biblioteca pode fornecer orientação no gerenciamento de dados.”

Em resposta a *e-Science* e a crescente necessidade de usar dados, percebe-se que as bibliotecas universitárias estão contratando equipes especializadas tais como a figura do *data librarian* (bibliotecários de dados) ou *data specilists* (especialistas em dados), ou ainda treinando bibliotecários de referência para conduzirem o processo da GDP, envolvendo-se na coleta de dados, na compilação e acesso a fontes de dados (CALZADA PRADO; MARZAL, 2013).

Witt *et al.* (2009) ao traçarem o perfil e os dados enviados por pesquisadores para informar sobre como a biblioteca universitária orienta sobre dados de pesquisa, observaram que tanto o corpo docente e o corpo discente careciam de habilidades em Gestão e Curadoria de Dados. Ao passo que a pesquisa orientada a dados se torna uma crescente, conhecer o processo de GDP a torna fundamental para o pesquisador. Whitmire (2015) aborda que apesar desta crescente demanda por uma educação formal em GDP, ainda é incipiente em bibliotecas universitárias, embora os bibliotecários estejam liderando um papel de suporte no fornecimento de treinamento e apoio na GDP para professores e alunos de pós-graduação.

Por sua vez, de acordo com Tartarotti, Dal’avedove e Fujita (2019) ainda é incipiente no Brasil a figura do “bibliotecário de dados”, e ainda tímidas as iniciativas de suporte à pesquisa orientada por dados, bem como no desenvolvimento da competência em dados.

Existe atualmente uma demanda por suporte em GDP com a intenção de promover a competência em dados ao pesquisador. No entanto, não existe um modelo padrão de conteúdo para essas iniciativas. O que se observa é que cada biblioteca oferta um determinado conteúdo, que deve ser condizente com o objetivo do discente e da instituição que está ofertando (WHITMIRE, 2015).

Mas será possível obter um núcleo mínimo de conteúdo que possa ser replicado nas iniciativas desenvolvidas por bibliotecas no suporte a GDP?

Se sim, quais seriam os temas a serem utilizados de forma introdutória para ampliação da competência em gestão de dados?

Na consecução dos objetivos deste trabalho pretende-se investigar as ações de promoção de competência em dados oferecidas por bibliotecas universitárias, visando sistematizar o seu conteúdo, para que possa contribuir às

ações de competência em dados promovidas por bibliotecários. Observa-se por meio dos treinamentos ofertados que tipo de conteúdo são abordados nestas ações para que possam ser sintetizados. Para a execução da investigação, o recorte selecionado foram as bibliotecas universitárias dos Países Baixos (Holanda) e do Reino Unido, de universidades que sejam credenciadas pela *European University Association* (2020).

A metodologia concentra-se no mapeamento de conteúdo, tendo como ponto de partida as iniciativas promovidas por bibliotecas em competência em dados, encontradas nos portais institucionais e/ou a partir do contato prévio, solicitando às respectivas instituições os materiais utilizados nos treinamentos e *workshops* encontrados no processo de investigação. Ao final será proposto, por meio dos conteúdos levantados, uma categorização e classificação das temáticas para que possam contribuir às ações de competência em dados promovidas por bibliotecários aos pesquisadores. Sublinha-se que esta investigação propõe estimular o debate inicial sobre uma estrutura geral para a GDP, deixando em aberto a variação entre os domínios.

2 Gestão de Dados de Pesquisa e Competência em Dados

São processos da GDP tudo aquilo que se refere à manipulação, organização, documentação e agregação de valor no tocante a facilitar o compartilhamento dos dados. Eynden *et al.* (2011) destacam que a gestão dos dados de pesquisa é um trabalho em equipe entre os pesquisadores, principais responsáveis por gerenciar os dados de pesquisa, e as Instituições, por oferecerem apoio técnico e recursos humanos para as orientações sobre o ato de planejar as etapas anteriores e posteriores à criação dos dados, e domínio sobre as ferramentas de infraestrutura.

Os dados de pesquisa são dependentes de uma interpretação para que possam ser reusados em outros contextos. A garantia da compreensão no futuro está na gestão de dados que se faz por meio da atribuição de metadados e da *linkagem* com outros documentos que fornecem o contexto aos dados. A gestão de dados envolve também estudos acerca dos aspectos de segurança, preservação, acesso, compartilhamento, reuso, citação, dentre outros. São dependentes de ações planejadas que conferem integridade aos dados em sua oferta e acesso contínuo para serem descobertos e utilizados posteriormente.

Para tal, é imprescindível que os pesquisadores conheçam a comunidade a qual os dados se originam, o histórico da pesquisa, e o ciclo de vida dos dados de pesquisa. É conveniente identificar quais são os atores responsáveis por cada etapa de geração e transformação dos dados de pesquisa e quais são os papéis definidos e

acordados no processo de pesquisa. Os pesquisadores não estão sozinhos nesta empreitada, e podem contar com o apoio de bibliotecários, arquivistas, gestores em tecnologia, cientistas de dados, dentre outros. Saber localizar dados no tempo e no espaço à luz de um contexto oportuniza padrões adequados na digitação, na seleção dos formatos e documentação para a recuperação da informação (ABBOTT, 2008). A citação de autoria com referências e identificadores persistentes permite individualizar os dados, identificá-los, garantindo o reconhecimento da autoria e viabilizando as condições de reprodutibilidade dos dados e da pesquisa pela comunidade científica, possibilitando assim a oportunidade para a continuidade das pesquisas e para a construção coletiva da ciência.

Impulsionados pelos financiadores, os Estados Unidos, Reino Unido e Canadá estão exigindo cada vez mais dos pesquisadores dados de pesquisa abertos, o que requer o cultivo de novas responsabilidades com a gestão desses dados. De acordo com Koltay (2015), os bibliotecários se envolvem com a educação em dados para que estudantes e pesquisadores se tornem cientistas com conhecimento em dados.

Na perspectiva da pesquisa científica a competência em dados é um tipo de educação que direcionada aos pesquisadores constrói diálogos, treinamentos e reflexões sobre o uso, descoberta e toda a infraestrutura tecnológica, cultural, social e econômica orientada aos dados de pesquisa. Esta fase de entendimento sobre o que seria a competência em dados, bem como qual a sua importância, principalmente no contexto da pesquisa na segunda década do século XXI, marca a necessidade de discutir sobre o desenvolvimento de habilidades necessárias a dados de pesquisa, métodos e formatos.

Em pesquisa recente Ghodoosi *et al.* (2023) abordam que o termo *data literacy*, competência em dados, é recente na literatura e data do trabalho de Schield (2004) a partir das influências dos termos *information literacy* e *statistical literacy*. A emergência do termo encontra-se no bojo da crescente proliferação das tecnologias da informação e comunicação em organizações empresariais.

Contudo, ao longo das últimas décadas a inteligência artificial e as demandas sociais estão impulsionando a *data literacy* como um campo com contornos próprios, mas estabelece uma relação estreita com a competência em informação, uma vez que no escopo da competência em informação está uma expansão lógica e conceitual que caminha para a curadoria e gestão de dados (CARLSON *et al.*, 2013). Para Koltay (2015), o termo *data literacy* é definido em sua relação com a *information literacy* por meio do desenvolvimento de habilidades ao longo da formação em pesquisa.

Em busca de uma definição, Martin (2014, p. 1) ao citar a *wikipedia* aponta que a *data literacy*, é a “habilidade

de ler, criar e comunicar dados como informação” e que os pesquisadores estão sendo chamados a desenvolver tais habilidades e práticas para expandir os serviços ofertados pelas bibliotecas. No esboço de uma definição programática do conceito de *data information literacy* na tabela 1, tem-se as doze competências a serem desenvolvidas na GDP.

Tabela 1. *As 12 competências da data information literacy*

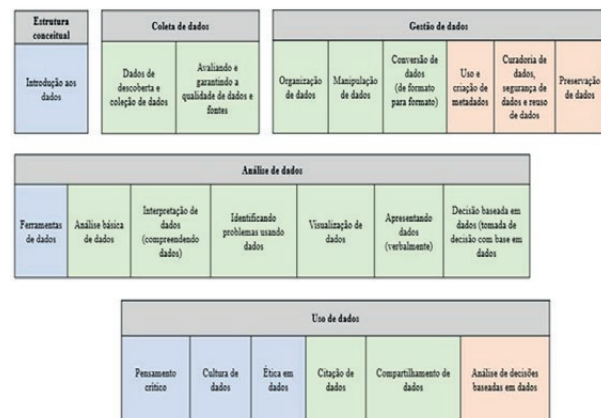
Processamento e análise de dados	Curadoria e reuso de dados
Gestão e organização de dados	Conversão e interoperabilidade de dados
Preservação de dados	Visualização de dados
Bases de dados e formato de dados	Descoberta e aquisição de dados
Ética e citação de dados	Metadados e descrição de dados
Qualidade e documentação de dados	Prática cultural

Fonte: Adaptado de Carlson *et al.* (2011; 2013; 2015, tradução nossa).

Para Koltay (2016), a competência em dados está descrita em três fases. A primeira fase está no uso das TICs, em parceria com especialistas de informação ou gerentes de dados, arquivistas e bibliotecários. A segunda fase é a conscientização sobre a criação, organização, validação, compartilhamento, armazenamento e curadoria de informações e/ou dados, a compreensão dos requisitos legais, éticos e de segurança, bem como a de metadados. A terceira fase consiste em aconselhar e educar os pesquisadores que possuem pouca experiência no assunto, estudantes e funcionários sobre todas as fases do processo e a necessidade de estar atento às mudanças.

Inclui-se a definição dada por Risdale, *et al.* (2015, p. 8) ao qual a “[...] competência em dados é a capacidade de coletar, gerenciar, avaliar e usar dados de maneira crítica [...]”. Na figura 1 apresenta-se a matriz para a competência em dados dividida pelos autores, em azul por ações relativas a competências conceituais, em verde, por competências essenciais e, em rosa, por competências avançadas.

Figura 1. *Matriz de competência em dados*



Fonte: Adaptado de Risdale, *et al.* (2015, p. 3, tradução nossa).

Destaca-se na Figura 1 que nas ações de competência conceitual em dados, pretende-se apresentar o conhecimento sobre os dados, a compreensão sobre os seus usos e empregos, problemas, desafios e o conhecimento sobre as suas ferramentas. Estas ações circundam o exercício do pensamento crítico, para entender a cultura, ou melhor, o ambiente e o histórico sobre os dados, ou o porquê, onde, como e quando utilizar dados.

Igualmente, pensar de forma crítica como trabalhar com dados, contribui para definição dos alicerces de uma proposta ética em dados em uma pesquisa orientada à dados, como por exemplo em saber padrões de citação correta de dados em determinado domínio e como compartilhá-los de maneira legal. São propostas de cunho teórico e prático, pensando no desenvolvimento de habilidades e métodos que estimulem o pensar criticamente, no reconhecimento da importância dos dados, e no oferecimento de “[...] suporte a um ambiente que promova o uso crítico de dados para aprendizagem, pesquisa e tomada de decisão em dados [...]” (RISDALE, *et al.*, 2015, p. 38).

Isto requer o comprometimento com as tarefas em dados e a colaboração mútua entre educadores, estudantes, desenvolvedores, pesquisadores, instituições, organizações, consumidores da informação em dados estruturados, dentre outros agentes da sociedade.

Risdale *et al.* (2015) apontam que há um esforço por parte das instituições de graduação e pós-graduação em incluir a competência em dados, ou noções sobre dados, dentro de seus currículos, de forma que a educação em dados possa ser significativa ao indivíduo. Neste sentido, uma proposta de ação de competência em dados requer a utilização de dados do mundo real, no desenvolvimento de cursos, disciplinas, e a intenção de um processo de aprendizagem ao longo da vida, ou seja, para além dos espaços tradicionais de aprendizagem.

3 Metodologia

A metodologia aplicada nesta pesquisa teve uma abordagem qualitativa, realizada a partir de um método exploratório e dedutivo. Configura-se como exploratória à medida que seu objetivo principal foi investigar as ações de promoção de competência em dados oferecidas em bibliotecas universitárias dos Países Baixos e Reino Unido.

3.1 Delimitação do campo

O campo de pesquisa configurou-se em ambiente *web*. As bibliotecas universitárias foram delimitadas no decorrer da investigação, por instituições de ensino superior de países europeus, que sejam referência na promoção de ações de competência em dados, no apoio e suporte aos pesquisadores acadêmicos.

A justificativa para o recorte ser a partir de instituições europeias se deu pela sua característica intrínseca de olhar para o pesquisador e para os desafios e oportunidades, trazidos pelas novas formas de coleta, armazenamento, manipulação e compartilhamento de dados de pesquisa, no movimento do acesso aberto. Um exemplo marco deste direcionamento está no relatório “*Science as an Open Enterprise: Open data for Open Science*” da *Royal Society* (2012), o qual aborda a forma como pesquisadores, universidades, agências de fomento, dentre outros organismos precisam se adaptar aos novos processos com a tecnologia concernente aos dados de pesquisa. Este relatório organiza uma pequena definição para os termos mais utilizados no “saber fazer” pesquisa em dados, que ainda geram dúvidas na comunidade científica.[1]

A União Europeia também possui o *General Data Protection Regulation* (GDPR) (2016), um regulamento geral sobre a proteção de dados que entrou em vigor em 2018 e orienta a pesquisa de dados nos países membros. Isto impactou diretamente no financiamento da pesquisa e na prestação de contas sobre a gestão dos dados de pesquisa e por conseguinte na demanda de apoio por parte dos pesquisadores.

Ao olhar para a Comissão Europeia e as exigências impostas pelos projetos financiados por ela, dentro do programa *Horizon 2020*, que teve por base anterior o projeto piloto *Open Aire 2015*, e do programa *Horizon* que o substituiu, como a criação do Plano de Gestão de Dados (PGD) de pesquisa alinhados à gestão de dados FAIR, salienta-se o compromisso da comunidade europeia com o compartilhamento de dados, mas também garantindo o resguardo necessário aos pesquisadores.

Deste modo, optou-se por investigar as bibliotecas universitárias do Reino Unido e Países Baixos, por estarem em fase de desenvolvimento mais avançado, se comparado a outros países europeus, sobre processos que envolvam a GDP e a competência em dados nas

instituições que promovem a pesquisa. O objetivo não é traçar um comparativo entre as iniciativas dos dois países referente ao apoio da biblioteca no processo da GDP e sim mapeá-los a fim de compreender em que esferas podem ser observados.

Como justificativa para escolha dos dois países, cabe ressaltar que o Reino Unido desenvolve desde 2005 o *Digital Curation Centre* (DCC), que extrapolou o suporte à pesquisa em instituições do Reino Unido e se tornou uma referência mundial. O DCC se propõe a ajudar nos desafios da gestão e curadoria de dados de pesquisa, bem como na construção da capacidade e habilidades para GDP.

Por sua vez, os Países Baixos, têm o financiamento do Ministério da Ciência ao GO FAIR, uma iniciativa global, também financiada pelos Ministérios da Ciência da Alemanha e França. O GO FAIR, tem um escritório em Leiden, e vem impulsionando, de maneira aberta, redes de implementação em instituições no mundo todo, a partir da adoção de princípios orientadores para permitir que dados sejam dados encontráveis, acessíveis, interoperáveis e reusáveis. Estes princípios foram discutidos primeiramente nos Países Baixos, por meio de um debate proposto pelo *Netherlands eScience Center and the Dutch Techcentre for the Life Sciences* (DTL) no *Lorentz Center* em Leiden. Desta discussão, foi delineada a *Joint Declaration Data Citation Principles* (JDDCP) do *Force 11*, que definiu o ecossistema FAIR, como a terminologia *Findable, Accessible, Interoperable e Reusable* (FAIR).

Os Países Baixos também contam com uma plataforma nacional de gestão de dados, a *Landelijk Coördinatiepunt Research Data Management* (LCRDM) (2019). Esta iniciativa é uma rede nacional de especialistas, a qual algumas universidades são adeptas, no domínio de investigação da GDP. Sua proposta é ser um elo entre a política de gestão de dados e a prática diária, já que exige nacionalmente uma prática de ação e coordenação conjunta.

Ademais, em uma pesquisa realizada em março de 2019, nas bases referenciais Scopus e WoS, identificou-se que o Reino Unido e os Países Baixos eram países de destaque na composição do descritor “*data literacy*” (COSTAL, 2019).

Para facilitar o processo de busca por bibliotecas universitárias do Reino Unido e dos Países Baixos, o recorte foi feito a partir do diretório de membros da *European University Association*, no ano de 2020 [2]. A escolha por esta instituição, tem por fator determinante o escopo categoria de membros, que inclui universidades, associações europeias ou outros organismos ativos no ensino superior ou da pesquisa.[3] *A European University Association* tem um diretório, disponível online, com uma lista completa de membros por país e seus respectivos organismos.

Deste modo, foram selecionadas as universidades do Reino Unido e Países Baixos que compõem a condição de membro na *European University Association* (2020). Após a delimitação das instituições, cada universidade foi acessada em seus endereços eletrônicos, que estavam disponíveis na *European University Association* (2020). O acesso aos portais das universidades ocorreu entre os meses de outubro a março de 2021, à procura das suas respectivas bibliotecas universitárias e ações para GDP.

3.2 Método de Análise

O estudo da análise das descrições dos portais, do conteúdo dos portais, materiais disponíveis e a construção de categorias foi orientado pelos princípios da homogeneidade, exaustividade, exclusividade, objetividade e pertinência (BARDIN, 1977). De acordo com Bardin (1977, p. 117, grifo nosso)

[...] as categorias são rubricas ou classes, as quais reúnem um grupo de [...] [unidades de registro] sobre um título genérico, agrupamento esse efetuado em razão dos caracteres [semânticos] em comum destes elementos.

Conforme Bardin (1977) a organização da análise do conteúdo das páginas e dos materiais disponibilizados realizou-se em três etapas:

- a) Pré-análise: que se refere a delimitação a partir da leitura flutuante das informações disponíveis online nos portais e da escolha do corpus documental. O universo documental foi delimitado à priori e são as ações de competência em dados promovidas pelas bibliotecas universitárias das instituições em questão no apoio à pesquisa. A preparação do material foi feita a partir de resumos que sintetizaram os assuntos encontrados, a fim de gerar indicadores de unidades de categorização para análise temática;
- b) a exploração do material: extraiu-se do resumo unidades temáticas do assunto;
- c) o tratamento e inferência dos resultados: a partir das informações obtidas, as unidades temáticas do assunto, foram agrupadas em categorias com seus respectivos eixos-temáticos. A categorização teve por função oferecer uma representação sintetizada dos conteúdos encontrados nos portais. Teve por fundamento a estrutura da gestão de dados apresentada no desenvolvimento do referencial teórico a partir de Risdale *et al.* (2015).

As categorias foram cunhadas a partir das ações em competência em dados, por intermédio do agrupamento de conteúdos sobre os *workshops* e guias de treinamento oferecidos por bibliotecas das instituições holandesas e britânicas. Por fim, com base na categorização dos assuntos e nos Cânones do método analítico sintético de Ranganathan (1967), foi proposto como produto desta pesquisa, uma proposta de classificação que pode contribuir no direcionamento de ações em competência

em dados promovidos por bibliotecas universitárias aos pesquisadores.

4. Resultados

As categorias reunidas nesta pesquisa, foram extraídas de unidades temáticas de assunto a partir de um resumo sobre a análise das ações encontradas em portais de bibliotecas universitárias britânicas e holandesas. Configuram-se como um mapa de tipologia da análise dos conteúdos de GDP encontrados nesses portais, a partir dos treinamentos desenvolvidos e do contato por e-mail, solicitando, caso possível, os materiais que porventura não estejam disponíveis em acesso aberto. Desta forma, pretende-se representar os conhecimentos pertencentes e explicitados por estas comunidades de análise: as ações de promoção de competência em dados oferecidas em bibliotecas universitárias dos Países Baixos e Reino Unido.

O (Apêndice A), apresenta as categorias explicitadas no referencial teórico, e as unidades temáticas propostas com base nos assuntos encontrados no mapeamento dos portais das bibliotecas. A base das categorias foi elencada a partir do modelo discutido no referencial teórico proposto por RISDALE *et al.* (2015) a matriz de competência em dados (Figura 1).

Sendo assim, (Apêndice A) a *estrutura conceitual* desta terminologia versa sobre a contextualização do que seriam dados de pesquisa, pois devido a sua complexidade esta é uma emblemática recorrente nos workshops, conteúdos e guias ofertados pelas bibliotecas. A *coleta de dados*, evidencia a necessidade representar um fluxo informacional para a descoberta de dados em fontes apropriadas, assim como deixar clara a sua proveniência, os métodos em seus meios legais e éticos, para que se possa ter uma garantia de serem encontrados por outro pesquisador ou serem reproduzidos no futuro.

A *gestão de dados* apresenta os aspectos do gerenciamento de todas as etapas do ciclo de vida dos dados de pesquisa. Os dados precisam ser descritos e ter uma documentação associada contendo seus métodos, ferramentas utilizadas, versões para que possam ser compreendidos no futuro. Deste modo, compreende as medidas tomadas para documentar desde o PGD ao reuso de dados. Prestar conta de todo esse ciclo é tarefa do pesquisador nos Países Baixos e no Reino Unido e, para tal, ele precisa compreender esse caminho para obter financiamento. Percebe-se que um PGD bem alinhado é primordial para uma pesquisa que necessita de recursos para sair do papel. Outro aspecto é dar segurança aos dados da pesquisa, principalmente se eles forem dados sensíveis, ou seja, pessoais. É necessário definir papéis estabelecendo quem acessa, quais dados serão arquivados, embargados, restritos, anonimizados e compartilhados.

Apontar soluções para *análise de dados* é um elo que vem sendo traçado entre a biblioteca, os departamentos de pesquisa e os serviços de Tecnologia da Informação. Traz orientações para converter dados brutos em dados processados que precisam de uma limpeza e exigem a capacidade de habilidades em computação e de opções específicas de armazenamento e proteção. Geralmente, requer um uso de um software especializado, se possível com código de acesso aberto para facilitar o compartilhamento e reuso. O conhecimento sobre esses softwares é oferecido em workshops e tutoriais em ambos os países, principalmente para o doutorado.

Por fim, a categoria *uso de dados* denota toda uma reflexão sobre a cultura da pesquisa nas bibliotecas universitárias, na própria universidade por meio dos grupos de departamentos, nas determinações nacionais e da União Europeia. Trazer exemplos e experiências com dados de pesquisa entrevistando e dando voz a pesquisadores e promovendo encontros é uma tarefa a ser direcionada. A noção sobre o uso de dados amplia as fontes de consulta sobre a GDP, permite disponibilizar e disseminar tutoriais ofertados e propõe de forma educativa pensar criticamente sobre as medidas acionadas antes, durante e depois da gestão de dados. Essa atividade é desenvolvida por meio da noção de contexto, da leitura, em especial de gráficos e tabelas, da adequação aos princípios éticos pautados nas políticas de GDP. Como estratégia educativa há um movimento ainda discreto por meio do suporte de GDP, que atua com atividades em sala de aula, ou por meio de *workshops*. A intenção é construir conhecimento prático e teórico para uma nova geração de pesquisadores com habilidades de manipulação e análise de dados.

As unidades temáticas apresentadas no (Apêndice A) podem servir como um guia de temas a serem trabalhados dentro de cada disciplina de um curso de GDP. Neste caso, cada categoria poderia se configurar como uma disciplina.

4.1 Proposta de classificação para competência em dados de pesquisa

Os princípios que regem esta sistematização proposta procuram obedecer aos encontrados nos cânones do método analítico sintético de Ranganathan (1967). A característica de divisão em renques tem por propósito reunir elementos que compartilhem características em comum em uma classe e de forma horizontal. Assim, o método utilizado é sobre o que dispõe a classificação em Renques a partir do Cânnon para Sequência Útil, que dispõe de uma sequência dos elementos em uma classe que deve ter utilidade aos propósitos daqueles ao qual se destina (CAMPOS; GOMES, 2008). Para Gomes e Campos (2019, p. 29):

Estes são, talvez, os mais úteis para um modelo de navegação. Considerando que a apresentação das informações em um portal se dá em uma tela, que é bidimensional, os princípios a seguir são relevantes para a

distribuição espacial dos elementos no interior de uma Faceta. Tais princípios estão relacionados ao Cânnon da Sequência Útil. Este cânnon fornece orientação para ordenação dos elementos no interior de uma Faceta. E é bastante útil para a ordenação das facetas no espaço de uma página Web. Se a dinâmica é útil na construção de um instrumento de indexação (tesauro, taxonomia), na organização dos elementos em uma página a ordenação é relevante, pela linearidade requerida em tal espaço (como em uma página escrita, por exemplo).

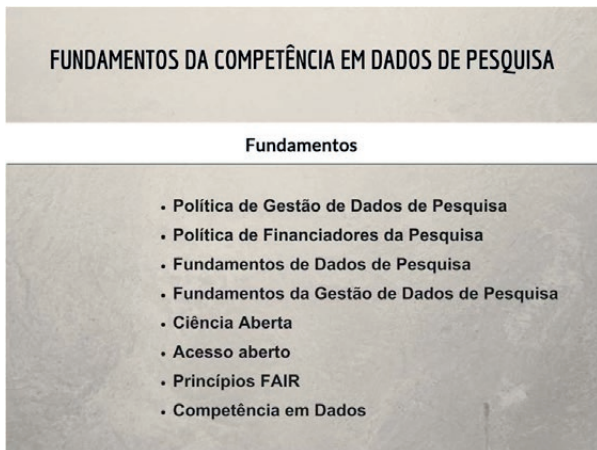
No Cânnon para Sequência Útil, Ranganathan propõe um conjunto de Princípios que servem para a ordenação em qualquer nível, desde a classe mais geral até a mais específica, sempre que haja um conjunto de elementos que requeiram alguma ordem. São eles:

1. Princípio do Posterior-no-Tempo;
2. Princípio do Posterior-na-Evolução;
3. Princípios da Contiguidade Espacial;
4. Princípios para Medida Quantitativa;
5. Princípio da Complexidade Crescente;
6. Princípio da Sequência Canônica;
7. Princípio da Garantia Literária;
8. Princípio da Ordem Alfabética.

(GOMES; MOTTA; CAMPOS, 2006, sem paginação).

Nesta classificação será utilizado o Princípio do Posterior no Tempo, que corresponde de forma ordenar fenômenos, processos e atividades. A escolha por este elemento se dá por ser a gestão de dados um processo, justificando assim a distribuição das ações de competências em dados por categorias que expressam a continuidade dessas ações ao longo do tempo. Para ordenar as ações dentro das categorias foi utilizado como critério, a forma como as atividades estavam apresentadas nos conteúdos analisados nos portais. É certo que na prática o processo da GDP depende de um fluxo de pesquisa, que, por sua vez, depende de uma comunidade com características próprias e que as posições das atividades variam nesta classificação de antes, durante e depois, mesmo elas apresentadas como uma proposta generalista. Não há um modelo a ser seguido e as ações podem ser ajustadas a qualquer momento em uma pesquisa. A exemplo, o PGD precisa ser feito em uma etapa anterior a pesquisa, mas pode ser ajustado conforme ela for avançando. No entanto, adotou-se essa classificação olhando para um entendimento geral, respectivo a quantidade de vezes em que foram ordenadas desta forma nos portais. Assim, a classificação foi dividida da seguinte forma: Os fundamentos da Competência em Dados de Pesquisa (figura 2) e a classificação das ações de competência em dados (figura 3).

Figura 2. Fundamentos da competência em dados de pesquisa



Fonte: Dados da pesquisa.

Os fundamentos (figura 2) referem-se à compreensão dos fluxos culturais, sociais e conceituais que envolvem a GDP. Calzada Prado e Marzal (2013) indicam que a competência em dados pode estimular a capacidade de definir com precisão a necessidade informacional, como localizar fontes de dados que sejam adequadas aos propósitos. A competência em dados apoia a avaliação crítica dessas fontes e as ideias contidas nelas. Os autores apontam outros papéis da competência em dados tais como a capacidade de gerenciar, analisar e sintetizar dados para apoiar argumentos ou gerar novas ideias. A habilidade de documentar as fontes utilizadas e a capacidade de registrar ou comunicar os resultados de dados de maneira ética. Nesta perspectiva as bibliotecas vêm oferecendo ações que discorrem sobre a política de dados adotada por uma sociedade e como elas influenciam a organização das políticas institucionais das universidades e dos financiadores da pesquisa. Em um nível mais conceitual apresentam os fundamentos de dados de pesquisa e a sua tipologia. Somado a isso, aborda os fundamentos da GDP e a sua relação com estratégias para tornar dados de pesquisa encontráveis, acessíveis, interoperáveis e reutilizáveis, com níveis de abertura e compartilhamento conforme as demandas da Ciência Aberta. A competência em dados pretende trazer à luz as próprias práticas adotadas de gerenciamento de dados de pesquisa durante o ciclo de vida dos dados de pesquisa.

Na figura 3 as ações de competência em dados de pesquisa estão classificadas em três processos: antes da gestão, durante a gestão e depois da gestão.

O planejamento da gestão de dados compõe este processo anterior a gestão, pois antecede a prática e lista todas as informações sobre os dados de um projeto. Fornece informações sobre o tipo de dados, o método de coleta, o formato e a documentação dos dados. Também inclui seções sobre as instalações usadas, razões legais ou éticas para compartilhar dados e sobre a forma como

os dados são compartilhados e preservados a longo prazo. As ações em competência em dados são voltadas para ajudar a descobrir ferramentas, oferecer modelos ou orientar individualmente os pesquisadores a criarem seus próprios PGD.

Figura 3. Classificação das ações de competência em dados



Fonte: Dados da pesquisa.

Durante o processo de GDP as ações auxiliam os pesquisadores na fase de coleta, análise e organização dos dados de pesquisa por meio dos serviços de dados de pesquisa. Fornece catálogos de fontes para descoberta de dados, catálogo de ferramentas de softwares de análise de dados, assim como cursos e treinamentos para o aprendizado sobre essas ferramentas. Do mesmo modo, incentiva a partir de tutoriais e cursos, métodos para organizar os dados de pesquisa de modo que eles não se percam, fiquem seguros e protegidos. A GDP envolve a criação, armazenamento seguro e preservação de dados de pesquisa, a estruturação e o monitoramento da pesquisa. Então, saber o quê, porque, quando, onde e como é crucial para controlar a pesquisa e obter resultados. Está em entender como os computadores podem nos ajudar a pensar, como podemos ajudar os computadores a pensar, que informações podem ser extraídas de dados processados. Procura investigar quais são as habilidades úteis para tornar dados de pesquisa documentados para que possam ser armazenados, compartilhados e publicados. Haendel, Vasilevsky e Wirz (2016) apresentam que o investimento na gestão de dados é um recurso que valoriza o trabalho dos cientistas. Isto ajuda a promover uma cultura com abordagem mais responsável em relação a colaboração entre os pesquisadores e a reprodutibilidade de suas pesquisas, pois dados compartilhados fomentam práticas produtivas poupando recursos. Assim sendo, após a GDP é preciso verificar se os dados refletem um compromisso com a preservação, gerenciamento e compartilhamento de acordo com os princípios éticos e legais para que possa ser compartilhado com segurança,

de modo que outros pesquisadores possam replicá-los dando visibilidade. Assim, ao final é possível selecionar os dados de pesquisa que precisam ser mantidos e destruir os que não precisam ser mantidos. Preservar os dados da pesquisa significa depositá-los em um repositório por área temática, institucional ou em arquivos de dados. Ao publicar dados no repositório ou arquivar e compartilhá-los com outras pessoas, podem ser feitas restrições de acesso, se apropriado.

Um conjunto de dados de pesquisa também pode ser publicado em um *data journals* (periódico de dados) revisado por pares. Os *data journals* têm por propósito motivar cientistas, já que reconhecem a sua capacidade de fornecer dados estruturados, uma vez esse reconhecimento estabelecido fomenta treinamento de competência em dados e informação, aliados a uma educação estatística e ética, a fim de possibilitar um novo “padrão cultural” (HAENDEL, VASILEVSKY; WIRZ, 2016). Os periódicos de dados ou diários de dados são publicações cujo objetivo principal é expor conjuntos de dados. Normalmente, um conjunto de dados pode ser publicado em um jornal de dados, que consiste em um resumo, introdução e descrição de dados com métodos e materiais, bem como uma breve conclusão sobre oportunidades de reutilização. Como benefícios de publicar dados de pesquisa em um artigo de dados, inclui-se a revisão por pares do conjunto de dados, isto maximiza as oportunidades de reutilização do conjunto de dados e possibilita credibilidade e notoriedade no meio científico. Outra ação referente às questões éticas e legais é o tema da propriedade intelectual. Deve-se esclarecer quem é o detentor dos direitos autorais de um conjunto de dados, especialmente quando são usados dados existentes ou quando colabora com terceiros. O *Copyright* é uma forma de direito de propriedade intelectual que surge automaticamente se uma obra original é criada. Os direitos autorais não cobrem os fatos, ideias ou conceitos subjacentes, mas apenas a maneira particular como foram expressos. A produção protegida por direitos autorais da pesquisa pode incluir planilhas (e outras formas de dados originalmente selecionados e organizados), publicações, livros, relatórios e programas de computador. Geralmente as universidades possuem escritórios jurídicos que em parceria com as demandas do suporte à pesquisa podem ser consultores para se obter mais informações sobre direitos autorais e dados. No entanto, as questões sobre os direitos de exploração e reuso de dados podem ser ainda mais importantes do que as de propriedade. Quem pode usar os dados? Quem pode publicá-lo? Quem pode fornecê-lo a terceiros? São questões que podem ser esclarecidas em um plano de gestão de dados, por exemplo. É preciso declarar todos os acordos entre as partes interessadas em um plano de gestão de dados e negociar os termos de processamento, disseminação e reutilização. Quando chegar a hora de compartilhar os dados da pesquisa é preciso atribuir a

licença mais apropriada, deixando claro as condições de uso O licenciamento e as condições de reuso, quando bem documentadas promovem o valor da citação e o uso de identificadores persistentes (DOI) para conjuntos de dados, possibilitando que o crédito adequado seja dado por meio de citação. As citações aos dados podem aumentar o impacto acadêmico dando visibilidade a pesquisa. Saber qual a forma mais adequada de apresentar resultados de dados de pesquisa ajuda a amparar uma tomada de decisão, e quando direcionadas para um público global, possibilita a atração de novos parceiros de pesquisa dentro e fora da universidade.

5 Conclusão

Na tentativa de identificar como bibliotecários de dados promovem ações em competência em dados no auxílio à pesquisa e quais são os conteúdos explorados, por intermédio dos do agrupamento de conteúdos sobre os workshops e treinamentos ofertados pelas bibliotecas das instituições holandesas e britânicas foi proposto um quadro de temáticas abordadas nas ações de competência elaborado e organizado em categorias. Como resultado da sistematização, propôs-se duas classificações: 1) sobre os fundamentos da competência em dados de pesquisa e 2) ações de competência em dados. Os princípios que regem esta sistematização procuram obedecer aos encontrados nos cânones do método analítico sintético de Ranganathan (1967), a saber, o Cânon para Sequência Útil por meio do Princípio Posterior no Tempo.

Nos resultados da classificação observou-se que as ações de competência em dados direcionam-se para estruturar primeiramente um entendimento sobre temas que circundam a cultura institucional a qual a biblioteca universitária faz parte. Esses temas abordam as políticas de GDP e dos financiadores da pesquisa refletindo sobre as conformidades legais e éticas. Outros temas são sobre a definição de dados de pesquisa, dos fundamentos da GDP, do acesso aberto e de dados abertos, bem como uma introdução aos Princípios FAIR e como a competência em dados pode ajudar na compreensão dos aspectos relacionados a GDP.

Utilizou-se como referencial teórico a competência em informação, a competência em dados. Identificou-se que a competência em informação e a competência em dados possuem laços estreitos uma vez que possuem aspectos semelhantes. No entanto, para que um profissional da informação possa desenvolver ações de competência em dados, ou estar envolvido em qualquer habilidade em pesquisa com dados é preciso que ele tenha primeiro desenvolvido habilidades em competência em informação (KOLTAY, 2015). Isto significa que o sucesso ou insucesso do desenvolvimento da competência em dados vai depender primeiramente do quanto este profissional é envolvido com as práticas de competência em informação.

O bibliotecário de dados, embora ainda seja uma figura que cabe uma definição mais aprofundada, é aquele que possui afinidade com a competência em informação e com dados de pesquisa. Nas bibliotecas universitárias britânicas e holandesas o seu papel está tomando forma e vem impulsionando a criação de suportes de GDP, junto com outros parceiros para buscar construir por meio de ações, diálogos e reflexões uma rede de apoio. Como parceiros temos os *data steward* [4], os departamentos de pesquisa, suportes jurídicos de segurança da informação e de TI.

A competência em dados também pode ser pensada sob a perspectiva da educação em informação e em dados na medida em que é necessário um olhar crítico e ético que incorpora as fases e responsabilidades do processo de pesquisa científica, inclusive no âmbito da corresponsabilidade e da solidariedade. Evidencia-se que o escopo da competência em dados está para além da prática e do uso das ferramentas, mas inclui uma dimensão histórica do processo, do seu papel e uma outra, a dimensão crítica dos processos, dos propósitos, das ferramentas, do uso de dados de pesquisa, sua relação com a sociedade, dentre outros aspectos.

Uma próxima etapa dessa investigação, pode ser a construção de uma ementa, um glossário e posteriormente um tesouro dos conceitos abordados. As teorias da Organização do Conhecimento, especialmente a Teoria do Conceito de Dalhberg (1978) são sempre úteis para organização desses três propósitos, pois oferecem base lógica e conceitual para a distribuição de conteúdo. A definição e a estrutura conceitual poderão não apenas validar a estrutura lógico-conceitual da ementa, como se transformar em um material útil para aqueles que quiserem começar a promoção da competência em dados através da ministração de treinamentos.

Notas

- [1] São os chamados “*Data Terms*” (ROYAL SOCIETY, 2012, p. 12), os termos mais utilizados na pesquisa em dados e que carecem de definições e maiores esclarecimentos aos pesquisadores. Este relatório da Royal Society (2012) “[...] é um importante documento para a prática de dados científicos abertos [...]” (HENNING, *et al.*, 2019, p. 393).
- [2] A *European University Association* é uma organização sem fins lucrativos que representa mais de 800 universidades e associações de reitores em 48 países europeus. Esta organização desempenha um papel crucial na influência das políticas da União Europeia (UE) em educação superior, pesquisa e inovação.
- [3] Outras associações como a *iSchool* e a *CODATA* foram pensadas para compor esta investigação, mas não apresentaram números expressivos a serem analisados, de universidades do Reino Unido e dos Países Baixos.
- [4] Termo ainda sem tradução para o português. Salienta-se que diferentemente da realidade brasileira, a formação em

Biblioteconomia nos países analisados no escopo deste trabalho se dá via pós-graduação. Portanto, o *data steward* é um profissional que pode não possuir uma educação formal em Biblioteconomia, embora ele faça uma ligação entre o bibliotecário e o pesquisador.

Referências

- ABBOTT, D. **What is digital curation?** Edinburgh, UK: Digital Curation Centre, 2008.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Martins Fontes, 1977.
- BORGMAN, C. L. Research data: who will share what, with whom, when, and why? *In*: CHINA-NORTH AMERICAN LIBRARY CONFERENCE, 5., 2010, Beijing, PEK. **Proceedings** [...]. Beijing, PEK: CALA, 2016. p. 1-21.
- BUYS, C. M.; SHAW, P. L. Data management practices across an institution: survey and report. **Journal of Librarianship and Scholarly Communication**, [S. l.], v. 3, n. 2, p. 1-24, 2015. DOI: <https://doi.org/10.7710/2162-3309.1225>.
- CALZADA PRADO, J; MARZAL, M. A. Incorporating data literacy into information literacy programs: core competencies and contents. **Libri**, [S. l.], v. 63, n. 2, p. 123-134, 2013. DOI: <https://doi.org/10.1515/libri-2013-0010>.
- CAMPOS, M. L. A; GOMES, H. E. Taxonomia e classificação: o princípio de categorização. **DataGramZero – Revista de Ciência da Informação**, Paraíba, v. 9, n. 4, 2008.
- CARLSON, J.; FOSMIRE, M.; MILLER, C. C.; NELSON, M. S. Determining data information literacy needs. *In*: CARLSON, J. R.; JONHSTON, L. R. **Data information literacy: librarian, data, and the education of a new generation of researchers**. Indiana: Purdue University Press, 2015. p. 11-34.
- CARLSON, J.; FOSMIRE, M.; MILLER, C. C.; NELSON, M. S. Determining data information literacy needs: a study of students and research faculty. **Libraries and the Academy**, Baltimore, v. 11, n. 2, p. 629-657, 2011.
- CARLSON, J.; JOHNSTON, L.; WESTRA, B.; NICHOLS, M. Developing an approach for data management education: a report from the data information literacy project. **International Journal of Digital Curation**, [S. l.], v. 8, n. 1, p. 204–217, 2013. DOI: 10.2218/ijdc.v8i1.254.
- COSTAL, M. **Competência em informação para gestão de dados de pesquisa: um olhar voltado para a formação do (a) bibliotecário (a)**. 2019. Trabalho de Conclusão de Curso (Bacharelado em Biblioteconomia e Gestão de Unidades de Informação) – Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, 2019.
- DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Brasília, DF, v. 7, n. 2, 1978. DOI: 10.18225/ci.inf.v7i2.115.
- EUROPEAN COMMISSION. **H2020 programme: guidelines on FAIR data management in horizon 2020**. Bruxelas, Bélgica: EU, 2016.

- EUROPEAN UNION. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016. On the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). **Official Journal of the European Union**: EU, 2016. p. 1-88.
- EUROPEAN UNIVERSITY ASSOCIATION. **Member directory**. Bruxelas, BE: EUA, 2020.
- EYNDEN, V.; CORTI, L.; WOOLLARD, M.; BISHOP, L.; HORTON, L. **Managing and data sharing**: best practice for researchers. Colchester: University of Essex, 2011.
- GHODOOSI, B.; WEST, T.; LI, Q.; TORRISI-STEELE, G.; DEY, S. A systematic literature review of data literacy education. **Journal of Business & Finance Librarianship**, [S. l.], v. 28, n. 2, p. 122-127, 2023. DOI: <https://doi.org/10.1080/08963568.2023.2171552>.
- GOMES, H. E.; CAMPOS, M. L. A. **A organização do conhecimento na web**: contribuições de Shiyali Ramamrita Ranganathan e Ingetraut Dahlberg. Niterói: IACS/UFF, 2019. (Cadernos Acadêmicos).
- GOMES, H. E.; MOTTA, D. F.; CAMPOS, M. L. A. **Revisitando Ranganathan**: classificação na rede. Rio de Janeiro: Conexão Rio, 2006. Disponível em: <http://www.conexaorio.com/biti/revisitando/revisitando.htm#relevante>. Acesso em: 14 mar. 2021.
- HAENDEL, M. A.; VASILEVSKY, N. A.; WIRZ, J. A. Dealing with data: a case study on information and data management literacy. In: HAGEN-MCINTOSH, J. **Information and data literacy**: the role of the library. Oakville: Apple Academic Press, 2016. p. 3-12.
- HENNING, P. C.; RIBEIRO, C. J. S.; SANTOS, L. O. B. S.; SANTOS, P. X. Go fair e os princípios fair: o que representam para a expansão dos dados de pesquisa no âmbito da ciência aberta. **Em Questão**, Porto Alegre, v. 25, n. 2, p. 389-412, 2019. DOI: 10.19132/1808-5245252.389-412. Disponível em: <https://brapci.inf.br/index.php/res/v/113770>. Acesso em: 25 maio 2023.
- KOLTAY, T. Are you ready? tasks and roles for academic libraries in supporting Research 2.0. **New Library World**, Bingley, v. 117, n. 1/2, p. 94-104, 2016.
- KOLTAY, T. Data literacy: in search of a name and identity. **Journal of Documentation**, Bingley, v. 71, n. 2, p. 401-415, 2015. DOI: <https://doi.org/10.1108/JD-02-2014-0026>.
- LANDELIJK COÖRDINATIEPUNT RESEARCH DATA MANAGEMENT (LCRDM). **Positioning paper voor 2019 en verder**. Nederlands, NL: Publicatiedatum, 2019. DOI: <https://doi.org/10.5281/zenodo.3335410>.
- MARTIN, E. R. What is data literacy? **Journal of eScience Librarianship**, [S. l.], v. 3, n. 1, 2014. DOI: <https://doi.org/10.7191/jeslib.2014.1069>.
- RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.
- RIDSDALE, C.; ROTHWELL, J.; SMIT, M.; ALI-HASSAN, H.; BLIEMEL, M.; IRVINE, D.; KELLEY, D.; MATWIN, S.; WUETHERICK, B. **Strategies and best practices for data literacy education**: knowledge synthesis report. Halifax: Dalhousie University, 2015. DOI: 10.13140/RG.2.1.1922.5044.
- ROYAL SOCIETY. **Science as an open enterprise**: open data for open science. London, UK: Royal Society, jun. 2012.
- SCHIELD, M. Information literacy, statistical literacy and data literacy. **Iassist Quarterly Journal (IQ)**, v. 28, n. 2/3, p. 6-11, 2004. Disponível em: https://iassistquarterly.com/public/pdfs/iqvol282_3shields.pdf. Acesso em 26 maio 2023.
- SCHUMACHER, J.; VANDECREEK, D. Intellectual capital at risk: data management practices and data loss by faculty members at five american universities. **International Journal of Digital Curation**, [S. l.], v. 10, n. 2, p. 96-109, 2015. DOI: <https://doi.org/10.2218/ijdc.v10i2.321>
- TARTAROTTI, R. C.; DAL'VEDOVE, P. R.; FUJITA, M. S. L. Biblioteconomia de dados em repositórios de pesquisa: perspectivas para a atuação bibliotecária. **Informação & Informação**, Londrina, v. 24, n. 3, p. 207-226, 2019. DOI: <http://dx.doi.org/10.5433/1981-8920.2019v24n3p207>.
- VILAR, P.; ZABUKOVEC, V. Research data management and research data literacy in Slovenian science. **Journal of Documentation**, Bingley, v. 75, n. 1, p. [1-21], 2019. DOI: <https://doi.org/10.1108/JD-03-2018-0042>.
- WHITMIRE, A. L. Implementing a Graduate-Level Research Data Management Course: Approach, Outcomes, and Lessons Learned. **Journal of Librarianship and Scholarly Communication**, [S. l.], v. 3, n. 2, p. 1-22, 2015. DOI: <https://doi.org/10.7710/2162-3309.1246>
- WITT, M.; CARLSON, J.; BRANDT, S.; CRAGIN, M. Constructing data curation profiles. **International Journal of Digital Curation**, [S. l.], v. 4, n. 3, p. 93-103, 2009. DOI: <https://doi.org/10.2218/ijdc.v4i3.117>.

Apêndice A – Categorias da GDP em bibliotecas britânicas e holandesas

<i>Categorias / Disciplina</i>	<i>Unidade Temática / Temas de estudo e ensino</i>
Estrutura conceitual	Dados de pesquisa; Conjunto de dados de pesquisa;
Coleta de dados	Proveniência de dados; Descoberta de dados; Protocolos de coleta de dados; Método de coleta de dados; Ferramentas de coleta de dados;
Gestão de dados de pesquisa	Curadoria digital, Plano de gestão de dados; Gestão de custos; Organização de dados de pesquisa; Proteção de dados; Preservação de dados; Nomenclatura de arquivos; Formatos de arquivo; Conversão de dados; Controle de versão; Criptografia de dados; Preservação de dados; Verificação de dados; Autenticação de dados; Consentimento informado; Processamento de dados; Documentação de dados, (documentação descritiva, documentação administrativa, documentação estrutural); Padrão de metadados; Metadados (metadados descritivos, metadados administrativos, metadados estruturais); Identificadores persistentes; Seleção de dados; Embargo de registros; Descarte de registros; Arquivamento de dados (arquivamento de médio prazo, arquivamento de longo prazo); Armazenamento de dados; Segurança de dados; Classificação de dados (disponibilidade, integridade e confidencialidade); Ferramenta de classificação de dados;
Análise de dados	Mineração de dados; Software de análise de dados; Modelagem de dados; Simulação de dados; Visualização de dados; Análise de dados quantitativos; Análise de dados qualitativos; Limpeza de dados; Anonimização de dados; Pseudonimização de dados; Crítica de dados;
Uso de dados	Citação de dados; Fontes de dados; Banco de dados; Repositório de dados de pesquisa; Portais de dados de organizações; Mecanismos de busca de dados; Compartilhamento de dados; Publicação de dados; Ética em dados; Propriedade intelectual; Licença de dados; Competência em informação; Reuso de dados; Ciência aberta; Acesso aberto; Princípios FAIR; Política de dados de pesquisa; Apresentando resultados de dados (<i>data storytelling</i>);

Fonte: as autoras 2023.

Proposta de ontologia para gestão documental arquivística de acervo acadêmico

Proposal of an ontology for archival document management of academic collections

Luiz Rogério Lopes Silva (1), Rodrigo Eduardo Botelho-Francisco (2)

(1) Universidade Federal do Paraná, Curitiba, luiz.lopes@ufpr.br. (2) rodrigobotelho@ufpr.br

Resumo

Objetivo: Propor um modelo conceitual baseado em ontologias para a gestão documental arquivística de acervos físicos de Instituições de Ensino Superior. **Metodologia:** A metodologia escolhida foi a Ontoforinforscience, respeitando-se as etapas propostas e utilizando-se a ferramenta Onto4AllEditor para criar e editar a ontologia. A aquisição de conhecimento sobre o domínio é oriunda da análise documental de um acordo estabelecido entre a Universidade Federal do Paraná (UFPR) e o Ministério da Educação (MEC) para a guarda e manutenção do acervo acadêmico de duas faculdades descredenciadas, originalmente estabelecidas em Curitiba-PR. **Resultados:** Uma ontologia de domínio composta por três subontologias (planejamento, tratamento documental e salvaguarda), 29 classes e três propriedades relacionais. O modelo, que visa facilitar a criação de sistemas de gerenciamento e busca mais eficientes e precisos, contribui para a integridade e a acessibilidade dos acervos acadêmicos, bem como o uso prático da ferramenta pela equipe de trabalho contratada pelo acordo UFPR/MEC. **Conclusão:** A ontologia está em sua versão de teste, com possibilidade de ampliação de termos candidatos. Representa, através da definição de conceitos e seus relacionamentos, as etapas do projeto de pesquisa e serve à negociação de significados do campo. Reforça, dessa forma, seu potencial de adequação para a gestão documental arquivística de acervo acadêmico.

Palavras-chave: Ontologia. acervo Acadêmico. Gestão arquivística. Organização do conhecimento. Gestão documental.

Abstract

Objective: To propose a conceptual model based on ontologies for archival document management of physical collections in Higher Education Institutions. **Methodology:** The chosen methodology was Ontoforinforscience, following the proposed steps and using the Onto4AllEditor tool to create and edit the ontology. Knowledge acquisition about the domain comes from documentary analysis of an agreement established between the Federal University of Paraná (UFPR) and the Ministry of Education (MEC) for the custody and maintenance of the academic collection of two discredited colleges, originally located in Curitiba-PR. **Results:** A domain ontology composed of three sub-ontologies (planning, document treatment, and safeguarding), 29 classes, and three relational properties. The model, which aims to facilitate the creation of more efficient and accurate management and retrieval systems, contributes to the integrity and accessibility of academic collections, as well as the practical use of the tool by the team hired under the UFPR/MEC agreement. **Conclusion:** The ontology is in its testing version, with the possibility of expanding candidate terms. By defining concepts and their relationships, it represents the stages of the research project and serves the negotiation of meanings in the field. It reinforces its potential for suitability in the archival document management of academic collections.

Keywords: Ontology. Academic collection. Archival management. Knowledge organization. Records management.

Resumen

Objetivo: Proponer un modelo conceptual basado en ontologías para la gestión documental archivística de colecciones físicas en Instituciones de Educación Superior. **Metodología:** La metodología elegida fue Ontoforinforscience, siguiendo los pasos propuestos y utilizando la herramienta Onto4AllEditor para crear y editar la ontología. La adquisición de conocimiento sobre el dominio proviene del análisis documental de un acuerdo establecido entre la Universidad Federal de Paraná (UFPR) y el Ministerio de Educación (MEC) para la custodia y mantenimiento de la colección académica de dos universidades desacreditadas, originalmente ubicadas en Curitiba-PR. **Resultados:** Una ontología de dominio compuesta por tres subontologías (planificación, tratamiento documental y salvaguarda), 29 clases y tres propiedades relacionales. El modelo, que tiene como objetivo facilitar la creación de sistemas de gestión y búsqueda más eficientes y precisos, contribuye a la integridad y accesibilidad de las colecciones académicas, así como al uso práctico de la herramienta por parte del equipo contratado bajo el acuerdo UFPR/MEC. **Conclusión:** La ontología se encuentra en su versión de prueba, con posibilidad de ampliar los términos candidatos. Mediante la definición de conceptos y sus relaciones, representa las etapas del proyecto de investigación y sirve para la negociación de significados en el campo. Refuerza su potencial de idoneidad en la gestión documental archivística de colecciones académicas.

Palabras clave: Ontología. Colección académica. Gestión archivística. Organización del Conocimiento. Gestión documental.

1 Introdução

As ontologias provaram ser ferramentas eficientes e poderosas para reunir e compartilhar conhecimento, uma vez que fornecem especificações explícitas de conceituações. Nas ciências da Computação e da Informação, este tipo de instrumento é visto como uma espécie de metamodelo no qual se pode confiar para promover melhorias nos modelos conceituais (ALMEIDA, 2020), bem como uma ferramenta para auxiliar a Organização do Conhecimento (MAZZOCCHI, 2018).

Como artefato, as ontologias são utilizadas em uma série de aplicações, incluindo sistematização de conceitos e definições consistentes, com possibilidade de inferências sobre diferentes domínios (CAMPOS, 2010), ou ainda no processo de desenvolvimento de produtos, de forma a colaborar com o tratamento de problemas relacionados à interoperabilidade semântica e para negociar significados em determinado campo de conhecimento (GUIZZARDI *et. al.*, 2022).

Neste trabalho, apresenta-se uma abordagem que aproxima as ontologias da gestão documental arquivística (GDA), especificamente das etapas de planejamento, tratamento documental e salvaguarda de acervo físico procedente de Instituições de Ensino Superior (IES). O objetivo é propor um modelo conceitual baseado em ontologias sobre o domínio “Gestão de documentos de Instituições de Ensino Superior”. A aquisição de conhecimento sobre o domínio é oriunda da análise documental de acordo estabelecido entre a Universidade Federal do Paraná (UFPR) e o Ministério da Educação (MEC) para a guarda e manutenção do acervo acadêmico de duas faculdades descredenciadas, originalmente estabelecidas em Curitiba-PR.

O trabalho se justifica em dois aspectos: (a) uma ontologia da GDA de acervo acadêmico pode contribuir para a conformidade legal com a Lei nº 8.159/1991 – Lei de Arquivos (BRASIL, 1991), que estabelece as diretrizes para a gestão documental no âmbito público. Ademais, esta ontologia pode fornecer uma estrutura formal para a definição de conceitos e entidades envolvidas no contexto de acervo acadêmico, garantindo a consistência e a conformidade da gestão documental e tornando mais fácil para as instituições acadêmicas atenderem às regulamentações legais de forma efetiva; (b) a ontologia pode ser útil nas atividades práticas dos acervos acadêmicos, ajudando em sua organização, descrição e preservação.

O artefato apresentado neste artigo está em versão de testes, com a possibilidade de ampliação de termos, conceitos e relações. Espera-se que ontologia sirva como um modelo de referência para projetos e processos que visam o correto tratamento de documentos, assim como sua salvaguarda. Para isso, apresenta-se um

primeiro conjunto de definições formais dos conceitos e relações que compõem o acordo MEC/UFPR, no intuito de criar sistemas de gerenciamento e busca mais eficientes e precisos, assim como garantir sua integridade e acessibilidade para interessados no tratamento deste tipo de acervo.

2 Gestão de documentos arquivísticos

Um primeiro exercício teórico deste trabalho está em compreender as aproximações e distinções entre “Gestão de documentos” e “organização do conhecimento arquivístico”, para então estabelecer um entendimento sobre “Gestão de documentos arquivísticos”. Desse modo, o conceito de gestão de documentos adotado é o da Lei 8.159/1991, que define a Gestão de documentos como o conjunto de procedimentos técnicos que envolvem produção, tramitação, uso, avaliação e arquivamento dos documentos em fase corrente e intermediária, com o objetivo de eliminá-los ou recolhê-los para guarda permanente.

No que tange à Organização do conhecimento arquivístico, a abordagem de Silva e Tognoli (2022) e Barros e Souza (2019) foi considerada relevante porque evidencia que o princípio da proveniência é a base da organização arquivística e que qualquer esforço neste sentido deve considerar os documentos produzidos ou recebidos por uma mesma entidade produtora, o chamado Fundo Arquivístico (FA).

Barros e Sousa (2019) defendem que, para organização do conhecimento arquivístico, o contexto deve ser o elemento central. Nesta linha de pensamento, Tognoli (2021) argumenta que os documentos devem ser organizados de acordo com o conjunto documental ao qual pertencem, respeitando-se dois princípios fundamentais da teoria e prática arquivística: proveniência e ordem original. Os autores defendem que a organização do conhecimento arquivístico deve garantir o acesso à informação contida nos documentos arquivísticos, utilizando para tal estratégias de organização, descrição e representação desses documentos respeitando sua estrutura original e contexto de produção.

Lehmkuhl e Silva (2017) destacam a importância da informação orgânica arquivística e sua relação com os contextos e funções arquivísticas. Alencar e Cervantes (2019) apontam que as atividades de organização e representação da informação arquivística devem considerar sua origem social, administrativa, econômica, política e cultural, bem como sua circunscrição no tempo e no espaço.

Em vista do que foi exposto, neste trabalho entende-se a GDA como uma prática que envolve desde a criação até a destinação final dos documentos, passando pela

triagem, higienização, desinfestação, classificação, digitalização, descrição e armazenamento adequado dos acervos. Todo o processo deve levar em consideração o valor histórico, probatório ou informativo dos documentos, com atenção à sua ordem original e cuidado para que documentos de diferentes produtores não sejam misturados.

2.1 Uso de ontologias em gestão documental arquivística

Barros e Sousa (2019) destacam a importância da classificação arquivística funcional na organização de acervos documentais e ressaltam que essa abordagem oferece possibilidades de normalização e incorpora aspectos semânticos em sua estruturação. No entanto, os autores consideram que a classificação funcional não é suficiente por si só e pode ser complementada por meio de outras formas estruturadas e relacionadas de sistemas de organização do conhecimento, como as taxonomias, ontologias e índices.

Em relação à ontologia, Barros e Sousa (2019) destacam que ela pode ser utilizada como um instrumento misto que vai além de um ciclo documental fechado e pode ser aplicada em todas as funções arquivísticas. Nesse sentido, a construção da ontologia se dá a partir do próprio desenvolvimento da classificação arquivística funcional, à medida que se constituem informações sobre o órgão produtor no contexto da categorização taxonômica. Dessa forma, a ontologia permite uma melhor compreensão das relações hierárquicas de forma contextual e possibilita o reuso em diferentes contextos, além da busca por meio de sistemas de recuperação.

Por fim, os autores destacam que a utilização da ontologia na GDA possibilita uma melhor organização dos acervos e permite que os documentos sejam encontrados com mais facilidade e rapidez. Além disso, a ontologia permite a identificação de relações entre os documentos, bem como sua contextualização, o que pode ser fundamental para a tomada de decisões em diferentes áreas do conhecimento. Assim, para uma gestão documental arquivística eficiente e eficaz, é fundamental que a ontologia seja utilizada em conjunto com outras formas estruturadas de sistemas de organização do conhecimento.

Um exemplo da utilização de ontologias na GDA de acervos acadêmicos pode ser vislumbrado no trabalho de Shu-Jiun Chen (2019), que as utiliza para representar o fundo *Cheng-po's Paintings and Documents* (CCP) na base de dados dos arquivos do Instituto de História de Taiwan, garantindo a interoperabilidade entre diferentes sistemas e instituições.

Outro exemplo está no trabalho de Low *et al.* (2022), que abordam a aplicação de ontologia como uma ferramenta para a representação semântica de documentos arquivísticos. A pesquisa destes autores explora a ontologia BFO (*Basic Formal Ontology*) como

base para a criação de uma ontologia específica para o domínio arquivístico, reutilizando a *Information Artifact Ontology* (IAO) a fim de melhorar a interoperabilidade entre sistemas e a recuperação de informações.

A *Information Artifact Ontology* (IAO), por sua vez, “descreve artefatos relacionados à comunicação e ao registro de informação” (ALMEIDA, 2020, p. 305). Ela é organizada em torno de vários conceitos fundamentais, como “artefato da informação”, “característica do artefato”, “representação do artefato”, “função do artefato” e assim por diante. Cada conceito é definido de forma clara e precisa, permitindo que os usuários compreendam as relações entre os diferentes artefatos da informação.

Em suma, entende-se que a utilização de ontologias na GDA pode favorecer inferências automáticas nos processos de organização e recuperação de recursos informacionais e contribuir para que os princípios arquivísticos sejam honrados.

2.2 Gestão de documentos arquivísticos de acervo acadêmico

A gestão de documentos arquivísticos em IES é um processo técnico-administrativo que envolve políticas, regras e procedimentos para determinar como os documentos devem ser produzidos, tramitados, acessados, arquivados e preservados pelos órgãos acadêmicos e administrativos. Para que esta gestão seja eficiente, é necessário o desenvolvimento de políticas e procedimentos específicos, como a definição de responsabilidades, criação de planos de classificação e tabelas de temporalidade, realização de avaliações periódicas e adoção de medidas de segurança e preservação adequadas.

No Brasil, os critérios para o arquivamento de documentos produzidos pelas instituições de ensino superior, com o objetivo de garantir a preservação, a organização e a disponibilidade desses documentos para consulta, são estabelecidos pela Portaria nº 1224/2013 do Ministério da Educação (MEC, 2013). Esta preconiza que os documentos físicos sejam armazenados de forma adequada, em ambiente seguro e com condições ideais de conservação, de modo a garantir a integridade e a durabilidade dos documentos ao longo do tempo. Além disso, a Portaria recomenda que as instituições adotem políticas de gestão documental que contemplem o controle de acesso e a identificação dos responsáveis pelo acervo físico.

No geral, pode-se dizer que a portaria do MEC enfatiza a importância do tratamento adequado dos documentos produzidos pelas instituições de ensino superior, incluindo documentos em formato digital e documentos físicos, de forma a garantir sua disponibilidade para consulta futura e a preservação da memória institucional. Para responder a esta demanda, as IES devem entender e atender a algumas etapas da GDA,

incluindo o planejamento, tratamento e a salvaguarda dos acervos.

Neste trabalho, compreende-se o planejamento a partir de um conjunto de atividades anterior ao tratamento documental, o que inclui, por exemplo, um diagnóstico com a situação arquivística das entidades produtoras e dos documentos, bem como um plano de trabalho composto pelas atividades de pré-tratamento, tratamento e salvaguarda do acervo.

O pré-tratamento considera aspectos como definição dos recursos humanos, tecnológicos e operacionais para que o tratamento documental ocorra de forma adequada e eficiente. Isso envolve a identificação, organização, descrição, codificação/indexação e o armazenamento. Já a salvaguarda de acervos visa proteger os documentos físicos contra danos e perdas. Nesta etapa, insere-se a digitalização dos documentos, processo que permite a conversão de documentos físicos para o formato digital e amplia suas possibilidades de acesso.

No caso das IES, a adoção da digitalização e descrição dos acervos pode ser uma estratégia valiosa para promover o acesso, facilitando que pesquisadores, estudantes e outros interessados encontrem e usem os materiais. No entanto, é importante ressaltar que a digitalização também apresenta desafios e problemas, especialmente quando se trata de preservar o valor histórico e legal dos documentos originais. Nesse sentido, é crucial garantir que a integridade e autenticidade do conteúdo original sejam mantidas, além de observar as leis e regulamentações aplicáveis.

3 Procedimentos metodológicos

Neste trabalho, para a criação da ontologia da GDA de acervos acadêmicos foi escolhida a Ontofoinformatics, metodologia que objetiva permitir que especialistas em Organização do Conhecimento e profissionais da informação superem dificuldades do jargão técnico, bem como questões lógicas e filosóficas que envolvem o desenvolvimento de ontologias (MENDONÇA; CASTRO, 2021).

A Ontofoinformatics é composta por oito etapas, sendo que a primeira consiste na especificação do domínio ou escopo geral da ontologia, do propósito geral, das classes de usuários, do uso pretendido, do tipo da ontologia e do grau de formalidade. Nesta etapa também é importante definir o escopo de cobertura e os atributos necessários da computação, como cardinalidade, equivalência, disjunção e herança múltipla.

A segunda etapa trata do detalhamento do método de aquisição e extração de conhecimento, assim como da seleção dos materiais de referência do domínio. Na terceira etapa faz-se a construção do dicionário de conceitos e verbos, tabela de valores e propriedades, além da representação dos modelos conceituais.

Na quarta etapa, incentiva-se o uso de ontologias de fundamentação como ponto de partida para o desenvolvimento ontológico. Na quinta etapa formaliza-se a ontologia, ou seja, é descrito o significado de cada propriedade a ser definida e se estabelecem diretrizes para essa definição. Nesta fase também se constrói, por meio de linguagem lógica, definição ou restrição formal de uma classe ou relação. E por fim são identificados os diferentes tipos de relação ontológica.

Na sexta etapa, estabelece-se um conjunto de critérios avaliativos do conteúdo ontológico, os quais envolvem validação e verificação, e elabora-se instrumento de validação para ser encaminhado a um painel de especialistas no domínio. Na sétima etapa, produz-se então o documento final com o conteúdo da ontologia e, finalmente, na oitava etapa, realiza-se a disponibilização da ontologia em formato e meio eletrônico, em linguagem lógica.

Neste artigo o foco está nas etapas de Especificação, Aquisição de Conhecimento, Conceitualização, Fundamentação e Formalização. Para execução do trabalho foi utilizado o *software* Onto4AllEditor, considerado neste trabalho por oferecer recursos visuais e intuitivos para a modelagem: permite aos usuários se concentrarem na definição dos conceitos e relações do domínio em questão, sem demandar preocupação com detalhes técnicos. Além disto, a ferramenta permite validação e verificação da ontologia, garantindo que as definições estejam corretas e consistentes.

3.1 O contexto estudado

O domínio em questão envolve o acervo de duas faculdades privadas do Estado do Paraná que já tiveram suas atividades encerradas e, com isso, a transferência de seus acervos acadêmicos para a Universidade Federal do Paraná (UFPR).

A transferência de acervo para a UFPR, bem como a responsabilidade pela sua guarda e manutenção, foi estabelecida em portaria do MEC quatro anos depois do encerramento das atividades das organizações estudadas. A resolução veio ao encontro da portaria número 315/2018, art. 38, do próprio MEC, que considera os arquivos universitários como patrimônio cultural e documental das IES e que devem ser tratados como tal, ou seja, devem ser preservados e disponibilizados para consulta. A resolução estabelece também que os arquivos universitários devem seguir as normas e diretrizes do CONARQ, responsável pela formulação da política nacional de arquivos públicos e privados.

Em abril de 2023, com os acervos já em posse da UFPR, realizou-se contagem provisória a qual estimou uma quantidade de três milhões de folhas nos acervos das faculdades em tela. Os mesmos são objeto de projeto de pesquisa e extensão que visa tratá-los, digitalizá-los e salvaguardá-los, seguindo as normas e diretrizes do

CONARQ. Nesse contexto, entende-se que uma aproximação participante pode comprovar (ou não) as informações presentes na ontologia proposta, bem como manter a relevância do instrumento e evitar inconsistências e redundâncias. O projeto, por sua vez, está em andamento, o que significa que a ontologia aqui apresentada pode sofrer alterações conforme os desdobramentos das etapas.

Ainda sobre este isto, é importante ressaltar que a UFPR, assim como diversas universidades públicas no Brasil, já era responsável pelo registro de diplomas de cursos superiores, processo que garante a prova e validade dos documentos em território nacional, conforme estabelece a Portaria nº 1.095, de 25 de outubro de 2018 (MEC, 2018). Assim, vislumbra-se que o tratamento, salvaguarda e disponibilização digital dos acervos em tela garantirão, como planejado, a devida análise processual para emissão de diplomas e outros documentos pertinentes de discentes outrora matriculados nas faculdades em questão.

3.2 Especificação

A especificação de uma ontologia é um processo iterativo e contínuo, já que novos requisitos e necessidades podem surgir ao longo do tempo. Na primeira versão da ontologia apresentada neste trabalho, o *template* de especificação disponibilizado pelo software Onto4allditor foi suficiente para descrever informações sobre descrição, domínio, propósito geral, perfil de usuários, intenção de uso, tipo da ontologia, grau de formalidade, delimitação do escopo de cobertura, questões de competência e colaboradores.

De modo geral, a ontologia tem como escopo a “gestão documental arquivística de acervo acadêmico”. O recorte compreende o tratamento, salvaguarda e digitalização do acervo das faculdades estudadas. O objetivo é fornecer um vocabulário bem fundamentado e uma linguagem comum para o domínio, permitindo a compreensão e compartilhamento de conhecimento de forma eficiente e precisa entre os profissionais da área, incluindo a equipe que trabalha no projeto. Além disso, a ontologia visa contribuir para a interoperabilidade entre sistemas e para a possibilidade de reuso de conhecimento.

Trata-se de uma ontologia de domínio, com médio rigor formal, ou semi-formal, com maior expressividade em relação a versões como OWL-Lite e menor formalização que linguagens como a FOL (First-Order Logic ou Lógica de Primeira Ordem).

3.3 Aquisição de conhecimento

A criação de uma ontologia requer a aquisição de conhecimento sobre o domínio em questão, que é organizado e representado por meio de classes e relações ontológicas. Para a criação do vocabulário que sustenta esta ontologia foram utilizadas diversas fontes de

conhecimento, como a enciclopédia da ISKO, artigos científicos, documentos oficiais e entrevistas. Entende-se que as múltiplas fontes de conhecimento são essenciais para garantir a qualidade e relevância da ontologia, uma vez que cada fonte contribui de forma única para a compreensão do domínio em questão.

Contudo, vale ressaltar que não se trata de um artefato estático, mas com a capacidade de ser atualizado e aprimorado conforme novas informações e conhecimentos são adquiridos. As classes e relações aqui apresentadas são genuinamente ontológicas desde que limitadas ao escopo do estudo proposto. Recomenda-se que a ontologia seja periodicamente revisada e atualizada, a fim de manter sua relevância e precisão para o domínio estudado.

3.4 Reuso de ontologias

Na primeira versão desta ontologia não foram reusadas ontologias de fundamentação. Para as próximas versões a modelagem conceitual será embasada na UFO-A e UFO-B, criadas por Guizzard e sua equipe.

A UFO-A tem como objetivo fornecer uma base consistente para o desenvolvimento de outras espécies de ontologias e linguagens ontológicas. Ela é composta por classes como *endurant*, *endurant universal*, *kind*, *subkind*, entre outras. Já a UFO-B adiciona os aspectos perdurantes e é formada por classes como *perdurant*, *role*, *event*, *among others*.

Inicialmente, entende-se que as classes *endurant universal*, *kind*, *subkind* e *perdurant* podem ser reusadas. A classe *endurant universal*, por exemplo, é utilizada para representar entidades universais que possuem propriedades e características próprias, como instituições de ensino, empresas e organizações governamentais. Já as classes *kind* e *subkind* são usadas para representar variações de uma mesma entidade, como diferentes tipos de documentos arquivísticos ou diferentes versões de um mesmo documento. Por fim, a classe *perdurant* representa entidades que têm existência temporal e que mudam ao longo do tempo, como eventos e processos relacionados à GDA.

3.6 Conceitualização e Formalização

Na etapa de conceitualização foram identificados e analisados os conceitos candidatos às classes da ontologia, utilizando-se de conhecimentos da área de arquivologia e de análise documental sobre o acervo das IES estudadas e de profissionais da área de arquivologia. Além disso, a organização deste conhecimento permitiu identificar as relações e propriedades da ontologia. Os modelos conceituais gráficos podem ser vistos na Seção 4 “Criando a ontologia” e facilitam a visualização e compreensão dos conceitos e suas interações.

Para a formalização, como informado anteriormente, foi usado o software Onto4allditor. Nele, as propriedades descritivas *identificador de classe (IRI)*, *label*, *domain* e

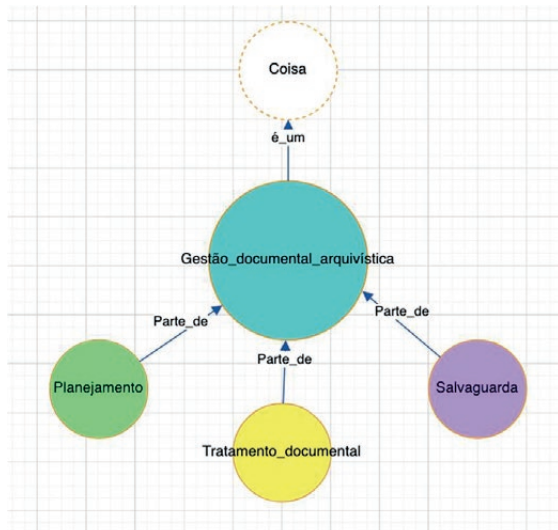
range são preenchidas automaticamente. As propriedades “*retrocompatíveis com*” (*backwardCompatibleWith*), “*termo alternativo*” (*equivalentTo*), “*sub_propriedade_de*” (*subpropertyOf*), “*inverso_de*” (*inverseOf*) e “*desarticulado_com*” (*disjointWith*) não foram formalizadas porque a aquisição de conhecimento não revelou tais características. A propriedade “*sinônimos*” foi atribuída somente às classes que continham termo equivalente.

No processo, prezou-se pelos seguintes princípios ontológicos: única definição para cada classe, não-circularidade, formalização apenas de condições necessárias e suficientes que compõem a definição de classe. Qualquer outra informação complementar foi incluída como comentários (*comment*) ou anotações. Tal procedimento responde ao princípio do vocabulário mínimo e ao princípio da autorreferência (GUIZZARDI, 2000). O vocabulário mínimo deve ser o menor possível e não apresentar ambiguidades, enquanto a autorreferência indica que a definição de um termo deve, sempre que possível, ser feita utilizando outros termos do dicionário.

4 Criando a ontologia

A ontologia proposta neste trabalho está dividida em três subontologias: planejamento, tratamento documental e salvaguarda. A representação visual disposta na Figura 1 facilita aos usuários entenderem a ontologia como um todo e localizarem rapidamente os conceitos que procuram. Ao dividir um domínio complexo em subdomínios menores e mais específicos colabora-se para sua compreensão e, por conseguinte, para o gerenciamento de informações. A macrovisão da ontologia está no apêndice A.

Figura 1. *Superclasses da ontologia*



Fonte: Os autores (2023).

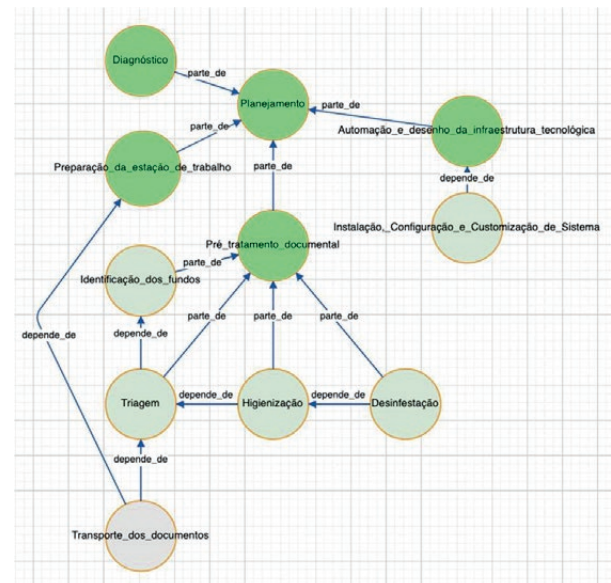
Os mapas conceituais de cada subontologia mostram de forma clara quais são os conceitos que compõem as subontologias e como eles se relacionam entre si. Os mapas destacam os principais conceitos e mostram como eles se encaixam em um todo coeso, sem se ater a detalhes técnicos. A representação visual pode ser adaptada para diferentes necessidades, desde diagramas simples para usuários iniciantes até gráficos mais complexos para especialistas em um campo, além de ser útil para identificar lacunas na ontologia e orientar seu desenvolvimento futuro.

4.1 Sub-ontologia de planejamento

A subontologia de planejamento envolve as classes de diagnóstico, identificação e reconhecimento dos fundos, preparo do local de trabalho, automação e desenho da infraestrutura tecnológica, transporte dos documentos, instalação, configuração e customização de Sistema, desinfestação, triagem, higienização e desinfestação.

O diagnóstico é uma ferramenta utilizada para conhecer a situação arquivística das instituições e dos documentos, envolvendo o levantamento de informações sobre a volumetria, forma, suporte, tipo, função, cronologia, identificação original das inscrições, localização e condições de armazenamento e conservação dos documentos. A preparação do local de trabalho inclui a composição física do ambiente, estruturação do espaço físico, mobiliários, ferramentas, infraestrutura tecnológica e materiais necessários para o planejamento e execução do tratamento documental. A automação dos processos envolve o planejamento e execução da infraestrutura tecnológica necessária, incluindo a instalação e configuração de computadores, scanners e impressoras.

Figura 2. *Subontologia de planejamento*



Fonte: Os Autores (2023).

A classe “Instalação, Configuração e Customização de Sistema” é uma etapa prévia ao início do tratamento documental e está relacionada à implementação de um sistema de gestão arquivística de documentos (SIGAD) adequado ao e-ARQ Brasil. A classe de “instalação e configuração” estabelece uma relação ontológica de dependência com “automação e desenho”.

Do mesmo modo, a classe “transporte dos documentos” depende de “triagem” e “preparação da estação”, referindo-se ao processo de transporte físico das caixas e dos documentos para o local de trabalho. O transporte deve atender às necessidades de manutenção das caixas, sem prejuízos para o tratamento documental, sobretudo no que toca à conservação e dissociação dos documentos/caixas.

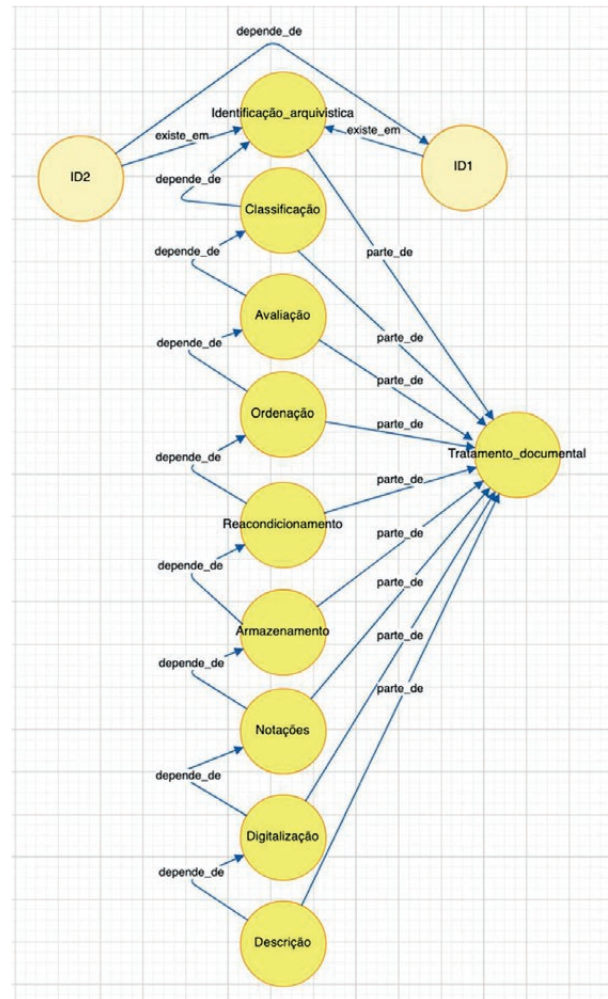
A classe “desinfestação” consiste na preservação dos documentos pela eliminação de agentes biológicos que possam prejudicá-los, oferecendo uma relação de dependências com as classes “triagem” e “higienização”. A triagem é um processo inicial de avaliação e seleção de materiais, visando identificar quais itens do acervo devem ser preservados e quais podem ser descartados ou transferidos para outras instituições. A higienização, por sua vez, envolve a remoção de sujeiras, poeiras, fungos, bactérias e outros contaminantes que possam danificar os materiais do acervo físico, mantendo a integridade física dos materiais e evitando a propagação de doenças ou infestações por pragas. Por fim, a desinfestação é um processo de tratamento químico ou físico que tem como objetivo eliminar pragas e insetos que possam danificar os materiais do acervo. Este processo é importante para evitar a propagação de pragas que possam destruir o acervo ao longo do tempo.

4.2 Sub-ontologia de tratamento documental

Aqui, destacam-se as seguintes classes: identificação arquivística, classificação, avaliação, ordenação, recondicionamento, armazenamento, notações dos invólucros e caixas, digitalização e descrição. A relação ontológica entre estas classes é de dependência, ou seja, a relação “*depende_de*” indica que uma classe depende de outra para poder funcionar corretamente e que mudanças em uma classe podem afetar diretamente o comportamento das classes que dependem dela.

A “identificação arquivística” é composta por ID1 e ID2. A primeira trata da validação da separação dos fundos e documentos de atividade meio e fim, levantamento das funções e atividades dos documentos e identificação dos vínculos desses documentos com o órgão produtor. A ID2 é dependente ontologicamente da primeira, uma vez que tem como objetivo identificar as espécies e tipos documentais. Ambas identificam as características intrínsecas e extrínsecas dos documentos com base na diplomática arquivística.

Figura 3. Subontologia de tratamento documental



Fonte: Os Autores (2023).

Subsequentemente, existem outras relações de dependência: a “classificação” e “avaliação” dependem da “identificação arquivística”, uma vez que a identificação das espécies e tipos documentais é fundamental para representar o documento de acordo com a atividade e função que o gerou. Neste caso, a Tabela de Temporalidade é utilizada para atribuir valor para destinação final dos documentos, podendo ser a eliminação ou a guarda permanente.

A classe de “ordenação” depende da identificação arquivística, classificação e avaliação, pois é preciso identificar a tipologia documental e a função do documento antes de estabelecer um sistema de ordenação. A ordenação também pode influenciar na classificação e avaliação, já que documentos mal ordenados podem dificultar a identificação da função e atividade que o gerou, resultando em um valor equivocado para sua destinação final.

Já as classes de “recondicionamento” e “armazenamento” também dependem da identificação arquivística, da classificação e avaliação, pois é

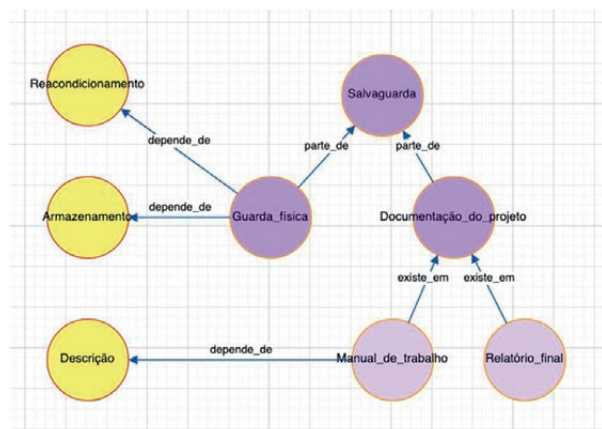
necessário conhecer as funções e atividades dos documentos para definir um local físico apropriado para o armazenamento e garantir a sua proteção contra agentes externos. Além disso, a “notação dos invólucros e caixas” também é uma classe dependente da identificação física dos documentos, que pode ser utilizada em conjunto com o sistema de armazenamento para facilitar a localização arquivística.

Por fim, as classes de “digitalização” e “descrição” dependem de todas as outras anteriores, pois é necessário que os documentos estejam identificados, classificados, ordenados e armazenados de forma adequada para que possam ser digitalizados e descritos de maneira eficiente. A descrição arquivística, por sua vez, é importante para a criação de instrumentos de pesquisa que orientam o usuário/pesquisador na recuperação da informação, permitindo que os documentos sejam encontrados com facilidade e rapidez.

4.3 Sub-ontologia de salvaguarda

Na subontologia de salvaguarda de documentos, a classe "guarda física" depende da classe "reacondicionamento e armazenamento", pois é necessário que os documentos estejam adequadamente protegidos e acondicionados antes de serem transportados para o local de guarda definitiva. Conceitualmente, trata-se do transporte físico das caixas e dos documentos para o local de guarda "definitiva", respeitando a ordenação, a notação e o reacondicionamento previamente realizados.

Figura 4. Subontologia de salvaguarda



Fonte: Os Autores (2023).

As classes “manual de trabalho” e “relatório final” existem em “Documentação do Projeto”, pois tem como objetivo sintetizar e registrar todas as etapas do projeto, desde o planejamento até a salvaguarda. Além disso, a classe “Manual de Trabalho” depende da classe “Descrição”, pois é necessário descrever de forma clara e objetiva todas as etapas do projeto para que o manual seja elaborado de forma adequada.

4.4 Termos candidados para ampliação da ontologia

No sentido de ampliar a ontologia e promover um conhecimento mais holístico do domínio, alguns termos são candidatos: *plano de preservação digital, política de preservação digital, estratégia de preservação digital, modelo de metadados, repositório digital, formato de arquivo digital, certificação digital, migração digital, emulação digital, autenticidade digital, descrição arquivística, código de classificação, tabela de temporalidade, tabela de destinação, autoridade arquivística, autoridade de assunto, identificador persistente, temporalidade, destinação, autoridade, identificador persistente, segurança física, segurança digital, gestão de riscos, gestão de incidentes, recuperação de desastres, continuidade de negócios, auditoria de segurança, gestão de autenticidade*. Os termos candidatos possivelmente incluirão novos conceitos, relações ou propriedades, mas deverão ser analisados com rigor a fim de identificar a relevância, a precisão e a coerência com o restante da ontologia.

5 Considerações finais

Este trabalho apresentou a criação de uma ontologia para a GDA de acervo acadêmico, com foco nas subontologias de planejamento, tratamento documental e salvaguarda. A ontologia foi desenvolvida a partir da análise documental de projeto de pesquisa e extensão que visa a guarda e manutenção do acervo acadêmico de duas IES descredenciadas no Estado do Paraná.

As subontologias foram criadas a partir da análise de conceitos e relações presentes na literatura sobre GDA, bem como os documentos oficiais do projeto de pesquisa e extensão. O artefato desenvolvido possui 29 classes com 03 relações (*parte_de*, *existe_em*, *depende_de*) entre si e tem o intuito de facilitar a criação de sistemas de gerenciamento e busca mais eficientes e precisos, que garantam a integridade e a acessibilidade dos acervos acadêmicos para as gerações presentes e futuras.

No entanto, é importante ressaltar que o projeto analisado ainda está em andamento, o que significa que ainda é possível ampliar e aprimorar a ontologia aqui apresentada. Além disso, como limitação da pesquisa, destaca-se que a ontologia desenvolvida é específica para a GDA de acervo acadêmico, não contemplando outros tipos de acervos.

Dessa forma, trabalhos futuros incluem a continuidade do desenvolvimento da ontologia para outras subontologias, além da análise de outros acervos, como museológicos e bibliográficos, e a realização de testes e validação da ontologia em ambientes reais de GDA de acervo acadêmico.

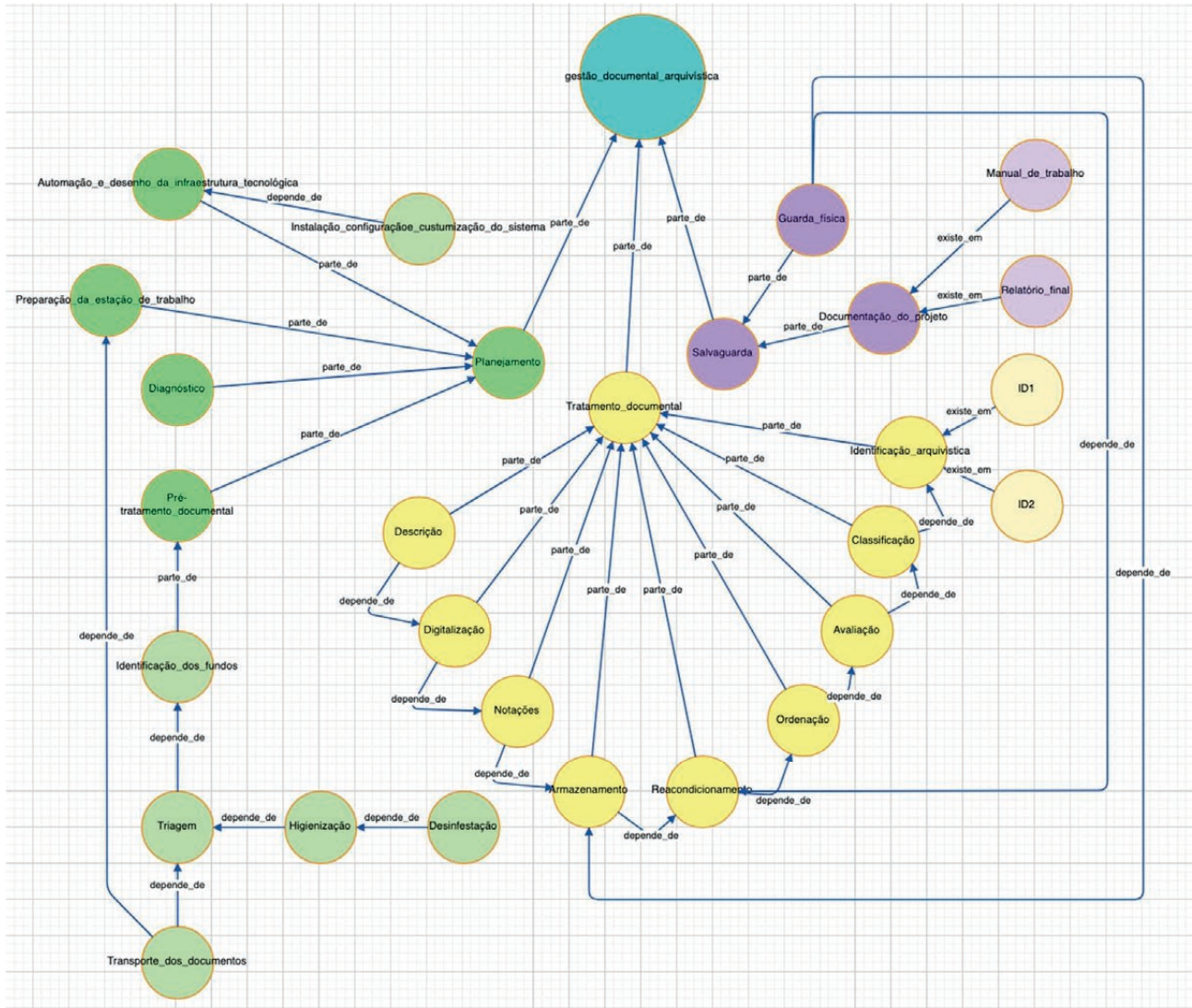
Notas

- [1] O presente trabalho foi realizado com apoio no âmbito de convênio celebrado entre a UFPR e a Fundação da Universidade Federal do Paraná para o desenvolvimento da Ciência, Tecnologia e da Cultura.

Referências

- ALENCAR, M. F.; CERVANTES, B. M. N. Organização e representação do conhecimento arquivístico: em busca de um método para construção de tesouro funcional. **Em Questão**, Porto Alegre, p. 368-389, 2019.
- ALMEIDA, M. B. **Ontologia em Ciência da Informação. Teoria e Método**. Curitiba: CRV, 2020. v. 1, (Representação do Conhecimento em Ciência da Informação).
- BARROS, T. H. B.; SOUSA, R. T. B. Archival science and knowledge organization: mapping methodological relationships. **Knowledge Organization**, [s. l.], n.46, v. 7, p. 493-502, 2019. Acesso em: 05 abr. 2023.
- BRASIL. **Lei nº 8159, de 8 de janeiro de 1991**. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/18159.htm. Acesso em: 20 abr. 2023.
- CAMPOS, M. L. A. O papel das definições na pesquisa em ontologia. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, p. 220-238, 2010.
- CHEN, S-J. Semantic enrichment of linked archival materials. **Knowledge Organization**, [s. l.], v. 46, n. 7, p. 530-547, 2019.
- GUIZZARDI, G. **Desenvolvimento para e com reuso: um estudo de caso no domínio de vídeo sob demanda**. 2000. Dissertação (Mestrado em Informática) – Universidade Federal do Espírito Santo, Vitória, 2000.
- GUIZZARDI, G.; BENEVIDES, A. B.; FONSECA, C. M.; PORELLO, D.; ALMEIDA, J. P.; SALES, T. P. UFO: Unified foundational ontology. **Applied Ontology**, [s. l.], v. 17, n. 1, p. 167-210, 2022.
- LEHMKUHL, C. S.; SILVA, E. C. L. A representação da informação arquivística nos registros civis. In: CONGRESSO BRASILEIRO EM ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO (ISKO), 4., 2017, Recife. **Anais [...]**. Recife: ISKO, 2017. p. 138-145.
- LÖW, M. M.; ROCHA, R. P.; ABEL, M.; GARCIA, L. F. Ontologia e documento arquivístico: análise ontológica para representação semântica do documento arquivístico em BFO. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 27 n. 1, p. 1-27, 2022.
- MAZZOCCHI, F. Knowledge organization system (KOS): an introductory critical account. **Knowledge Organization**, Würzburg, v. 45, n. 1, p. 54-78, 2018.
- MEC. **Portaria MEC nº 1223, de 18 de dezembro de 2013**. Disponível em: <https://www.gov.br/conarq/pt-br/legislacao-arquivistica/portarias-federais/portaria-mec-no-1-224-de-18-de-dezembro-de-2013>. Acesso em: 20 abr. 2023.
- MEC. **Portaria nº 1095, de 25 de outubro de 2018**. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/47330359/. Acesso em: 20 abr. 2023.
- MENDONÇA, F. M.; CASTRO, L. P. OntoForInfoScience e Onto4ALLEditor: metodologia e editor de ontologias como facilitadores na construção de ontologias por especialistas do domínio e cientistas da informação. **Fronteiras da Representação do Conhecimento**, [s. l.], v. 1, n. 1, p. 145-173, 2021.
- SILVA, A. M. S.; TOGNOLI, N. B. A organização do conhecimento arquivístico: a emergência de uma comunidade discursiva brasileira. **Acervo: Revista do Arquivo Nacional**. [s. l.], v. 35 n. 2, 2022.
- TOGNOLI, N. B. Os desafios para a formação e o desenvolvimento profissional na Organização do Conhecimento à luz da agenda 2030. In: CONGRESSO ISKO ESPANHA-PORTUGAL, 5., 2021, Lisboa. **Anais [...]**. Lisboa: Centro de Estudos Clássicos, 2021. p. 39-52.

Apêndice A – Macrovisão da ontologia



Fonte: Os Autores (2023).

Prospecção de termos candidatos no domínio do Transtorno do Espectro Autista

Prospecting candidate terms in the autism Spectrum Disorder Domain

Larissa Kawana Passos Benatto Perandre Pereira (1), Benjamin Luiz Franklin (2)

(1) Universidade Estadual de Londrina, Rodovia Celso Garcia Cid, PR-445 Km 380 – Campus Universitário, Londrina – PR, 86082-700, larissa.benatto@uel.br, (2) belfra@uel.br

Resumo

Introdução: É fundamental o destaque da relevância de um vocabulário controlado para a indexação e recuperação eficiente de informações sobre o Transtorno do Espectro Autista (TEA), tendo em vista que ainda não se tem algo como tal. A estrutura semântica auxilia na padronização de termos e na distinção de homografias, aprimorando a qualidade da literatura científica sobre o TEA. O uso de um vocabulário controlado é fundamental para obter resultados precisos na busca por informações e, por isso, essa pesquisa visa o uso das metodologias dispostas para a realização da parte anterior a construção deste vocabulário. **Objetivo:** Realizar uma busca minuciosa para encontrar termos que possam ser utilizados como candidatos para compor os sistemas de organização do conhecimento, com o intuito de viabilizar a redução dos custos necessários para o desenvolvimento e manutenção de um vocabulário controlado que possa ser aplicado futuramente em áreas e estudos científicos. **Metodologia:** São incluídas análise de conteúdo, análise de dados, análise exploratória e análise temática. Além disso, utiliza-se o modelo entidade-relacionamento para recuperar informações e criar consultas de agregação em formato de tabela, as quais serão apresentadas por meio de gráficos ilustrativos. **Resultados:** Após o tratamento inicial dos dados, fez-se um banco de dados e foram aplicadas as consultas de agregação, demonstrando os significativos resultados por meio de tabelas, gráficos e um diagrama do modelo entidade-relacionamento. **Conclusão:** Conclui-se que esta pesquisa tem um papel fundamental no desenvolvimento de Sistemas de Organização do Conhecimento (SOC), já que utiliza recursos computacionais de ponta disponíveis (atualmente). Por meio das metodologias empregadas, é possível prospectar termos candidatos de forma mais precisa e eficiente, contribuindo para a construção de um vocabulário controlado padronizado e de qualidade. Isso, por sua vez, resulta em uma indexação e recuperação mais eficientes de informações sobre o Transtorno do Espectro Autista, além de reduzir custos e promover avanços na área de estudos científicos relacionados.

Palavras-chave: Prospecção de Termos Candidatos. Modelo Entidade-Relacionamento. Análise de Dados. Transtorno do Espectro Autista.

Abstract

Introduction: Fundamentally, it is important to highlight the relevance of controlled vocabulary for indexing and efficient retrieval of information related to Autistic Spectrum Disorder (ASD), considering there is no available controlled vocabulary for its purpose. The semantic structure helps standardize terms and distinguish homographs, improving the quality of the scientific literature on ASD. The usage of controlled vocabulary is essential in obtaining precise results in the search for information; therefore, this research aims to use available methodologies in order to carry out the "prior" parts in constructing the said vocabulary. **Objective:** Conduct a meticulous search to find terms that can be used as candidates to compose a knowledge organization system, aiming to enable the reduction of costs necessary for the development and maintenance of a controlled vocabulary which otherwise could be applied in future scientific areas of study. **Methodology:** Includes content analysis, data analysis, exploratory analysis, and thematic analysis. Moreover, the entity-relationship model will be applied in retrieving information and creating aggregation queries in a table format. It will also be developed and presented through illustrative graphics. **Results:** In the "prior" part of the initial treatment of data, a database was developed, and applied aggregation queries, thus demonstrating the significant results through tables, graphs, and a diagram of the entity-relationship model. **Conclusion:** It can be concluded that this research has a fundamental role in the development of future Knowledge Organization Systems since it uses the currently available, state-of-the-art, computational resources. Throughout the used methodologies, it is possible to prospect candidate terms more precisely and efficiently, which in turn provides confident control and quality. Correlated, it results in more efficient indexing and retrieval of information about Autism Spectrum Disorder, additionally, promoting reduced costs and advancements in the related scientific areas of study.

Keywords: Prospection of Candidate Terms. Entity-Relationship Model. Data Analysis. Autism Spectrum Disorder.

Resumen

Introducción: Es fundamental destacar la relevancia de un vocabulario controlado para la indexación y recuperación eficiente de informaciones acerca del Trastorno del Espectro del Autismo (TEA), dado que tal recurso todavía no existe. La estructura semántica ayuda en la estandarización de términos y en la distinción de homógrafos, mejorando la calidad de la literatura científica sobre el TEA. El uso de un vocabulario controlado es crucial para que sea posible obtener resultados precisos en la recuperación de la información, y, por lo tanto, esta investigación tiene como objetivo utilizar las metodologías disponibles para la parte anterior de la construcción de este vocabulario. **Objetivo:** Realizar una búsqueda exhaustiva para encontrar términos que puedan ser utilizados como candidatos para componer los sistemas de organización del conocimiento, con el intuito de viabilizar la reducción de los costos necesarios para el desarrollo y mantenimiento de un vocabulario controlado que pueda ser aplicado en áreas y estudios científicos futuros. **Metodología:** Incluirá el análisis de contenido, análisis de datos, análisis exploratorio y análisis temático. Además, se utilizará el modelo entidad-relación para recuperar información y crear consultas de agregación en formato de tabla, las cuales se presentarán mediante gráficos ilustrativos. **Resultados:** Después del procesamiento inicial de los datos, se creó una base de datos y se aplicaron consultas de agregación, demostrando resultados significativos con el uso de tablas, gráficos y de un diagrama del modelo entidad-relación. **Conclusión:** Se concluye que esta investigación desempeña un papel fundamental en el desarrollo de futuros Sistemas de Organização del Conocimiento, ya que utiliza recursos computacionales de última generación actualmente disponibles. Con las metodologías empleadas es posible prospectar términos candidatos de manera más precisa y eficiente, contribuyendo a la construcción de un vocabulario controlado, estandarizado y de calidad. Esto, a su vez, resulta en que la indexación y la recuperación de informaciones acerca del Trastorno del Espectro del Autismo sean más eficientes, además de reducir costos y promover avances en el área de estudios científicos relacionados.

Palabras-clave: Prospección de términos candidatos. Modelo Entidad-Relación. Análisis de datos. Trastorno del Espectro del Autismo.

1 Introdução

O Transtorno do Espectro Autista (TEA) foi identificado e nomeado apenas no século XX, por meio dos trabalhos dos psiquiatras Leo Kanner, austríaco, em 1943, e Hans Asperger, alemão, em 1944. Antes disso, havia uma crença generalizada de que crianças não poderiam ter distúrbios mentais, o que gerava muito preconceito e exclusão social para as pessoas com deficiência mental, que eram frequentemente internadas em hospitais. Além disso, termos pejorativos eram comumente utilizados para descrever essas pessoas. Atualmente, são adotados vocabulários controlados para garantir a utilização apropriada de termos em instituições e centros de informação, melhorando a qualidade da indexação e facilitando o armazenamento e a recuperação de documentos para os usuários. Desde o século passado, os estudos sobre o TEA se multiplicaram, devido à sua crescente relevância e incidência na sociedade.

Este presente artigo destaca a importância da criação de um vocabulário controlado para auxiliar na indexação e recuperação eficaz de informações relacionadas ao Transtorno do Espectro Autista (TEA) nas áreas de estudo e na sociedade científica como um todo. Essa estrutura semântica ajuda a controlar sinônimos, padronizar termos e diferenciar homografias, reunindo termos com o mesmo significado. A indexação é considerada a parte mais importante para obter resultados precisos em uma estratégia de busca. De acordo com Fujita (2003, p. 61), a indexação — que consiste na criação de índices — é uma prática antiga no tratamento de documentos e é conhecida como o processo de análise de conteúdo desses documentos.

Complementar à prática de indexação, existe o estudo sobre a Terminologia que, para Cabré (1995), este

estudo como objeto, busca diferentes distinções, tais quais podem ser averiguadas pela linguística como um conjunto de termos caracterizados por seus subconjuntos dentro dos componentes léxicos da gramática em um idioma específico. Dando linha a este seguimento, o processo terminológico se dá por um conjunto de etapas que envolve primeiramente a delimitação do domínio, subdomínio, classe e subclasse, categorias, termos (termos candidatos) e a definição do termo dentro de um contexto na literatura formal.

O conhecimento da funcionalidade do domínio comum de termos técnicos e científicos, aliciado à convicção de que a formação dos estudantes de Letras, em especial, os tradutores, em muito, se beneficiaria com os estudos de Terminologia, aceitamos o desafio para qual fomos chamadas: organizar e divulgar terminologias de áreas de interesse da sociedade [...]. (KRIEGER; MACIEL, p. 13, 2001)

A combinação dos métodos terminológicos, a indexação apropriada e as propriedades dos recursos computacionais resultam num conjunto de benefícios aos pesquisadores; dentre eles, pode-se citar a diminuição na construção, desenvolvimento e manutenção dos possíveis vocabulários controlados.

Em suma, os vocabulários controlados podem ser ferramentas que dão garantia literária e qualidade em informações relativas ao TEA ou ao tema específico de cada vocabulário, sendo preciso considerar os desafios que virão a ser enfrentados na construção, desenvolvimento, manutenção e atualização, tais como os custos altos do manejo. Por isso, traz-se a prospecção de termos candidatos na aplicação das seguintes metodologias como uma resolução deste problema, acarretando o barateamento dos custos de todo um processo.

Estes procedimentos podem trazer o benefício na identificação de potenciais termos úteis que caracterizam o Transtorno do Espectro Autista, que podem ser amplamente utilizados por toda a comunidade científica e a comunidade discursiva do TEA.

A incorporação dos futuros vocabulários controlados que possam ser resultantes da aplicação dessas metodologias garante que informações relacionadas ao TEA sejam mais precisas e acessíveis a toda sociedade, bem como o reforçamento da inclusão de pessoas autistas nas diferentes perspectivas representadas por esta comunicação.

2 Desenvolvimento

2.1 Metodologia

A elaboração de um vocabulário controlado não apenas auxilia nos processos documentais, mas, até o momento presente, não há um vocabulário controlado disponível para este tema. A criação de tal vocabulário, utilizando técnicas de análise exploratória, resulta em uma redução nos custos tanto para sua construção quanto para sua manutenção contínua. O objetivo desta pesquisa consiste em elencar a prospecção de termos candidatos no domínio do Transtorno do Espectro Autista e intencionalmente relacioná-los de forma que sejam agrupados na aplicação de algoritmos, considerando que existem diversos termos relacionados ao tema que fazem parte do vocabulário utilizado pela comunidade discursiva em questão e pelos pesquisadores.

Comunidades discursivas são determinados grupos focais que partilham de um interesse em comum. Aqui irá ser trabalhado diretamente sob os interesses da comunidade discursiva do Transtorno do Espectro Autista e os registros referentes ao tema. Esses registros referem-se aos documentos criados e compartilhados para ou pela própria comunidade discursiva entre os membros e simpatizantes, podendo ser registros informais ou formais, que incluem textos, artigos, reuniões, palestras e afins. Os registros da comunidade discursiva permitem o compartilhamento e disseminação da informação, estabelecendo valores à comunidade e servindo de fonte de informação para pesquisadores; no entanto, é válido ressaltar que esses registros nem sempre são acessíveis ou compartilhados livremente além da comunidade.

Este artigo utiliza os preceitos da análise de conteúdo, que aplica métodos objetivos, sistemáticos e quantitativos. Segundo Bardin (2004), a análise de conteúdo é “um conjunto de instrumentos metodológicos cada vez mais subtis em constante aperfeiçoamento, [...] até a extração de estruturas traduzíveis em modelos baseados na hermenêutica”.

Essa metodologia de pesquisa visa analisar dados de forma que os envolvam em um método sistemático para identificar e classificá-los.

Assim, este estudo adotará os princípios da análise temática e análise exploratória, seguindo duas abordagens da análise temática: a análise indutiva, que se baseia nos dados coletados pela pesquisa e não busca emitir opiniões ou ser influenciada pelo pesquisador; e a análise teórica, que explora os dados analisados e pode ser influenciada pelo analisador (ROSA; MACKEDANZ, 2021, p. 11, apud BRAUN; CLARKE, 2006).

A ciência de dados é responsável pelo tratamento de dados brutos através da coleta, análise e interpretação destes dados para gerar conhecimento através da análise exploratória a partir de uma hipótese inicial, conhecido como *insight*, e adquirir informações relevantes. De acordo com Oliveira e McDonell (2018, p. 4), esta ciência pode ser definida como “um processo de transformação de dados em informação, que são usadas para apoiar a tomada de decisões”. Para isso, existe o *workflow* da ciência de dados, que se configura numa estrutura útil para orientar todos os processos na análise dos dados, composto por várias etapas. Inicia-se com a coleta dos dados, sendo necessário obter dados confiáveis e substanciais de fontes públicas ou privadas. Oliveira e McDonell realçam a importância de se entender o contexto em que a informação está inserida, sendo significantes para a pesquisa. Seguindo este pensamento, a próxima etapa se configura na preparação dos dados. Isso inclui a limpeza, a seleção de variáveis e a escolha de um formato adequado para a manipulação destes dados. A etapa de análise de dados consiste na aplicação da estatística para que a pergunta-chave da pesquisa seja respondida. A comunicação dos resultados busca representar a pesquisa de forma objetiva e clara, sendo assim demonstrada da forma mais viável como um gráfico, um relatório ou até mesmo um artigo científico. Utilizar o *workflow* da ciência de dados é essencial, pois garante que as etapas sistemáticas de um processo sejam consistentes e rigorosas para a resolução de um problema.

A análise exploratória é uma importante ferramenta na pesquisa científica de várias áreas do conhecimento. Ela permite o uso de métodos estatísticos para investigar uma hipótese na pesquisa qualitativa, utilizando um processo indutivo ou dedutivo. Com isso, é possível avaliar a concordância entre as análises usadas como metodologia de pesquisa. A análise exploratória ajuda a identificar padrões, tendências e relações entre as variáveis e auxilia na escolha das técnicas estatísticas apropriadas para tratar os dados, resultando em precisão e confiança em sua conclusão, utilizando essa abordagem que contribui para o avanço do conhecimento científico.

Também será abordada a técnica da modelagem entidade-relacionamento (MER). É uma técnica que consiste em representar entidades de forma semântica e visual.

De acordo com Elmarsi e Navathe (2010), a modelagem consiste em três componentes principais: entidades, atributos e o relacionamento entre elas. A modelagem entidade-relacionamento é muito utilizada na área de sistemas de informação na projeção de banco de dados, pois permite uma representação formal na relação entre dados.

A metodologia utilizada conta com o suporte de softwares e programas livres dedicados, que auxiliam em todas as etapas da prospecção de termos candidatos para futuro desenvolvimento de vocabulários controlados. Processo de curadoria e análise dos dados são fundamentais para garantir a qualidade dos vocabulários controlados, sendo realizados de forma rigorosa e criteriosa, ajudando a eliminar erros e imprecisões. Essas etapas incluem a obtenção, tratamento e curadoria dos dados, além da análise dos mesmos. A utilização dessas ferramentas é fundamental para garantir maior precisão e eficiência no processo de construção dos vocabulários controlados. Esses softwares e programas livres específicos são projetados para ajudar na identificação e organização de termos relacionados ao TEA, tornando o processo de prospecção mais ágil e eficiente. O uso de tecnologias modernas ajuda a garantir a atualização constante dos vocabulários controlados, garantindo que as informações estejam sempre corretas e atualizadas, possibilitando uma ampliação significativa do vocabulário controlado, tornando-o mais completo e preciso. Para isso, utilizam-se os seguintes softwares:

- HTTrack Website Copier — tem como objetivo retirar todas as informações de um site, sendo elas informações que estão de livre acesso aos usuários ou até mesmo informações que são bloqueadas por *bots*, para que nem todos tenham acesso. Dentre as informações possíveis, pode-se citar arquivos em HTML, PDF, imagens, áudios, entre outros. Foi utilizado na extração do corpus documental que será fonte de transformação da proposta de vocabulário controlado;
- Open Refine — permite a manipulação de uma quantidade exagerada de dados, compilando-os e padronizando de forma útil a uma análise. Ele se evidencia em meio a outros, pois consegue trabalhar com uma gama alta de dados difíceis de serem analisados. Desta forma, criam-se colunas, que separam os documentos do corpus como em uma forma prévia de um banco de dados, podendo retirar do texto através de funções todas as informações pertinentes — siglas, acrônimos, datas. Também realiza a função de extração dos

dados em alguns formatos, podendo ser em HTML, .xls, .xlsx, SQL e outros;

- LibreOffice Banco de Dados — permite a criação e gerenciamento de bancos de dados de forma fácil e gratuita. Com ele, é possível criar tabelas, formulários, consultas e relatórios para organizar e manipular dados de maneira eficiente.

Para que estes processos se desenvolvam, é necessária a utilização da análise de dados; uma abordagem capaz de coletar e processar informações de dados brutos. Essa metodologia é utilizada em várias áreas do conhecimento e, para que seja efetivada, existem algumas etapas que podem ser seguidas, sendo elas: a preparação dos dados, exploração, análise descritiva, análise inferencial, qualitativa, interpretação e conclusão dos resultados.

Houve várias tipologias para classificar e identificar tipos de estratégias de métodos mistos que os desenvolvedores podem usar em seu estudo de métodos mistos proposto. Creswell e Plano Clark (2007) identificam 12 sistemas de classificação retirados dos campos de avaliação, enfermagem, saúde pública, política e pesquisa em educação e pesquisa social e comportamental. Nestas classificações, os autores usam termos diversos para seus tipos de designs e há uma quantidade substancial de sobreposição nas tipologias. (CRESSWELL, 2009, p. 206)

Na fase de preparação dos dados, verifica-se a qualidade e a organização da informação e os transforma em um formato apropriado para essa análise. Neste caso, utilizou-se o software HTTrack, para realizar essa coleta, e de um conversor de PDF, a fim de transformá-los em TXT. Na exploração dos dados, identificam-se de maneira sistemática os padrões e relações entre estes dados e, na análise descritiva, tem-se como finalidade descrever e resumir os aspectos de um conjunto de dados massificado; para isso, foi utilizado o software Open Refine, que dispôs as informações de dados brutos em um formato de planilha, e a partir da aplicação de uma função, foi solicitada a demonstração das facetas onde se encontravam as siglas do corpus analisado. Na etapa inferencial, criam-se hipóteses para a identificação de variáveis, analisando-as de forma mais geral, e desta forma, aplica-se a etapa qualitativa, onde organizam-se, classificam-se e categorizam-se os dados. Nesta parte da pesquisa, foram excluídas siglas impertinentes e irrelevantes à prospecção.

Será seguido o *workflow* da análise dos dados. Cada software utilizado nesta pesquisa atua diretamente em cada etapa desse *workflow*. Usa-se o HTTrack como coletor de dados dentro de um determinado site. Neste caso, usa-se o link do corpus, que é colocado dentro do programa e aplicam-se as funções que permitem a coleta destes dados, e é nesse momento que o pesquisador escolhe qual o tipo de dado ele gostaria de obter, podendo ser imagens, PDFs, arquivos em

HTML, MP3; tudo depende da necessidade do pesquisador. Para a realização da segunda etapa, é necessário que o pesquisador escolha qual tipo de dado vai ser tratado e como será trabalhado. Nesse caso, foram utilizados PDFs dos periódicos do Canal Autismo, convertidos em TXT, para que estes dados ficassem mais leves e fossem mais fáceis de manipular. Assim sendo, a segunda e terceira etapa são progressivas, isto é, a segunda etapa finaliza uma parte do procedimento e a terceira etapa continua este processo. Para dar essa continuidade, utiliza-se o software Open Refine, que emprega o corpus em TXT em formato de tabela, como uma forma prévia de um banco de dados. De forma inicial, as colunas que ele nos resulta são: *files*, relativo aos documentos inseridos; *column*, relativo às linhas de cada coisa escrita no documento, e o número de linhas. Aqui, escolhem-se quais informações de que pesquisador necessita, aplicando funções para a retirada da informação destes textos, resultantes em uma outra coluna *x* baseada na coluna solicitada. Como exemplo, aplica-se uma função baseada na coluna de texto solicitando que sejam retiradas as siglas iniciadas de A – Z dentro de parênteses ou não. Aplicando as facetas falsas, resultariam nas linhas do texto e do documento onde se encontram essas siglas. Formalmente, utiliza-se o Libre Office como último software para o tratamento dos dados na agregação de consultas, resultando em um banco de dados, e que, por fim, apresentado nessa pesquisa, também resulta em uma tabela e um diagrama correspondentes aos resultados de forma clara e objetiva.

Foi escolhido prospectar siglas como termos candidatos, pois elas podem ser incorporadas no discurso em uma forma privilegiada de extração destes termos com baixo custo e sem ter o vocabulário controlado à disposição, justamente porque se destacam, e isso faz que os procedimentos fiquem mais viáveis ao pesquisador, isto é, textos são compostos por muitas palavras, sendo difíceis de manipulá-los manualmente para que se encontrem termos relevantes. Além disso, termos possuem variações de significados e sinônimos, o que dificulta a identificação deles. Dessa forma, utilizar siglas na prospecção de termos candidatos faz que se economize tempo e recursos, tornando as operações automatizadas. Isso implica que não necessariamente se precisa de supermáquinas para realizar essa atividade e muito menos de um conhecimento prévio.

2.2 Resultados e Discussão

A Revista do Canal Autismo foi selecionada como fonte para a extração do corpus desta pesquisa. A revista é uma publicação trimestral disponível online e de forma gratuita, que tem como objetivo divulgar informações sobre o autismo para a comunidade

discursiva autista, familiares, pesquisadores e profissionais da área.

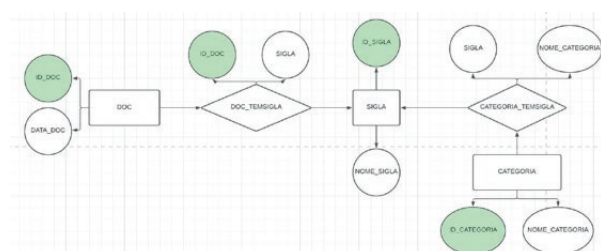
Em 23 de abril de 2010 [...] criaram a Revista Autismo, gratuita, impressa e digital. Foram publicadas edições nos anos de 2010 a 2013. De 2014 a 2018, houve uma pausa nas publicações impressas, pois até então, a revista funcionava como um projeto de uma ONG, sem anúncios publicitários e vivendo somente de doações. (REVISTA AUTISMO, 2010-).

O intuito da revista é promover a inclusão e valorização dos autistas, difundindo o conhecimento sobre o tema. Com a escolha dessa fonte, obteve-se um corpus rico em informações atualizadas e relevantes sobre o autismo, que podem ser utilizadas para análises de termos candidatos e construção de vocabulários controlados. A disponibilidade online e gratuita da revista também permite que mais pessoas possam ter acesso ao conhecimento sobre o autismo, contribuindo para a conscientização e desmistificação do transtorno.

Através da análise de conteúdo, recuperou-se o modelo entidade-relacionamento e, a partir disso, foram feitas consultas de agregação no banco de dados do Libre Office que consolidaram as categorias destes termos. Destas listas de termos, foram excluídas siglas que representassem regiões geográficas, universidades e organizações.

A construção de um diagrama MER é baseado em um conjunto de regras, tais como a identificação de uma chave primária para cada identificação de cada entidade. Através do diagrama a seguir, pode-se observar precisamente as entidades, suas características e suas relações.

Figura 1. Diagrama do Modelo Entidade-Relacionamento



Fonte: Elaborado pelos autores.

Este é um diagrama de modelo entidade-relacionamento, que exhibe as entidades e suas atribuições, bem como as tabelas que estabelecem relações entre elas. Na tabela DOC, é possível notar a presença de uma chave primária denominada ID_DOC e uma coluna que indica as datas dos documentos. Essa tabela está conectada a uma tabela de relação chamada DOC_TEMSIGLA, que possui o ID_DOC como chave primária e a coluna de siglas.

Tabela 1. *Diagnóstico e Sintoma*

CID	Classificação Internacional de Doenças
DI	Déficit Intelectual
PEA	Perturbação do Espectro Autista
TDAH	Transtorno do Déficit de Atenção e Hiperatividade
TEA	Transtorno do Espectro Autista
TGD	Transtornos Globais do Desenvolvimento
TOC	Transtorno Obsessivo-Compulsivo
TP	Transtorno da Personalidade

Fonte: Elaborado pelos autores.

Tabela 2. *Intervenção*

AAF	Abordagem de Análise Funcional
ABA	Análise do Comportamento Aplicada
AFF	Abordagem de Família e Foco
CAA	Comunicação alternativa e Ampliada
EA	Educação para Autistas
TANGG	Terapia Assistida por Animais com Necessidades Especiais

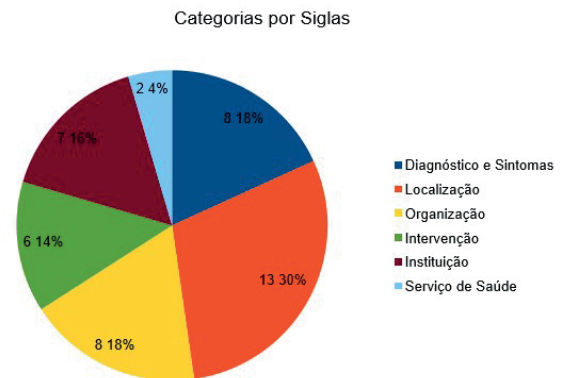
Fonte: Elaborado pelos autores.

Tabela 3. *Serviço de Saúde*

MSP	Ministério da Saúde
SUS	Sistema Único de Saúde

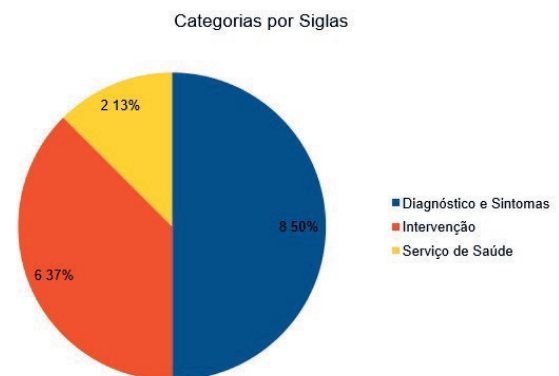
Fonte: Elaborado pelos autores.

Essas três tabelas acima são o resultado das consultas de agregação do banco de dados feito. É consistente em três categorias principais: Diagnósticos e Sintomas, Intervenção e Serviço de Saúde. Cada um desses tópicos possui siglas seguidas de suas definições. Para diagnósticos e sintomas, as siglas de relevância são aquelas que remetem aos diagnósticos clínicos com base no Transtorno do Espectro Autista. Na tabela de intervenção, as siglas selecionadas fazem parte na metodologia utilizada pelos profissionais da saúde no tratamento de pacientes com TEA e, por fim, na tabela de Serviço e Saúde, foram selecionadas siglas relativas ao cenário dos serviços ofertados que tratam diretamente das questões do TEA.

Figura 2. *Gráfico inicial sem limpeza dos dados*

Fonte: Elaborado pelos autores.

O gráfico nos mostra a representação em porcentagem da quantidade de siglas relacionadas a cada categoria. Foi gerado a partir de um resultado geral e prematuro sem passar pela limpeza dos dados inúteis ao resultado final. No gráfico, é possível observar que as siglas prospectadas em organizações, instituições e regiões geográficas são compatíveis com a maioria da prospecção realizada, mas, ainda assim, não são dados que representam o domínio e nem as categorias necessárias. Após passar pela correção de pertinência das siglas prospectadas, tem-se o seguinte gráfico:

Figura 3. *Gráfico final das consultas de agrupamento*

Fonte: Elaborado pelos autores.

O segundo gráfico nos mostra que a categoria que contempla o maior número de siglas prospectadas do corpus é a de Diagnóstico e Sintomas. Consequentemente, a categoria Intervenção, que representa os métodos utilizados para o tratamento de autistas, corresponde à segunda posição com o maior número de siglas prospectadas. E, por último, mas não menos importante, vem a categoria de Serviços de Saúde, correspondendo a 2,3% das siglas prospectadas.

Em relação às siglas, foi identificado que a categoria Diagnósticos e Sintomas representa 8,50% do total, com um total de oito siglas. Já Intervenção corresponde a 6,37% e possui seis siglas, enquanto Serviço de

Saúde corresponde a 2,13%, com apenas duas siglas. Para evitar duplicação de siglas, não é necessário considerar a frequência com que elas aparecem no corpus, já que esse dado não é relevante na construção de um vocabulário controlado. Isso pode ser problemático quando se trata de termos candidatos, uma vez que as palavras têm sinônimos que podem levar à perda de tempo na pesquisa. Dessa forma, é importante eliminar as siglas duplicadas e utilizar termos, siglas ou palavras em sua forma completa para a garantia da eficiência e o encontro da informação com maior facilidade e precisão.

3 Conclusão

Atualmente, os pesquisadores possuem ferramentas poderosas para aprofundar suas investigações em diversas áreas do conhecimento, graças aos avanços tecnológicos e às metodologias disponíveis para a coleta, análise e disponibilização de recursos computacionais. À medida que há atualizações no mundo da tecnologia, a distribuição de poder computacional aponta uma vasta gama de variedades de desempenho, facilitando a análise de dados para que não se faça isso manualmente, pois permite que os algoritmos tratem de dados massivamente, identificando padrões e tendências em grandes conjuntos de dados.

Um das áreas que podem utilizar esse benefício é a linguística e a ciência da informação, que podem desfrutar as metodologias aplicáveis na prospecção de termos candidatos para o desenvolvimento de vocabulários controlados, a fim de auxiliar na indexação e recuperação eficiente de informações e para dar visibilidade às comunidades discursivas que posteriormente estarão relacionadas aos temas escolhidos para esta construção. Essas análises permitem que os pesquisadores identifiquem padrões e tendências em um grande volume de dados, contribuindo para uma visão mais ampla e precisa sobre os temas de estudo.

De fato, é relevante que os pesquisadores estejam abertos às novas metodologias e às novas tecnologias, e que desfrutem de tal eminência, levando em consideração que estes pesquisadores devam ter uma visão crítica na utilização destes recursos, evitando a reprodução de estigmas que possam comprometer a qualidade da aplicação e da usabilidade das análises.

Mediante isto, os procedimentos de coleta, análise e disponibilização dos recursos computacionais, indicase a essencialidade deles para a pesquisa científica contemporânea, oferecendo vantagens na construção e manutenção dos vocabulários controlados, a fim de garantir a melhoria nas pesquisas científicas.

Referências

- CABRÉ, Maria Teresa. La terminología hoy: concepciones, tendencias y aplicaciones. **Ciência da Informação**. v. 24, n. 3, 1995. Disponível em: <https://revista.ibict.br/ciinf/article/view/567/568>. Acesso em: 20 abr 2023.
- CRESWELL, John W. Writing. Evaluating Mixed Methods Research. *In*: CRESWELL, John W. Writing. **Research design: qualitative, quantitative, and mixed methods approaches**. Los Angeles: SAGE, 2014. cap. 10, p. 213-240. Disponível em: <http://www.ceil-conicet.gov.ar/wp-content/uploads/2015/10/Creswell-Cap-10.pdf>. Acesso em: 19 abr. 2023.
- ELMASRI, Ramez; NAVATHE, Shankant B. **Sistemas de banco de dados**. 6. ed. São Paulo: Pearson, 2011. 1168 p. Disponível em: https://tonysoftwares.com.br/attachments/article/5297/Sistema_de_banco_de_dados_Navathe.pdf. Acesso em: 19 abr 2023.
- FUJITA, Mariângela Spotti Lopes. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**: Campinas, v. 1, n. 1, p. 60 – 90, jul./dez., 2003. ISSN: 1678-765X. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2089>. Acesso em: 20 abr 2023.
- MAIA, Regina M.C.S; ALVARENGA, Lídia. Teoria da Classificação Facetada e Contribuições para o Modelo Entidade Relacionamento. **DataGramZero - Revista de Ciência da Informação**, Rio de Janeiro, v. 5, n. 1, fev. 2004. Disponível em: http://www.brapi.inf.br/_repositorio/2010/07/pdf_18564.pdf. Acesso em: 19 abr 2023.
- MEDRI, Waldir. **Análise exploratória de dados**. 2011. Lato Sensu (Especialização em Estatística) – Centro de Ciências Exatas, Departamento de Estatística, Universidade Estadual de Londrina, Londrina, 2011. Disponível em: <https://docs.ufpr.br/~benitoag/apostilamedri.pdf>. Acesso em: 19 abr 2023.
- OLIVEIRA, Paulo Felipe de; GUERRA, Saulo; MCDONNELL, Robert. **Ciência de dados com R: introdução**. Brasília: Editora IBPAD, 2018. 240p. Disponível em: https://edisciplinas.usp.br/pluginfile.php/7376061/mod_resource/content/1/Cien%20de%20Dados%20com%20R%20-%20Intro.pdf. Acesso em: 20 abr. 2023.
- REIS, Marcelo Menezes; LINO, Manoel de Oliveira. **Introdução e análise exploratória de dados**. Disponível em: https://www.inf.ufsc.br/~marcelo.menezes.reis/Caps1_e_2.pdf. Acesso em: 15 abr 2023.
- REVISTA AUTISMO. São Paulo: Canal Autismo, 2010-. ISSN: 2596-0539. Disponível em: <https://www.canalautismo.com.br/revista/>. Acesso em: 19 abr 2023.

ROSA, Liane Serra da; MACKEDANZ, Luiz Fernando. A análise temática como metodologia n pesquisa qualitativa em educação em ciências. Blumenau: **Revista Atos de Pesquisa em Educação**, 2021, v. 16, ISSN 1809-0354. Disponível em: <https://proxy.furb.br/ojs/index.php/atosdepesquisa/article/download/8574/4963/#:~:text=Portanto%2C%20a%20an%C3%A1lise%20tem%C3%A1tica%20envolve,os%20padr%C3%B5es%20repetidos%20de%20significado>. Acesso em: 15 abr 2023.

SANTOS, Raquel Santana. Linguística de corpus e análise crítica do discurso: um diálogo possível. **Prisma.com**, [S.l.], v. 10, n. 19, p. 131-160, 2010. Disponível em: <http://aleph.letras.up.pt/index.php/prismacom/article/view/1870>. Acesso em: 19 abr 2023.

TEIXEIRA, Edson Belo. A análise de dados na pesquisa científica: importância e desafios em estudos organizacionais. **Desenvolvimento em Questão**, Ijuí, v. 1, n. 2, p. 177-201, 2011. DOI: 10.21527/2237-6453.2003.2.177-201. Disponível em: <https://revistas.unijui.edu.br/index.php/desenvolvimentoemquestao/article/view/84>. Acesso em: 19 abr 2023.

A relevância da análise de domínio: um estudo da classificação das atividades-fim nas Instituições Federais de Ensino Superior - IFES

The relevance of domain analysis: a study of the classification of core activities in Federal Institutions of Higher Education

Rosale de Mattos Souza (1), Rosa Zuleide Lima de Brito (2) Andrea Gonçalves dos Santos (3)

(1) Universidade Federal do Estado do Rio de Janeiro, Brasil, rosale.m.souza@unirio.br (2) Universidade Federal da Paraíba, Brasil, rosazuleide@hotmail.com (3) Universidade Federal do Rio Grande, Brasil, andreasantos@furg.br

Resumo

Este estudo interdisciplinar entre a Arquivologia, Organização do Conhecimento e a Ciência da Informação visa analisar o Código de Classificação referente às Atividades-Fim das Instituições Federais de Ensino Superior – IFES à luz da Teoria da classificação e da Análise de Domínio onde observaram-se inconsistências e lacunas existentes na Portaria AN/MJ nº 92/2011. O Grupo de Trabalho (GT) de classificação do Comitê Nacional dos Arquivistas das IFES (CNIFES) desde 2018 se reúne a fim de estudar e atualizar o código, visando sua otimização, para melhorar a lógica, a racionalização, a comunicação e promover a facilitação do uso da Tabela de Temporalidade de Documentos e Gestão de Documentos nas Instituições. Como metodologia utilizou-se o diálogo da Organização do Conhecimento, da Teoria da Classificação e Classificação Arquivística, verificando classes, sub-classes, atividades, conceitos, termos descritores, atos e documentos produzidos e recebidos por estas instituições. Como metodologia empírica o uso do Tesouro do CIBEC/INEP como parâmetro para análise dos termos e conceitos; analisando itens apontados pelo levantamento do GT realizado junto às IFES em 2018-2019. Apresenta-se propostas de termos, conceitos, novos códigos e justificativas na atualização. Como resultado, o GT - Classificação promoveu um relatório circunstanciado e encaminhado à Coordenação de Gestão de Documentos – COGED, do Arquivo Nacional. Como conclusões, sugerimos a revisão e atualização do Código de Classificação referente às Atividades-Fim das IFES; que a classificação arquivística e o domínio de conhecimento estejam dentro das agendas de pesquisas da Arquivologia, da Organização do Conhecimento e da Ciência da Informação.

Palavras-chave: Classificação. IFES. Organização do Conhecimento. Atividades-Fim. Análise de domínio.

Abstract

This interdisciplinary study between Archivology, Knowledge Organization and Information Science aims to analyze the Classification Code referring to core activities of Federal Institutions of Higher Education - IFES in the light of Classification Theory and domain analysis where they observed there are inconsistencies and gaps in Ordinance AN/MJ nº. 92/2011. The Classification Working Group (GT) of the IFES National Committee of Archivists (CNIFES) has been meeting since 2018 in order to study and update the code, aiming at its optimization, to improve logic, rationalization, communication and promote facilitation the use of the Table of Temporality of Documents and Document Management in the Institutions. As a methodology, the Knowledge Organization dialogue, Classification Theory and Archival Classification were used, verifying classes, sub-classes, activities, concepts, descriptive terms, acts and documents produced and received by these institutions. As an empirical methodology, the use of the CIBEC/INEP thesaurus as a parameter for analyzing terms and concepts; analyzing items pointed out by the GT survey carried out with the IFES in 2018-2019. Proposals for terms, concepts, new codes and justifications for updating are presented. As a result, the GT - Classification promoted a detailed report forwarded to the Document Management Coordination – COGED, of the National Archive. As conclusions, we suggest reviewing and updating the Classification Code referring to the core activities of the IFES; and that the archival classification and domain of knowledge are within the research agendas of Archivology, Knowledge Organization and Information Science.

Keywords: Classification. IFES. Knowledge Organization. Core-Activities. Domain analysis.

Resumen

Este estudio interdisciplinario entre Archivología, Organización del Conocimiento y Ciencias de la Información tiene como objetivo analizar el Código de Clasificación referente a las actividades finalísticas de las Instituciones Federales de Educación Superior - IFES a la luz de la teoría de la clasificación y el análisis del dominio donde se observaron inconsistencias y vacíos en la Ordenanza. AN/MJ N° 92/2011. El Grupo de Trabajo de Clasificación (GT) del Comité Nacional de Archiveros de IFES (CNIFES) se reúne desde 2018 con el fin de estudiar y actualizar el código, visando su optimización, mejorar la lógica, la racionalización, la comunicación y promover la facilitación del uso del Mesa de Temporalidad de Documentos y Gestión Documental en las Instituciones. Se utilizó como metodología el diálogo Organización del Conocimiento, la Teoría de la Clasificación y la Clasificación

Arquivística, verificar clases, subclases, actividades, conceptos, términos descriptivos, actos y documentos producidos y recibidos por estas instituciones. Como metodología empírica, el uso del tesauro CIBEC/INEP como parámetro de análisis de términos y conceptos; analizando los ítems señalados por la encuesta GT realizada con el IFES en 2018-2019. Se presentan propuestas de términos, conceptos, nuevos códigos y justificaciones para su actualización. En consecuencia, el GT - Clasificación promovió un informe detallado remitido a la Coordinación de Gestión Documental – COGED, del Archivo Nacional. Como conclusiones sugerimos revisar y actualizar el Código de Clasificación referente a las actividades finales de las IFES; y que la clasificación archivística y el dominio del conocimiento están dentro de las agendas de investigación de la Archivología, la Organización del Conocimiento y las Ciencias de la Información.

Palabras clave: Clasificación. IFES. Organización del Conocimiento. Actividades finales. Análisis de Dominio.

1 Introdução

A universidade pública federal, como qualquer instituição, foi criada para desempenhar finalidades, que neste caso específico, fazem parte das atividades-fim de ensino, pesquisa e extensão, oferecendo ensino superior de graduação e pós-graduação gratuitos e de qualidade à população. Desta forma, produz documentos e um domínio de conhecimento específico, o da educação de ensino superior, com seu contexto, atores, discursos, documentos, termos e informações.

Defendemos aqui a importância que a análise de domínio ajuda a perceber o discurso do ensino superior (graduação e pós-graduação): contexto, atores, suas funções, relações orgânicas numa perspectiva arquivística. Se bem que o levantamento da produção documental, assim como as tarefas operacionais nos arquivos repercutem não só na classificação, assim como em outras atividades e campos do ensino superior.

Para seu funcionamento, demandam atividades administrativas, Atividades-meio, que dão suporte às atividades-fim, que por sua vez se constituem nas atividades de cunho didático-pedagógicas e acadêmicas de ensino, pesquisa e extensão. Para que estas se realizem, a produção de documentos se faz necessária, uma vez que, são registros fundamentais das ações e negociações realizadas no cotidiano das instituições. O conhecimento arquivístico especializado é exigido do profissional arquivista, desde a produção documental, como a primeira etapa da gestão de documentos, até chegar às etapas de funções arquivísticas de classificação, avaliação, destinação, manutenção e uso, ou seja, no que tange a elaboração de instrumentos de Gestão de Documentos, como são os casos do Código de Classificação de Documentos – CCD, que na verdade deveria ser chamada de Plano de Classificação e a Tabela de Temporalidade e Destinação de Documentos – TTDD. De acordo com a Lei nº 8.159, de 8 de janeiro de 1991, a gestão de documentos, consiste em um

[...] conjunto de procedimentos e operações técnicas referentes à sua produção, tramitação, uso, avaliação e arquivamento em fase corrente e intermediária, visando a sua eliminação ou recolhimento para guarda permanente (BRASIL, 1991, s/p)

A mesma lei ainda dispõe que a gestão de documentos é dever do Poder Público, e que todos têm direito a receber dos órgãos públicos informações de seu interesse particular ou de interesse coletivo ou geral contidas em documentos de arquivos.

O problema de pesquisa trata sobre as inconsistências verificadas entre a teoria e a prática, que podem ser observadas nos códigos de classificação existentes em instituições arquivísticas, em particular nas universidades e instituições de ensino superior brasileiras, cujas metodologias não apresentam de forma clara, fundamentos teóricos científicos consistentes em suas elaborações.

Verificaram-se muitas solicitações de alterações na estrutura das classes e sub-classes no Código (Plano) de Classificação das IFES, associado à Portaria 92, de 2011, do Arquivo Nacional, da parte dos representantes das instituições Federais de Ensino. Assim, a partir do levantamento realizado em 2018, incluindo as sugestões da criação de novos descritores com as respectivas justificativas, que foram discutidas pelas autoras e levadas para apreciação em reuniões com os demais membros do grupo.

Com base no exposto, este trabalho tem como objetivo geral apresentar os resultados e propostas das análises realizadas no Código de Classificação referente às Atividades- Fim das IFES no Brasil, em particular os itens das subclasses na classe 100; que se subdivide nos seguintes objetivos específicos: Refletir sobre classificação arquivística e Teoria da Classificação; identificar o contexto institucional e o domínio de conhecimento em instituições de ensino superior; revisar e atualizar o Código de Classificação.

No que diz respeito ao aspecto teórico-metodológico, buscamos as fontes bibliográficas e o método exploratório, nos baseamos nos principais teóricos sobre estudo de análise de domínio, Teoria da Classificação, na Ciência da Informação e na Arquivologia.

Quanto à metodologia empírica, estabelecemos os seguintes passos: a partir dos resultados de entrevistas com representantes das IFES no período de 2018-2019, realizada pelo então GT-classificação da rede ARQUIFES, foram levantados os itens relacionados às classes e subclasses da classe 100 que apresentavam problemas de melhor escolha de termos descritores, atividades, e espécies documentais relacionados à

Portaria 92, de 2011, do Código de Classificação das IFES, com base no Thesaurus Brased do INEP sobre Educação, legislação específica e experiências práticas das arquivistas no domínio de narrativas, negociações, e discursos da área de educação de ensino superior, e utilização do método descritivo, a fim de alcançar os objetivos propostos.

Nessa perspectiva, registramos nesse artigo, a experiência obtida como membros do Grupo de Trabalho de Classificação e Temporalidade de documentos – GT Classificação da rede ARQUIFES (Rede Nacional de Arquivistas das IFES), criada com o objetivo de contribuir para a revisão do Código, a fim de fundamentar as demandas levantadas por meio de um questionário aplicado junto aos arquivistas das IFES que utilizam o instrumento criado e consolidado pela Coordenação de Documentos – COGED, do Arquivo Nacional e com representantes de Instituições Federais de Ensino Superior - IFES entre o período de 2006 a 2010, culminando com a Portaria AN/MJ nº. 92/2011.

2 Arcabouço teórico- metodológico: Organização do Conhecimento e Classificação Arquivística

Sobre a classificação de documentos arquivísticos em universidades e outras instituições de Ensino superior, pode tratar-se de uma temática ainda em busca de uma fundamentação teórica consistente, como aponta Smit (2018) que escreve sobre a classificação e representação por assunto (temática), vocabulário controlado e controle de vocabulário. Da mesma forma, Souza (2022) trata da classificação funcional (por funções) e discute o que é função, pois carece de melhor definição.

Sobre essa fragilidade, Sousa (2002) em seus escritos recomenda que se deve recorrer à Teoria da Classificação e Organização do Conhecimento, criadas por Dahlberg nos anos 1970 e 1980, já consolidadas na Biblioteconomia e por meio da Ciência da Informação, como interlocutora preferencial para esse exercício, por dois motivos fundamentais: a proximidade do objeto de estudo e a possibilidade de operação com elementos balizadores da Teoria da Classificação, da Representação da Informação e da Teoria do Conceito.

Nessa direção, Carlan e Medeiros (2011, p.55), afirmam que

[...] no contexto da Ciência da Informação, os Sistemas de Organização do Conhecimento - SOC ou esquemas de representação do conhecimento são instrumentos que fazem a tradução dos conteúdos dos documentos originais e completos, para um esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, conseqüentemente, facilitar a recuperação das informações contidas nos documentos. (CARLAN E MEDEIROS 2011, p.55)

Desta forma, o diálogo entre a área da organização do conhecimento e a classificação se aproximam do ponto de vista teórico e das atividades práticas.

Existe uma epistemologia da Organização do Conhecimento (OC) que circunda emblemáticos autores e pesquisadores, tais como, Ingetraut Dahlberg, que foi uma das fundadoras da Organização do Conhecimento nos anos 1970 e 1980, preocupada com estudos sobre classificação, termos e conceitos. Sabemos também que a área de organização do conhecimento tem suas raízes históricas na base teórica da classificação, que remonta desde a antiguidade nos filósofos Aristóteles e Porfírio.

Segundo Hjørland (2008), a OC se relaciona em atividades tanto em bibliotecas como em arquivos, visto que este campo de estudo está preocupado com a descrição, indexação, classificação, sistemas de organização do conhecimento usados para organizar e disponibilizar documentos, por meio de representações, palavras, aspectos semânticos e conceitos, levando em consideração as estruturas coletivas, a organização e significados sociais.

Para a elaboração do saber é necessário códigos linguísticos para a formação de conceitos, Dahlberg (1978) define conceito como:

[...] a unidade de conhecimento que surge pela síntese dos predicados necessários relacionados com determinado objeto e que, por meio de sinais linguísticos, pode ser comunicado (DALBERHG, 1978, p. 104).

É possível estabelecer que a comunicação utiliza a linguagem como uma ferramenta para a representação do saber.

Na Arquivologia os princípios arquivísticos devem ser considerados, em particular na função classificação, como o princípio de respeito aos fundos ou princípio da proveniência, criado por Nataly de Wally, em 1841, e que distingue a Arquivologia de outras áreas de conhecimento.

[...] O respeito aos fundos – para adotar aqui sua definição mais simples, deixando de lado todos os problemas de interpretação que abordaremos no decorrer deste trabalho – consiste em manter grupados, sem misturá-los a outros, os arquivos (documentos de qualquer natureza) provenientes de uma administração, de uma instituição ou de uma pessoa física e jurídica: é o que se chama de fundo de arquivos dessa administração, instituição ou pessoa. [...] (DUCHEIN, 1986, p.14)

Quanto ao Princípio da Proveniência, temos o seguinte conceito:

[...] a maior parte das definições de respeito aos fundos, se reportam, na verdade, a esta noção, a tal ponto que o equivalente da expressão respeito aos fundos, nos países de língua germânica e em vários outros é princípio da proveniência.” O fundo é definido como o conjunto de arquivos cuja proveniência é a mesma. (DUCHEIN, 1986, p. 23).

Já no que diz respeito à ordem original e ao princípio da territorialidade chegamos aos seguintes conceitos: o da Ordem Original é o princípio que mantém a ordem (classificação) em que os documentos foram organizados originalmente pelos órgãos produtores, ou seja, por aspecto cronológico, alfabético, geográfico, temático (por assunto), espécie documental, e etc. No que se refere ao Princípio da Territorialidade, este respeita o local geográfico de produção e criação dos documentos. E estes princípios estão associados aos princípios da Proveniência e de Respeito aos Fundos Documentais.

Conforme Sousa (2002) a função arquivística classificação de documentos é considerada matricial para o fazer arquivístico, e também pode se constituir na explicitação de funções ou estruturas dentro de um contexto de produção documental. É a representação do mapeamento deste contexto produtivo que subsidia as demais etapas da gestão de documentos. Assinala-se ainda que os Princípios arquivísticos acima expostos nos séculos XIX e XX foram primaciais para a organização do conhecimento na Arquivologia. A tradução e divulgação de normas internacionais de descrição arquivísticas, tais como a Norma Internacional de Descrição Arquivística - ISAD - G e a Norma Nacional de Descrição Arquivística – NOBRADE, pelo Arquivo Nacional do Brasil também foram decisivas no avanço da OC no Brasil.

Em suma, como apontado por Tognoli, Silva e Silva (2019), embora os princípios e conceitos devam ser respeitados da Arquivística – como o conceito de fundo, o princípio da proveniência e o conceito de documento de arquivo, a área se desloca da norma de custódia e preservação, para outras abordagens, tais como:

[...] direção a uma abordagem que privilegia, além da custódia, o teor social implicado nos processos de produção, organização e uso de documentos de arquivo – ressaltando, aqui a importância da interlocução com outras disciplinas, como a OC (TOGNOLI; SILVA; SILVA, 2019, p. 55).

Na classificação arquivística consideramos os seguintes elementos:

[...] vejam como elementos – função, organização e assunto- entram na classificação de documentos públicos. Visando ao estudo desse ponto, os métodos de classificação podem ser divididos em três: funcional, organizacional e por assunto (SCHELLENBERG, 2002, p.88).

Os documentos, uma vez agrupados por atividades, podem ser, além disso, agrupados de acordo com as funções e suas competências. Os grupos funcionais são as classes maiores normalmente criadas para a classificação de documentos de uma entidade (SCHELLENBERG, 2002, p. 90).

A Arquivologia, no que diz respeito a descrição, ainda possui a visão de uma série de atividades ligadas ao registro de aspectos objetivos do documento. A investigação da linguagem artificial e natural utilizada na arquivística é necessária para potencializar as normas de descrição, sistemas de recuperação da informação, bases de dados e instrumentos de pesquisa, além do atendimento ao usuário.

Se a estrutura orgânica se reflete num esquema de classificação, as classes primárias, em geral, representam os principais elementos organizacionais da repartição. A divisão em classes organizacionais é possível e aconselhável somente em governos de organização estável, e cujas funções e processos administrativos sejam bem definidos (SCHELLENBERG, 2002, p. 91).

Por exemplo, qual é a missão e quais são as funções da universidade pública federal no Brasil: Ensino, Pesquisa e Extensão? Se os documentos se referem a um campo especial de pesquisa, como por exemplo “química agrícola”, os cabeçalhos de assuntos ou tópicos devem corresponder às subdivisões lógicas daquele campo específico.

[...] Os cabeçalhos para arquivos extraídos de uma análise puramente lógica dos assuntos, abrangendo um campo de conhecimento humano, comparam-se àqueles sob os quais o material de bibliotecas é classificado (SCHELLENBERG, 2002, p. 93).

Desta forma, pode-se inferir que os processos, trâmites, funções, subfunções, atividades, atos, tipologia de documentos, tipos de documentos, assuntos são todos elementos necessários para a construção de uma taxonomia e de uma classificação na arquivística, levando em consideração a classificação funcional, a organizacional e a temática.

2.1 A relevância da análise de domínio para as funções arquivísticas

As funções arquivísticas, que são executadas no decorrer das atividades do arquivista, foram listadas por Rousseau e Couture (1998) em: produção, avaliação, aquisição, conservação, classificação, descrição e difusão. O cumprimento de cada uma dessas funções deve ser o objetivo da instituição arquivística, já que a sua implementação contribui para a gestão documental de forma adequada. Assim, cada uma delas contribui para a execução da próxima ou ainda depende da anterior para ser executada.

Para Hjørland (2004, *apud* NEGREIROS, 2008), um domínio pode ser:

[...] uma disciplina, um campo escolar. Pode ser ainda uma comunidade discursiva conectada a um partido político, à religião, ao comércio, ou a um lazer (HJORLAND, 2004, *apud* NEGREIROS, 2008, p. 81).

Nesse sentido, o autor explica que Hjørland (2004) relaciona e mostra como o domínio pode melhorar as práticas informacionais dentro da CI, prover melhor os serviços de informação e integrar diversos campos de pesquisa, como a recuperação da informação.

Conforme Danuello (2007) é por meio da análise de domínio torna-se possível verificar o que é efetivamente relevante ou significativo em um determinado campo, de tal modo que aspectos como tendências, padrões, processos, agentes e seus relacionamentos possam ser identificados e analisados. Assim a análise de domínio se caracteriza pelo estudo dos aspectos teóricos de dado entorno, constituindo um meio para a geração de novo conhecimento acerca da interação da comunidade científica com a informação (SMIRAGLIA, 2013).

Nesse sentido, Guimarães (2014) afirma que:

[...] o domínio, objeto de análise, decorre da aplicação de um princípio inerente à própria organização do conhecimento – a categorização – na medida em que a partir da identificação de um conjunto de traços comuns (que gera um critério ou diferença) é possível reunir coisas semelhantes e separar coisas diferentes, uma vez que coisas semelhantes tendem a se comportar de maneira semelhante: *Qui se rassemble se semble*. (GUIMARÃES, 2014, p. 16)

Na CI, segundo Guimarães (2014), o conceito de análise de domínio foi inicialmente utilizado por Hjørland e Albrechtsen (1995) como uma nova perspectiva de abordagem para a investigação da área, com grande ênfase no contexto (perspectiva sociológica) e menos no indivíduo (perspectiva cognitivista). Assim,

[...] a melhor maneira de se entender a informação na CI é estudar os domínios de conhecimento relacionados com suas comunidades discursivas, que são distintos grupos sociais sincronizados em pensamento, linguagem e conhecimento, constituintes da sociedade moderna” (NASCIMENTO; MARTELETO, 2004, p. 1)

Por sua vez, Hjørland (2002) defende que os recursos informacionais devem ser tratados (identificados, descritos, organizados e disseminados) para servirem a um objetivo específico. Para Negreiros (2008)

[...] Isto corrobora com a idéia de que, em arquivos, prima-se pela organização dos documentos para uma posterior utilização da informação, tendo como base o contexto da instituição que os produziu e aos objetivos, funções e atividades que os originaram. Muito embora os objetivos de documentos arquivísticos sirvam determinantemente como provas de ações, isso não os redime de serem insumos para a tomada de decisão das instituições e de servirem à pesquisa e a história (NEGREIROS, 2008, p. 82).

Significa dizer que os documentos e as informações ali existentes, representam um domínio específico por formarem um acervo específico, produzido um contexto específico.

Hjørland (2002) apresenta onze abordagens para a análise de domínio, sendo: produção de guias de literatura; Construção de classificações e tesouros especializados; Indexação e recuperação da informação especializada; Estudos empíricos de usuários; Estudos bibliométricos; Estudos históricos; Estudos de documentos e gêneros; Estudos epistemológicos e críticos; Estudos terminológicos, linguagens para propósitos determinados, semântica de bases de dados e estudo dos discursos; Estrutura e instituições da comunicação científica, e; Cognição científica, conhecimento perito e inteligência artificial.

Estas abordagens estão relacionadas à análise de domínio, cada uma com aplicações e metodologias de estudo e entendimento de campos específicos do conhecimento. Negreiros (2008) explica que

[...] A instituição na qual o arquivo e seu conjunto de documentos está inserido será como um domínio a ser analisado para que o processo de organização e recuperação possa ser realizado de acordo com as necessidades específicas desse domínio e a solução de sistema in-formatizado de gerenciamento de documentos, siga o mesmo caminho (NEGREIROS, 2008, p. 83).

Os procedimentos metodológicos utilizados na elaboração de instrumentos de gestão documental e de recuperação da informação encontram alinhamento com os elencados por Hjørland (2002), porém eles se constituem documentos únicos e específicos para cada instituição, como no caso das IFES relacionados com as funções arquivísticas desempenhadas, citadas por Rousseau e Couture (1998). Sendo que cada uma das tipologias de instrumentos de gestão ou de recuperação de documentos e informações em arquivos servem a um objetivo específico, podem ser identificados: Código (ou Plano) de Classificação de Documentos; Tabela de Temporalidade e Destinação de Documentos; Guia; Inventário; Catálogo; Índice, dentre outros.

Hjørland (2002) aponta que a maioria das pesquisas em classificação estão direcionadas para os esquemas de classificação universais (como os bibliográficos), e que a indexação e a recuperação da informação embora sejam sempre específicas, os estudos da CI têm ignorado o modo como domínios diversos demandam os diferentes sistemas para organizar e recuperar documentos.

Nesse sentido, a Teoria da Análise de Domínio deve estudar seu contexto geral, os contextos institucionais, a análise de discurso, a terminologia e o vocabulário controlado para a correta elaboração dos códigos de classificação e conseqüentemente, as tabelas de temporalidade e destinação de documentos. Corroborando, Hjørland (2002) ainda defende que focar mais fortemente as necessidades de domínios diferentes pode tornar os estudos mais realistas e os profissionais mais relevantes.

3 A metodologia empírica

No primeiro momento foram distribuídas as tarefas por códigos 100 (Ensino superior), 200 (Pesquisa), 300 (Extensão), 400 (Educação básica e profissional) e 500 (Assistência estudantil) entre membros do GT classificação da rede ARQUIFES, visando o estudo ponto a ponto das sugestões, acréscimos e mudanças propostas pelos arquivistas no período 2018-2019.

Vale ressaltar que a classe analisada nesse estudo é a 100 (Ensino superior), que representa a primeira função ou atividade fim das IFES, e suas subclasses:

Tabela 1. *Subclasses da Classe 100 – Ensino Superior*

110	Normatização. Regulamentação
120	Cursos de graduação (inclusive na modalidade à distância)
130	Cursos de pós-graduação <i>strito sensu</i> (inclusive na modalidade à distância)
140	Cursos de pós-graduação <i>lato sensu</i> (inclusive na modalidade à distância)
190	Outros assuntos referentes ao ensino superior

Fonte: Dados da pesquisa.

Os membros participantes do sub grupo, responsáveis pela reunião dos formulários referentes à classe 100, se encarregaram de realizar a revisão e análise das sugestões, acréscimos e mudanças propostas pelos arquivistas no período de 2018-2019. As discussões apontaram a fragilidade da metodologia adotada na criação das classes e subclasses adotadas para a classe 100, onde não se identificam de forma clara, quais critérios foram atribuídos em determinados níveis criados na elaboração do código de classificação.

Foram discutidas novas propostas de assuntos/tipos documentais que poderiam ser inseridas e novas reestruturações na classe, que foi apresentada ao GT e encaminhadas em relatório circunstanciado à Coordenação de Gestão de Documentos – COGED, do Arquivo Nacional. Também foi utilizada como metodologia empírica o uso de termos e conceitos do Tesouro do CIBEC/INEP como parâmetros de análise.

4 Discussões e resultados

Deve-se considerar que o referido Código de classificação das IFES no Brasil é a primeira versão, publicado há mais de dez anos, onde ainda não foi realizada nenhuma atualização. Também é importante mencionar que a dinâmica e os processos das instituições, assim como o conhecimento muda e evolui, fazendo com que os instrumentos de gestão de documentos precisem ser revisados e atualizados periodicamente, como é o caso dos citados nesse estudo.

O Código de Classificação de Documentos (CCD) de arquivo, pela atribuição da codificação numérica, é adotada pelo Conselho Nacional de Arquivos (CONARQ), por meio do método de numeração decimal. Tomando como base a mecânica do sistema de classificação decimal, desenvolvido por Dewey, com a divisão dos conjuntos documentais do órgão, produzidos em razão de suas funções e atividades, em dez classes e, sucessivamente, em dez subclasses, dez grupos e dez subgrupos, usando-se a notação decimal.

Na estruturação da hierarquia das classes, subclasses, grupos e subgrupos, a maioria das atividades desempenhadas utiliza expressões que denotam “ações” ou “conteúdos de ações”, as quais refletem o conjunto de competências e atribuições necessárias para o funcionamento e atuação da instituição, identificados por meio do levantamento documental.

Assim, o código reúne os documentos de acordo com o ato realizado e destes conjuntos documentais com a atividade que os gerou, terminando por juntá-los sob a respectiva função, na qual implica a competência do órgão produtor.

À época da elaboração do Código e sua respectiva Tabela de Temporalidade de Documentos - TTDD, houve consulta ao *Thesaurus* Brasileiro da Educação - Brased, do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP (INEP, 2020), além do estudo de atos normativos específicos, dos procedimentos utilizados em cada Instituição, assim como também a pesquisa na legislação de ensino superior (ARQUIVO NACIONAL, 2011).

Ao longo das discussões realizadas em várias reuniões do Grupo de Trabalho foram levantados alguns problemas de ordem teórico-prático relevantes: o desconhecimento dos critérios adotados na metodologia que foi utilizada para a divisão dos níveis hierárquicos (onde o ensino é dividido, por exemplo); a característica inflexível para a criação de novas classes no nível hierárquico; vale ressaltar que um grupo pequeno de representantes das IFES participou no processo de desenvolvimento do código, desde 2011 até a presente data não foram realizadas chamadas públicas à comunidade arquivística para a sua revisão e atualização, e o uso de assuntos como forma de estabelecer classes se mostra ineficiente e muitas vezes ambíguo na hora de realizar a classificação e consequente avaliação, pois não são o fiel reflexo das informações institucionais. Desta forma, sugere-se a classificação funcional e por tipologia, o que garante mais precisão e consistência no Plano de Classificação.

Aliado aos problemas citados, podemos acrescentar que o seu correspondente da Atividade-Meio, o Código de Classificação e Tabela de Temporalidade e Destinação de documentos relativos às Atividades-Meio do Poder Executivo Federal foi atualizado e revisto em 2020, e publicado por meio da Portaria AN

nº 47, após 19 anos de sua criação. Ficamos nos questionando se estas atualizações estão acompanhando as novas demandas do Governo Federal, da Advocacia Geral da União - AGU, do Tribunal de Contas da União – TCU, no que se referem à gestão de riscos, Lei Geral de Proteção de Dados Pessoais, Governo Digital, etc.

4.1 Propostas para atualização de código de classificação de documentos e Tabela de Temporalidade e Destinação de documentos, relativos às atividades-fim

As análises apresentadas foram realizadas com base nos dados coletados a partir da aplicação de questionários aplicados aos representantes das IFES no período de 2018 – 2019. Houve a apresentação de itens que sugeriam a modificação de códigos, termos (descritores), explicações, e apresentações de propostas de mudanças, relacionadas com as funções, sub-funções, atividades, negociações, atos, explicações e justificativa de mudanças de descritores. No GT IFES classificação, do qual fazemos parte, houve divisão das classes entre os membros do grupo, nos cabendo analisar os dados da classe 100 do código. Então a partir desses dados fizemos as propostas que virão a seguir.

Ao todo, foram sugeridas 4 alterações de descritor, 9 alterações de notas explicativas, 2 alterações de descritores e notas explicativas, 7 inclusões de descritores e 2 exclusões de descritores, por considerá-los incompletos, onde não refletem os assuntos, nem documentos produzidos na atualidade.

Uma das propostas do grupo para alteração do descritor que se refere à denominação da subclasse 120 - Cursos de graduação, inclusive à distância modificado para 120 - Promoção do Ensino Superior de Graduação. Esta sugestão pela alteração de denominação a deixa mais genérica, pois as funções e atividades abrangidas por esta contemplam não apenas os “Cursos de Graduação”, mas várias outras unidades e subunidades (Pró-Reitorias, Departamentos Didáticos, Direções de Centro). Nesse sentido, a palavra promoção não altera ou acrescenta às atividades administrativas relacionadas com o ensino de graduação, dá maior entendimento à atividade geral.

Com o mesmo objetivo, alteração do descritor, o grupo apresentou a sua proposta para o código 125.6 - Monitorias | Estágios não obrigatórios | Programas de iniciação à docência. Este descritor reúne os documentos referentes às solicitações de bolsas individuais de monitorias, nas instituições de apoio e fomento ao ensino e pesquisa, aos estágios não obrigatórios e aos Programas de iniciação à docência. Assim, o grupo sugeriu que o termo do código 125.6 - Atividades de complementação e qualificação da formação acadêmica, esta modificação na nomenclatura do descritor irá englobar as atividades de

ensino assim como também as que complementam e qualificam a formação. Ou seja, colocar os documentos de monitorias, estágios não obrigatórios, estágio obrigatório e programas de iniciação à docência juntos. É importante salientar que, no caso dos estágios obrigatórios, são considerados disciplinas, porém se utilizam de documentos e procedimentos similares com os não obrigatórios em muitas instituições.

Uma das propostas que altera o descritor e a nota explicativa refere-se ao descritor 125.11 - Processo de seleção (vestibular) onde nas subdivisões de, este descritor classificam-se documentos referentes a cada etapa do processo de seleção de ingresso ao ensino de graduação. A proposta do grupo foi: 125.11 - Processo de seleção onde nas subdivisões deste descritor classificam-se documentos referentes a cada etapa e tipos do processo de seleção (vestibular, SISu, ENEM, EAD, vagas complementares do SISU, processos especiais, ocupação de vagas ociosas, vagas complementares, cotas, comunidades quilombolas, comunidades indígenas).

O descritor 125.21 – Matrícula | Registro teve proposta de acréscimo de documentos. Originalmente este descritor agrupa documentos referentes ao processo de cadastramento inicial do aluno para estabelecer a sua vinculação à instituição e ao seu respectivo curso e a proposta complementa a nota. Nesse sentido, a proposta foi complementar com a frase: [...] documentos solicitados para matrícula, autodeclaração de raça, aferição de renda. O mesmo acontece com o descritor 125.421 – Expedição que tratam de documentos referentes à solicitação de expedição do diploma acrescentaria a frase “assim como a expedição de 2º via de diploma”.

No que se refere à inclusão de descritores, podemos citar aos Programas de ensino e projetos de ensino, junto com os desdobramentos de cada uma: Proposição e Avaliação | Resultados. Esta sugestão deve-se ao conjunto de ações e documentos oriundos de programas de ensino, que são criados e aprovados nas instituições como os programas de pesquisa e extensão.

4.2 Percepções sobre as propostas para atualização dos instrumentos de Gestão de Documentos, relativos às Atividades-Fim

Devemos considerar também que se trata ainda da primeira versão do referido Código de classificação, publicada há mais de dez anos, onde ainda não foi realizada nenhuma atualização. Também é importante mencionar que a dinâmica e os processos das instituições, assim como o conhecimento muda e evolui, fazendo com que os instrumentos de gestão de documentos precisem ser revisados e atualizados periodicamente, como é o caso dos citados nesse estudo.

Ao longo das discussões realizadas em várias reuniões do Grupo de Trabalho foram levantados alguns problemas de ordem teórico-prático relevantes: o desconhecimento dos critérios adotados na metodologia que foi utilizada para a divisão dos níveis hierárquicos (onde o ensino é dividido, por exemplo); a característica inflexível para a criação de novas classes no nível hierárquico; se bem um grupo pequeno de representantes das IFES participou no processo de desenvolvimento do código, desde 2011 até a data não foram realizadas chamadas públicas à comunidade arquivística para a sua revisão e atualização, e; o uso de assuntos como forma de estabelecer classes se mostra ineficiente e muitas vezes ambíguo na hora de realizar a classificação e consequente avaliação, pois não são o fiel reflexo das informações institucionais, sugere-se a classificação funcional e por tipologia, o que garante mais precisão e consistência no Plano de Classificação.

5 Conclusão

Portanto, este trabalho mostra os resultados da revisão e atualização de dito instrumento Código de Classificação das IFES, que consta na Portaria 92, de 2011, em particular a classe 100 e itens em suas subclasses, por julgarmos que há necessidade de ser incluída a análise dos termos, conceitos e análise dos domínios de conhecimento na produção e trâmite de documentos das universidades e institutos de educação superior no país. Apresentamos os resultados e propostas das análises realizadas no Código de Classificação referente às Atividades-Fim das IFES no Brasil, em particular os itens das subclasses na classe 100; refletimos sobre classificação arquivística e Teoria da Classificação; identificamos o contexto institucional e o domínio de conhecimento em instituições de ensino superior. E procuramos apresentar propostas de revisão e atualização do Código de Classificação de Atividades-Fim, em particular os itens das subclasses relacionados à classe 100, conforme o que foi apresentado neste artigo de forma mais detalhada anteriormente na seção quatro.

Vimos que a classificação e a representação da informação não são apenas atividades técnicas, mas ações mentais e físicas dentro de um contexto cultural e social, e se manifestam de forma simbólica, semântica e ideológica (de valores, tendências linguísticas, políticas e culturais). Constatamos neste estudo a importância da interdisciplinaridade e do diálogo entre a classificação e representação da informação na Biblioteconomia, Arquivologia, Diplomática, Tipologia Documental e a Ciência da Informação.

Para isso, deve-se pensar a gestão de documentos e a recuperação da informação de acordo com o contexto de produção dos documentos e a organização que os

originou, onde a análise de domínio colaborará na preservação, sobretudo, do princípio de proveniência.

É destacado que é incipiente o número de instituições arquivísticas brasileiras que utilizam indexação, representação temática e vocabulários controlados para o desenvolvimento de estudos sobre classificação. Revela-se que a reflexão e investigação desses processos técnicos se fazem mais que necessárias.

É imprescindível inserir como parte da agenda de pesquisas dos arquivistas a temática da Organização do Conhecimento, incluindo estudos sobre Teoria do conceito, análise de domínio, teoria da classificação e representação da informação, classificação arquivística. Desta forma, tentamos contribuir para a fundamentação científica na elaboração de taxonomias (hierarquia de termos), códigos e planos de classificação arquivísticos, cuja metodologia utilizada seja apresentada de forma clara, contribuindo para que futuras adaptações ou alterações no (código) plano de classificação de atividades-fim das IFES possam ser efetuadas de forma planejada, com menos inconsistências e incertezas.

Referências

- ARQUIVO NACIONAL. Portaria n 47, de 14 de fevereiro de 2020. Dispõe sobre o Código de Classificação e Tabela de Temporalidade e Destinação de Documentos relativos às Atividades-Meio do Poder Executivo Federal. **Diário Oficial da União**. Publicado em: 20/02/2020 | Edição:36| Seção: 1| Página: 74. Disponível em <https://www.in.gov.br/web/dou/-/portaria-n-47-de-14-de-fevereiro-de-2020-244298005>. Acesso em 07 fev. 2022.
- BRASIL. **Lei nº 8.159, de 8 de janeiro de 1991**. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8159.htm. Acesso em 07 fev. 2023
- CARLAN, Eliana; MEDEIROS, Marisa Bracher Basílio. Sistemas de Organização do conhecimento na visão da Ciência da Informação. **RICI: Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 4, n. 2, p. 53-73, 2012. DOI: 10.26512/rici.v4.n2.2011.1675. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/1675>. Acesso em: 28 maio. 2023.
- CONSELHO NACIONAL DE ARQUIVOS. Portaria n 92, 23.09.2011. **Código de classificação de Atividades-Fim e Tabela de Temporalidade de Documentos das Instituições Federais de Ensino Superior**. Brasília, Arquivo Nacional, 2011. Disponível em <https://www.gov.br/conarq/pt-br/legislacao-arquivistica/portarias-federais/portaria-no-92-de-23-de-setembro-de-2011>. Acesso em 07 fev. 2022.
- DANUELLO, J. C. **Produção científica docente em tratamento temático da informação no Brasil: uma abordagem métrica como subsídio para a análise do domínio**. 2007. 122f. Dissertação (Mestrado) – UNESP, Marília, 2007.

- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, n.33, v.1, p.11-19. 2006.
- DAHLBERG, I. Teoria do Conceito. *In: Ciência da Informação*. Rio de Janeiro, 7(2): 101-107, 1978.
- DUCHEIN, M. O respeito aos fundos em arquivística: princípios teóricos e problemas práticos. **Arquivo & Administração**, v. 10-14, n. 2, 1986. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/49818>. Acesso em: 26 mar. 2023.
- GUIMARÃES, J. A. C. Análise de domínio como perspectiva metodológica em organização da informação. *In: Ciência da Informação*, [S. l.], v. 43, n. 1, 2015. Disponível em: <https://revista.ibict.br/ciinf/article/view/1415>. Acesso em: 25 mar. 2023.
- HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, 35 (3/2), 86-101. Sociedade Brasileira para Organização do Conhecimento. 2008 Disponível em: <https://isko.org.br/grupos-de-pesquisa/Em ISKO BRASIL>. Acesso em: 24 jun. 2021.
- HJØRLAND, B. Domain analysis in information science: eleven approaches – traditional as well as innovative. **Journal of Documentation**, London, v. 58, n. 4, p. 422-462, 2002. Disponível em: https://www.researchgate.net/publication/249366184_Domain_analysis_in_information_science_Eleven_approaches_-_Traditional_as_well_as_innovative/link/00b49539471ae89904000000/download Acesso em 25 mar. 2023
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA - INEP. **Thesaurus brasileiro de educação-BRASED**. [2020]. Disponível em: <<http://www.inep.gov.br/pesquisa/thesaurus>> Acesso em: 24 maio 2023 Acesso em: 15 mar. 2005.
- JOHNSON, K.; STEVENS, S. (Ed.). **Exploring interactions of people, places and information**. Fredericton: University of New Brunswick, 2011. 6p. (Proceedings of Annual CAIS/ACSI Conference, 39., 2011, Fredericton, Canada.)
- NASCIMENTO, D. M.; MARTELETO, R. M. A “informação construída” nos meandros dos conceitos da Teoria Social de Pierre Bordieu. **DataGramZero**, v. 5, n. 5, out./2004. Disponível em: <https://brapci.inf.br/index.php/res/v/5679>. Acesso em: 22 mar. 2023.
- NEGREIROS, L. R. Uma compreensão dos fundos arquivísticos sob o enfoque da análise de domínio. *In: Informação & Informação*, [S. l.], v. 13, n. 2, p. 79–95, 2008. DOI: 10.5433/1981-8920.2008v13n2p79. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/1818>. Acesso em: 26 mar. 2023.
- POMBO, O. Da classificação dos seres à classificação dos saberes. Leituras. **Revista da Biblioteca Nacional de Lisboa**, Lisboa, n. 2, p. 19-33, 1998.
- ROUSSEAU, J.; COUTURE, C. **Os fundamentos da disciplina Arquivística**. Lisboa: Publicações Dom Quixote, 1998.
- SCHELLENBERG, T. R. Princípios de Classificação. *In: SCHELLENBERG, T. R. Arquivos modernos: princípios e técnicas*. Rio de Janeiro: FGV, 2002.
- SMIRAGLIA, R. P. (2013). Domain Coherence Within Knowledge Organization: People, Interacting Theoretically, Across Geopolitical and Cultural Boundaries. **Proceedings of the Annual Conference of CAIS / Actes Du congrès Annuel De l'ACSI**. <https://doi.org/10.29173/cais601>
- SMIT, J. W. Vocabulário controlado e controle de vocabulário em arquivos. **Acervo - Revista do Arquivo Nacional**, v. 31, n. 3, p. 46-56, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/107070>. Acesso em: 18 abr. 2023.
- SOUSA, R. T. B. **As bases do processo classificatório em Arquivística: um debate metodológico**. São Paulo: Associação dos Arquivistas de São Paulo, 2002.
- SOUSA, R. T. B. A classificação funcional de documentos de arquivo é uma abstração intelectual ou um instrumento prático? **Acervo**, Dossiê Organização do Conhecimento. Rio de Janeiro, v. 35, n. 2, p. 1-21, maio/ago. 2022
- TOGNOLI, N. B.; SILVA, A. M. S.; SILVA, A. P. Organização do Conhecimento e Arquivologia: uma análise de domínio nos periódicos KnowledgeOrganization e Scire. **Informação & Informação**, Londrina, v. 24, n. 3, p. 52 – 77, set./dez. 2019. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/38726>. Acesso em: 02 fev. 2022.

A representação da informação arquivística nos arquivos estaduais brasileiros

The representation of archival information in brazilian state archives

Eliezer Pires da Silva (1), Priscila Ribeiro Gomes (2)

(1) Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Avenida Pasteur - até 213/214 Botafogo, 22290240 - Rio de Janeiro, RJ – Brasil, eliezepires@gmail.com, (2) pri.unirio@yahoo.com.br

Resumo

Os arquivos públicos são fontes muito importantes para a pesquisa histórica e para a construção de memória e identidade dos diferentes grupos sociais. Essas instituições preservam documentos e informações que são fundamentais para a compreensão da história, da cultura e da sociedade em geral. Os arquivos públicos também têm a função de divulgar e difundir informações, contribuindo para a democratização do acesso ao conhecimento. O **objetivo** foi investigar a disponibilização pelos arquivos estaduais de instrumentos de pesquisa e outros meios de representação da informação arquivística que permitem aos usuários terem acesso a informações sobre o patrimônio cultural e histórico do país. A metodologia da pesquisa tem uma proposta qualitativa com abordagem exploratória. A abordagem metodológica reconhece o cenário da representação da informação arquivística nos arquivos estaduais brasileiros combinando uma articulação conceitual entre a Organização do Conhecimento e a Arquivologia com um levantamento empírico dos sites dos arquivos estaduais e suas ferramentas de recuperação da informação. Ao combinar essas duas perspectivas foi possível obter uma compreensão mais abrangente sobre como os arquivos públicos estão organizando e representando suas informações. Os **resultados** apontam para a necessidade de avanço nos instrumentos de pesquisa disponíveis pelas instituições que anunciam ter essa finalidade para os acervos que organizam e preservam. A **conclusão** é que atualmente contamos com 24 arquivos públicos estaduais, 19 possuem seus sites, 16 disponibilizam algum instrumento de pesquisa e 12 oferecem bases de dados para pesquisa online.

Palavras-chave: Representação da informação em arquivos. Arquivo público. Instrumento de pesquisa em arquivos.

Abstract

Public archives are very important sources for historical research and for building the memory and identity of different social groups. These institutions preserve documents and information that are fundamental for understanding history, culture and society in general. Public archives also have the function of disseminating and disseminating information, contributing to the democratization of access to knowledge. The **objective** was to investigate the availability of research instruments by state archives and other means of representing archival information that allow users to have access to information about the cultural and historical heritage of the country. The **methodology** adopted to recognize the scenario of representation of archival information in Brazilian state archives combined a conceptual articulation between the Organization of Knowledge and Archival Science with an empirical survey of state archives websites and their information retrieval tools. By combining these two perspectives, it was possible to gain a more comprehensive understanding of how public archives are organizing and representing their information. The **results** point to the need to advance the research tools available by the institutions that announce that they have this purpose for the collections they organize and preserve. The **conclusion** is that we currently have 24 state public archives, 19 have their websites, 16 provide some research tool and 12 offer databases for online research.

Keywords: Representation of information in archival. Public archive. Archival research instrument.

Resumen

Los archivos públicos son fuentes muy importantes para la investigación histórica y para la construcción de la memoria y la identidad de los diferentes grupos sociales. Estas instituciones conservan documentos e información que son fundamentales para comprender la historia, la cultura y la sociedad en general. Los archivos públicos también tienen la función de difundir y difundir la información, contribuyendo a la democratización del acceso al conocimiento. El **objetivo** fue investigar la disponibilidad de instrumentos de investigación por parte de los archivos estatales y otros medios de representación de información arquivística que permitan a los usuarios tener acceso a información sobre el patrimonio cultural e histórico del país. La **metodología** adoptada para reconocer el escenario de representación de la información arquivística en los archivos estatales brasileños combinó una articulación conceptual entre la Organización del Conocimiento y la Arquivística con un levantamiento empírico de los sitios web de los archivos estatales y sus herramientas de recuperación de información. Al combinar estas dos perspectivas, fue posible obtener una comprensión más completa de cómo los archivos públicos organizan y representan su información. Los **resultados** apuntan para la necesidad de

avanzar en las herramientas de investigación disponibles por parte de las instituciones que anuncian tener este propósito para las colecciones que organizan y preservan. La **conclusión** es que actualmente contamos con 24 archivos públicos estatales, 19 tienen sus sitios web, 16 brindan alguna herramienta de investigación y 12 ofrecen bases de datos para la investigación en línea.

Palabras clave: Representación de la información en archivo. Archivo público. Herramienta de investigación de archivo.

1 Introdução

Entendemos a Organização do Conhecimento como campo em que as informações e os conhecimentos são estruturados, classificados e organizados para fins de recuperação e acesso eficiente. Isso inclui contextos como bibliotecas, arquivos, museus, por exemplo. Nesse horizonte, bibliotecários, arquivistas e cientistas da informação, utilizam técnicas e ferramentas específicas para representar informações, a fim de torná-las mais acessíveis e úteis para seus usuários. Utilizam-se vocabulários controlados, além de padrões e normas para a descrição e recuperação da informação, conforme fica demonstrado em Silva e Orrico (2013); Orrico e Silva (2019); Silva e Pereira (2019); Freitas e Silva (2023).

Nesta pesquisa estamos abordando uma dimensão aplicada da Organização do Conhecimento com base no cenário dos instrumentos de recuperação da informação nos arquivos estaduais brasileiros. Esses arquivos são importantes fontes para a pesquisa histórica e a construção de memória e identidade dos diferentes grupos sociais, servem para a proteção de direitos, para além de divulgar e difundir informações para a sociedade em geral.

Nosso objetivo é reconhecer o cenário da representação da informação arquivística no marco dos arquivos estaduais brasileiros, considerando seus instrumentos de pesquisa disponibilizados na internet. A perspectiva metodológica adotada envolveu tanto uma articulação conceitual sobre interseccionalidades entre a Organização do Conhecimento e a Arquivologia, quanto o levantamento empírico dos sites dos arquivos estaduais com suas ferramentas de recuperação da informação.

Os arquivos públicos estaduais no Brasil possuem múltiplas funções envolvendo gestão de documentos, preservação de memória coletiva, acesso ao patrimônio documental sob sua guarda, apoio aos estudos retrospectivos e o potencial pedagógico do arquivo, incluindo metodologias de educação patrimonial enquanto meio para promover a conscientização da sociedade sobre a importância da preservação do patrimônio, conforme destaca Rodrigues e Gomes (2021), além da promoção da transparência e proteção de direitos. Essas instituições arquivísticas interessam a uma ampla gama de usuários, incluindo pesquisadores, historiadores, estudiosos, profissionais de diversas áreas e cidadãos em geral, na medida em que é a efetividade de um direito fundamental nas sociedades democráticas.

Além desta introdução esta comunicação conta com seguinte estrutura: metodologia, discussão sobre a representação da informação nos arquivos, o cenário dos instrumentos de pesquisa nos arquivos públicos estaduais no Brasil, considerações conclusivas e referências citadas.

2 Metodologia

A pesquisa tem uma proposta qualitativa com abordagem exploratória. Para atingir seus objetivos faz uso de dois eixos metodológicos: da técnica de pesquisa bibliográfica e análise dos sites dos arquivos estaduais.

O levantamento bibliográfico foi feito em bases de dados de referência da produção científica em Arquivologia e Organização do Conhecimento no Brasil, são elas: base da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) e Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (Brapci).

O levantamento empírico sobre os arquivos estaduais foi realizado por intermédio de buscas pelo Google no mês de abril de 2023, com os termos arquivo estadual e o nome de cada uma das 27 unidades federativas do país. Não foi possível identificar a existência de arquivos estaduais em três unidades da federação: Rondônia, Roraima e Tocantins. Outros quatro estados, apesar da possibilidade de identificar a existência da instituição, não foram encontrados seus sites: Acre, Amapá, Paraíba, Rio Grande do Norte e Sergipe.

Adotamos a perspectiva de Corujo e Freitas (2021) em que a organização do conhecimento tem uma relação direta com os arquivos, uma vez que a organização, a classificação e a descrição de conteúdos e contextos estão no centro do processamento técnico de acervos arquivísticos. De acordo com Barros e Sousa (2020) a organização do conhecimento auxilia na representação da informação arquivística por meio de sistemas que permitam a recuperação de informações contidas nos documentos de arquivo, possibilitando que sejam compreendidas e interpretadas em seu contexto histórico e social para a construção de compreensões mais amplas e significativas sobre a sociedade.

3 Representação da informação em arquivos

O processamento técnico arquivístico é um conjunto de atividades que visam organizar e descrever um acervo

arquivístico para facilitar seu acesso e uso pelos usuários. Essas atividades incluem a identificação, classificação, descrição e indexação dos documentos de arquivo. Isso é feito por meio da criação de instrumentos de pesquisa, tais como catálogos, inventários e guias, que efetivam a representação da informação nos arquivos, conforme afirmam Linden e Vianna (2019).

De acordo com Santos e Silva (2016) a construção dos instrumentos de representação da informação arquivística são sustentadas por diversos elementos teórico-conceituais. Os arquivos são considerados fontes de memória coletiva, pois guardam informações sobre a história e a cultura de uma sociedade. Dessa forma, o acesso aos arquivos é fundamental para a construção da identidade e da memória social. Para Silva e Pereira (2019) os arquivos são também ferramentas essenciais para a promoção da transparência e accountability das instituições públicas e privadas, permitindo o acesso às informações e documentos produzidos e recebidos por essas instituições.

Yakel (2003) discute as várias formas de representação arquivística e a importância dessas representações na mediação do acesso aos documentos de arquivo. Essa autora aborda a complexidade da representação arquivística e como essa complexidade está relacionada com a diversidade de usuários e usos dos documentos de arquivo.

Para Feitas e Silva (2023) a representação da informação em arquivos apoia-se na descrição dos documentos em termos de seus conteúdos e contextos e está relacionada aos instrumentos de recuperação da informação que precisam ser claros, precisos e consistentes, bem como incluir o uso de padrões e normas para assegurar a interoperabilidade e a compatibilidade com outros sistemas de informação. Para Vital e Medeiros (2019) é importante considerar as diferentes necessidades e expectativas dos usuários dos arquivos e adaptar as estratégias de representação da informação de acordo com essas demandas.

A promoção do acesso aos acervos arquivísticos pelos usuários é um desafio importante para os arquivos brasileiros, e várias estratégias devem ser adotadas para alcançar esse objetivo. Nesse sentido, os arquivos podem promover a divulgação de seus acervos por meio de eventos, exposições, palestras, publicações e outras iniciativas que permitam que o público conheça os documentos disponíveis. Há também alguns arquivos brasileiros que oferecem o acesso remoto aos seus acervos por meio de plataformas online, permitindo que usuários acessem os documentos digitais e digitalizados de qualquer lugar do mundo.

Deve-se observar que essas estratégias têm seus limites, uma vez que nem todos os arquivos têm recursos para implementá-las. Portanto, a promoção do

acesso aos acervos arquivísticos pelos usuários deve ser planejada, levando em consideração a realidade institucional de cada arquivo e de seus acervos.

Os instrumentos de pesquisa disponibilizados pelas instituições arquivísticas estaduais podem variar de acordo com o seu contexto. As principais ferramentas de pesquisa utilizadas são:

- Inventários: instrumento que contém informações detalhadas sobre o conteúdo do acervo arquivístico, organizado geralmente em ordem alfabética, cronológica ou temática.
- Catálogos: instrumento que reúne informações sobre os documentos existentes no acervo de forma organizada em diferentes categorias, como data, local, tipo de documento.
- Guias: instrumento que fornece informações gerais sobre o acervo, bem como orientações para os usuários que pretendem pesquisar documentos específicos.
- Bases de dados: sistemas de informações sobre acervos, permitindo que os usuários realizem pesquisas por palavras-chave, datas e tipos de documentos.

É importante ressaltar que a disponibilidade desses instrumentos de pesquisa pode variar de acordo com o estado de organização dos documentos e desempenham um papel fundamental na mediação entre os usuários e os acervos, conforme Fujita e Rodriguez (2018).

Para Lehmkuhl e Silva (2017) há uma tendência recente no campo arquivístico, que questiona a visão tradicional de que a descrição e organização dos arquivos são tarefas meramente técnicas e objetivas. Ao contrário, a representação da informação arquivística se insere em processos sociais e históricos que envolvem múltiplos atores e interesses. A adoção de tecnologias digitais e a possibilidade de colaboração em rede permitem ampliar as formas de participação e construção coletiva da representação da informação nos arquivos, incluindo a participação de diferentes grupos sociais e a criação de comunidades de prática. Essa perspectiva amplia o debate sobre a representação da informação arquivística, que deixa de ser vista como uma atividade neutra e técnica, e passa a ser compreendida como um processo social, histórico e político.

Os instrumentos de pesquisa que indexam os documentos de arquivo podem tanto colocar em evidência quanto fazer desaparecer documentos diante dos usuários. Para Llanes-Padrón e Moro-Cabero (2017) a representação da informação em um instrumento de pesquisa pode ser influenciada por diferentes fatores, como a visão de mundo dos arquivistas, a relevância percebida pelos usuários e a disponibilidade de recursos para o trabalho de indexação.

As instituições arquivísticas têm a função social de garantir o acesso aos documentos de sua custódia por meio de instrumentos que representam o conteúdo desses registros de forma contextualizada. Os instrumentos de pesquisa concretizam a finalidade dos arquivos de promover o acesso à informação arquivística, garantindo que os documentos sejam organizados e representados para a divulgação de seus conteúdos.

Atualmente assistimos aos debates sobre o Records in Contexts - Conceptual Model, ou RIC-CM. Trata-se de um modelo conceitual de descrição arquivística adequado ao século XXI. Esse documento expressa anos de trabalho do pelo EGAD, Expert Group on Archival Description, vinculado ao ICA (Conselho Internacional de Arquivos). Para Orrico, Silva e Pereira (2019, p.9),

A abordagem multinível de descrição dos acervos arquivísticos convive com a emergência de um modelo multidimensional como expressão de superação de uma entidade única e completa a ser descrita pelo arquivista. Desloca-se de uma hierarquia para uma lógica de rede ou facetas na descrição arquivística. Desse modo, permite-se que seja possível descrever os documentos e as complexas realidades contextuais em que se inserem.

Com o RIC-CM tenta-se trazer uma abordagem menos hierarquizada e mais multidimensional. O EGAD embora tenha mudado de nomes, durante o passar do tempo ele já conta com mais de três décadas de estudos sobre a descrição arquivística que resultou na criação dos 4 padrões da descrição arquivística:

- ISAD/G: General International Standard Archival Description, para acervos;
- ISAAR/CPF: International Standard Archival Authority Record for Corporate Bodies, Persons and Families, para produtores de arquivos.
- ISDF: International Standard for Describing Functions, para as funções dos produtores de arquivos;
- ISDIAH: International Standard for Describing Institutions with Archival Holdings, para as instituições custodiadoras de acervos arquivísticos.

A articulação conceitual sobre interseccionalidades entre a Organização do Conhecimento e a Arquivologia permite entender que ambas as áreas possuem preocupações semelhantes quanto à organização, representação e acesso à informação. A Organização do Conhecimento é responsável por estudar e desenvolver técnicas para a organização e representação da informação em diferentes contextos, enquanto a Arquivologia se dedica à representação da informação a partir de documentos arquivísticos.

Ao relacionar essas áreas, é possível entender a importância da organização e representação adequada dos documentos arquivísticos para garantir o acesso e a preservação de informações relevantes para a sociedade. A articulação entre essas áreas permite o desenvolvimento de técnicas e ferramentas para a representação da informação arquivística, facilitando a busca e recuperação de informações em acervos arquivísticos.

4 Cenário dos arquivos estaduais brasileiros

Os arquivos estaduais no Brasil são instituições responsáveis pela gestão, preservação e disponibilização de documentos de valor histórico, cultural e administrativo produzidos pelos órgãos públicos estaduais. Além de serem a autoridade sobre a gestão de documentos nessa esfera de governo, esses arquivos têm como missão garantir o acesso aos documentos públicos e a promoção do conhecimento da história e da cultura de cada estado brasileiro, conforme Fonseca (1998).

Os arquivos estaduais apresentam uma institucionalidade diversa no contexto brasileiro. Alguns são subordinados às secretarias estaduais de cultura, de educação, de gestão. As suas atribuições envolvem participar da gestão de documentos no âmbito deste estado em questão, contemplando ainda a organização, descrição, conservação e difusão dos seus acervos, bem como a realização de exposições, palestras e eventos educativos, e o desenvolvimento de pesquisas e estudos.

Para Jardim e Fonseca (2004) os usuários dos arquivos públicos estaduais no Brasil podem incluir pesquisadores acadêmicos, historiadores, genealogistas, jornalistas, estudantes, advogados, entre outros profissionais. Também podem ser membros da sociedade em geral que buscam informações sobre sua história pessoal, familiar ou comunitária, bem como pessoas que procuram informações para fins administrativos, como a obtenção de certidões ou comprovantes.

Deveriam existir 27 arquivos públicos estaduais no Brasil, considerando os 26 estados mais o Distrito Federal, um para cada unidade federativa do país. Cada uma dessas instituições precisa ser responsável pela gestão e preservação dos documentos de valor histórico, cultural e administrativo produzidos pelos órgãos públicos estaduais, além de promover o acesso à informação e à pesquisa do patrimônio documental recolhido.

Identificamos informações sobre a existência de 24 arquivos estaduais. Dos 19 casos em que a instituição possui site, 12 disponibilizam uma base de dados para recuperação da informação no seu acervo. 16 instituições disponibilizam ao menos algum tipo

tradicional de instrumento de pesquisa (inventário, catálogo, guia).

É interessante observar que a disponibilização de instrumentos de pesquisa tradicionais é importante para facilitar o acesso à informação contida nos acervos dos arquivos estaduais. Porém, a disponibilização de uma base de dados online pode facilitar ainda mais o acesso à informação, uma vez que permite que os usuários realizem pesquisas precisas em um grande volume de documentos a partir de palavras chaves.

Quando uma instituição arquivística disponibiliza seus tradicionais instrumentos de pesquisa em seu site, isso significa que ela está disponibilizando informações sobre os documentos que fazem parte de seu acervo. Esses instrumentos de pesquisa permitem que pesquisadores e outros interessados possam encontrar e usar informações valiosas para suas atividades de pesquisa e educação.

Na análise das bases de dados disponibilizadas pelos arquivos estaduais brasileiros identificamos 6 instituições que utilizam o ICA-AtoM (Access to Memory), que é um sistema baseado em software livre que foi desenvolvido pelo Conselho Internacional de Arquivos (ICA). Ele é projetado especificamente para ajudar os arquivos públicos a gerenciar e fornecer acesso a seus acervos. Ele oferece recursos padronizados de descrição arquivística, pesquisa online, controle de acesso, integração com outros sistemas e uma comunidade ativa de usuários.

Há necessidade de avanço nos instrumentos de pesquisa disponíveis pelas instituições que anunciam ter essa finalidade para os acervos que organizam e preservam. Abaixo apresentamos o quadro que apresenta esse cenário dos arquivos públicos estaduais brasileiros.

Quadro 1: *Instrumentos de pesquisa online dos arquivos estaduais brasileiros.*

UF	Instituição	Site	Base online	Listas
AC	Divisão de Arquivo Público do Estado do Acre	Não	Não	Não
AL	Arquivo Público de Alagoas	Sim	Não	Sim
AP	Arquivo Público do Estado do Amapá	Não	Não	Não
AM	Arquivo Público do Estado do Amazonas	Sim	Não	Não
BA	Arquivo Público do Estado da Bahia da Fundação Pedro Calmon	Sim	Sim	Sim

CE	Arquivo Público do Estado do Ceará	Sim	Não	Sim
DF	Arquivo Público do Distrito Federal	Sim	Sim	Sim
ES	Arquivo Público do Estado do Espírito Santo	Sim	Sim	Sim
GO	Arquivo Histórico Estadual de Goiás	Sim	Não	Não
MA	Arquivo Público do Estado do Maranhão	Sim	Sim	Sim
MT	Arquivo Público de Mato Grosso	Sim	Sim	Sim
MS	Arquivo Público Estadual do Mato Grosso do Sul	Sim	Não	Não
MG	Arquivo Público Mineiro	Sim	Sim	Sim
PA	Arquivo Público do Estado do Pará	Sim	Não	Não
PB	Arquivo Público do Estado da Paraíba	Não	Não	Não
PR	Arquivo Público do Estado do Paraná	Sim	Sim	Sim
PE	Arquivo Público de Pernambuco	Sim	Sim	Sim
PI	Arquivo Público do Estado do Piauí	Sim	Não	Sim
RJ	Arquivo Público do Estado do Rio de Janeiro	Sim	Sim	Sim
RN	Arquivo Público Estadual do Rio Grande do Norte	Não	Não	Sim
RS	Arquivo Público do Estado do Rio Grande do Sul	Sim	Sim	Sim
RO	Arquivo do Estado de Rondônia - Não identificado	Não	Não	Não
RR	Arquivo do Estado de Roraima- Não identificado	Não	Não	Não
SC	Arquivo Público do Estado de Santa Catarina	Sim	Sim	Sim
SP	Arquivo Público do Estado de São Paulo	Sim	Sim	Sim
SE	Arquivo Público de Sergipe	Não	Não	Não
TO	Arquivo do Estado de Tocantins- Não identificado	Não	Não	Não

Fonte: Elaboração dos autores.

Nos arquivos públicos estaduais brasileiros as ações e discursos de divulgação dos acervos apresentam algumas características marcantes, tais como: relevância histórica dos documentos que compõem seus acervos; emergente abordagem pedagógica em suas ações, buscando transmitir aos usuários informações sobre os arquivos de forma acessível e didática; crescente utilização de tecnologias digitais para disponibilizar seus acervos de forma online; aumento da relação com a comunidade em que estão inseridas, realizando ações de divulgação e de educação patrimonial, incluindo visitas guiadas, palestras e eventos culturais. Tais ações visam, segundo Pinto e Gomes (2020), promover uma aproximação entre o cidadão e o arquivo, uma vez que possibilitam o acesso e compreensão da importância do acervo, despertando um senso de pertencimento e identidade aos envolvidos.

Ao examinar os instrumentos de pesquisa e outros meios de identificação da descrição arquivística nos arquivos públicos, podemos identificar diferentes abordagens para a realização dessa função arquivística.

Por exemplo, alguns arquivos públicos estaduais podem optar por publicar inventários, catálogos e guias impressos ou digitais, que fornecem informações descritivas sobre seus acervos, como nomes de autores, títulos de documentos, datas e outras informações relevantes. Esses instrumentos de pesquisa geralmente são organizados de forma sistemática e podem ser usados por pesquisadores e público em geral para localizar documentos específicos em um acervo.

Outros arquivos públicos adotam uma abordagem mais avançada e desenvolvem sistemas de busca para recuperação dos documentos custodiados pelas instituições. Esses sistemas geralmente são baseados em tecnologias digitais e permitem que os usuários pesquisem e acessem informações em uma interface de usuário intuitiva. Eles também podem fornecer recursos avançados, como controle de acesso, pesquisa por palavra-chave e resultados de pesquisa personalizados.

A análise dos instrumentos de recuperação da informação e demais meios de representação da informação arquivística nos arquivos públicos estaduais nos permite reconhecer diferentes abordagens desde a publicação de instrumentos de pesquisa tradicionais até o desenvolvimento de sistemas de busca avançados para recuperação de informações.

5 Conclusão

A interseção entre Arquivologia e Organização do Conhecimento pode trazer muitos avanços no campo teórico e prático. A aplicação de conceitos e técnicas de Organização do Conhecimento em documentos de arquivo pode ajudar a melhorar a descrição,

recuperação e acesso à informação, tornando-a mais precisa, completa e contextualizada. Por sua vez, a Arquivologia pode contribuir com a compreensão das particularidades dos documentos de arquivo, suas relações com as instituições produtoras e detentoras. A troca de conhecimentos e práticas entre esses campos pode trazer inovações e soluções para os desafios enfrentados na representação da informação arquivística.

Segundo o levantamento empírico realizado, há uma variação significativa no cenário da representação da informação arquivística nos arquivos estaduais brasileiros. Dos 24 estados identificados, 19 possuem site institucional e, destes, 12 disponibilizam uma base de dados para recuperação da informação no seu acervo. Além disso, 16 instituições disponibilizam ao menos algum tipo tradicional de instrumento de pesquisa (inventário, catálogo, guia).

No entanto, também foi identificado que três unidades da federação (Rondônia, Roraima e Tocantins) não possuem arquivos estaduais identificados e outros quatro estados (Acre, Amapá, Paraíba, Rio Grande do Norte e Sergipe) não possuem seus sites institucionais disponíveis.

Além disso, mesmo entre os arquivos que possuem sites e instrumentos de pesquisa disponíveis, há uma variação na qualidade e na acessibilidade desses instrumentos. Alguns arquivos apresentam bases de dados bem estruturadas e fáceis de utilizar, enquanto outros oferecem apenas inventários em formato de PDF, por exemplo, dificultando a busca por informações específicas.

Portanto, é possível perceber que há um caminho a ser percorrido pelos arquivos estaduais brasileiros em relação à representação da informação arquivística e à disponibilização de seus acervos na internet, mas também há iniciativas importantes sendo realizadas em diferentes partes do país.

A construção de novos modelos de representação da informação arquivística com ênfase nos usos para a construção de memória e identidade de vários grupos sociais é uma questão que envolve diversos aspectos teóricos e metodológicos.

Em termos teóricos, é importante considerar que a representação da informação arquivística não é um processo neutro, mas sim influenciado por diversos fatores, como a cultura, a história, as relações de poder e os interesses políticos e econômicos. Nesse sentido, é fundamental que os modelos de representação da informação arquivística levem em conta a diversidade cultural e a pluralidade de perspectivas, de forma a possibilitar a representação de diferentes vozes e memórias.

Além disso, é necessário considerar que os arquivos são espaços de construção da memória coletiva e da

identidade de diferentes grupos sociais. Nesse sentido, é importante que os modelos de representação da informação arquivística sejam capazes de valorizar as experiências e saberes desses grupos, bem como de reconhecer a importância de suas lutas e resistências na construção da história e da cultura brasileira.

Do ponto de vista metodológico, a construção de novos modelos de representação da informação arquivística pode envolver a utilização de diferentes técnicas e abordagens, como a análise de gênero, a análise interseccional, a história oral, entre outras. Essas técnicas e abordagens podem ser utilizadas para identificar e analisar as lacunas existentes nos modelos de representação da informação arquivística atuais, bem como para identificar novas formas de representação que possam levar em conta a diversidade cultural e a pluralidade de perspectivas.

Por fim, é importante destacar que a construção de novos modelos de representação da informação arquivística com ênfase nos usos para a construção de memória e identidade de vários grupos sociais requer a participação ativa desses grupos no processo. É fundamental que os grupos sociais sejam reconhecidos como sujeitos ativos na construção e preservação de suas memórias e identidades, e que tenham voz na definição dos modelos de representação da informação arquivística que os representam.

Referências

- BARROS, Thiago Henrique Bragato; SOUSA, Renato Tarciso Barbosa de. Archival science and knowledge organization: mapping methodological relationships. **KNOWLEDGE ORGANIZATION**, v. 46, n. 7, p. 493-501, 2020.
- CORUJO, Luís Miguel Nunes; FREITAS, Maria Cristina Vieira de. Organização, gestão e representação da informação e do conhecimento na arquivística: interações e interlocuções. **PontodeAcesso**, v. 15, n. 3, 2021.
- FONSECA, Maria Odila. Informação, arquivos e instituições arquivísticas. **Arquivo e Administração**, v. 1, n. 1, p. 33, 1998.
- FREITAS, Thais Rodrigues de; SILVA, Eliezer Pires da. Os instrumentos de pesquisa nos arquivos. **Logeion: Filosofia da Informação**, v. 9, n. 2, p. 246-257, 2023.
- FUJITA, Mariângela Spotti Lopes; RODRIGUEZ, Sonia Maria Troitinho. Política de indexação em arquivos de instituições de saúde. **Informação em Pauta**, v. 3, n. 1, p. 95-116, 2018.
- JARDIM, José Maria; FONSECA, Maria Odila. Estudos de usuários em arquivos: em busca de um estado da arte. **DataGramZero-Revista de Ciência da Informação**, v. 5, n. 5, 2004.
- LINDEN, Leolíbia Luana; VIANNA, William Barbosa; BRÄSCHER, Marisa. Aspectos relevantes da arquivística contemporânea no âmbito da organização do conhecimento no Brasil. **Brazilian Journal of Information Science**, v. 13, n. 2, p. 48-58, 2019.
- LEHMKUHL, Camila Schwinden; SILVA, E. C. L. A representação da informação arquivística nos registros civis. In: **IV Congresso Brasileiro em Organização e Representação do Conhecimento (ISKO)**. 2017. p. 138-145.
- LLANES-PADRÓN, Dunia; MORO-CABERO, Manuela. Records in contexts: A new model for the representation of archival information in semantic web scenarios. **El Profesional de la Información**; Vol. 26, Núm. 3, 2017.
- ORRICO, Evelyn Goyannes Dill; SILVA, Eliezer Pires da. Divulgação científica nos arquivos do Brasil: representação arquivística na construção da memória e identidade. **Em Questão**, p. 256-277, 2019.
- ORRICO, Evelyn Goyannes Dill; SILVA, Eliezer Pires da; PEREIRA, Diogo Baptista. Os instrumentos de representação da informação arquivística para a construção de memória e identidade. Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação, In: **XX ENANCIB**, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/123658>. Acesso em: 17 abr. 2023.
- PINTO, Luíza; GOMES, Priscila. Ações educativas em Arquivos: diálogos possíveis na Arquivologia pós-moderna. **Revista do Arquivo Geral da Cidade do Rio de Janeiro**, n. 19, p. 253-267, 2020.
- PEREIRA, Diogo Baptista; SILVA, Eliezer Pires da. Funções arquivísticas: caracterizando finalidades de instituições de arquivo. **ÁGORA: Arquivologia em debate**, v. 29, n. 58, p. 1-22, 2019.
- RODRIGUES, Fernanda; GOMES, Priscila. Arquivologia e educação. **P2P & INOVAÇÃO**, v. 7, p. 63-87, 2021.
- SANTOS, Cleber Belmiro dos; SILVA, Eliezer Pires da. O acesso e a divulgação de acervos arquivísticos na Casa de Oswaldo Cruz. In: **XVII Encontro Nacional de Pesquisa em Ciência da Informação**. 2016.
- SILVA, Eliezer Pires da; ORRICO, Evelyn Goyannes Dill. A NORMALIZAÇÃO DA DESCRIÇÃO ARQUIVÍSTICA E A ORGANIZAÇÃO DO CONHECIMENTO. In: **I Congresso ISKO Espanha e Portugal/XI Congresso ISKO Espanha**. 2013.
- VITAL, L. P.; MEDEIROS, M. B. B. Representação temática de documento arquivístico: em busca de especificidades. **Em Questão**, Porto Alegre, v. 25, n. 1, p. 345-367, 2019.
- YAKEL, Elizabeth. Archival representation. **Archival Science**, v. 3, p. 1-25, 2003.

Representação de figurino: a Moda como informação

Costume representation: Fashion as information

Maria Cecília J. Barros (1), Rosali Fernandez de Souza (2), Naira Silveira (3)

(1) PPGCI-IBICT/UFRJ, Rua Lauro Muller, Rio de Janeiro, cecilia.barrosufrj@gmail.com, (2) rosalfs@gmail.com, (3) naira.silveira@unirio.br

Resumo

A presente pesquisa visa contribuir para o desenvolvimento de uma abordagem para a Organização e Representação Documentária de figurinos. Busca analisar os registros bibliográficos de figurinos em museus, identificando os elementos descritivos presentes; comparar os elementos descritivos dos registros de figurinos com os documentos normativos. A pesquisa de caráter exploratório, de modo a viabilizar ao pesquisador uma abordagem objetiva quanto ao problema de pesquisa. Foi realizado um levantamento por meio de vocabulários controlados, e o resultado da consulta terminológica foi submetido em bases nacionais e internacionais. Analisou-se uma peça representativa para a Moda: vestido de luto. Foi feita uma consolidação dos dados resultantes dessa pesquisa. Com base nos elementos descritivos expostos, estabeleceu-se uma proposta de representação de figurino segundo documentos normativos. Como resultado, observou-se que os registros constantes nas instituições selecionadas: Victoria & Albert, Museu do Traje, Kyoto Costume Institute e FIT Museum, apesar de tratarem da mesma peça, possuem elementos descritivos destoantes uns dos outros. Conclui-se, portanto, que é importante a interoperabilidade entre os sistemas de informação, respeitando as características próprias de cada instituição e de seus usuários. E, que se faz necessário o estudo das normativas, da teoria e do contexto social de modo a determinar os elementos essenciais para a representação de figurinos em coleções.

Palavras-chave: Organização e Representação do Conhecimento. Representação Descritiva. Figurino.

Abstract

This research aims to contribute to the development of an approach for the Documentary Organization and Representation of costumes. It seeks to analyze the bibliographic records of costumes in museums, identifying the descriptive elements present; compare the descriptive elements of costume records with normative documents. The research was exploratory in nature in order to allow the researcher an objective approach to the research problem. A search was conducted using controlled vocabularies, and the result of the terminological consultation was submitted to national and international databases. A representative fashion item was analyzed: a mourning dress. A consolidation of the resulting data was made. Based on the exposed descriptive elements, a proposal was established for the representation of the costume according to the normative documents. As a result, it was observed that the records in the selected institutions: Victoria & Albert, Costume Museum, Kyoto Costume Institute and FIT Museum, despite dealing with the same piece, have descriptive elements that differ from each other. It is concluded, therefore, that interoperability between information systems is important, respecting the characteristics of each institution and its users. And, that it is necessary to study the norms, the theory and the social context, in order to determine the essential elements for the representation of costumes in collections.

Keywords: Knowledge Organization and Representation. Descriptive Representation. Costumes.

Resumen

Esta investigación pretende contribuir con el desarrollo de un enfoque para la Organización y Representación Documental de trajes. Busca analizar los registros bibliográficos de trajes en museos, identificando los elementos descriptivos presentes, comparando los elementos registrados con los documentos normativos. La investigación fue de carácter exploratorio, con el fin de permitir al investigador un abordaje objetivo del problema en análisis. Se realizó una búsqueda utilizando vocablos controlados, y el resultado se remitió a bases de datos nacionales e internacionales. Se analizó una pieza de moda representativa: un vestido de luto. Se hizo una consolidación de los datos resultantes. A partir de los elementos descriptivos expuestos, se estableció una propuesta de representación del traje según los documentos normativos consultados. Como resultado, se observó que los registros de las instituciones seleccionadas: Victoria & Albert, Costume Museum, Kyoto Costume Institute y FIT Museum, a pesar de ser la misma pieza, presentan elementos descriptivos diferentes entre sí. Se concluye, por lo tanto, que es importante la interoperabilidad entre los sistemas de información, respetando las características de cada institución y de sus usuarios. Además, es necesario estudiar las normas, la teoría y el contexto social para determinar los elementos esenciales en la representación de los trajes de las colecciones.

Palabras clave: Organización y representación del conocimiento. Representación Descriptiva. El traje.

1 Introdução

Uma das principais dificuldades encontradas no campo de acervos está pautada nas coleções têxteis preservadas em museus brasileiros. Há falta de cursos de formação na área e, por consequência, uma carência de profissionais especializados que deem conta de representações adequadas para recuperação de informação especializada. Ademais, há pouca literatura especializada em língua portuguesa, o que agrava a falta de apoio e incentivo para os desenvolvimentos de pesquisas sobre o tema.

A Moda evidencia não somente as relações sociais, culturais e econômicas das classes em sociedade, como também constitui ferramenta essencial para a construção de poder do ser social. Tal conceito transborda as fronteiras disciplinares, estabelecendo uma pluralidade de abordagens interpretativas para o mesmo objeto. Ou seja, o campo da Moda pode ser visto como uma área transdisciplinar à medida que pode ser definida pela conjunção de várias disciplinas simultaneamente.

Os acervos de figurino, aliado ao modelo de representação documentária, são de suma importância para efetivação de seu uso. Nesse contexto, a relevância do tema para a Ciência da Informação se justifica à medida que na organização de todo tipo de acervo em um sistema de recuperação da informação, se faz necessário delinear um conjunto de ações pautadas no tripé: preservação, pesquisa e comunicação, respaldadas em paradigmas da Organização e Representação da Informação e do Conhecimento.

A necessidade de intensificar a exploração de recursos de representação desse tipo de documento (figurino), acrescida de estudos de ressignificação e apropriação de formatos de entrada de dados existentes, para a representação de figurino, justificam, contextualizam e agregam relevância científica a esta pesquisa. Os espaços institucionais que abrigam tais acervos são espaços, onde são alocadas as inscrições do conhecimento, constituídos de objetos de memória atribuídos a valores simbólicos, estéticos e culturais com os quais a sociedade celebra a memória para a construção de identidades.

Deste modo, apresenta-se uma investigação de acervos de figurino, no escopo da Organização e Representação do Conhecimento, visando a contribuir para o desenvolvimento de uma abordagem adequada para representação da informação. Para isso, são utilizados fundamentos das áreas da Ciência da Informação, Biblioteconomia e Museologia, buscando identificar semelhanças e diferenças entre as representações, além de contribuir com a caracterização destes acervos como documento.

2 Desenvolvimento

“A Moda é fruto do amadurecimento da afirmação do eu, da valorização social do indivíduo e da sua personalidade”. (BALDINI, 2006, p. 34) Pode-se afirmar que a Moda se desenvolveu graças a uma demanda contínua de uma inovação estética evidenciada, sobretudo, pela aparência, de modo a representar seu “lugar social” dentre as sociedades modernas.

2.1 Moda e Sociedade

A palavra Moda refere-se ao “modo”, maneira, comportamento, uso, hábito ou forma de agir característica de um determinado meio ou de uma determinada época: costume. Corroborando com a ideia, João Braga (2011, p. 38) afirma que:

[...] a palavra “Moda”, esta tem sua origem etimológica na língua latina e deriva de *modus*, que significa “modo, maneira”. Em algumas línguas neolatinas como o português, o espanhol e o italiano, originaram a palavra “Moda”; em francês, outra neolatina, deu a palavra *mode*. Percebe-se que o radical é exatamente o mesmo e, neste entendimento, Moda, antes de ser Moda, como o senso comum assim o entende, é modo, é maneira. É uma maneira de portar-se é um modo de vestir-se.

Dentre os possíveis fatores que fizeram o homem cobrir o corpo, pode-se citar: fatores culturais: pudor; fatores físicos: proteção e fatores estéticos: para diferenciar-se em sociedade. Embora a palavra Moda remeta à cobertura do corpo, roupas e adornos, o termo é muito mais abrangente do que o conceito de uma série de elementos visuais e comportamentais que a defina, individualize e contextualize seus indivíduos em um meio social. Para Lipovetsky (1989, p. 24):

[...] a Moda não tem conteúdo próprio; forma específica da mudança social, ela não está ligada a um objeto determinado, mas é, em primeiro lugar, um dispositivo social caracterizado por temporalidade particularmente breve, por reviravoltas mais ou menos fantasiosas, podendo, por isso, afetar esferas muito diversas da vida coletiva.

A Moda reflete não somente, acerca dos modos e de suas civilizações, assim como seus códigos. Adquirindo, assim, papel relevante no meio social e na comunicação subjetiva com o mundo exterior. Ou seja, a Moda adquire papel de comunicação simbólica, uma forma de sua posição e status social. A Moda sempre foi instrumento de estratificação social, as técnicas empregadas na confecção denunciam quem são os portadores de poder.

Com isso as classes inferiores no intuito de legitimar seu lugar em sociedade e obter respeitabilidade entre as demais classes sociais, iniciam uma verdadeira busca pela nobreza. Imitam as maneiras de vestir e ser e seu comportamento das classes superiores. Para Mello e Souza (1987, p. 130), neste período, a Moda é vista como “[...] um dos instrumentos mais poderosos de integração e desempenha uma função niveladora, ao permitir que o indivíduo se confunda com o grupo e desapareça num

todo maior que lhe dá apoio e segurança”, alimentando, assim, o desejo material de elevar-se na escala social, e atingir, ao menos esteticamente, o mesmo nível da elite.

Em contrapartida, para demarcar o distanciamento social, e ocultar vestígios comportamentais e de vestimentas apropriados pela burguesia, as classes superiores se veem obrigadas a inovar constantemente sua aparência, de modo a legitimar sua posição e não se assemelhar com os demais estamentos da sociedade. À medida que as demais classes reproduzem os padrões de consumo das classes mais altas da Europa, estas buscam se diferenciar de outros grupos. Acima de tudo, a burguesia precisava aparentar sua posição elevada aos demais: representando não apenas o patrimônio, mas um capital simbólico. No que tange o conceito de capital simbólico, o mesmo pode ser visto como o prestígio, honra e poder, mas palavras de Bourdieu e de Roche (1987, p.164) afirmam que:

O “capital simbólico” é, na verdade, um efeito da distribuição das outras formas de capital em termos de reconhecimento ou de valor social, é poder atribuído àqueles que obtiveram reconhecimento suficiente para ter condição de impor o reconhecimento.

Roche, dialoga com o conceito de capital simbólico ao declarar que a dialética da Moda sustenta uma habilidade ímpar de compreender as transformações que ocorrem em sociedade.

A partir do século XVII, especialmente após o grande movimento de reflexão religiosa nascido das reformas católicas e protestante, a roupa passa a ser o centro dos debates sobre riqueza e pobreza. [...] Na visão moral cristã, a Moda serviu de meio para avaliar a adaptação dos costumes às exigências éticas. Para economia, que priorizavam a utilidade e o motor de consumo, ela também foi, um século mais tarde, o exemplo habitual da produção humana que se almejava para a melhoria da sociedade e da vida (ROCHE,2007, p. 21)

Ainda de acordo com Roche (2007, p. 70)

A Moda representa a cartada simbólica na batalha das aparências numa sociedade em que a distribuição e difusão das riquezas estavam em processo de mudança, permitindo uma maior ou menor mobilidade social. Era uma questão para a nobreza e para burguesia, para as elites e para os que ainda não haviam chegado.

Se por um lado o conceito de capital simbólico foi construído progressivamente por Bourdieu (1987), ele está alicerçado a uma estrutura social de categorias. Logo, essa estratificação da sociedade, está evidenciada por meio do capital simbólico, no qual é determinado pelo o que é enaltecido pela sociedade vigente.

Bourdieu e Delsaut (1975) fazem um estudo sobre a estrutura do campo da alta costura em Paris, que apresenta distribuição desigual de uma espécie particular de capital entre as diferentes *maisons*, engendrando a concorrência no campo e as condições para a entrada de novos agentes nessa competição. Os autores referem-se a detentores de um capital de autoridade que conduzem o campo a produzir

capital simbólico que, por sua vez, é produtor de crenças. Por exemplo, semanários e revistas especializadas dedicavam páginas a diferentes costureiros, indicando não somente sua posição na distribuição de capital específico, mas representando parte do lucro simbólico e material que eles podiam obter na produção do campo, sendo o efeito de grife do costureiro um ato de magia. (CAMPOS; LIMA, 2018, p.110)

O livro *O Casaco de Marx: Roupas, Memória, Dor* de Peter Stallybrass (2016), explicita o que disse Bourdieu, ao se referir ao capital simbólico. O texto relata como um bem de consumo, como o casaco de Marx em plena Era Vitoriana, tinha um valor agregado superior ao simbólico ou a necessidade de cobrir o corpo. Stallybrass, narra como o casaco de Marx, bem como as vestes das filhas se tornaram patrimônio da família. Marx penhora diversas vezes o casaco para ter acesso ao Gabinete de Londres, mas em contrapartida era o casaco que o impunha respeito para estar naquele local, adquirindo, assim papel relevante, no meio social, e na comunicação subjetiva com o mundo exterior. Ou seja, a Moda adquire papel de comunicação simbólica, uma forma de sua posição e status social.

2.2 Organização e Representação do Conhecimento

A Organização do Conhecimento, enquanto campo de estudo, é pautada fundamentalmente na análise de características e propriedades dos conceitos de modo que as semelhanças e diferenças possam ser identificadas, determinando assim, os tipos de relacionamentos entre conceitos. Essa organização documental é possível com a utilização de processos técnicos que incluam atividades de classificação, a indexação e a catalogação. O itálico deve ser usado apenas para ressaltar termos em outras línguas, títulos das subseções, títulos na lista de referências e, ocasionalmente, para enfatizar palavras ou frases.

Buckland (1991, p. 1-2) em sua obra intitulada *Information as thing*, analisou os usos do termo informação na Ciência da Informação, segmentando-a em 3 aspectos:

Informação-como-processo: Quando alguém é informado, aquilo que conhece é modificado. Nesse sentido “informação” é “o ato de informar...; comunicação do conhecimento ou “novidade” de algum fato ou ocorrência; a ação de falar ou o fato de ter falado sobre alguma coisa” (2) Informação-como-conhecimento: “Informação” é também usado para denotar aquilo que é percebido na “informação-como processo”: o “conhecimento comunicado referente a algum fato particular, assunto ou evento; aquilo que é transmitido, inteligência, notícias”. (3) Informação-como-coisa: O termo “informação” é também atribuído para objetos, assim como dados para documentos, que são considerados como “informação”, porque são relacionados como sendo informativos, tendo a qualidade de conhecimento comunicado ou comunicação, informação, algo informativo.

A análise de Buckland levanta considerações importantes: se por um lado Buckland retoma o conceito de documento, ao afirmar que informação-como coisa pode ser dado aos objetos de caráter informativo, ou seja, que tem a qualidade de comunicar algo mesmo com a ausência da escrita; por outro indica o caráter subjetivo inerente a toda informação. Desta forma, Buckland corrobora a ideia ao afirmar que todo e qualquer objeto é informativo, de acordo com o enfoque que lhe é dado.

Sob essa perspectiva, todo documento pode ser considerado informacional, no sentido de que todo documento carrega em si “todos os meios que servem para informar e comunicar algo e que não tenham a escrita como principal meio de expressão” (SANTOS, 2007, p.57). Mais do que uma narrativa bibliográfica, tais atividades provocam questionamentos relacionados à natureza do acervo. Entender as diversas nuances de significado que tal documento possui, assim como seu caráter informacional, se torna necessário para uma pesquisa que tem como foco acervos tridimensionais. É imprescindível, para atingir tais objetivos, compreender, mesmo que de forma parcial, os diferentes aspectos que podem ser inferidos a partir do estudo de objetos. Deste modo, tal pesquisa, debruça-se quanto à investigação das possibilidades e, bem como a delimitação, dos objetos como fonte primária de informação e como objetos em si.

Sob a perspectiva de Otlet (1934), citado por Santos (2007, p. 57), em seu *Traité de Documentation*, obra sobre a organização e o acesso ao conhecimento, sugere que todo objeto pode ser considerado documento, ampliando assim, o sentido do termo documento, que segundo o Otlet, “são todos os meios que servem para informar e comunicar algo e que não tenham a escrita como principal meio de expressão”, como por exemplo: manuscritos, mas também arquivos, mapas, esquemas, ideogramas, diagramas, desenhos e reproduções dos mesmos, fotografias de objetos reais, entre outros. Contudo, a visão de Otlet quanto a documento circula entre as noções de livro, documentação e bibliografia, ou seja, quanto à documentação escrita.

Suzane Briet (1894-1989) citada por Pinheiro (2002, p. 3), em seu livro *Qu'est-ce la Documentation* (1951), considera o documento como “qualquer traço concreto ou simbólico preservado ou registrado com o propósito de representar, construir ou comprovar um fenômeno físico ou intelectual”. Deste modo, Briet ressignifica o conceito de documento cunhado por Otlet. Briet não somente amplia o entendimento de documento, para além do suporte escrito, como também corrobora com a ideia de Otlet ao conceituar documento para o campo da Ciência da Informação. Logo, pode-se afirmar que tanto o suporte escrito, como oral, visual e objetos podem ser considerados documentos sob o viés do espectador. Como dito por Briet, documento é tudo aquilo que pode ser utilizado como instrução ou prova de um fato

Na sociedade da informação a necessidade de aperfeiçoamento das técnicas utilizadas para a Organização da Informação e do Conhecimento ganham espaços cada vez maiores na busca por modelos de representação documentária, que visem a dar conta da recuperação e uso efetivo da vasta produção de informação e conhecimento produzidos na contemporaneidade. Nas visões de Café e Sales (2010, p. 117) percebe-se a preocupação com o tema:

Em uma visão mais ampla, podemos dizer que precisamos organizar para poder compreender o mundo e nos comunicarmos melhor. [...]. No contexto dos sistemas de informação, a função não é diferente. Organizamos um acervo para compreendê-lo melhor e assim podermos recuperar objetos informacionais, isto é, informações registradas nos mais variados suportes.

Bräscher e Café (2008, p. 5) definem a Representação Descritiva “como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico”. Para Silveira (2017, p. 7) essa descrição sob o aspecto bibliográfico expressa pelo processo de Representação Descritiva (Catalogação) onde são utilizados padrões, formatos de entrada e linguagens específicos, internacionalmente aceitos, e que têm como objetivo homogeneizar dados que sirvam como embasamento para sua recuperação e intercâmbio. No desenvolvimento deste estudo, os princípios da Catalogação, quanto a documentos normativos que nortearam a representação de um acervo constituído por documentos não tradicionais.

2.3 Documentos normativos

No linear da história da Biblioteconomia, alguns documentos desenvolvidos no âmbito da Representação Descritiva, se tornaram um marco para os profissionais da área. O principal objetivo desses documentos está pautado em fortalecer as diretrizes do Controle Bibliográfico Universal, viabilizar o intercâmbio internacional de dados e orientar os profissionais da informação no que tange os processos de Catalogação (SILVEIRA, 2017).

As características da Catalogação aliadas à aceção, sintaxe, semântica, semiótica, estabelecem juntas os fundamentos necessários para o desenvolvimento eficaz do processo comunicativo entre o usuário e o acervo representado no catálogo, onde o foco será sempre a recuperação relevante da informação, no caso desse estudo, a recuperação do acervo de figurino.

Em sentido amplo pode-se afirmar que a Catalogação se refere à representação como ‘algo no lugar de’, ou seja, faz-se uso de metadados para representação dos itens identificadores de objetos. Para tanto é necessário o uso de padrões internacionais que possibilitem o intercâmbio entre bases de dados de diferentes unidades de informação. Para o desenvolvimento desta pesquisa, selecionou-se os seguintes documentos normativos: a Declaração de Princípios Internacional de Catalogação

(ICP), a International Standard Bibliographic Description (ISBD) e quanto a especificidade desta pesquisa o ICOM Costume (International Council of Museums).

Deste modo, o modelo optou por termos e definições que, de maneira mais ampla, são aplicáveis a todos os registros bibliográficos. Em consequência, os elementos de dados que são vistos como especializados ou são específicos para certos tipos de recursos, geralmente não são representados no modelo

As diretrizes Internacionais de documentação sobre objetos museológicos, como por exemplo o ICOM, contemplam uma descrição das categorias de dados que podem ser utilizadas no desenvolvimento de registros catalográficos. Estas Diretrizes devem ser utilizadas como base para catalogação dos acervos museológicos, a fim de atender todas as disciplinas no âmbito de um museu.

O principal objetivo em aplicar os documentos normativos nos registros catalográficos visa permitir a pesquisa e apresentação de resultados de acordo com elementos considerados essenciais aos registros e figurinos, e que estejam pautados nos documentos normativos adotados para a análise da representação documentária de figurino

3 Análise e Resultado

A Representação Documentária realizada em acervos está pautada na descrição detalhada das peças, assim como no registro da biográfico de cada item e de suas características físicas. Tal descrição fundamenta-se como fonte de informação, e contribui para a identificação, localização e recuperação da peça nos acervos.

A seleção das peças partiu do contexto de figurino e impacto social, cada peça pode revelar informações específicas de contextos sociais diferentes. Optou-se pelo vestido vitoriano de luto, uma vez que a tendência de usar preto para designar o luto, foi instituída e popularizada pela Rainha Vitória, precursora da Moda Vitoriana. Logo, será realizada uma análise de cada registro bibliográfico individualmente e, posteriormente, uma análise comparativa entre os registros recuperados, de mesma tipologia.

Foram selecionadas 4 instituições que possuem acesso abertos aos registros bibliográficos a seus acervos. São elas: Victoria & Albert Museum, na Inglaterra; Fashion Institute of Technology Museum, nos EUA; Museu Nacional do Traje, em Portugal e Kyoto Costume Institute, no Japão.

Figura 1. *Queen Victoria Mourning Victoria & Albert Museum*



Fonte: Metropolitan Museum (2009)

A sociedade inglesa, durante a Era vitoriana, era pautada por regras de comportamento bastante rigorosas, conseqüentemente, os rituais de morte, nascimento e casamento. De modo geral, o ritual de luto, em particular, era mais rígido do que os demais. Desrespeitar as etapas do luto, bem como suas regras, eram vistos como atos imorais perante a sociedade.

A figura 1 exemplifica o luto extremo incorporado pela Rainha, presente nos rituais de morte e concedendo à Inglaterra, ao menos durante este período, as principais influências da Moda à época.

Jornais de costumes e manuais de etiqueta, muito comuns à época, traziam todas as recomendações e dicas a serem seguidas nesses momentos e eram muito populares entre a classe média. [...]. No entanto, nenhuma manifestação do luto era mais necessária e emblemática do que o luto vestimentar. Por meio dele, expressava-se imediatamente o apego ao morto e a tristeza pela perda: a dor pessoal tornada pública e visível, formando uma barreira simbólica entre o indivíduo e o seu meio – a imagem de austeridade que cobra distanciamento da mundanidade. Normalmente contava com dois estágios: luto fechado e meio-luto, cada um com suas regras próprias. Sua cor oficial era o preto, reconhecidamente a cor da ausência de luz e de vida. Os tecidos deveriam ser discretos como os de algodão ou lã, nunca brilhantes ou chamativos como o cetim, a seda e o veludo. (SCHIMITT, 2017, p.78)

Um dos aspectos mais característicos da sociedade vitoriana é o seu fascínio pela morte. Seja pelo luto

adotado pela Rainha Vitória por mais de anos, em decorrência da morte do príncipe consorte Albert (1861), ou pelo sentimento constante de perda causado pela onipresença da morte na Era Vitoriana. Contudo, a Moda foi muito generosa ao enaltecer o sentimento de luto instaurado na sociedade vitoriana. Foi durante a Era Vitoriana que o preto se tornou a cor tradicionalmente associada ao luto, popularizado pela Rainha Vitória.

Devido a sua relevância para a Moda, optou-se pelo vestido de luto Vitoriano, como peça para análise da descrição bibliográfica. Buscou-se nos acervos selecionados para a pesquisa o termo “Mourning Dress”, para as instituições de acervo em língua inglesa, e “vestido de luto”, para a instituição de língua portuguesa. Com base no resultado da busca, selecionou-se 4 peças que possuem as mesmas características que definem uma peça como vestido de luto. Todas as instituições retornaram positivamente na busca, ou seja, nas instituições houve recuperação do termo buscado.

A seguir, os resultados obtidos com a busca realizada nas instituições de Moda pesquisadas, assim como os campos de composição dos registros bibliográficos em relação aos documentos normativos listados.

Figura 2. *Queen Victoria Mourning Victoria & Albert Museum*



Fonte: Victoria & Albert (2021)

Observa-se que o registro ilustrado pela figura 2, tem um maior número de dados descritivos do que os anteriormente apresentados. Entre os campos informados descreve-se: Artist/Maker (Artista/criador);

Place of origin (Local de origem); Categories (Categoria); Object type (Tipo de objeto); Parts (Partes); Materials and techniques (Materiais e técnica); Brief description (Breve descrição); Physical description (Descrição Física); Dimensions (Dimensões); Marks and inscriptions (Marcas e inscrições); Credit line (Linha de crédito); Object history (Histórico do objeto); Summary (Resumo); Collection (Coleção); Accession number (Número de acesso).

Cabe ressaltar que apesar dos elementos descritivos não serem constituídos das áreas da ISBD, o registro do vestido de luto do V&A Museum menciona campos muitos semelhantes e incluem campos relacionados ao assunto, tais como Categories e Object type. Observa-se que esses descritores possuem hiperlink para dar acesso a outras peças classificadas pelos mesmos termos. Apesar de possuir algumas evidências que sugerem a aplicação de um vocábulo controlado, não foi localizado na homepage quaisquer documentos que reforcem esta afirmação.

Figura 3. *Mourning ensemble FIT Museum*



Fonte: FIT Museum (2015)

Dentre os campos que compõem o registro da figura 3, evidenciam-se: People and Brands (Pessoas e marcas); Retailer (Varejista); Medium (Materiais); Date (data); Country (País); Object number (Número do objeto). Observa-se que neste registro, pertencente a instituição, o FIT Museum, a descrição tiveram início pelos campos de autoria do vestido. Durante o período de produção do vestido, 1907, já estava instituída a utilização de etiquetas e marcas de autoria e responsabilidade pela criação e fabricação da peça. Este período é marcado pelo auge da 2ª Revolução Industrial, o que contribuiu com o desenvolvimento da indústria têxtil.

Figura 4. *Vestido de luto - Museu Nacional do Traje*



Fonte: Museu Nacional do Traje (2021)

Ao analisar a figura do vestido de luto do Museu Nacional do Traje de Portugal ilustrado pela figura 4, observa-se a utilização dos seguintes campos de descrição bibliográfica: N.º de Inventário; Supercategoria; Categoria; Denominação; Autor; Datação; Matéria; Técnica; Dimensões (cm); Descrição; Incorporação; Origem / Historial. Nota-se que a disposição dos elementos descritivos dos registos inicia-se pela última área da ISBD, a área 8, neste caso, o N.º de Inventário. Por se tratar de uma ficha de inventário, a organização dos elementos é encabeçada pelo campo mais relevante a ser representado, ou seja, a localização do item no acervo.

No que tange o objeto em questão, o vestido de luto vitoriano, a área 5 de descrição física da ISBD é um dos campos essenciais para a descrição bibliográfica, em virtude da sua constituição física que viabiliza a reprodução fidedigna do item, assim como reconstituir dados originais do objeto.

Embora o ISBD não englobe elementos que compõem o assunto, os campos Supercategoria e Categoria, refere-se ao conteúdo temático da peça. Contudo não foi identificado quaisquer documentos que orientem quanto a terminologia adotada para a representação documentária do acervo.

Figura 5. *Jacket Skirt - Kyoto Costume Institute*



Fonte: Kyoto Costume Institute (2021)

A figura 5 representa o registro do vestido Vitoriano de Luto do Kyoto Costume Institute, composto por 2 peças que dão a visão geral de um vestido. O termo que designa a imagem Jacket, Skirt, confirma esta percepção, e supre o conceito de título para determinar a expressão pela qual o vestido deve ser atribuído. A utilização de duas peças para composição do vestido, também se evidencia pela área 8 da ISBD. Neste caso, delimitada pelo campo Inventory number, ao determinar dois números de inventário a este registro: um para a jaqueta, e outro para a saia.

Observa-se que alguns elementos dos registros bibliográficos, que constituem o conjunto de elementos explicitados pela International Standard Bibliographic Description (ISBD) alinhado às regras de descrição do ICOM, são presentes nas descrições bibliográficas dos vestidos de luto.

A seguir será apresentado as áreas que compõem a ISBD e que constam nos registros bibliográficos dos vestidos de luto ora apresentados. Esta comparação busca evidenciar as áreas da ISBD ou por declaração *ipsis litteris* da referida área, ou por inscrição subentendida.

Área 1 área do título e da responsabilidade – No que tange o emprego de título, nota-se que eles não são delimitados por campos específicos. Mas que encabeçam os registros de modo a orientar o usuário quanto ao conteúdo. Quanto à autoria dos registros, ora citados, verifica-se que mesmo determinado campo delimitador para esta informação, nenhum dos registros para o vestido de luto determinou autor, designer ou “costureiro”. O que, de fato, é um dado relevante uma vez que o período da Era Vitoriana, está sobreposto ao início da Alta Costura. E tendo em vista o luxo aplica

aos vestidos, há de se determinar o estilista ou Maison a qual pertence. Quanto ao elemento de autoria, presente na área da ISBD, consta nos registros do Victoria & Albert, FIT Museum e Museu Nacional do Traje. O Kyoto Costume Institute (KCI), não consta menção a este elemento. A área do título por sua vez é mencionada de forma implícita nos registros ora citados, com exceção do Museu Nacional do Traje que possui campo delimitador para este elemento.

Área 2 área da edição – Não é mencionada em nenhum dos registros evidenciados neste estudo. Contudo pode-se ponderar que a edição cabe a uma revisão do item já confeccionado. O documento, por se tratar de um figurino histórico, não considera relevante indicar possíveis alterações feitas por outros costureiros. Na área da Moda esta ação é conhecida como réplica ou, no caso de uso, revivalismo histórico.

Área 4 área da publicação – De acordo com a ISBD, o local pertencente a área 4, está atrelado ao local de publicação do editor ou distribuidor (IFLA, 2007, p.48). Para a descrição de figurinos, o local descrito deve ser o local da Maison de seu criador, junto com os termos Marca e etiqueta, que determinam quem foi o produtor e distribuidor da peça. O ICOM menciona em seu documento, Diretrizes do Comitê de Indumentária, que o local descrito na representação deve ser o local de aquisição, uma vez que na Moda o fator relevante é o estilista da peça. Discriminada pelos elementos: fabricante, distribuidor, local e data. Observa-se que apenas no V&A todos os elementos são informados pelo registro bibliográfico. O FIT Museum menciona todos os elementos, com exceção do fabricante, o Museu Nacional do Traje, consta apenas a data de produção. Os elementos de data e local, estão implícitos no registro ora citado pelo Kyoto Costume Institute. Ademais, nota-se que as regras do ICOM definem apenas a indicação do local e data nos registros para esse tipo de objeto.

Área 5 área da descrição física – este é um dos campos essenciais para a descrição bibliográfica e representação do item, uma vez que se trata de uma peça de produção artística e cultural. O que viabiliza a reprodução da peça, assim como conservar e salvaguardar dados da peça original. Por meio da área 5 da ISBD, podem ser descritos dados como a técnica empregada, a composição física dos tecidos, estabelecer a colorimetria original da peça, o estado físico da peça, etc. Ou seja, a área da descrição física reúne todas as características que tornam o objeto em si, em uma vestimenta. É por meio da descrição física, que se pode, também, aferir dados quanto a apresentação do item, plano de exposição, tipo de luz, ângulo etc. Cada um desses elementos deve ser descrito de modo a representar o item no momento da catalogação. Essas informações que permitirão a eficaz recuperação do item no acervo. Composta pelos elementos: material/técnica, formato, Colorimetria e dimensões. Todos os elementos informados são descritos pelo registro do V&A Museum, Museu

Nacional do Traje e Kyoto Costume Institute, com exceção do FIT Museum onde o registro aponta apenas o tipo de tecido.

Área 6 área da série/coleção - pode ser atribuída ao termo Coleção, pois possui descritores que delimitam em qual segmento do acervo esta peça está alocada. A área 6 pode ser aplicada quando o recurso descrito pertence a um recurso bibliográfico maior, ou seja, quando dado item pertence a um mesmo agrupamento de peças que possuem características em comum. Neste caso, somente o Victoria & Albert aponta em seu registro a coleção à qual o objeto pertence.

Área 7 área das notas - Esta área não só esclarece como também amplia o entendimento das outras áreas, presente em todos os registros à presente pesquisa quanto aos vestido de luto

Área 8 área do Número Internacional Normalizado - Trata-se do número de localização da peça no acervo. Também indicado por todos os registros ora mencionados.

Assunto - Embora a ISBD não englobe elementos que compõem o assunto, entretanto apontado pelo ICOM como elemento essencial, os campos Supercategoria (Museu Nacional do Traje) e Categoria (Victoria & Albert), refere-se ao conteúdo temático da peça A catalogação de assunto, por exemplo, é o processo responsável por atribuir descritores ao documento de modo estabelecer os pontos de acesso temático. Determinar os assuntos que compõem o item são fundamentais pois facilitam a descrição, busca, localização e acesso a documentos que tratam de temática semelhante. Viabilizando a identificação do item dentro de uma determinada coleção ou acervo.

Em paralelo, o documento normativo do ICOM Costume (International Council of Museums) estabelece diretrizes quanto à catalogação de objetos museológicos. Estas Diretrizes podem ser utilizadas como base para catalogação dos acervos museológicos, a fim de atender peculiares temáticas de representação dos objetos no âmbito de um museu.

Cabe ressaltar que as Diretrizes para Documentação Museológica abarcam não somente o processo de registro catalográfico, organização e representação dos objetos, assim como as características físicas, dados administrativos, problemas resultantes de conservação e restauro, e toda informação histórica que o item possa expressar.

Acredita-se que tais documentos normativos, que abarcam modelos conceituais da Representação Descritiva, apresentam subsídios para a identificação de elementos essenciais na organização e representação de registros bibliográficos para o acervo de figurino.

4 Conclusão

Com base nos Princípios Internacionais de Catalogação (2016), pode-se afirmar que os registros atendem o princípio da significação, uma vez que menciona itens essenciais para determinar a diferenciação entre as peças. Ou seja, as informações que compõem o registro bibliográficos foram suficientes para individualizá-los quanto aos demais registros do acervo. Os registros possuem relevância pelo resumo histórico que descreve e o contextualiza, dentro do acervo. Este princípio diz que os registros bibliográficos devem refletir os elementos relevantes para os usuários, de modo a permitir a distinção entre as entidades, o que foi alcançado.

Como já mencionado, os campos que compõem a área 5 da ISBD, assim como a área de notas que descrevem a peça são essenciais não somente para interpretação do usuário quanto a peça exibida, mas também para contextualização da peça na história da Moda. A área da descrição física, se torna essencial para a interpretação da peça pelo usuário, haja vista que contém campos essenciais e relevantes para a descrição bibliográfica. Observa-se, também, que os campos onde as instituições não delimitam os elementos da peça em questão, são supridos pelas informações contidas na Área de notas. Logo, a relevância da Área de notas, para o registro bibliográfico, atribui à peça o detalhamento e a especificação das particularidades extrínsecas e intrínsecas que compõem o item, ampliando e enriquecendo a sua descrição.

Por sua vez, o vestido de luto do Museu Nacional do Traje, trata-se de um vestido de luto vitoriano, com todas as regras de conduta e vestimentas. A descrição física ilustra alguns itens de luxo na constituição do vestido, tais como o emprego de seda, cetim, renda, vidrilhos e cauda. A utilização desses itens se justifica na Área de notas, ao informar que apesar de ser um vestido de luto, em decorrência do falecimento de seu pai, foi utilizado durante o casamento da proprietária do vestido, caracterizando o meio-luto.

Como mencionado anteriormente, cada instituição estrutura seus dados de forma a viabilizar a interpretação do usuário. Contudo verifica-se que o registro ora apresentado pela instituição V&A Museum, é mais extenso que os demais. No decorrer da estrutura do registro o mesmo faz alguns resumos históricos de modo a situar o usuário quanto aos precedentes do item. O registro do V&A Museum é minuciosa ao narrar o ciclo de vida que as peças percorreram, desde sua criação, função e depósito na instituição, tornando esses dados relevantes e essenciais para percepção da construção das vestimentas. Ademais, pode-se ressaltar também que as imagens que ilustram os registros elucidam em que momento foram criados e para qual ocasião.

Quanto às figuras do vestido de luto, nota-se que a imagem ilustrada pelo FIT Museum, por exemplo,

evidencia um luto fechado, sóbrio, de modo a transmitir o pesar e a tristeza causada pela perda. Todavia, confeccionado para uma mulher jovem ou adulta, fica notório pela marcação acentuada da silhueta. Em contraponto, a figura do Kyoto Costume Institute, apesar de também se tratar de um luto fechado, foi fabricado para uma mulher idosa. A cintura larga, a saia em amplitude maior, o tecido pesado, estimam o sentido de corpo vedado ressaltados pelos rituais de luto na Era Vitoriana, para mulheres de idade mais avançada.

A estrutura dos registros bibliográficos do Kyoto Costume Institute evidencia a disposição da Área de notas, sempre encabeçando os elementos descritivos dos itens selecionados. Essa organização se faz relevante para os usuários do KCI uma vez que o campo de notas contextualiza a Moda Ocidental em consonância com período da História em que se encontra o item. Cabe ressaltar que o campo de notas esclarece e expande a descrição nas Áreas 1-6 da ISBD e podem relacionar-se com qualquer aspecto do registro. As notas podem também referir-se à história bibliográfica do registro e indicar relações com outros elementos da descrição. Tornando-se um campo relevante para compreensão do item pelo usuário.

Logo, pode-se concluir que não é possível determinar qual dentre os documentos normativos selecionados, melhor atende aos registros bibliográficos no âmbito da Moda. Consequentemente, que se faz necessária uma convergência desses documentos de modo a determinar os elementos essenciais para a descrição bibliográfica, e sua eficaz recuperação. Representar adequadamente os registros de Moda, se torna relevante uma vez que preconizam a Moda como informação. Tendo em vista as nuances que abarcam cada documento normativo, a partir de seus elementos constituintes, foi possível apresentar uma proposta de registro catalográfico com os elementos considerados essenciais para descrição bibliográfica no âmbito da Moda.

Neste contexto, "[...] o aperfeiçoamento de um sistema descritivo universal de vestuário e acessórios que dê conta da diversidade de objetos, favorece o entendimento de profissionais de nacionalidades e culturas variadas". (BENARUSH, 2014, p.6). Logo, pode-se inferir que os elementos essenciais que compõem a proposta de catalogação de figurino são relevantes para a busca e recuperação de informações ao usuário, viabilizando o acesso de forma eficaz aos documentos.

A partir do desenvolvimento da pesquisa, observou que a Moda é um campo que propicia estudar e narrar os fatos históricos sob o viés cultural, social e de comportamento. Portanto, o figurino é considerado um documento e, como tal, deve ter elementos que possam representar os aspectos descritivos e temáticos, possibilitando o acesso à informação. Ao analisar os elementos presentes nos registros dos figurinos dos

acervos de museus renomados, observou-se que cada instituição optou por elementos de representação próprios, embora existam muitos elementos em comum. A diferença entre a presença e ausência de elementos pode ser considerada como consequência de características próprias de cada acervo ou que os elementos são poucos estudados de modo colaborativo e integrado.

Os resultados das análises revelaram que as diversas instituições que possuem acervos de figurino utilizam representações generalistas para catalogar itens compostos em sua maioria por materiais têxteis. A pluralidade de itens que englobam a totalidade de campos que descrevem e constituem o acervo de Moda, evidenciou a necessidade de uma representação que abrangesse as nuances que compõem o figurino. Os registros bibliográficos analisados na presente pesquisa, denotaram que representar as particularidades do figurino, pode suprir as necessidades informacionais do usuário, bem como viabilizar a correta e precisa representação dessa informação.

Constatou-se, também, que a aplicação dos documentos normativos, abrangem não somente as bibliotecas, assim como podem ser utilizados em qualquer unidade de informação e podem ser aplicados em outros suportes. Neste contexto, é possível se desenvolver e estruturar um sistema descritivo que dê conta das nuances e variedades que abarcam esse objeto, definido como figurino. Os parâmetros utilizados para desenvolvimento dos campos que delimitam a descrição bibliográfica, investigados no âmbito da Organização e Representação de documentos, proporcionou a identificação de características das peças analisadas e a respectiva interpretação em seus contextos sociais ou simbólicos, inerentes aos aspectos da cultura material no tempo e no espaço.

A colaboração entre os campos da Biblioteconomia, Ciência da Informação, Museologia e Moda poderá contribuir em benefício do estabelecimento e consolidação das práticas de Representação Descritiva e Representação Temática, especificamente no caso de acervos de figurino e outros acervos de Moda em geral.

Referências

- BALDINI, M. **A invenção da Moda**: As teorias, os Estilistas, a História. 1. ed. Lisboa: Edições 70, 2006. 142p.
- BENARUSH, M. K. **Termos básicos para catalogação de vestuário**. Rio de Janeiro: Secretaria de Estado de Cultura, 2014.
- BOURDIEU, P. **Choses dites Paris**: Minuit, 1987.
- BRAGA, J. **Reflexões sobre Moda**. v.1. São Paulo, Editora Anhembi Morumbi, 2011.
- BRASCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento. In: ENANCIB, IX. 2008, São Paulo. **Diversidade cultural e políticas de informação**. São Paulo: USP, 2008. Disponível em: <http://enancib.ibict.br/index.php/enancib/ixenancib/paper/vicw/3016/2142>. Acesso em: 27 dez.2020.
- BUCKLAND, M. K. Information as thing. **Journal of the American Society for Information Science (JASIS)**, v.45, n.5, p.351-360, 1991.
- CAFE, L.M.A.; SALES, R. (2010). Organização da informação: conceitos básicos e breve fundamentação teórica. En: Robreo, J.; Brascher, M. (Orgs). **Passeios no Bosque da Informação**: estudos sobre Representação e Organização da Informação e do Conhecimento. Brasília,DF: Ibict
- CAMPOS, P. H. F.; LIMA, R. D. C. P. Capital simbólico, representações sociais, grupos e o campo do reconhecimento. **Cadernos de Pesquisa**, v. 48, n. Caderno de Pesquisa. 2018, p. 100–127, jan. 2018.
- FASHION INSTITUTE OF TECHNOLOGY MUSEUM. **Mourning ensemble**. New York, 2021. Disponível em: [http://fashionmuseum.fitnyc.edu/view/objects/asitem/search\\$0040/0?t:state:flow=7a31fa04-1036-45a4-a573-2ce8a00304d8](http://fashionmuseum.fitnyc.edu/view/objects/asitem/search$0040/0?t:state:flow=7a31fa04-1036-45a4-a573-2ce8a00304d8). Acesso em 31 maio. 2021.
- IFLA. **Descripción Bibliográfica Internacional Normalizada (ISBD)**. Espanha, 2007. Tradução em espanhol realizada pela Comissão de Tradução da Biblioteca Nacional de Espanha. Disponível em: <https://archive.ifla.org/ubcim/nd1/isbdM-pt.pdf> Acesso em 19 jun. 2021
- KYOTO COSTUME INSTITUTE. **Jacket Skirt**. Kyoto, 2021. Disponível em: https://www.kci.or.jp/en/archives/digital_archives/1850s_1860s/KCI_094 Acesso em 31 mai. 2021.
- LIPOVETSKY, G. **O Império do Efêmero**: a Moda e seu destino nas sociedades modernas. São Paulo: Cia. das Letras, 1989.
- MELLO E SOUZA, G. **O espírito das roupas**: a Moda no século XIX. São Paulo: Companhia das letras, 1987, p.130.
- METROPOLITAN MUSEUM OF NEW YORK. **Queen Victoria Mourning**. 2009. Disponível em: <https://www.metmuseum.org/art/collection/search/155839>. Acesso 01 jul. 2021
- MUSEU NACIONAL DO TRAJE. **Traje de luto**. Lisboa, 2021. Disponível em: <http://www.matriznet.dgpc.pt/MatrizNet/Objectos/ObjectosConsultar.aspx?IdReg=1109073> Acesso em 31 mai. 2021.
- PINHEIRO, L. V. R. Gênese da Ciência da Informação ou sinais anunciadores da nova área. In: **O campo da Ciência da Informação**: gênese, conexões e especificidades. João Pessoa, UFPB, 2002. p.61-86.
- ROCHE, D. **A Cultura das aparências**: uma história da indumentária, séculos XVII e XVIII. São Paulo, Senac, 2007.
- SANTOS, M. J. V. C. A representação da informação em Arquivos: viabilidade de uso de padrões utilizados na Biblioteconomia. **Acervo**: Revista do Arquivo Nacional. Rio de Janeiro, v.20, n. 1-2, p.57-66, jan/dez 2007.
- SCHMITT, Juliana Luiza de Melo. A dor manifesta: vestuário de luto do século XIX. In: Congresso Internacional de História. 8, 2017, Maringá-PR, **Anais...Maringá:UEM**,

2017. p.77-81. Disponível em:
<https://docplayer.com.br/44361476-A-dor-manifesta-vestuario-de-luto-no-seculo-xix.html> Acesso em: 25 mai. 2021.

SILVEIRA, Naira Christofolletti. **Um diálogo sobre instrumentos de representação descritiva. Encontro de Representação Documental** (UFSCar), n.1, 2017. Disponível em:

<http://www.telescopium.ufscar.br/index.php/enredo/enredo/aper/viewFile/108/97>. Acesso em: 20 jan. 2021.

STALLYBRASS, P. **O casaco de Marx: Roupas, memória, dor**. 5 ed. Autêntica: São Paulo, 2016.

VICTORIA & ALBERT MUSEUM. **Wedding Dress**. Londres, 2021. Disponível em:
<https://collections.vam.ac.uk/item/O78857/wedding-dress-unknown/> Acesso em: 02 jun. 2021.

Representação e Organização do Conhecimento dos dados legislativos na Câmara dos Deputados do Brasil: análise e classificação em direção à Agenda 2030

Representation and Knowledge Organization of legislative data in the Chamber of Deputies of Brazil: analysis and classification towards the 2030 Agenda

Carla Maria Martellote Viola (1), Luana Farias Sales (2)

(1) Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e a Escola de Comunicação (ECO) da Universidade Federal do Rio de Janeiro (UFRJ), Rua Lauro Muller, 455 - 4º andar - Botafogo - RJ - CEP 22290 – 160, viola.carla@gmail, (2) luanafsales@gmail.com.

Resumo

Objetivo: Analisar e classificar os dados legislativos encontrados no Portal da Câmara dos Deputados, para verificar em que medida os Objetivos de Desenvolvimento Sustentável da Agenda 2030 são contemplados na sua disponibilização. **Metodologia:** Bibliográfica para compreender as dimensões aplicadas aos dados legislativos e suas expressões de conhecimento. Complementa-se com a pesquisa documental, para ordenar as orientações que abrangem a análise da taxonomia com vistas à Agenda 2030. No que tange aos objetivos, adota-se a investigação exploratória, que abrange a literatura e o Portal da Câmara dos Deputados, e descritiva, pela caracterização dos dados legislativos. Intenta-se, ainda, uma abordagem qualitativa que revela a representação e organização desses dados. **Resultados:** A partir do modelo utilizado, os dados legislativos, quanto à origem, são para pesquisas governamentais e de registros; ao grau de processamento, são dados brutos; à abordagem, são quantitativos; à natureza, são textuais; ao tipo de coleta, é automático; ao nível de sensibilidade, são extensíveis; à materialidade, são digitais; e à perenidade, são episódicos institucionais. Além destas, outras classes são identificadas. **Conclusão:** A disponibilização dos dados legislativos na interface Dados Abertos do Portal da Câmara dos Deputados possibilita a disseminação da informação, o monitoramento e a prestação de contas requeridos pela Agenda 2030. Os diversos formatos, em que os 30 conjuntos de dados são apresentados, permitem amplo acesso e alta qualidade de uso. Além de que, por serem atualizados anualmente e estarem presentes em um portal governamental, são considerados atuais e confiáveis, de acordo com os Objetivos de Desenvolvimento Sustentável.

Palavras-chave: Dados legislativos. Câmara dos Deputados. Representação e Organização do Conhecimento. Objetivos de Desenvolvimento Sustentável. Agenda 2030.

Abstract

Objective: To analyze and classify the legislative data found on the Chamber of Deputies' Portal to verify to what extent the Sustainable Development Goals of the 2030 Agenda are contemplated in its availability. **Methodology:** Bibliographic to understand the dimensions applied to legislative data and their expressions of knowledge. It is complemented with documentary research to order the guidelines that cover the analysis of the taxonomy with a view to the 2030 Agenda. Regarding the objectives, an exploratory investigation is adopted that covers the literature and the Portal of the Chamber of Deputies and descriptive by the characterization legislative data. There is also a qualitative approach that reveals the representation and organization of these data. **Results:** From the model used, the legislative data as to the origin are for government research and records; the degree of processing is raw data; to the approach are quantitative; to nature are textual; the type of collection is automatic; sensitivity level are extensible; to materiality they are digital and to perpetuity they are institutional episodic. In addition to these, other classes are identified. **Conclusion:** Making legislative data available in the Open Data interface of the Chamber of Deputies Portal enables the dissemination of information, monitoring and accountability required by the 2030 Agenda. access and high quality of use. In addition, because they are updated annually and are present on a government portal, they are considered up-to-date and reliable in accordance with the Sustainable Development Goals.

Keywords: Legislative data. Chamber of Deputies. Representation and Knowledge Organization. Sustainable Development Goals. 2030 Agenda.

Resumen

Objetivo: Analizar y clasificar los datos legislativos que se encuentran en el Portal de la Cámara de Diputados para verificar en qué medida se contemplan en su disponibilidad los Objetivos de Desarrollo Sostenible de la Agenda 2030. **Metodología:** Bibliográfica para comprender las dimensiones aplicadas a los datos legislativos y sus expresiones de conocimiento. Se complementa con una investigación documental para ordenar los lineamientos que abarquen el análisis de la taxonomía con miras a la Agenda 2030. En

cuanto a los objetivos se adopta una investigación exploratoria que abarca la literatura y el Portal de la Cámara de Diputados y descriptiva por la caracterización de los datos legislativos. También existe un enfoque cualitativo que revela la representación y organización de estos datos. **Resultados:** Del modelo utilizado, los datos legislativos en cuanto al origen son para investigación y registros gubernamentales; el grado de procesamiento son datos sin procesar; al enfoque son cuantitativos; a la naturaleza son textuales; el tipo de cobro es automático; el nivel de sensibilidad es extensible; a la materialidad son digitales ya la perennidad son episódicos institucionales. Además de estos, se identifican otras clases. **Conclusión:** La disponibilidad de datos legislativos en la interfaz de Datos Abiertos del Portal de la Cámara de Diputados permite la difusión de la información, el seguimiento y la rendición de cuentas que exige la Agenda 2030. Acceso y alta calidad de uso. Además, debido a que se actualizan anualmente y están presentes en un portal del gobierno, se consideran actualizados y confiables de acuerdo con los Objetivos de Desarrollo Sostenible.

Palabras clave: Datos legislativos. Cámara de los Diputados. Representación y Organización del Conocimiento. Metas de Desarrollo Sostenible. Agenda 2030.

1 Introdução

Os dados legislativos são gerados por parlamentares, pelos funcionários e pelas atividades da Câmara dos Deputados. O processo legislativo, que compreende a elaboração, análise e votação de vários tipos de propostas: leis ordinárias, medidas provisórias, emendas à Constituição, decretos legislativos e resoluções, entre outras (Brasil, 2023a), é a principal fonte de dados legislativos, contudo, outros procedimentos realizados e adotados na Câmara dos Deputados também produzem dados. Esses dados estão disponibilizados no Portal da Câmara dos Deputados, no serviço de Dados Abertos (BRASIL, 2023b). Esta interface é a principal e mais abrangente fonte de dados legislativos que o Portal da Câmara dos Deputados disponibiliza para as várias audiências interessadas em fazer consulta, uso e pesquisa.

Paralelamente a essa iniciativa legislativa brasileira de disponibilizar dados legislativos, o documento “Transformando Nosso Mundo: A Agenda 2030 para o Desenvolvimento Sustentável”, das Nações Unidas, alerta na introdução que a *disseminação da informação* e das tecnologias da comunicação e interconectividade global têm um grande potencial para acelerar o progresso humano, para eliminar o fosso digital e para o desenvolvimento de sociedades do conhecimento (UNITED NATIONS, 2015, grifo nosso).

Seu Objetivo de Desenvolvimento Sustentável 16 reforça a necessidade de assegurar o *acesso público à informação* e proteger as liberdades fundamentais, em conformidade com a legislação nacional e os acordos internacionais (UNITED NATIONS, 2015, grifo nosso).

Além disso, o item 18, do Objetivo de Desenvolvimento Sustentável 17, que trata dos dados, monitoramento e prestação de contas, reforça a necessidade de disponibilidade de dados de *alta qualidade, atuais e confiáveis, desagregados* por renda, gênero, idade, raça, etnia, status migratório, deficiência, localização geográfica e outras características relevantes em contextos nacionais (UNITED NATIONS, 2015, grifo nosso).

Soma-se a tais intenções a prevalência dos parlamentos nacionais no auxílio para alcançar os Objetivos de

Desenvolvimento Sustentável, que está estampada no item 45 da Introdução da Agenda 2030, que reconhece o *papel essencial dos parlamentos nacionais*, através da promulgação de legislação e adoção de orçamentos, bem como seu papel na garantia da responsabilização para a implementação efetiva dos compromissos internacionais (UNITED NATIONS, 2015, grifo nosso).

Seguir esse documento norteador das Nações Unidas se deve ao fato do Brasil ser um dos 193 Estados-membros signatários da Agenda 2030 e ter se comprometido a tomar medidas ousadas e transformadoras para promover o desenvolvimento sustentável nos próximos anos. Essa agenda global é um plano de ação para as pessoas, o planeta e a prosperidade, que busca fortalecer a paz universal (UNITED NATIONS, 2015).

Nessa conjectura, questiona-se: (1) Como se caracterizam os dados legislativos existentes na Câmara dos Deputados na disponibilização dos seus dados legislativos? (2) Como os dados legislativos são identificados, apresentados, classificados e disponibilizados? (3) De que forma a disponibilização dos dados legislativos pode contribuir para os Objetivos de Desenvolvimento Sustentável da Agenda 2030?

Para tal, o objetivo da pesquisa é analisar e classificar os dados legislativos encontrados no Portal da Câmara dos Deputados, para verificar em que medida os Objetivos de Desenvolvimento Sustentável da Agenda 2030 são contemplados na sua disponibilização.

A pesquisa se justifica pela necessidade que o governo brasileiro tem de apresentar resultados para as Nações Unidas em prol dos Objetivos de Desenvolvimento Sustentável da Agenda 2030.

Para tal, utiliza-se, quanto aos procedimentos, o método bibliográfico, para compreender as dimensões aplicadas aos dados legislativos e suas expressões de conhecimento. Complementa-se com a pesquisa documental, para ordenar as orientações que abrangem a análise através do uso de uma taxonomia com vistas à Agenda 2030. No que tange aos objetivos, adota-se a investigação exploratória, que abrange a literatura e o Portal da Câmara dos Deputados, e descriptiva, pela caracterização dos dados legislativos. Intenta-se, ainda,

uma abordagem qualitativa que revele a representação e organização desses dados.

A Ciência da Informação se notabiliza nesta pesquisa por oferecer atributos que possibilitam a Representação e Organização do Conhecimento dos dados de qualquer área, a partir da identificação, padronização e classificação.

2 Taxonomia e padronização

A taxonomia é a forma mais antiga usada para a identificação, classificação e organização. Surgiu com a classificação biológica e, com o tempo, se tornou muito empregada na Biologia e na Lógica (CAPELLARI, 2008).

No contexto da atualidade, as taxonomias são estruturas classificatórias que têm por finalidade servir de recurso para a representação e a organização da informação. No campo da Ciência da Informação, segundo Maculan e Aganette (2018, p.1), “Vickery (1960) foi o primeiro a definir uma taxonomia como um instrumento para organizar o conjunto de entidades de um domínio”.

Dessa forma, nas ciências da documentação e da informação, os táxons, que são as mínimas estruturas, seriam os dados ou documentos e têm disposição hierárquica e horizontal. É possível, por meio de uma taxonomia, criar categorias dentro da classificação, a qual estabelece relações de semelhanças ou de interdependência. (CURRÁS, 2010).

Por conseguinte, Gomes (2014, p. 5) elucida que:

[...] a taxonomia no ambiente digital agrega informação e conhecimento. Seu uso em tal ambiente é parte da infraestrutura de conhecimento e de informação e resgata a importância da classificação como meio de organização de conhecimento [...].

Além de que, a construção das estruturas das linguagens documentárias, no que concerne à qualidade de um determinado mecanismo de busca, é absolutamente fundamental, bem como é salutar, também, que sejam observados os níveis de eficiência e eficácia ao longo do processo de busca e recuperação da informação desse mecanismo.

Segundo Vital e Café (2011, p. 125):

Taxonomias estão voltadas para a organização das informações em ambientes específicos, visando à recuperação eficaz e, para isso, estabelecendo parâmetros em todo o ciclo de produção da informação, onde profissionais distribuídos por espaços físicos distintos participam do processo de criação do conhecimento de forma organizada.

A partir desses entendimentos, é preciso pensar, também, na padronização. Uma mesma instituição, entidade, ou organismo, que dispõe de diversos serviços de informação, por setores, por exemplo, deve padronizar

seus mecanismos de busca, a fim de evitar intercorrências que possam afetar a qualidade da experiência de seus/suas usuários/as.

Gomes (2014, p. 8) aborda o uso de taxonomias para o desenvolvimento dos mecanismos de busca de informação, e elenca algumas questões que devem ser observadas neste processo:

[...] se serviços essenciais de um governo organizam diferentes taxonomias sob diferentes aspectos, isso pode acarretar, entre outros: impossibilidade de usar um dado já existente em outra organização, dificuldade de intercâmbio de informação, falta de acesso a outros sistemas ou a partes de outros sistemas, incompatibilidade de sistemas de informação, dificuldade de estabelecer uma linguagem comum.

Dessa forma, as taxonomias precisam acompanhar a natureza dinâmica e as variações linguísticas que ocorrem no fluxo de informações das instituições, e, conseqüentemente, estar em constante atualização e padronização nos diversos setores.

De acordo com Campos e Gomes (2007, p.5):

As taxonomias, como estruturas classificatórias, representam os propósitos de organização intelectual de um dado contexto. Neste sentido, são diferentes, dependendo do tipo de organização e de informações que pretendem representar.

Nessa linha de raciocínio, para uma navegação em que os termos se apresentam de forma estruturada e coerente, o uso de taxonomias deve ser adotado. Dessa forma, a classificação possibilita o acesso através de classes e subclasses, atendendo a quantidade de níveis necessária para maior especificidade dos assuntos que agreguem informações aos documentos de dados de determinada base.

Para Campos e Gomes (2007, p.9) alguns critérios gerais devem ser adotados na elaboração de taxonomias:

Comunicabilidade - o termo empregado deve representar a linguagem utilizada pelo usuário;

Utilidade - o nível de especificidade dos termos deve expressar um agrupamento de documentos e não um único documento, ou seja, o termo só é útil quando representativo para um conjunto de documento;

Estimulação - uso de termos que induzem o usuário a continuar a navegação pelo sistema;

Compatibilidade - os termos empregados representem o campo que se está ordenando, fazendo parte das atividades e funções da organização.

Para Gilchrist e Kibby (2000), a taxonomia é, em sentido amplo, a criação da estrutura (ordem) e de rótulos (nomes) que ajudam a localizar informações relevantes. Num sentido mais específico, é a ordenação e a rotulagem dos dados, que permitem o gerenciamento sistemático da informação primária.

Diante dessas considerações, elege-se um modelo de taxonomia para reflexões.

2.1 Um modelo-base de taxonomia para dados

Os pesquisadores Luana Sales e Fernando Sayão (2019, 2020) apresentam uma proposta de taxonomia para os dados de pesquisa que permite reflexões para se pensar os dados legislativos.

A partir de uma lacuna na caracterização dos dados de pesquisa que representa a ausência de um instrumento terminológico e classificatório que fundamente a gestão e a curadoria desses dados, Sales e Sayão (2019), sob a abordagem teórico-metodológica da Organização do Conhecimento, propõem uma taxonomia para a classificação de tipos dados de pesquisa com o fim de ajudar as audiências, que os utilizam, a gerenciá-los e curá-los de forma efetiva.

Sayão e Sales (2020) estabelecem uma base terminológica funcional para dados de pesquisa que considera os aspectos lógicos, que abrangem a discussão do que é o objeto para definição de dado de pesquisa, e os aspectos ontológicos, que abordam a relação do objeto com o mundo, no caso, a relação do dado de pesquisa no ambiente científico.

Inspirados na Teoria do Conceito de Dahlberg (1978) e na Teoria de Eugene Wüster (1981), os autores definem o dado de pesquisa como:

[...] todo e qualquer tipo de registro coletado, observado, gerado ou utilizado no âmbito da pesquisa científica, que pode ser interpretado, tratado e aceito como evidência pela comunidade científica e necessário para analisar, validar e produzir resultados de pesquisa (SAYÃO; SALES, 2020, p. 32).

A partir desse conceito de dados de pesquisa, entende-se os dados legislativos como todo e qualquer dado gerado ou usado no ambiente legislativo, tratado ou aceito como evidência pelas diversas audiências que dele possam se utilizar para propor/produzir iniciativas governamentais, políticas públicas e privadas e pesquisas acadêmicas.

Posteriormente, os autores elegem as características relevantes encontradas nos dados de pesquisa em relação à ambiência científica e propõem uma taxonomia. Oito classes são elencadas com subclasses e algumas com mais uma divisão de classes.

Sales e Sayão (2019) apresentam a seguinte taxonomia:

- Quanto à origem dos dados de pesquisa, estes podem ser: (1) dados de pesquisa que estão divididos em três subclasses: (a) observacionais, (b) experimentais e (c) computacionais; (2) dados para pesquisa que também apresentam uma divisão tripartite: (a) governamentais, (b) registros e (c) transacionais;
- Quanto ao grau de processamento, encontram-se: (1) os dados brutos/primários/crus; (2) os dados intermediários/pré-processados/secundários, estes

se dividem em (a) limpos, (b) analisados, (c) anonimizados e (d) derivados; e por fim, (3) os dados finais/processados/terciários, sob a divisão (a) publicados/arquivados e (b) alta densidade;

- Quanto à abordagem, os dados são: (1) qualitativos e (2) quantitativos;
- Quanto à natureza, a seguinte divisão é evidenciada: (1) número, (2) multimídia, (3) software, (4) visualização, (5) textual, (6) artefato e (7) processo. Estas 7 (sete) subclasses se subdividem em outras classes mais específicas;
- Quanto ao tipo de coleta, esta está dividida em: (1) automática e (2) manual;
- Quanto ao nível de sensibilidade, identifica-se os dados (1) não sensíveis/ostensivos, (2) confidenciais, (3) dados pessoais, divididos em: (a) anonimizados e (b) sensíveis;
- Quanto à materialidade, têm-se (1) os dados digitais e (2) os dados físicos, que se particionam em: (a) amostra, (b) espécime e (c) artefato.
- Quanto à perenidade, a partição é em (1) dados canônicos/referenciais e (2) episódicos, que se dividem nas seguintes coleções: (a) individuais, (b) comunitárias e (c) institucionais.

Contudo, a utilização de dados não fica só na ambiência da pesquisa, o uso de dados é bem mais amplo. A era da datificação também é uma realidade para governos e para as esferas públicas. Atualmente os dados são os protagonistas para propostas governamentais, iniciativas parlamentares e decisões de políticas públicas.

Nessa perspectiva, prossegue-se com os aspectos normativos dos conteúdos informacionais que representam e organizam o conhecimento dos dados legislativos na Câmara dos Deputados e que servem como parâmetros para a aplicabilidade da taxonomia proposta por Sales e Sayão (2019) aos dados legislativos, que são apresentados nos resultados.

2.2 Conteúdos informacionais da Câmara dos Deputados

Na estrutura administrativa da Câmara dos Deputados está o Centro de Documentação e Informação, que coordena a gestão da informação, garantindo a qualidade da informação e seu acesso, cuida das atividades de documentação, informação e pesquisa, da arquitetura de informação, da edição de livros e da gestão de relacionamento com o público externo (BRASIL, 2023c).

No concernente à Política de Gestão de Conteúdos Informacionais, da Câmara dos Deputados, esta foi instituída pelo Ato da Mesa nº 46, de 16 de julho de 2012.

Esse documento observava que o modelo a ser implantado, dentre outras características, deveria:

- (1) ter relação orgânica entre os conteúdos informacionais corporativos, por meio da aplicação de instrumentos de classificação e taxonomias corporativas;
- (2) realizar tratamento dos conteúdos informacionais, que englobasse sua indexação, sua classificação taxonômica e sua descrição;
- (3) gerir os instrumentos de classificação e demais taxonomias corporativas;
- (4) participar da revisão e do aperfeiçoamento do Plano de Classificação Funcional, do Tesouro e outras taxonomias corporativas; e,
- (5) gerenciar, quando couber, taxonomias específicas relacionadas a seus processos de trabalho, observando diretrizes, princípios, objetivos e requisitos desta política e de sua regulamentação (BRASIL, 2012).

Em seguida, foi estabelecida a Política de Indexação de Conteúdos Informacionais da Câmara dos Deputados pelo Ato da Mesa nº 80, de 31 de janeiro de 2013, que é regida pelos princípios da transparência, da efetividade, da eficiência, da qualidade, da imparcialidade e da acessibilidade. Seus objetivos são:

- (1) garantir a excelência da indexação de conteúdos informacionais, com vista à recuperação tempestiva de informações relevantes e pertinentes aos processos de trabalho da Câmara dos Deputados e às demandas dos cidadãos;
- (2) definir e implantar parâmetros corporativos de qualidade da indexação; e,
- (3) estabelecer práticas uniformes de indexação para os diferentes tipos documentais (BRASIL, 2013).

O mesmo ato aborda aspectos sobre o Tesouro da Câmara dos Deputados (Tecd), que é o instrumento de linguagem documentária, que reflete e controla a terminologia dos domínios temáticos relevantes para a Câmara dos Deputados e é composto de lista estruturada de termos, além de glossários, siglários, diretórios e produtos de controle terminológico. Seus objetivos são:

- (1) reduzir a ambiguidade no processo de indexação;
- (2) aumentar a precisão da recuperação da informação nos sistemas e repositórios de informação da Câmara dos Deputados (BRASIL, 2013).

As normativas instruem o tratamento dos dados legislativos disponibilizados pela Câmara dos Deputados, que são descritos a seguir.

3 Resultados

A partir das taxonomias propostas por Sales e Sayão (2019), as seguintes características são identificadas nos dados legislativos analisados e classificados na interface Dados Abertos do Portal da Câmara dos Deputados (Quadro 1).

Quadro 1. *Classe e subclasses dos dados legislativos*

<i>Classes</i>	<i>Subclasses</i>
Quanto à origem	dados para pesquisa governamentais e de registros
Quanto ao grau de processamento	dados brutos/primários/crus
Quanto à abordagem	dados quantitativos
Quanto à natureza	dados textuais
Quanto ao tipo de coleta	automática
Quanto ao nível de sensibilidade	não sensíveis/ostensíveis
Quanto à materialidade	dados digitais
Quanto à perenidade	episódicos institucionais

Fonte: Autoria própria (2023).

Acrescente-se que além das classes

e subclasses utilizadas com base nas pesquisas realizadas pelos autores citados, os dados legislativos disponíveis na interface Dados Abertos apresentam outras classificações.

Como exemplo, apresentam-se as seguintes classes:

- Quanto ao formato - os dados legislativos são representados por 5 (cinco) formatos: csv, xlsx, ods, json e xml;
- Quanto à periodicidade – os dados legislativos estão registrados por ano;
- Quanto à atualização - os dados legislativos são atualizados anualmente.
- Quanto aos tipos de conjuntos de dados disponíveis para consulta – os dados legislativos estão divididos em 30 conjuntos de dados. São estes:

1. Despesas pela Cota para Exercício da Atividade Parlamentar;
2. Proposições;
3. Classificação temática das proposições;
4. Autores das Proposições por ano de apresentação;
5. Frentes Parlamentares;
6. Deputados das frentes;
7. Legislaturas;

8. Mesas diretoras;
9. Órgão da Câmara;
10. Deputados membros dos órgãos;
11. Deputados;
12. Ocupações dos Deputados;
13. Profissões dos Deputados;
14. Eventos;
15. Eventos e seus órgãos realizadores;
16. Deputados presentes em cada evento;
17. Requerimentos de realização de eventos;
18. Votações realizadas a cada ano;
19. Orientações de bancadas por votação;
20. Voto de cada parlamentar;
21. Proposição objeto de cada votação;
22. Proposições afetadas por votação;
23. Funcionários da Câmara;
24. Licitações – Dados básicos;
25. Contratos por licitação;
26. Itens licitados;
27. Pedidos de licitações;
28. Propostas por licitação;
29. Tesouro - termos, notas e substituições;
30. Tesouro - categorias e subcategorias dos termos.

4 Conclusão

A Câmara dos Deputados é a casa legislativa federal integrante do Congresso Nacional brasileiro, composta por representantes do povo, eleitos pelo sistema proporcional em cada Estado da Federação, em cada Território e no Distrito Federal (BRASIL, 2023d).

Portanto, entende-se que os dados legislativos são dados gerados no âmbito legislativo, e especificamente nesta pesquisa, disponíveis na interface Dados Abertos do Portal da Câmara dos Deputados, e fazem parte dos dados governamentais brasileiros. São dados considerados públicos e ostensíveis a toda a sociedade.

A disponibilização dos dados legislativos na interface Dados Abertos do Portal da Câmara dos Deputados possibilita a disseminação da informação, o monitoramento e a prestação de contas requeridas pela Agenda 2030.

Os diversos formatos, em que os 30 conjuntos de dados estão representados, permitem amplo acesso e alta

qualidade de uso. Além de que, por serem atualizados anualmente e estarem presentes em um portal governamental, são considerados atuais e confiáveis, de acordo com os Objetivos de Desenvolvimento Sustentável da Agenda 2030.

Em síntese, a Câmara dos Deputados está em consonância com a Agenda 2030 quanto à importância dos parlamentos nacionais colaborarem com seus fundamentos e objetivos, inclusive quanto à transparência de ações, atuações e informações em prol dos Objetivos de Desenvolvimento Sustentável.

Por fim, alerta-se que, quanto à desagregação dos dados intentada pela agenda global, futuros estudos são indicados para análise mais aprimorada dos conjuntos de dados legislativos fornecidos na interface Dados Abertos da Câmara dos Deputados.

Referências

BRASIL. Câmara dos Deputados. **Ato da Mesa nº 46**, de 16 de julho de 2012. Institui a Política de Gestão de Conteúdos Informativos da Câmara dos Deputados e dá outras providências. 2012. Disponível em: <https://www2.camara.leg.br/legin/int/atomes/2012/atodamesa-46-16-julho-2012-773824-publicacaooriginal-137093-cd-mesa.html>. Acesso em: 16 abr. 2023.

BRASIL. Câmara dos Deputados. **Ato da Mesa nº 80**, de 31 de janeiro de 2013. Dispõe sobre a Política de Indexação de Conteúdos Informativos, o Tesouro da Câmara dos Deputados e dá outras providências. 2013. Disponível em: <https://www2.camara.leg.br/legin/int/atomes/2013/atodamesa-80-31-janeiro-2013-775250-publicacaooriginal-138914-cd-mesa.html>. Acesso em: 16 abr. 2023.

BRASIL. Câmara dos Deputados. **Entenda o Processo Legislativo**. 2023a. Disponível em: <https://www.camara.leg.br/entenda-o-processo-legislativo/>. Acesso em: 16 abr. 2023.

BRASIL. Câmara dos Deputados. **Dados Abertos**. 2023b. Disponível em: <https://dadosabertos.camara.leg.br/swagger/api.html#staticfil>. Acesso em: 16 abr. 2023.

BRASIL. Câmara dos Deputados. **Estrutura Organizacional**. 2023c. Disponível em: <https://www2.camara.leg.br/a-camara/estruturaadm/camara-dos-deputados>. Acesso em: 19 abr. 2023.

BRASIL. Congresso Nacional. **Glossário de Termos Legislativos**. 2023d. Disponível em: <https://www.congressonacional.leg.br/legislacao-e-publicacoes/glossario-legislativo>. Acesso em: 19 abr. 2023.

CAMPOS, Maria Luiza Almeida; GOMES, Hagar Espanha. **Taxonomia e classificação**: a categorização como princípio. *In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO*, 8., 2007, Salvador. (Disponível em: <http://hdl.handle.net/20.500.11959/brapci/172684>). Acesso em: 19 abr. 2023.

CAPELLARI, Renato Soares. Breve histórico da taxonomia e da sistemática. **Revista Simbio-Logias**. v.1, n.1, p. 221-222, maio 2008. Disponível em:

https://www.ibb.unesp.br/Home/ensino/departamentos/educacao/breve_historico_pedro.pdf. Acesso em: 19 abr. 2023.

CURRÁS, Emília. **Ontologias, taxonomias e tesouros em teoria de sistemas e sistemática**. Tradução: Jaime Robredo. Brasília, DF: Thesaurus, 2010. 182 p.

DAHLBERG, Ingtraut. A referent-oriented, analytical concept theory for interconcept. **International Classification**, Frankfurt, v. 5, n.3, p.142-150, 1978.

GILCHRIST, Alan; KIBBY, Peter. **Taxonomies for business: access and connectivity in a wired world**. London: TFPL, 2000.

GOMES, Hagar Espanha. Taxonomia e a Web, Construção e Uso. In: GOMES, Hagar Espanha (coord.). **Biblioteconomia, Informação & Tecnologia da Informação**. 2014.

Disponível em:

<http://www.conexaoario.com/bit/taxonomianaweb.pdf>. Acesso em: 16 abr. 2023.

MACULAN, Benildes Coura M. S.; AGANETTE, Elisângela Cristina. A Teoria da Classificação Facetada na Construção de Taxonomias Facetadas. In: SEMINAR ON ONTOLOGY RESEARCH IN BRAZIL, 11, 2018 - ONTOBRAS.

Proceedings [...]. São Paulo, 2018. Disponível em:

<https://ceur-ws.org/Vol-2228/paper2.pdf>. Acesso em: 16 abr. 2023.

SALES, Luana Farias; SAYÃO, Luis Fernando. Uma proposta de taxonomia para dados de pesquisa. **Revista Conhecimento em Ação**, v. 4, n. 1, p. 31-48, 2019.

Disponível em:

<https://revistas.ufrj.br/index.php/rca/article/view/26337/14573>. Acesso em: 19 abr. 2023.

SAYÃO, Luis Fernando; SALES, Luana Farias. Afinal, o que é dado de pesquisa? **BIBLOS - Revista do Instituto de Ciências Humanas e da Informação**, v. 34, n. 2, 2020.

Disponível em:

<https://www.seer.furg.br/biblos/article/view/11875>. Acesso em: 19 abr. 2023.

UNITED NATIONS. General Assembly. **Transforming Our World: The 2030 Agenda for Sustainable Development**. A/70/L.1. Seventieth session. 18 September 2015. Disponível em:

http://www.un.org/ga/search/view_doc.asp?symbol=A/70/L.1&referer=/english/&Lang=E. Acesso em: 19 abr. 2023.

VITAL, Luciane Paula; CAFÉ, Ligia Maria Arruda.

Ontologias e taxonomias: diferenças. **Perspectivas em**

Ciência da Informação, v.16, n.2, p.115-130, abr./jun. 2011.

Disponível em: <https://doi.org/10.1590/S1413-99362011000200008>. Acesso em: 16 abr. 2023.

WÜSTER, Eugen. L'étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, l'informatique et les sciences des chose.

In: RONDEAU, Guy; FELBER, H. (org.). **Textes choisis de terminologie**. Québec: Girssem, 1981. p. 57-114.

Os sistemas de organização do conhecimento nos currículos dos programas de pós-graduação em Ciência da Informação nas Regiões Sul e Sudeste do País

The knowledge organization systems in the curriculum of the graduate programs in Information Science in the Southern and Southeast Regions of the Country

Clarice Luzia Rodrigues Casoni (1), Raimunda Fernanda dos Santos (2),
Maria Elisabeth Baltar Carneiro de Albuquerque (3)

(1) Universidade Estadual de Londrina (UEL), Londrina - PR, clarice.luzia.casoni@uel.br. (2) Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro - RJ/ Universidade Estadual de Londrina (PPGCI/UEL), Londrina-PR, fernanda@facc.ufrj.br. (3) Universidade Federal da Paraíba (UFPB), João Pessoa - PB, ebaltar2007@gmail.com

Resumo

Estuda aspectos relativos aos Sistemas de Organização do Conhecimento (SOC) no contexto da Ciência da Informação. Objetiva analisar os componentes curriculares sobre Sistemas de Organização do Conhecimento nos Programas de Pós-Graduação em Ciência da Informação das regiões Sul e Sudeste do País. Objetiva-se especificamente: mapear as disciplinas voltadas para os Sistemas de Organização do Conhecimento; apresentar um panorama dos conteúdos contemplados nesses componentes curriculares; elencar os principais tipos de SOC estudados; investigar se as novas possibilidades de representação do conhecimento por meio da Folksonomia encontram-se contempladas nas estruturas curriculares analisadas. Utiliza como metodologia as pesquisas bibliográfica, documental, exploratória, descritiva com abordagem quali-quantitativa mediante um estudo de caso nos Programas de Pós-Graduação (*lato sensu* e *stricto sensu*) da área de Ciência da Informação de instituições federais e estaduais localizadas nas regiões Sul e Sudeste do Brasil. Constata que dos 15 (quinze) cursos de Pós-Graduação, apenas oito possuem disciplinas sobre a temática, o que equivale a 53,3% dos cursos ofertados na área citada anteriormente. Identifica que todos os componentes curriculares ofertados são de caráter eletivo e apresentam teorias e questões basilares para o processo de construção de um instrumento de controle terminológico. Taxonomias, tesouros, ontologias, tabelas de classificação, glossários, mapas conceituais, *topic maps*, vocabulários controlados e Folksonomia foram os tipos de Sistemas de Organização do Conhecimento citados nas ementas dos componentes curriculares analisados. Conclui-se que apenas o PPGCI IBICT/UFRJ e o PPGCI/UEL registraram em suas disciplinas a moderna concepção da participação das comunidades virtuais colaborativas no processo de representação do conhecimento.

Palavras-chave: Sistemas de Organização do Conhecimento; Programas de Pós-Graduação; Ciência da Informação.

Abstract

Study aspects related to Knowledge Organization Systems (SOC) in the context of Information Science. Analyze the curricular components on Knowledge Organization Systems in Graduate Programs in Information Science in the South and Southeast regions of the country. Specifically aims to: map the disciplines focused on Knowledge Organization Systems; present an overview of the content covered in these curricular components; list the main types of SOC studied; investigate whether the new possibilities of knowledge representation through Folksonomy are contemplated in the remaining curricular structures. Uses bibliographical, documental, exploratory, descriptive research with a qualitative and quantitative approach as a methodology through a case study in Postgraduate Programs (*lato sensu* and *stricto sensu*) in the area of Information Science of federal institutions and located in the South and Southeastern Brazil. Note that of the 15 (fifteen) Postgraduate courses, only eight have disciplines on the subject, which is equivalent to 53.3% of the courses offered in the aforementioned area. Identifies that all curricular components offered are elective and present theories and basic questions for the construction process of a terminological control instrument. Taxonomies, thesauri, ontologies, classification tables, glossaries, conceptual maps, topic maps, controlled controlled and Folksonomy were the types of Knowledge Organization Systems cited in the menus of the curricular components analyzed. Concluded that only the PPGCI IBICT/UFRJ and the PPGCI/UEL registered in their disciplines the modern conception of the participation of collaborative virtual communities in the knowledge representation process.

Keywords: Knowledge Organization Systems; Graduate Programs; Information Science.

Resumen

Estudia aspectos relacionados con los Sistemas de Organización del Conocimiento (SOC) en el contexto de las Ciencias de la Información. Tiene como objetivo analizar los componentes curriculares sobre Sistemas de Organización del Conocimiento en Programas de Posgrado en Ciencias de la Información en las regiones Sur y Sudeste del país. Específicamente tiene como objetivo: mapear las disciplinas enfocadas en los Sistemas de Organización del Conocimiento; presentar una visión general del contenido cubierto en estos componentes curriculares; enumerar los principales tipos de SOC estudiados; investigar si las nuevas posibilidades de representación del conocimiento a través de la Folksonomía están contempladas en las estructuras curriculares analizadas. Utiliza como metodología la investigación bibliográfica, documental, exploratoria, descriptiva con enfoque cualitativo y cuantitativo a través de un estudio de caso en los Programas de Posgrado (lato sensu y stricto sensu) en el área de Ciencias de la Información de instituciones federales y estatales ubicadas en el Regiones del Sur y Sudeste de Brasil. Señala que de los 15 (quince) cursos de Posgrado, sólo ocho tienen disciplinas sobre el tema, lo que equivale al 53,3% de los cursos ofrecidos en la referida área. Identifica que todos los componentes curriculares que se ofrecen son optativos y presentan teorías y cuestiones básicas para el proceso de construcción de un instrumento de control terminológico. Taxonomías, tesauros, ontologías, tablas de clasificación, glosarios, mapas conceptuales, mapas temáticos, vocabularios controlados y Folksonomía fueron los tipos de Sistemas de Organización del Conocimiento citados en los programas de los componentes curriculares analizados. Se concluye que sólo el PPGCI IBICT/UFRJ y el PPGCI/UEL registraron en sus disciplinas la concepción moderna de la participación de comunidades virtuales colaborativas en el proceso de representación del conocimiento.

Palabras clave: Sistemas de Organización del Conocimiento; Programas de postgrado; Ciencias de la Información.

1 Introdução

A representação se configura como uma operação presente nas práticas sociais e em diferentes ambientes de construção e transfiguração da sociedade. As relações entre a humanidade e o conhecimento evidenciam a necessidade das pessoas em registrar, representar e organizar o conhecimento com vistas à recuperação, preservação, acesso, uso e apropriação da informação.

A dimensão multifacetada da representação desafia aqueles que se propõem a estudá-la no contexto da Ciência da Informação, em especial na Organização e Representação da Informação e do Conhecimento, sob as perspectivas epistemológica e pragmática acerca do acesso às informações a partir dos diferentes dispositivos, plataformas e contextos. Nesse entendimento, a informação e o conhecimento ancorados no desenvolvimento tecnológico tornam-se elementos imprescindíveis para as organizações e instituições, contribuindo para os processos, produtos, instrumentos, ferramentas e serviços oferecidos à sociedade.

Para Dahlberg (1978), a Organização do Conhecimento é a ciência em que a condição *sine qua non* consiste na sistematização de conceitos (unidades do conhecimento) e a sua estruturação em Sistemas de Organização do Conhecimento (SOC). De acordo com Carlan (2010) tais sistemas são estruturados por meio da representação do conhecimento, a qual contempla métodos aplicados ao tratamento de documentos a partir de um esquema estruturado sistematicamente cujo foco é a representação do conteúdo com a finalidade principal de organizar a informação e o conhecimento e, posteriormente, viabilizar a recuperação da informação.

Portanto, a Organização do Conhecimento contempla abordagens teóricas e práticas com vistas a análise e a sistematização de conceitos em diferentes tipos de Sistemas de Organização do Conhecimento, visando a

“construção de modelos de mundo que se constituem em abstrações da realidade” (BRASCHER; CAFÉ, 2008, p. 6).

Embora os Sistemas de Organização do Conhecimento sejam concebidos, estruturados e baseados em teorias e metodologias clássicas instituídas há séculos na área de Organização e Representação do Conhecimento, torna-se imprescindível abordar questões contemporâneas considerando as novas metodologias, tecnologias e possibilidades de representação do conhecimento alinhadas às demandas da sociedade para a formação de profissionais e pesquisadores qualificados.

Considerando que a Ciência da Informação contempla, dentre outras questões, investigações sobre os processos, instrumentos e produtos relacionados à Organização e Representação do Conhecimento, surge o seguinte questionamento: De que maneira os Sistemas de Organização do Conhecimento têm sido estudados no contexto da Pós-Graduação em Ciência da Informação *lato sensu e stricto sensu*?

Visando responder o questionamento apresentado anteriormente, esta pesquisa tem como objetivo geral: analisar os componentes curriculares sobre Sistemas de Organização do Conhecimento nos Programas de Pós-Graduação em Ciência da Informação das regiões Sul e Sudeste do país. Objetiva-se especificamente: a) mapear as disciplinas voltadas para os Sistemas de Organização do Conhecimento; b) apresentar um panorama dos conteúdos contemplados nesses componentes curriculares; c) elencar os principais tipos de SOC estudados; d) investigar se as novas possibilidades de representação do conhecimento por meio da Folksonomia encontram-se contempladas nas estruturas curriculares analisadas.

De acordo com Torres, Almeida e Simões (2017) os Sistemas de Organização do Conhecimento têm despertado o interesse de pesquisadores provenientes de

diferentes áreas do conhecimento, dentre elas: a Ciência da Informação, a Ciência da Computação, a Filosofia, a Comunicação, por exemplo. Com isso, torna-se importante contribuir para a formação qualificada de pesquisadores e profissionais da informação para atuar no processo de análise, modelagem, construção, atualização e implementação de SOC em diferentes contextos.

Diante disso, a relevância deste estudo decorre, em linhas gerais, da necessidade de investigar se os currículos dos Programas de Pós-Graduação em Ciência da Informação nas modalidades acadêmica e profissional contemplam questões epistemológicas e/ou pragmáticas relacionadas aos conceitos, teorias, métodos, padrões e ferramentas de construção e atualização de instrumentos de representação do conhecimento em ambientes tradicionais e contemporâneos.

Além disso, acredita-se que esta pesquisa contribui para a Organização do Conhecimento, sobretudo no que diz respeito aos estudos sobre SOC, tendo em vista que se busca conhecer como se dá a formação de profissionais e pesquisadores no campo.

Nesta pesquisa foi estabelecido um recorte das regiões Sul e Sudeste do país devido à sua relação com a localização geográfica do VI e VII Congresso Brasileiro em Organização do Conhecimento (ISKO Brasil 2021 e 2023), além de que as análises a serem apresentadas se configuram como parte dos resultados da Dissertação de uma das autoras.

Para dar seguimento a essas considerações, a seguir são discutidas questões relativas aos Sistemas de Organização do Conhecimento.

2 Sistemas de Organização do Conhecimento

Com o desenvolvimento das tecnologias de informação e comunicação aplicadas à organização e representação da informação em contextos tradicionais e contemporâneos, emergiram outros tipos de instrumentos de representação do conhecimento complementares aos sistemas de classificação visando atender às demandas de organização em diferentes domínios do conhecimento, dentre esses instrumentos é possível citar: as taxonomias, os cabeçalhos de assunto, as ontologias e os mapas conceituais. Esses instrumentos são denominados na literatura da área de Organização do Conhecimento como *Knowledge Organization Systems* (KOS) ou Sistemas de Organização do Conhecimento, em português, e tem como foco tratar o conteúdo dos objetos informacionais para fins de sua recuperação, acesso e uso.

Para Brascher e Carlan (2010) os Sistemas de Organização do Conhecimento (SOC) se configuram como objetos de interesse particular para a Ciência da infor-

mação, considerando a sua contribuição para a organização do conhecimento e a recuperação da informação.

Os SOC são instrumentos que realizam a tradução dos conteúdos dos objetos informacionais para um esquema conceitual estruturado sistematicamente, tendo como principais objetivos: delimitar o uso de termos; modelar conceitos e relações semânticas em domínios específicos; facilitar a recuperação da informação.

Os conceitos são delineados de acordo com o domínio em que estão inseridos e são, portanto, os componentes essenciais de um instrumento de representação do conhecimento. No que concerne às relações semânticas entre os conceitos, a ISO 25964-1 aponta variados tipos, dentre elas estão as relações de equivalência (entre termos sinônimos); hierárquicas (por meio dos níveis de superordenação e subordinação) e associativas (com termos que possuem algum tipo de ligação que não de superordenação e de subordinação). Esses três tipos de relações são considerados como os tipos básicos utilizados na estruturação de um SOC (LIMA; MACULAN, 2017).

De acordo com Carlan (2010) esses instrumentos foram originados em padrões que visam a interoperabilidade entre sistemas e usuários, contribuindo para minimizar o hiato semântico no que concerne à comunicação entre essas duas esferas.

Para a construção de um SOC consistente faz-se necessária a realização de um percurso teórico-metodológico fundamentado em princípios basilares como as teorias da classificação, a teoria do conceito, princípios da linguística documentária e da terminologia, bem como estudos acerca das relações semânticas entre conceitos.

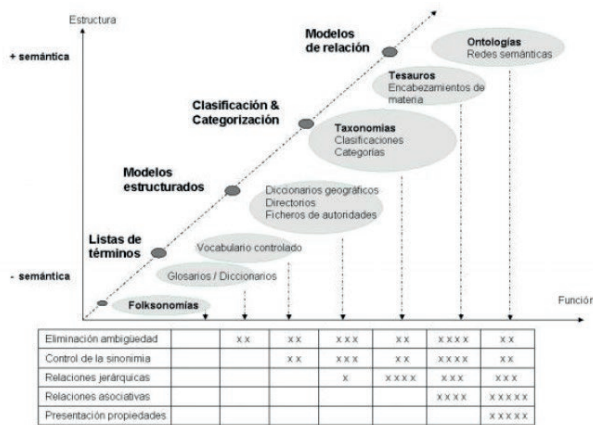
Zeng (2008) apresenta uma lista de Sistemas de Organização do Conhecimento por meio das suas tipologias, a saber: lista de termos (dicionários, glossários e anêis de sinônimos, por exemplo); modelos semelhantes a metadados (arquivos de autoridade, diretórios, *gazetteers*); classificação e categorização (cabeçalhos de assunto, esquemas de categorização, taxonomias e esquemas de classificação, por exemplo); modelos de relações (tesauros, redes semânticas, ontologias).

Soler Monreal e Gil Leiva (2010) ilustram algumas dessas tipologias, partindo das mais simples até as mais complexas, por meio da linha progressiva, em que as Folksonomias se configuram como SOC com baixo grau de formalização semântica e as ontologias com maior grau nesse sentido.

O uso de instrumento de controle terminológico para o processo de tradução dos conteúdos dos documentos em termos se configura como uma atividade importante para a recuperação da informação. Todavia, faz-se necessário atentar para os desafios que surgem em relação a proeminência de discursos hegemônicos e regulatórios instituídos nessas metalinguagens, bem como se

atentar às dinâmicas colaborativas de representação do conhecimento em ambientes digitais, necessitando assim que novos processos, metodologias e instrumentos sejam desenvolvidos para essa finalidade.

Figura 1. Estruturas e estratégias para a Organização do Conhecimento



Fonte: Soler Monreal e Gil Leiva (2010, p.634).

Concorda-se com Moura (2018, p. 12) quando a autora afirma:

A constatação de que os dispositivos de representação da informação podem viabilizar a circulação e a proeminência de discursos hegemônicos, autoritários e regulatórios trouxe inquietações ao campo da organização do conhecimento. Estas inquietações relacionam-se especialmente com as lógicas estruturantes e as possíveis assimetrias dispostas como discurso organizador ofertado no SOC.

Seguindo essa linha de raciocínio, as pesquisadoras Santos e Albuquerque (2021) apresentam em sua pesquisa, instrumentos cujo foco é agregar as potencialidades da Folksonomia e de outros Sistemas de Organização do Conhecimento (ontologia, taxonomia, tesouro, por exemplo), objetivando a produção significativa de metadados semânticos para fins de representação e recuperação de conteúdos na *web*. Esses instrumentos são denominados pelas autoras de Sistemas híbridos de Organização do Conhecimento e visam explorar as potencialidades da Folksonomia com o suporte dos formalismos e métodos utilizados na construção de instrumentos de controle terminológico, gerando novas perspectivas para auxiliar a representação e o compartilhamento de conhecimentos em ambiente *web*.

Diante do exposto, é importante enfatizar que os aspectos epistemológicos e pragmáticos acerca dos Sistemas de Organização do Conhecimento, bem como os seus diferentes tipos e níveis de complexidade precisam ser trabalhados de maneira verticalizada na formação de futuros pesquisadores e profissionais da informação. Esta pesquisa tem como foco investigar esses e outros aspectos relacionados à formação desses sujeitos no campo da Ciência da Informação.

Para dar seguimento a essas considerações, a seguir são apresentados os procedimentos metodológicos empreendidos para a realização desta pesquisa.

3 Procedimentos metodológicos

Visando atender o seu objetivo geral, foram realizadas as pesquisas bibliográfica, documental, exploratória e descritiva com abordagem quali-quantitativa para a recuperação, análise e mapeamento dos componentes curriculares sobre Sistemas de Organização do Conhecimento nos currículos dos cursos dos Programas de Pós-Graduação em Ciência da Informação das Regiões Sul e Sudeste do país.

No que diz respeito à pesquisa bibliográfica, foram realizadas buscas na Base de Dados em Ciência da Informação (BRAPCI), Portal de Periódicos CAPES e Biblioteca Digital Brasileira de Teses e Dissertações em março de 2021 aplicando os seguintes termos de busca: “Organização do Conhecimento”, “Representação do Conhecimento” e “Sistemas de Organização do Conhecimento”. Essas fontes de informação foram selecionadas nesta pesquisa devido à possibilidade de acesso aos artigos e produções científicas publicadas nos principais periódicos, eventos científicos nacionais e repositórios da área de Ciência da Informação do Brasil.

Para a recuperação das produções científicas sobre o tema não foi delimitado o recorte temporal, no intuito de recuperar todas as pesquisas publicadas acerca da temática. Posteriormente, foram consideradas apenas aqueles trabalhos que possuíam relação direta com o enfoque temático desta pesquisa.

Além disso, foi realizada uma pesquisa documental na Plataforma Sucupira, sistema de coleta de informações, análises e avaliações utilizadas como base padronizadora do Sistema Nacional de Pós-Graduação brasileira (SUCUPIRA, 2023). As buscas foram realizadas nessa plataforma em março de 2023 visando identificar os Programas de Pós-Graduação na área de Ciência da Informação (nas modalidades profissional e acadêmica) cujos cursos são avaliados e reconhecidos de Universidades Federais e Estaduais das Regiões Sul e Sudeste do país.

Na primeira etapa da busca na plataforma Sucupira foram selecionados os cursos avaliados e reconhecidos da área de avaliação “Comunicação e Informação”, especificando a Ciência da Informação como área do conhecimento vinculada, para tanto foram recuperados: Mestrados acadêmicos, Mestrados profissionais e Doutorado acadêmico. Não foi identificado nenhum curso de Doutorado na área na modalidade profissional.

Com a busca realizada anteriormente foram recuperados 40 (quarenta) cursos de Pós-Graduação distribuí-

dos em todas as Instituições de Ensino Superior do País (públicas e particulares). Desse total, foram selecionados apenas os cursos vinculados às Instituições de Ensino Superior Estaduais e Federais das Regiões Sul e Sudeste do País, totalizando 15 (quinze) cursos que estão distribuídos em 12 (doze) instituições de Ensino Superior do país. Esse último número se deve ao fato de que a Universidade de São Paulo (USP), a Universidade Federal de Minas Gerais (UFMG) e a Universidade Federal do Estado do Rio de Janeiro (UNIRIO) possuem mais de um curso de Mestrado (acadêmico e/ou profissional) ou mais de um Doutorado acadêmico vinculado. A UFMG, por exemplo, possui dois mestrados acadêmicos e dois doutorados acadêmicos sendo distribuídos em dois programas com nomes distintos na mesma instituição.

Após esse procedimento, foram realizadas buscas nos sites oficiais dos Programas de Pós-Graduação em Ciência da Informação em questão com o objetivo de analisar as suas respectivas grades curriculares e verificar se são ofertados componentes curriculares obrigatórios ou eletivos sobre os Sistemas de Organização do Conhecimento. Tendo em vista que na área de Organização do Conhecimento os SOC também são denominados de linguagens documentárias ou instrumentos de representação do conhecimento, foi realizada uma busca nas grades curriculares tentando identificar disciplinas sobre “Sistemas de Organização do Conhecimento”, “linguagens documentárias”, “instrumentos de representação do conhecimento”, além da busca pelo nome de SOC como “tesauros”, “ontologias”, “taxonomias”, “mapas conceituais”, “Folksonomia” e demais tipos listados no referencial teórico desta pesquisa.

O Quadro 01 (Apêndice A) apresenta as instituições, os nomes e a modalidade dos oito cursos de Pós-Graduação que se constituíram como *corpus* final da presente investigação.

Como procedimento para a realização da análise dessas pesquisas foi utilizada a análise de conteúdo por meio de uma leitura das ementas das disciplinas ofertadas nos Programas supracitados, possibilitando tornar compreensíveis os principais aspectos enfatizados acerca dos Sistemas de Organização do Conhecimento, bem como o caráter da disciplina (obrigatória ou eletiva) e a sua carga horária.

A partir dessa lógica, a seguir são apresentados os principais resultados da pesquisa.

4 Análise e discussão dos resultados

Após a análise das grades curriculares e das ementas das disciplinas propostas, observou-se que dos 15 (quinze) cursos de Pós-Graduação na área de Ciência da Informação de Universidades Federais e Estaduais

de Ensino localizadas nas regiões Sul e Sudeste do país, apenas oito possuem disciplinas sobre Sistemas de Organização do Conhecimento, o que equivale a 53,3% dos cursos ofertados na área citada anteriormente.

Todos os cursos que possuem disciplinas sobre os Sistemas de Organização do Conhecimento estão cadastrados na plataforma Sucupira na modalidade acadêmica, ou seja, não foram identificadas disciplinas ofertadas em programas cujos cursos são da modalidade profissional.

Embora os cursos de mestrado profissional na área de Ciência da Informação abordem temáticas relacionadas à Organização e Representação da Informação e do Conhecimento, proporcionando significativas contribuições para a formação profissional e científica de seus egressos, não foi notada a presença de disciplinas sobre Sistemas de Organização do Conhecimento. Observa-se, portanto, a necessidade de se promover no âmbito da pós-graduação *lato sensu* das regiões Sul e Sudeste do país, um conjunto de reflexões e ações que priorizem os SOC enquanto metalinguagens que instituem discursos e práticas nos mais diversificados contextos institucionais.

No total, foram analisadas ementas dos cursos oferecidos por sete instituições de ensino, sendo cinco federais e duas estaduais, são elas: Universidade Federal de Santa Catarina (UFSC), Universidade Federal do Rio de Janeiro (UFRJ), Universidade Federal Fluminense (UFF), Universidade Federal de Minas Gerais (UFMG), Universidade Federal do Rio Grande do Sul (FURG), Universidade Estadual de Londrina (UEL) e Universidade Estadual Paulista (UNESP). Diante disso, observa-se que os cursos estão predominantemente localizados na Região Sudeste do país.

No que concerne aos componentes curriculares, observou-se que os mesmos apresentam diferentes denominações, como por exemplo: instrumentos terminológicos de representação e recuperação da informação e do conhecimento, sistemas de organização do conhecimento, linguagem documentária. Além disso, também foram identificados componentes curriculares cujo foco está voltado para tipos específicos de Sistemas de Organização do Conhecimento, dentre eles: Folksonomia, taxonomia e ontologia. O Quadro a seguir apresenta de maneira detalhada os nomes dos componentes curriculares identificados:

Diante do exposto, foi possível perceber que todos os componentes curriculares ofertados que possuem relação com a temática deste estudo são de caráter eletivo e a sua carga horária é diversificada, variando entre 30 (trinta), 60 (sessenta) e 90 (noventa) horas.

Apesar de o Programa de Pós-Graduação em Gestão & Organização do Conhecimento da Universidade Federal de Minas Gerais (PPGGOC/UFMG) ter como foco

a Organização do Conhecimento, não foram identificadas disciplinas sobre SOC como componentes curriculares obrigatórios integrantes do seu currículo.

Quadro 2. Componentes curriculares identificados na pesquisa

PROGRAMA	DISCIPLINA	MODALIDADE E CARGA HORÁRIA
PPGCI/UFF	Instrumentos terminológicos de representação e recuperação da informação e do conhecimento	ELETIVA (60H)
PPGCI IBICT/UFRJ	Sistemas de Organização do Conhecimento	ELETIVA (60H)
PPGCI/UFMG	Linguagem documentária como instrumento de comunicação da informação	ELETIVA (60H)
PPG-GOC/UFMG	Sistemas de Organização do Conhecimento	ELETIVA (60H)
PGCIN/UFSC	Sistemas de Organização do Conhecimento	ELETIVA (30H)
PPG-CIN/UFRGS	Sistemas de Organização do Conhecimento aplicados a arquivos e bibliotecas	ELETIVA (60H)
	Terminologia, taxonomia e ontologia na Ciência da Informação	ELETIVA (60H)
PPGCI/UNESP	Sistemas de Organização do Conhecimento: aspectos teóricos e metodológicos	ELETIVA (90H)
PPGCI/UDEL	Folksonomia e representação colaborativa da informação	ELETIVA (30H)

Fonte: Elaborado pelas autoras.

Além disso, verificou-se que apenas um programa oferece duas disciplinas sobre a temática, nos demais foi identificada a oferta apenas de um componente curricular.

Conforme destacado no referencial teórico desta pesquisa, para a construção de um SOC consistente é necessária a utilização de princípios teóricos e metodológicos. Nessa perspectiva, constatou-se que as ementas das disciplinas analisadas apresentam teorias e questões basilares para o processo de construção de um instrumento de controle terminológico, são elas: teoria da classificação e a abordagem facetada, teoria do conceito, análise de domínio e terminologia. Esses e outros aspectos podem ser visualizados nas ementas das disciplinas disponíveis no Quadro 03 (Apêndice B).

Taxonomias, tesouros, ontologias, tabelas de classificação, glossários, mapas conceituais, *topic maps*, vocabulários controlados e Folksonomia foram os tipos de Sistemas de Organização do Conhecimento citados nas ementas dos componentes curriculares analisados.

Dentre os Programas de Pós-Graduação analisados, constatou-se que apenas o PPGCI IBICT/UFRJ e o PPGCI/UDEL registraram na ementa de suas disciplinas a moderna concepção da participação das comunidades virtuais colaborativas no processo de representação do conhecimento.

Em linhas gerais, foi possível observar que as disciplinas tratam sobre os Sistemas de Organização de maneira genérica, apresentando princípios teóricos e metodológicos para a sua construção e utilização. Embora duas disciplinas apresentem no título tipos específicos de SOC, como é o caso da “Terminologia, taxonomia e ontologia na Ciência da Informação” (oferecida pelo PPGCIN/UFRGS) e “Folksonomia e representação colaborativa da informação” (ofertada pelo PPGCI/UDEL), elas também contemplam diferentes tipos de SOC no conteúdo programático e na sua ementa.

Além disso, verificou-se que a disciplina ofertada pelo PPGCI/UDEL apresenta enfoque inovador no que diz respeito à inclusão de novas possibilidades de representação por meio da Folksonomia e dos Sistemas híbridos de Organização do Conhecimento. Esses últimos se configuram como instrumentos que integram as potencialidades da Folksonomia com o suporte dos formalismos e métodos dos instrumentos de controle terminológico para a representação de recursos na *web*.

Por fim, constatou-se que 75% dos Programas de Pós-Graduação em Ciência da Informação das regiões Sul e Sudeste do Brasil registraram em suas ementas que para além das bases conceituais, teóricas e metodológicas também são contempladas as abordagens práticas relacionadas aos SOC, como é o caso do PPGCI/UFF, PPGCI IBICT/UFRJ, PPGGOC/UFMG, PPGCI/UFRGS, PPGCI UEL e PPGCI/UNESP.

É importante ressaltar que as ementas se configuram como documentos que destacam os pontos essenciais sobre determinado(s) assunto(s) a serem apresentados em cada componente curricular de um curso. Dessa forma, as questões práticas que envolvem os SOC podem ou não ser contempladas de maneira informal pelos docentes que ministram as disciplinas em questão.

5 Considerações finais

Os Sistemas de Organização do Conhecimento se configuram como metalinguagens baseadas em princípios teóricos e metodológicos, visando instituir discursos e práticas em diferentes cenários informacionais.

Em linhas gerais, foi possível constatar que os SOC têm sido discutidos pela maioria (53,3%) dos cursos de Pós-Graduação em Ciência da Informação das Universidades Federais e Estaduais localizadas nas regiões Sul e Sudeste do país. Porém, identificou-se que apenas os cursos na modalidade acadêmica contemplam disciplinas exclusivamente dedicadas ao assunto e em caráter eletivo. Tal fato evidencia a necessidade de se promover no âmbito da Pós-Graduação *stricto e lato sensu* um conjunto de abordagens epistemológicas e pragmáticas que priorizem os SOC enquanto instrumentos de representação do conhecimento, sobretudo no contexto dos ambientes colaborativos, das tecnologias semânticas, da sociedade em rede e dos dados abertos interligados.

Dentre os Programas de Pós-Graduação analisados, constatou-se que apenas dois deles (PPGCI IBICT/UFRJ e o PPGCI/UEL) registraram na ementa de suas disciplinas a moderna concepção da participação das comunidades virtuais colaborativas no processo de representação do conhecimento, aspectos emergentes e necessários a serem discutidos no âmbito da Ciência da Informação brasileira.

Por fim, considera-se importante atentar-se para a formação de futuros profissionais e pesquisadores no que concerne aos aspectos epistemológicos e pragmáticos que envolvem a criação, a atualização e a implementação de Sistemas híbridos de Organização do Conhecimento, considerando as potencialidades da Folksonomia com o suporte dos formalismos e métodos dos instrumentos de controle terminológico para representar recursos em ambientes web.

Referências

BRASCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento? *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB*, 9., 2008, São Paulo. *Anais [...]* São Paulo: ECA/USP, 2008.

BRÄSCHER, M.; CARLAN, E. Sistemas de organização do conhecimento: antigas e novas linguagens. *In: ROBREDO,*

J.; BRÄSCHER, M. (Orgs.). **Passaios no Bosque da Informação**: estudos sobre Representação e Organização da Informação e do Conhecimento. Brasília DF: IBICT, 2010, Cap. 8, p. 147-176.

CARLAN, E. **Sistemas de Organização do Conhecimento**: uma reflexão no contexto da Ciência da Informação. 2010. Dissertação (Mestrado em Ciência da Informação). Universidade de Brasília, 2010.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Brasília, v. 7, n. 2, 1978. Disponível em: <https://revista.ibict.br/ciinf/article/view/115>. Acesso em: 2 dez. 2022.

LIMA, G. A.; MACULAN, B. C. M. S. Estudo comparativo das estruturas semânticas em diferentes Sistemas de Organização do Conhecimento. **Ciência e Informação**, Brasília, DF, v. 46, n.1, p. 60-72, jan./abr. 2017.

MOURA, M. A. Organização social do conhecimento e performatividade de gênero: dispositivos, regimes de saber e relações de poder. **Liinc em Revista**, Rio de Janeiro, v. 14, n. 2, 2018.

OORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Plataforma Sucupira**. Disponível em: <https://sucupira.capes.gov.br/sucupira/>. Acesso em: 20 mar. 2023.

SANTOS, R. F.; ALBUQUERQUE, M. E. B. C. Aplicabilidade da Folksonomia na construção e atualização de Sistemas híbridos de Organização do Conhecimento. **Ponto de Acesso**, Salvador, v. 15, n.3, 2021.

SOLER MONREAL, C.; GIL LEIVA, I. Posibilidades y límites de los tesauros frente a otros sistemas de organización del conocimiento: folksonomías, taxonomías y ontologías. **Revista Interamericana de Bibliotecología**, [S.l.], v.33, n.2, p.361-377, jul./dic. 2010.

TORRES, S.; ALMEIDA, M. B.; SIMÕES, M. G. M. Princípios para a modelagem de domínios em Sistemas de Organização do Conhecimento. *In: CONGRESSO ISKO ESPANHA-PORTUGAL*, 3., Coimbra, 2017. *Anais [...]*, Coimbra, 2017.

ZENG, M. L. Knowledge Organization Systems (KOS). **Knowledge Organization**, [S.l.], v.35, n.2/3, p.160-182, 2008.

Apêndice A

Quadro 1. *Cursos de Pós-Graduação em Ciência da Informação com disciplinas sobre SOC*

<i>REGLÃO</i>	<i>INSTITUIÇÃO</i>	<i>NOME DO CURSO</i>	<i>MODALIDADE</i>
Sul	Universidade Federal de Santa Catarina (PGCIN/UFSC)	Ciência da Informação	Mestrado e Doutorado Acadêmicos
Sul	Universidade Estadual de Londrina (PPGCI/UEL)	Ciência da Informação	Mestrado e Doutorado Acadêmicos
Sul	Universidade Federal do Rio Grande do Sul (PPGCIN/UFRGS)	Ciência da Informação	Mestrado Acadêmico
Sudeste	Universidade Federal Fluminense (PPGCI/UFF)	Ciência da Informação	Mestrado e Doutorado Acadêmicos
Sudeste	Universidade Federal do Rio de Janeiro (PPGCI BICIT/UFRJ)	Ciência da Informação	Mestrado e Doutorado Acadêmicos
Sudeste	Universidade Federal de Minas Gerais (PPGCI/UFMG)	Ciência da Informação	Mestrado e Doutorado Acadêmicos
Sudeste	Universidade Federal de Minas Gerais (PPGGOC/UFMG)	Gestão & Organização do Conhecimento	Mestrado e Doutorado Acadêmicos
Sudeste	Universidade Estadual Paulista (PPGCI/UNESP)	Ciência da Informação	Mestrado e Doutorado Acadêmicos

Fonte: Elaborado pelas autoras.

Apêndice B

Quadro 3. Componentes curriculares identificados na pesquisa

PROGRAMA	DISCIPLINA	EMENTA
PPGCI/UFF	Instrumentos terminológicos de representação e recuperação da informação e do conhecimento	Investiga princípios teóricos e metodológicos para a elaboração de instrumentos terminológicos no contexto da representação e da recuperação da informação, como: taxonomias, tesouros, ontologias, incluindo sistemas inteligentes, Objetiva analisar os diversos contextos informacionais e o desempenho de ferramentas terminológicas como instrumentos de comunicação, de acesso à informação e de descoberta de conhecimento. Inclui estudos sobre a compatibilidade terminológica.
PPGCI IBICT/UFRJ	Sistemas de Organização do Conhecimento	Princípios teóricos e metodológicos de elaboração de linguagens de representação e recuperação da informação. Tipologias de sistemas de organização do conhecimento e representação de informação. Análise de desempenho de esquemas e tabelas de classificação, taxonomias, tesouros, vocabulários controlados, ontologias, terminologias e glossários em função de usos específicos. A moderna concepção da participação das comunidades virtuais colaborativas
PPGCI/UFMG	Linguagem documentária como instrumento de comunicação da informação	Função documentária como elemento caracterizador da Ciência da Informação, no contexto das problemáticas contemporâneas de acesso à informação e de seu uso qualificado, explorando o papel das linguagens documentárias como instrumentos de comunicação da informação. Organização da informação como campo de conhecimento.
PPGGOC/UFMG	Sistemas de Organização do Conhecimento	Metodologias de aplicação, tratamento e organização da informação. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. Teoria e prática da construção e utilização de tesouros em ambiente digital. Sistemas de organização do conhecimento como componentes de projeto da arquitetura da informação em instituições.
PGCIN/UFSC	Sistemas de Organização do Conhecimento	Fundamentos teóricos dos sistemas de organização do conhecimento: Teorias da Classificação, Teoria do Conceito, Análise de Domínio, outras teorias. Sistemas de Organização do Conhecimento: funções e características.
PPGCIN/UFRGS	Sistemas de Organização do Conhecimento aplicados a arquivos e bibliotecas	Os sistemas de representação e organização do conhecimento (SOC) representam um avanço teórico metodológico para a Ciência da informação. A representação e organização do conhecimento é uma área que pode contribuir e muito para a construção de metodologias para a organização de arquivos e bibliotecas. Assim, a organização do conhecimento contribui especialmente no que se refere ao tratamento temático da informação e nas atividades de representação da informação. Neste sentido, essa disciplina visa apoiar o estudo e compreensão de textos recentes a respeito das temáticas relacionadas a representação, organização, acesso, classificação, descrição, ou seja, as atividades relacionadas a criação de sistemas de representação. Busca fomentar essas discussões e aprofundar os conhecimentos relacionados a essas áreas. Parte-se de uma reflexão inicial afim das origens e procedimentos, avançando para as diferentes vertentes na atualidade da área.
	Terminologia, taxonomia e ontologia na Ciência da Informação	Bases conceituais, teóricas e metodológicas para elaboração de linguagens de representação e recuperação da informação. Estudo da terminologia aplicada a Ciência da Informação, com foco na variação terminológica para organização de informação. Utilizando tabelas de classificação, taxonomias, tesouros, vocabulários controlados, ontologias, terminologias e glossários nos diversos contextos informacionais.
PPGCI/UNESP	Sistemas de Organização do Conhecimento: aspectos teóricos e metodológicos	Representação e organização da informação e do conhecimento. Sistemas de organização do conhecimento: taxonomias, sistemas de classificação, mapas conceituais, topic maps, tesouros e ontologias. A teoria da classificação e a abordagem facetada aplicada à construção de sistemas de organização do conhecimento. Teoria do conceito e relações conceituais. Aspectos linguísticos e terminológicos relacionados aos sistemas de organização do conhecimento. Aspectos de normalização aplicados aos sistemas de organização do conhecimento.
PPGCI/UDEL	Folksonomia e representação colaborativa da informação	Folksonomia e representação colaborativa da informação: abordagens epistemológicas e pragmáticas. Modelos e sistemas colaborativos de representação da informação. Sistemas híbridos de Organização do Conhecimento (integração das Folksonomias com instrumentos como tesouros, ontologias, taxonomias, etc.). Enfoques e perspectivas de estudos nacionais e internacionais sobre Folksonomia na Ciência da Informação.

Fonte: Elaborado pelas autoras.

Casoni, Clarice Luzia Rodrigues; Santos, Raimunda Fernanda dos Santos; Albuquerque, Maria Elisabeth Baltar Carneiro de. Os sistemas de organização do conhecimento nos currículos dos programas de pós-graduação em Ciência da Informação nas regiões sul e sudeste do país. *Organização e Representação do Conhecimento em diferentes contextos: desafios e perspectivas na era da digitalização*, 2023.

Thesa 2.0: a interação dos tesouros em tempos de dados ligados

Thesa 2.0: thesaurus interaction in linked data times

Rene Faustino Gabriel Junior (1), Rita do Carmo Ferreira Laipelt (2)

(1) Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, rene.gabriel@ufrgs.br

(2) rita.laipelt@ufrgs.br

Resumo

Introdução: A perspectiva de implementação de novos recursos tecnológicos suscita algumas reflexões sobre a metodologia de elaboração de tesouros e seus usos. **Objetivo:** O objetivo deste trabalho é apresentar as mudanças que verificamos nos tesouros com a implantação de recursos como o *linked data*, RDF e a explicitação de relações semânticas. Para isso, descreve-se as características e funcionalidades do Thesa, um software para a construção de tesouros ricos semanticamente devido ao uso e ampliação dos modelos Skos Reference e Prime. **Metodologia:** A metodologia utilizada foi composta por uma revisão bibliográfica relacionada as ferramentas de construção de tesouros e documental com base nas descrições técnicas das normas ISO e NIZO vigentes, e da ontologia Skos como base para fundamentar a construção do Thesa 2.0. **Resultados:** A experiência no uso da primeira versão do Thesa possibilitou identificar necessidades pragmáticas para melhorar a descrição entre conceitos, possibilitando a interação contextualizada entre diferentes sistemas de recuperação de informação. No estudo observou-se a necessidade de ampliação do modelo Skos Reference com a incorporação da explicitação da propriedade skos:related, com a sua contextualização de aplicação. E a utilização do Skos: Primer para a integração entre diferentes tesouros (*schemas*). **Conclusão:** Com a ampliação e explicitação da relação semântica implementadas no Thesa 2.0 é possível afirmar que tesouros podem ser semanticamente tão fortes quanto uma ontologia.

Palavras-chave: Thesa. Tesouros semânticos. Elaboração de tesouros. Ontologias. Dados ligados.

Abstract

Introduction: The perspective of implementing new technological resources raises some reflections on the methodology for creating thesauri and their uses. **Objective:** The objective of this work is to present the changes that we verified in the thesauri with the implementation of resources such as linked data, RDF, and the explanation of semantic relations. For this, the characteristics and functionalities of Thesa are described, a software for the construction of semantically rich thesauri due to the use and expansion of the Skos Reference and Prime models. **Methodology:** The methodology used consisted of a bibliographical review related to thesaurus and documentary construction tools based on the technical descriptions of the current ISO and NIZO standards, and of the Skos ontology as a basis to support the construction of Thesa 2.0. **Results:** The experience in using the first version of Thesa made it possible to identify pragmatic needs to improve the description between concepts, enabling contextualized interaction between different information retrieval systems. In the study, it was observed the need to expand the Skos Reference model with the incorporation of the explicitness of the skos:related property, with its application context. And the use of Skos: Primer for the integration between different thesauri (*schemas*). **Conclusion:** With the expansion and explanation of the semantic relationship implemented in Thesa 2.0, it is possible to state that thesauri can be semantically as strong as an ontology.

Keywords: Thesa. semantic Thesauri. Elaboration of thesauri. Ontologies. Linked data.

Resumen

Introducción: La perspectiva de implementar nuevos recursos tecnológicos plantea algunas reflexiones sobre la metodología de creación de tesouros y sus usos. **Objetivo:** El objetivo de este trabajo es presentar los cambios que verificamos en los tesouros con la implementación de recursos como datos enlazados, RDF y la explicación de relaciones semánticas. Para ello, se describen las características y funcionalidades de Thesa, un software para la construcción de tesouros semánticamente ricos gracias al uso y expansión de los modelos Skos Reference y Prime. **Metodología:** La metodología utilizada consistió en una revisión bibliográfica relacionada con tesouros y herramientas de construcción documental a partir de las descripciones técnicas de las normas ISO y NIZO vigentes, y de la ontología Skos como base para sustentar la construcción de Thesa 2.0. **Resultados:** La experiencia en el uso de la primera versión de Thesa permitió identificar necesidades pragmáticas para mejorar la descripción entre conceptos, posibilitando la interacción contextualizada entre diferentes sistemas de recuperación de información. En el estudio se observó la necesidad de ampliar el modelo de Referencia Skos con la incorporación de la explicitud de la propiedad skos:related, con su contexto de aplicación. Y el uso de Skos: Primer para la integración entre diferentes tesouros (esquemas). **Conclusión:** Con la expansión y explicación de la relación semántica implementada en Thesa 2.0, es posible afirmar que los tesouros pueden ser semánticamente tan fuertes como una ontología.

Palabras clave: Thesa Tesouros semânticos. Elaboración de tesouros. Ontologías. Datos vinculados.

1 Introdução

É inegável que os avanços tecnológicos recentes têm impactado todas as áreas de organização do conhecimento. No que se refere à elaboração de tesouros os impactos e aplicações são imensuráveis. A perspectiva de implementação de novos recursos tecnológicos ao mesmo tempo que suscita algumas reflexões sobre a metodologia de elaboração de tesouros e seus usos também nos permitem demonstrar na prática a evolução desses Sistemas de Organização do Conhecimento (SOC).

O objetivo deste trabalho é apresentar as mudanças que verificamos nos tesouros com a implantação de recursos como o linked data, RDF e a explicitação de relações semânticas as quais a Norma ISO 25.964-1 (INTERNATIONAL..., 2011, p. 67), denomina como “[...] relacionamentos customizados [...]”. Para atingir esses objetivos descrevemos as características e funcionalidades de um software para a construção de tesouros ricos semanticamente devido ao uso e ampliação dos modelos Skos Reference e Primer, o Thesa.

As alterações implementadas no Thesa, e os tesouros resultantes do uso desses recursos confirmam a perspectiva apresentada por Moreira (2019) de que, do ponto de vista da Ciência da Informação, não existem diferenças substanciais entre tesouros e ontologias. Para o autor, ao considerarmos os aspectos estrutural e funcional dos SOC, o conjunto de elementos que os aproxima é maior do que aqueles que os diferencia.

Na literatura é recorrente a afirmação de que as ontologias se diferenciam dos tesouros devido a maior riqueza e flexibilidade no relacionamento entre os termos viabilizada pela explicitação e formalização desses relacionamentos semânticos. E que, em função dessas características, ontologias são SOC semanticamente mais fortes do que os tesouros. (SALES; CAFÉ, 2008; LANZA, 2023).

Como fio condutor de nossas reflexões, apresentamos neste texto novos recursos implementados no software Thesa que ajudam a demonstrar na prática que a distância entre tesouros e ontologias diminui a cada dia.

2 Organização da informação

Os tesouros são ferramentas de organização da informação que têm como objetivo principal facilitar a recuperação da informação. Eles são estruturas vocabulares controladas, que reúnem um conjunto de termos que representam conceitos relacionados a uma área específica do conhecimento (SALES; CAFÉ, 2008).

Os tesouros são uma importante ligação com a Web Semântica, uma vez que permitem a organização

estruturada e semântica do conhecimento especializado, possibilitando a integração e o compartilhamento de dados de forma mais inteligente e interoperável.

2.1 Tesouros, Ontologias e Linked Data

É recorrente na literatura sobre Sistemas de Organização do Conhecimento a comparação entre tesouros e ontologias. E nestas comparações também são recorrentes os indicativos de que as ontologias são SOC semanticamente mais fortes do que os tesouros. No entanto, é importante considerar, que a função e a utilização de cada uma dessas ferramentas são diferentes. Por isso, podemos apenas indicar para qual situação e propósito um tesouro ou uma ontologia são mais adequados. O sucesso da escolha está em saber identificar quando um sistema de informação precisa de um tesouro e/ou quando necessita de uma ontologia (KLESS; *et al.*, 2015).

É preciso lembrar que os termos em um tesouro são pensados para representar informação e os termos em uma ontologia são para recuperar informação.

Assim do ponto de vista morfológico e sintático, um termo em um tesouro não pode ser quebrado/generalizado considerando sua hierarquia. É preciso usar o termo completo, visto que ele deve significar por si só, mesmo quando separado da estrutura hierárquica que o contextualiza, é por isso, também, que não se deve quebrar o sintagma terminológico. (LAIPALT; VAN DERLAAN, 2020). Por isso, por exemplo, ao subordinarmos o termo limão siciliano ao termo limão em um tesouro, não podemos quebrar o sintagma e mencionar apenas siciliano. É preciso usar o termo completo, limão siciliano, pois siciliano sozinho fica ambíguo e perde parte do seu significado. Consequentemente, não seria possível usar apenas a expressão siciliano para indexar uma obra sobre limão siciliano.

No caso de uma ontologia, por ser um artefato, o termo fica vinculado a estrutura. Embora não seja recomendado quebrar o sintagma terminológico, não teria problema utilizar apenas parte do termo, visto que ele não perderia o sentido dentro da estrutura que o sustenta. Ou seja, a estrutura da ontologia mantém o contexto do termo. Isso possibilita que o sentido do termo permaneça mesmo que ocorra a quebra de um sintagma terminológico.

Tesouros e ontologias são ferramentas que embora tenham o mesmo propósito final de melhorar a qualidade da recuperação da informação, realizam essa função de maneira diferente (SALES; CAFÉ, 2008; KLESS; *et al.*, 2015).

2.2 Tesouro semântico

O termo “thesaurus” tem suas raízes no grego e línguas latinas e significa “tesouro”. Esse termo tornou-se popular com a publicação de Peter Mark Roget do seu

dicionário analógico, em Londres, em 1852, intitulado *Thesaurus Of English Words And Phrases*. O tesouro de Roget estabeleceu um denominador comum para vocabulários que relacionam seus termos usando um certo tipo de relação semântica (MOREIRA; ALVARENGA; OLIVEIRA, 2004).

O tesouro semântico é um tesouro documentário qualificado pela explicitação das relações paradigmáticas e sintagmáticas, ou seja, pelo detalhamento dos tipos de relacionamentos entre os conceitos. O tesouro semântico, segue a linha europeia dos tesouros facetados (*Thesaurus facet*) baseados na teoria classificatória de Ranganathan (LANCASTER, 2004), a qual aplica o uso de categorias para organização dos conceitos em um domínio.

Esses aspectos também são destacados pela norma ISO 25964-1 e a ISO 25946-2 (INTERNATIONAL..., 2011; INTERNATIONAL..., 2013) ao apresentar os aspectos semânticos pela aplicação e ampliação dos relacionamentos customizados os quais, devido a explicitação das propriedades dos relacionamentos, possibilitam maior compreensão do sistema conceitual como um todo. Outro aspecto importante do tesouro semântico é a utilização dos recursos SKOS, que viabilizam a interoperabilidade e a ligação entre diferentes tesouros.

2.3 Simple Knowledge Organization System (SKOS)

O *Simple Knowledge Organization System* (SKOS) é um modelo para expressar a estrutura básica e conteúdo de sistemas de organização do conhecimento (CATARINO, 2014). É uma linguagem padronizada para descrever e representar conceitos e termos utilizados em sistemas de organização do conhecimento, como tesouros, listas de assuntos e taxonomias (ISAAC; SUMMERS, 2009).

A documentação do SKOS é dívida em SKOS Reference, uma especificação que define formalmente o modelo de dados do SKOS como uma ontologia OWL Full. E o SKOS Primer é um documento que fornece uma introdução ao SKOS e descreve como ele pode ser usado para representar esquemas de conceitos.

O SKOS, segundo Santos e Moreira (2018), é uma representação desses instrumentos, ou seja, um modelo de dados para desenvolvimento de vocabulários controlados em formatos interoperáveis entre máquinas com capacidade de processamento.

O SKOS foi projetado para fornecer um caminho de migração de baixo custo para portar sistemas de organização existentes para a Web Semântica. Ele pode ser usado sozinho ou em combinação com linguagens mais formais, como a Web Ontology Language (OWL, 2023). Alguns estudos são apresentados sobre o modelo de dados SKOS, como o de Ramalho (2015) ao considerar que o modelo representa um avanço

importante em relação às formas de representação de vocabulários no ambiente Web. Porém o autor destaca que apesar de oferecer uma sintaxe para declaração formal de sentenças, SKOS não proporciona subsídio computacional adequado para a realização de inferências avançadas para dar suporte a realização de inferências automáticas nos processos de organização e recuperação de informações.

3 Metodologia

A metodologia partiu de uma revisão bibliográfica relacionada as ferramentas de construção de tesouros e documental com base nas descrições técnicas das normas ISO 25964-1 e a ISO 25946-2 e da ontologia SKOS. Focando principalmente na identificação de subsídios para construção de um software que auxilia e facilita-se o ensino, pesquisa e criação de micro tesouros semanticamente ricos, com a incorporação da ontologia SKOS.

Constituído um *corpus* foram analisadas as contribuições para a construção de tesouros e reunidas em um documento com seus requisitos. Também foi analisado a documentação da Ontologia SKOS em relação a integração entre tesouros.

Com base no levantamento das demandas originárias da versão 1.0 do Thesa e das demandas identificadas na revisão de literatura, estabeleceu-se prioridades para implementação na versão 2.0. Com a categorização das prioridades, em alta, média e baixa, foram repassadas para a equipe de desenvolvimento do software a implementação na nova versão do Thesa. Realizando-se assim a produção da documentação da ontologia Thesa, do software e dos casos de uso.

Os resultados do estudo culminaram no desenvolvimento do Thesa 2.0 com algumas inovações descritas nos resultados.

4 O Thesa

O Thesa foi desenvolvido em 2017 pela Universidade Federal do Rio Grande do Sul (UFRGS), mais especificamente pelo curso de Biblioteconomia para o ensino da construção de tesouros. A ferramenta ganhou visibilidade e solicitações para uso por outras instituições de ensino e por bibliotecas no Brasil. Principalmente por estar disponível online, de forma gratuita e em português. Em paralelo, em 2018, foi aprovado o recurso CNPq (439277/2018-3) que foi empregado no aprimoramento da primeira versão do Thesa (GABRIEL JUNIOR; LAIPELT, 2017; GABRIEL JUNIOR; LAIPELT, 2019).

Com seu uso, sentiu-se a necessidade de inserir novos recursos incorporando novas propriedades descritivas, multilíngues e ampliação do esquema básico para o

avançado na ontologia Skos. Em relação ao SKOS, destaca-se a possibilidade de relacionar termos entre tesouros diferentes e de outros domínios.

Entre as qualidades da ferramenta, destaca-se sua interface amigável para uso por bibliotecários e usuários sem muitos conhecimentos de informática.

Até março de 2023, a ferramenta apresentava o armazenamento de 690 microtesouros com 38.436 conceitos e 56.499 termos, uma média de 55,7 conceitos e 81 descritores para cada microtesouros. Observou-se também que do total, 95 são abertos para consulta, o restante está restrito para o acesso apenas de seus autores (THESA, 2023).

A proposta do Thesa 2.0 alinha-se ao conceito de Web 2.0, utilizado para descrever a segunda geração da World Wide Web - tendência que reforça o conceito de troca de informações e colaboração dos internautas com sites e serviços virtuais. A ideia é que o ambiente on-line se torne mais dinâmico e que os usuários colaborem para a organização de conteúdo. No Thesa 2.0, os usuários criam os conceitos em seus tesouros podendo adicionar ou especificar as relações entre os conceitos. Indo além dos tradicionais TG (broader), TE (narrow) e TR (related).

A modelagem do Thesa com o SKOS estabelece um grande avanço na construção de tesouros, por possibilitar seu consumo por bibliotecários e usuários, integrando-se muitas vezes com os softwares de buscas dos catálogos ou bases de dados.

A base do Thesa foi construída no modelo básico da ontologia Skos (W3C, 2008), que atende plenamente os requisitos para construção e compartilhamento de tesouros. Entretanto foi necessário ampliá-la de forma a melhorar a descrição das relações associativas apresentadas no Quadro 1.

A W3C disponibiliza uma wiki com algumas ferramentas para trabalhar com o Skos (2010). Dentre essas ferramentas destacam-se as que possibilitam a integração com outros sistemas como o WordPress; ferramentas para validar o arquivo SKOS; construção de tesouros como o TemaTres (2023); ThManager (2023); e o SKOSMOS (2015). Sendo essa última utilizada por instituições europeias como o Le Centre national de la recherche Scientifique (CNRS) da França com o Loterre; O AGROVOC da FAO; a United States National Agricultural Library com o Agricultural Thesaurus; HASSET com The Humanities and Social Science Electronic Thesaurus dentre outras instituições.

Dos estudos da ontologia SKOS observou-se que o modelo fornece a estrutura básica e o conteúdo de esquemas de conceitos, como tesouros, esquemas de classificação, listas de títulos de assuntos, taxonomias, folksonomias e outros tipos semelhantes de vocabulário controlado. No SKOS básico (Reference), os conceitos

são identificados por meio de URIs, rotulados com notas e relacionados hierarquicamente ou por meio de redes de associação. Já no SKOS avançado (Primer), é possível mapear conceitos entre esquemas (microtesouros) e agrupá-los em coleções rotuladas ou ordenadas (derivando microtesouros), especificar relações entre rótulos de conceitos e estender o próprio vocabulário SKOS para atender a necessidades específicas de comunidades de prática ou combiná-lo com outros vocabulários de modelagem.

5 Resultados

Nas definições do SKOS não é possível estabelecer relações semântica mais fortes, entre as propriedades dos conceitos, explicitados apenas pelas propriedades tradicionais skos:broader; skos:narrow; e skos:related em um esquema (tesauro).

O SKOS (2008) permite que conceitos sejam compostos e publicados na *World Wide Web*, vinculados a dados na Web e integrados em outros esquemas conceituais, possibilitando a qualificação das relações paradigmáticas e sintagmáticas, constituindo a uma semântica forte dentro do Thesa.

Buscando o enriquecimento semântico e a interação no Thesa está proposto a ampliação das propriedades de relacionamento com skos:related.

A proposta do Thesa 2.0 é a definição de relações mais ricas, com a proposição de um modelo de referência apresentados no Quadro 1. Ressalta-se que o modelo é limitado, possibilitando o usuário a definição de propriedades customizadas conforme a necessidade do domínio que se está modelando.

Desta forma foi necessário ampliar as relações associativas no Thesa 2.0 com a incorporações de novas propriedades, apresentadas no Quadro 1, com as quais busca-se enriquecer a qualidade das descrições das relações entre os conceitos, entretanto mantendo a propriedade original skos:related. Ressalta-se ainda, que é possível o usuário criar propriedades conforme necessidade de descrição de seu domínio.

Quadro 1. *Propriedades ampliadas no Thesa para complementar a descrição do SKOS ampliada no Thesa 2.0*

<i>Propriedade</i>	<i>Descrição</i>
thesa:relatedPartOf inverso de thesa:relatedHasPart	meronímia (parte de),
thesa:relatedHasPart inverso de thesa:relatedPartOf	hierárquica (todo/parte)
thesa:hasAntonymy	antonímia (oposição conceitual)

Propriedade	Descrição
thesa:hasCoordinationof	coordenação (termos subordinados a um mesmo conceito)
thesa:hasPatientAction	ação/paciente ou objetivo
thesa:hasPropertiesThing	conceito ou coisa/propriedades,
thesa:hasActionAgent	ação ou coisa/contra agente,
thesa:hasSubpartOf	coisa/suas partes (se não ocorre a relação hierárquica)
thesa:hasFeedstockOf	matéria prima/produto
thesa:hasPropActionOf	ação/propriedade
thesa:hasFieldOfStudy	campo de estudo/objetos ou fenômenos estudados
thesa:hasCauseEffect	causa/efeito,
thesa:hasProcessInAgent	processo/agente
thesa:hasActionInProduct	ação/produto da ação,
thesa:hasProductFeature	características do produto(*)
thesa:relatedHasParentalAffinity	afinidade - parentesco (*)
thesa:relatedHasFriendAffinity	afinidade - amizade
thesa:relatedHasLoveAffinity	Afinidade - romance

Fonte: Elaborado pelos autores.

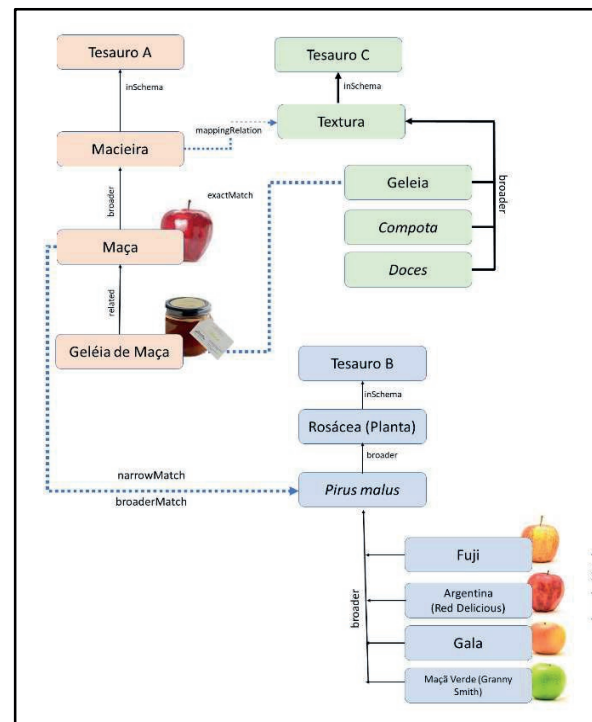
A limitação do relacionado está dentro do domínio do tesouro elaborado. Para possibilitar a ligação entre diferentes tesouros, utiliza-se o SKOS. Porém de forma pragmática, existe a necessidade de criar uma propriedade exata, próxima, similares e definições diferentes (exactMatch, closeMatch, mappingRelation, relatedMatch) dos conceitos intertesouros.

Ressalta-se que a ligação entre diferentes tesouros está condicionada ao uso da ontologia SKOS. Como o Thesa 2.0 utiliza o modelo Primer, este pode-se utilizar de conceitos de outros tesouros com SKOS *Reference* ou *Primer* (W3C, 2008).

Na figura 1, apresentamos um exemplo da aplicação do SKOS Primer para conectar diferentes tesouros. Como pode-se verificar, a ligação entre diferentes tesouros possibilita o enriquecimento dos tesouros interligados, possibilitando um aprofundamento do conhecimento sobre os termos de maneira colaborativa. Visto que, é possível visualizar outros termos e associações a partir de outros pontos de vista que não foram pensados na elaboração dos tesouros isolados. No exemplo que apresentamos, o Tesouro A foi construído a partir da perspectiva da Maça como produto, o Tesouro B foi construído a partir da Perspectiva da origem e nome científico das maçãs e o Tesouro C foi construído a partir da perspectiva de diferentes preparos e texturas. Com o uso da propriedade skos:exactMatch ligamos o termo

popular Maça do Tesouro A com o Termo científico *Pirus malus* do Tesouro B. Ao acessar o nome científico da maçã no Tesouro B, o usuário também terá acesso à informação de que existem outros tipos de maçã que se encontram listadas na estrutura do Tesouro B. Ao usarmos a propriedade skos:broadMatch para ligar o termo Geleia de Maça do Tesouro A ao termo genérico Geleia do Tesouro C evidenciamos pela análise da estrutura do Tesouro C que existem outros preparos que podem ser feitos com a maçã além da geleia, como compotas e doces. Ao usarmos a propriedade skos:mappingRelation ligamos o termo Macieira do Tesouro A ao termo Rosácea (Planta) do Tesouro B evidenciando que existem diversas famílias de plantas com flor e a Macieira é uma delas. Logo, as informações dos três tesouros se complementam e se enriquecem mutuamente ao agregar diferentes perspectivas e pontos de vista.

Figura 1. Relações Inter tesouros representados no SKOS



Fonte: Elaborado pelos autores.

As possíveis relações do SKOS estão representadas no Quadro 2. No Thesa 2.0 foram implementadas todas as relações intertesouros. Entretanto é necessário compreender a aplicações e uso dessas relações, por exemplo, a utilização do skos:closeMatch pode levar a uma ambiguidade semântica, isso porque a relação de equivalência ou similaridade não é sempre precisa ou unívoca. Em um SRI o closeMatch pode provocar uma imprecisão, com a recuperação de documentos irrelevantes para o usuário. Ou seja, pode levar a uma falta de confiança.

Quadro 2. Relações SKOS Primer em diferentes esquemas (multitesauros)

Propriedade	Descrição
skos:exactMatch	conceitos que são exatamente os mesmos em significado
skos:closeMatch	conceitos que têm um alto grau de similaridade, mas não são idênticos
skos:mappingRelation	conceitos similares, mas definidos de forma diferente, possam ser relacionados
skos:relatedMatch	conceitos que são relacionados ou conectados de alguma forma, mas que não são exatamente equivalentes
skos:broadMatch	Evidencia um conceito mais amplo a outro conceito mais específico
skos:narrowMatch	Evidencia um conceito mais específico a outro conceito mais amplo

Fonte: Elaborado pelos autores.

A interoperabilidade do SKOS pode ocorrer de duas maneiras, a primeira na troca de vocabulários controlados com outras ferramentas para este fim que utilize o mesmo protocolo, e segunda por meio da integração com sistemas de recuperação de informação (SRI).

A interoperabilidade do SKOS com os SRI é realizada principalmente por meio da utilização de recursos como o SPARQL (*SPARQL Protocol e RDF Query Language*), que é uma linguagem padrão para consulta e manipulação de dados RDF (*Resource Description Framework*). O RDF é uma estrutura de dados flexível para representar informações na Web Semântica, e é usado em conjunto com outras tecnologias como o OWL (*Web Ontology Language*) para criar ontologias (MILES; PÉREZ-AGÜERA, 2007).

Neste contexto o Thesa 2.0 buscou implementar algumas funcionalidades de importação de informação de outros sistemas como o TemaTres, SKOSMOS e outros software que atendam os padrões XML ou RDF em seus modelos de exportação. Das ferramentas de exportação do Thesa 2.0 destacam-se a possibilidade de enviar os dados nos formatos XML, RDF, TURTLE, CVS, JSON o que possibilita a integração com outros sistemas.

A utilização do RDF na representação do tesauro, possibilita a conexão com outros tesauros internos do Thesa, tesauros instanciados em outros servidores do Thesa externos, ou ainda com sistemas externos como o TemaTres, SKOSMOS ou ainda com provedores de informações como o WikiData, DBPedia ou outros

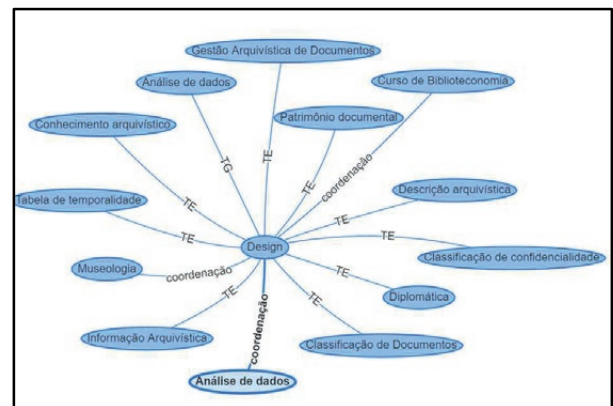
provedores que trabalham com dados ligados (*linked data*).

Uma funcionalidade importante implementada no Thesa 2.0 está a buscar por meio de *Application Programming Interface (API) Representational State Transfer (REST)*, possibilitando a escalabilidade e flexibilidade, independente de plataforma e linguagem. A API RESTful é facilmente integrável com outras aplicações e sistemas, o que a torna ideal para o desenvolvimento de aplicações distribuídas e para a exposição de serviços para a Web. O uso destas ferramentas possibilita que os sistemas possam realizar funções como o autocompletar, sugerindo termos padronizados para os usuários, possibilitando a padronização dos vocabulários utilizados.

Destaca-se que na versão do Thesa 2.0, bem como em sua versão anterior, essas ferramentas estão funcionais. Possibilitando a integração como os catálogos de bibliotecas. Atualmente essa integração está sendo testada na Brapci e no LattesData.

No desenvolvimento do Thesa tanto na versão 1.0 como na 2.0 observou-se a necessidade de uma representação gráfica dos conceitos representados. A figura 2 apresenta a os grafos do conceito Design.

Figura 2. Representação em Grafo do Thesa das relações paradigmáticas e sintagmáticas



Fonte: Elaborado pelos autores.

Outra premissa incorporada na versão 2 do Thesa está no uso de múltiplos termos preferenciais (*prefLabel*) para um mesmo conceito, desde que esses estejam em idiomas diferentes. Atribuições essas descritas no modelo SKOS, porém não permitidas nos tesauros anteriormente na metodologia de construção de tesauros.

Pelo Thesa trabalhar com um identificador persiste, a determinação de um termo preferencial pode ser alterada sem prejuízo a representação do conceito, pois a ela é realizada pela URI do conceito, e não pelo seu nome. A Figura 3 mostra uma representação em vários idiomas preferencial do termo Alcachofra.

Figura 3. Representação do conceito de Alcachofra no Thesa 2.0



Fonte: Elaborado pelos autores.

A figura 3 e 4 demonstram as funcionalidades de uso de imagem e vídeo como elemento de representação visual de um conceito. No caso da primeira figura, a apresentação de uma imagem de uma alcachofra. A Figura 4 faz parte do Glossário de Biblioteconomia traduzido para Língua Brasileira de Sinais – LIBRAS, vocabulário criado por Mariana Briese como parte de seu mestrado, para utilização pela comunidade surda, surdos, tradutores-intérpretes, estudantes de biblioteconomia e bibliotecários para possibilitar a comunicação nas bibliotecas. O Glossário consiste na descrição de um conceito, uma definição textual e um vídeo mostrando sua o conceito na linguagem de sinais. O glossário pode ser acessado pelo link <https://www.ufrgs.br/thesa/public/th/2> (THESA, 2023).

Figura 4. Tela de edição do conceito “Atraso” com incorporação de vídeo para surdo



Fonte: Mariana Briese (2023).

Com o desenvolvimento do Thesa 1.0 observou-se uma necessidade de ferramentas que possibilitem o profissional de informação construir vocabulários controlados, glossários, tesouros, que possam ser mais dinâmicos e que se integrem com outros sistemas de informação.

Na versão do Thesa 2.0 foram atendidas as demandas de incorporação vários idiomas, como termos preferenciais, a incorporação de elementos multimídias como imagens e vídeos, além do detalhamento de definições, notas de escopo, notações entre outros elementos descritivos. Ainda nesta versão foram implementadas propriedades de relacionamento entre

tesouros do próprio Thesa, bem como tesouros que utilizam outros sistemas baseados em SKOS.

6 Conclusão

O resultado do uso da primeira versão do Thesa possibilitou identificar uma necessidade de melhor descrição para a interação entre sistemas de recuperação de informação de forma integrada e automática.

O Thesa 2.0 busca concretizar um instrumento de construção de tesouros ricos semanticamente, plurilíngues, uso de multimídia como recurso complementar para a representar conceitos, e conexão com outros tesouros SKOS.

Com a ampliação e explicitação da relação semântica implementadas no Thesa 2.0 é possível afirmar que tesouros podem ser semanticamente tão fortes como uma ontologia.

Espera-se que o Thesa contribua para a discussão e uso de ferramentas SKOS no Brasil, além de contribuir para o ensino e pesquisa do tema, instigando bibliotecários e estudantes de pós-graduação a conhecer e estudar suas limitações e vantagens frente a instrumentos tradicionais de representação do conhecimento.

Referências

- BRIESE, M. **Glossário Digital de Biblioteconomia traduzido para a Língua Brasileira de Sinais (LIBRAS): uma necessidade emergente para inclusão social do surdo**. 2023. Dissertação (Mestrado em Letras) – Universidade Federal do Rio Grande, Rio Grande, 2023.
- CATARINO, M. E. Simple Knowledge Organization System: construindo sistemas de organização do conhecimento no contexto da Web Semântica. **Informação & Tecnologia (ITEC)**. Marília/João Pessoa, 1(1): 17-28, jan./jun., 2014. <https://brapci.inf.br/index.php/res/download/218240>. Acesso em: 2023.
- GABRIEL JUNIOR, R. F.; LAIPELT, R. C. F. Descrição das relações semânticas para aplicação em kos: uso do tesouro semântico aplicado (Thesa). **Revista P2P e Inovação**, [S.l.], v. 6, n. 1, p. 117-135, 2019. DOI: 10.21721/p2p.2019v6n1.p117-135. Disponível em: <https://revista.ibict.br/p2p/article/view/4946>. Acesso em: 23 mar. 2023.
- GABRIEL JUNIOR, R. F.; LAIPELT, R. C. F. Thesa: ferramenta para construção de tesouro semântico aplicado interoperável. **Revista P2P & Inovação**, Rio de Janeiro, [S.l.], v. 3, n. 2, p.124-145, Mar./Set. 2017.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-1:2011 Information and Documentation – Thesauri and Interoperability with Other Vocabularies – Part 1: Thesauri for Information Retrieval**. Geneva, Switzerland: ISO, 2011.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **ISO 25964-2:2013 Information**

and Documentation – Thesauri and Interoperability with Other Vocabularies – Part 2: Interoperability with other vocabularies. Geneva, Switzerland: ISO, 2013.

ISAAC, A.; SUMMERS, E. (ed.). **SKOS Simple Knowledge Organization System Primer.** [S.l.]: W3C Working Group Note, 18 Aug. 2009.

KLESS, D.; MILTON, S.; KAZMIERCZAK, E; LINDENTHAL, J. Thesaurus and ontology structure: Formal and pragmatic differences and similarities. *Jasist*, [S.l.], v. 66, n. 7, 2015. DOI: <https://doi.org/10.1002/asi.23268>. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.23268>. Acesso em: 23 mar. 2023.

LAPELT, R.C. F.; VAN DER LAAN, R. H. **Pré e pós-coordenação:** algumas reflexões. Organização da informação: textos didáticos. Porto Alegre: Editora da UFRGS, 2020. 88 pp 65-75, 2020.

LANCASTER, F. W. **Indexação e resumos.** 2. ed. rev. ampl. Tradução de A. A. Briquet de Lemos. Brasília: Briquet de Lemos, 2004. 452 p.

LANZA, C. **Semantic Control for the Cybersecurity Domain:** Investigation on the Representativeness of a Domain-specific Terminology Referring to Lexical Variation. CRC New York: Press, 2023. 170 p.

MILES, A.; PÉREZ-AGÜERA, J. R. SKOS: Simple Knowledge Organisation for the Web, **Cataloging & Classification Quarterly**, [S.l.], v. 43, n. 3-4, p. 69-83, 2007. DOI: 10.1300/J104v43n03_04. Disponível em: https://www.tandfonline.com/doi/citedby/10.1300/J104v43n03_04?scroll=top&needAccess=true&role=tab&aria-labelledby=cit. Acesso em: 23 mar. 2023.

MOREIRA, A.; ALVARENGA, L.; OLIVEIRA, A. P. Thesaurus and Ontology. *Knowl. Org.*, [S.l.], v. 31, n. 4, 2004.

MOREIRA, W. Tesouros e ontologias como modelos de sistemas de organização do conhecimento. *Brazilian Journal of Information Studies*, [S.l.], v. 13, n. 1, 2019. p.15-20.

RAMALHO, R. A. S. Análise do Modelo de Dados SKOS: Sistema de Organização do Conhecimento Simples para a

Web. *Informação & Tecnologia (ITEC)*, v. 2, p. 66-79, 2015.

SALES, R.; CAFÉ, L. Semelhanças e diferenças entre tesouros e ontologias. *DataGramaZero*, [S.l.], v. 9, n. 4, 2008.

SANTOS, J. C. F. D.; MOREIRA, W. SKOS: uma análise sobre as abordagens e suas as aplicações na ciência da informação. *Informação & Informação*, Londrina, v. 23, n. 3, p. 362-389, 2018. DOI: 10.5433/1981-8920.2018v23n3p362. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/29303>. Acesso em: 23 mar. 2023.

SKOS wiki tools. [S.l.]: W3C, 2010. Disponível em: <https://www.w3.org/2004/02/skos/wiki/Tools>. Acesso em: 23 mar. 2023.

SKOSMOS. **Open source web-based SKOS browser and publishing tool.** [S.l.]: National Library of Finland, 2015. Disponível em: <https://skosmos.org/>. Acesso em: 23 mar. 2023.

TEMA TRÊS. **Tesouro de demonstración y pruebas para el servidor de vocabularios controlados TemaTres.** [S.l.]: TemaTres, 2023. Disponível em: <https://www.r020.com.ar/tematres/demo/index.php>. Acesso em: 23 mar. 2023.

THESA. **Tesouro Semântico Aplicado.** Porto Alegre: UFRGS, 2023. Disponível em: <https://ufrgs.br/thesa>. Acesso em: 23 mar. 2023.

THMANAGER. **What is ThManager?** [S.l.]: University of Zaragoza; GeoSpatiumLab, 2023. Disponível em: <https://thmanager.sourceforge.io/index.html>. Acesso em: 23 mar. 2023.

W3C. [S.l.]: SKOS Simple Knowledge Organization System Primer, 2008. Disponível em: <https://www.w3.org/TR/2008/WD-skos-reference-20080829/skos.html>. Acesso em: 23 mar. 2023.

WEB ONTOLOGY LANGUAGE (OWL). [S.l.]: W3C, 2023. Disponível em: <https://www.w3.org/OWL/>. Acesso em: 23 mar. 2023.

O uso de taxonomias para o enriquecimento semântico de objetos

The use of taxonomy for the semantic enrichment of objects

Júnio Soares Dias (1), Claudiane Emanuele Nazário (2), Célia da Consolação Dias (3)

Universidade Federal de Minas Gerais (UFMG), Av. Pres. Antônio Carlos, 6627 – Pampulha-MG, Brasil, (1) juniodias@ufmg.br. (2) cen2001@ufmg.br (3) celiadias@eci.ufmg.br.

Resumo

A literatura da Ciência da Informação descreve a necessidade de estudos sobre o enriquecimento semântico de Sistemas de Organização do Conhecimento. Esta investigação apresenta uma análise primária de duas pesquisas em andamento sobre temáticas pertencentes à Organização do Conhecimento. O objetivo deste estudo foi iniciar a análise do estado da arte sobre a taxonomia, o enriquecimento semântico de objetos e o uso da taxonomia destinada ao enriquecimento semântico. Para isto, optamos por utilizar o Mapeamento Sistemático como metodologia de pesquisa nos últimos cinco anos de publicações em quatro bases de dados de pesquisas científicas. Os resultados obtidos foram uma coleta de dados, uma análise quantitativa e qualitativa. A análise quantitativa obtida representa as tendências da temática de pesquisa para o período investigado. Já a análise qualitativa representa o levantamento inicial do estado da arte para o estudo em questão. Concluiu-se que esses estudos obtidos sobre os Sistemas de Organização do Conhecimento e, principalmente, a taxonomia destinada ao enriquecimento semântico são incipientes e o seu desenvolvimento pode contribuir para a Ciência da Informação e outras áreas do conhecimento.

Palavras-chave: Taxonomia. Enriquecimento Semântico. Mapeamento Sistemático.

Abstract

The Information Science literature describes the need for studies about the semantic enrichment of Knowledge Organisation Systems. This research presents a primary analysis of two ongoing investigations on a theme belonging to the Knowledge Organisation. Our goal was to begin the analysis of the state of the art on taxonomy, the semantic enrichment of objects and the use of taxonomy intended for semantic enrichment. For this, we are close to use Systematic Mapping as a research methodology in the last five years of publications in four scientific research databases. The results obtained were a data collection, a quantitative and qualitative analysis. The quantitative analysis obtained represents the trends of the research theme for the investigated period. The qualitative analysis represents the initial survey of the state of the art for the study in question. We conclude that these studies on Knowledge Organisation Systems and especially the taxonomy intended for semantic enrichment are incipient and their development can contribute to Information Science and other knowledge areas.

Keywords: Taxonomy. Semantic Enrichment. Systematic Mapping.

Resumen

La literatura de Ciencias de la Información describe la necesidad de estudios sobre el enriquecimiento semántico de los Sistemas de Organización del Conocimiento. Esta investigación presenta un análisis primario de dos investigaciones en curso sobre temas pertenecientes a la Organización del Conocimiento. El objetivo de este estudio fue iniciar el análisis del estado del arte sobre la taxonomía, el enriquecimiento semántico de los objetos y el uso de la taxonomía destinada al enriquecimiento semántico. Para ello, hemos optado por utilizar el Mapeo Sistemático como metodología de investigación en los últimos cinco años de publicaciones en cuatro bases de datos de investigación científica. Los resultados obtenidos fueron una recopilación de datos, un análisis cuantitativo y cualitativo. El análisis cuantitativo obtenido representa las tendencias de la temática de investigación para el periodo investigado. Ya el análisis cualitativo representa la encuesta inicial del estado del arte para el estudio en cuestión. Concluimos que estos estudios sobre los Sistemas de Organización del Conocimiento y, principalmente, la taxonomía destinada al enriquecimiento semántico son incipientes y su desarrollo puede contribuir a la Ciencia de la Información y otras áreas del conocimiento.

Palabras clave: Taxonomía. Enriquecimiento Semántico. Mapeo Sistemático.

1 Introdução

O crescimento desordenado de dados e informações na *web* tem se tornado um tema recorrente na literatura da área de Biblioteconomia e Ciência da Informação (BCI). Existem repositórios de diferentes domínios do

conhecimento disseminados diariamente sem adotar qualquer padrão ou controle terminológico, o que dificulta sua interpretação por sistemas informatizados.

O enriquecimento semântico é um processo utilizado para atribuir maior significado aos dados e metadados,

através do uso da semântica atribuída por vocabulários pré-existent, sinônimos e informações de proveniência, visando facilitar a compreensão, a integração e o processamento dos dados por pessoas e máquinas (LIRA, 2014). Conforme apontado na pesquisa do referencial teórico, seção 2, há poucas discussões na BCI referentes ao processo de enriquecimento semântico, sobretudo relacionados à utilização de Sistemas de Organização do Conhecimento (SOC) tais como: taxonomias e ontologias para realizar o enriquecimento semântico e facilitar o processamento dos dados por homens e máquinas.

Na BCI, existem diversos esquemas de representação do conhecimento próprios para a representação da informação e que cumprem a função de comunicação entre os sistemas de informação. Nesta categoria de esquemas, podem ser citados as ontologias, os mapas conceituais e as taxonomias. Todos esses conceitos têm sua origem em outras áreas de conhecimento. As taxonomias têm sua origem na Biologia, e em relação à etimologia, o termo tem sua origem no grego *Táxis* que significa ordem e *Nomos* é igual à lei, norma ou regra. O conceito de taxonomia está relacionado com classificação e com sistemas de classificação. Para a compreensão do significado do termo resgata-se a sua definição presente na Norma ANSI/NISO Z39.19-2005 [1]. De acordo com a norma, a taxonomia é um conjunto de termos de um vocabulário controlado, organizados numa estrutura hierárquica, onde cada termo de uma taxonomia se encontra em mais do que uma relação com outros termos de outra taxonomia.

Esse artigo, cujo objetivo principal é avaliar as contribuições das taxonomias para o enriquecimento semântico de dados, é parte de duas pesquisas de doutorado mais abrangentes, ainda em desenvolvimento. Sendo assim, em busca do estado da arte sobre a taxonomia e o enriquecimento semântico, a questão de pesquisa para atingir esse objetivo estabelecido é: Em que medida o uso da taxonomia pode favorecer o enriquecimento semântico de objetos?

O presente artigo está dividido em cinco seções, a saber: A primeira seção contém os aspectos introdutórios e objetivos da pesquisa. A segunda seção é composta da exploração preliminar da literatura, contendo o referencial teórico, em suporte ao estudo apresentado. A seção três apresenta a metodologia utilizada para elaboração deste estudo. Já a seção quatro apresenta os resultados encontrados. Na última seção, são descritas as considerações finais e na sequência estão as referências bibliográficas usadas para a construção deste artigo.

2 Referencial teórico

Ao realizar uma revisão da literatura sobre o uso da taxonomia para o enriquecimento semântico, foi

permitido contemplar os principais conceitos que circundam o objeto deste projeto de investigação. Hjørland (2009) esclarece que os SOC, contextualizados pela própria área da BCI, são sistemas que podem organizar os conceitos pertencentes a um determinado domínio e as suas relações semânticas. Isto é, os SOC visam à organização da produção intelectual humana e são utilizados para representar e recuperar a informação (BRASCHER; CARLAN, 2010; DIAS, C. C. *et al.*, 2022).

Os SOC podem ser utilizados em diversas áreas do conhecimento, em inúmeras aplicações. Entre esses SOC e áreas do saber, podem-se citar como exemplos a taxonomia para a Ciência da Computação, nos processos da engenharia de requisitos (FAGUNDES *et al.*, 2020), para o Direito, nos argumentos jurídicos (LIMA; GOMES NETO, 2018), para a teoria de linguagem (AUSTIN, 2021) e para pesquisas científicas, nas temáticas da taxonomia corporativa e facetada (LOPES; AGANETTE; MACULAN, 2020), e na interoperabilidade (DIAS, C. C. *et al.*, 2022).

Ao considerar as pesquisas da Ciência da Informação, com foco na taxonomia e no enriquecimento semântico, podem-se citar as pesquisas de Maculan e Lima (2011), Aganette (2010), Aquino, Carlan e Brascher (2009), Barbosa (1972) e Ranganathan (1960) especificamente para a taxonomia. Além dos estudos de Lóscio, Burle e Calegari (2017), Lira (2014) e Sorrentino *et al.* (2013), que abordam o enriquecimento semântico.

Para a taxonomia como referencial teórico, a pesquisa de Barbosa (1972, p. 73) apresenta a necessidade de se “[...] considerar um documento, não apenas por seu aspecto físico mas também por seu conteúdo [...]”. Conforme apresentado pela autora, o aspecto do conteúdo levou à teoria de múltiplos aspectos ou facetas desenvolvida por Ranganathan (1960). Esse estudo aborda a taxonomia, que é constituída de um modelo de representação do conhecimento, discutido em Hjørland (1998).

A taxonomia constitui um modelo de representação do conhecimento, tomada de empréstimo pela Ciência da Informação, com origem na Biologia, como abordado no estudo conceitual de Aquino, Carlan e Brascher (2009). Para esses autores, o termo taxonomia

[...] derivou-se de um dos ramos da Biologia que trata da classificação lógica e científica dos seres vivos, fruto do trabalho do médico e botânico sueco Carolus Linnaeus, com sua *Biologia Sistemática* (AQUINO; CARLAN; BRASCHER, 2009, p. 204).

A associação de ideias de uma navegação facetada facilita a recuperação de informações, denominada de paradigma de navegação facetada (MACULAN; LIMA, 2011). As autoras afirmam que taxonomias proporcionam uma representação do relacionamento semântico dos objetos envolvidos na classificação.

A pesquisa de Aquino, Carlan e Brascher (2009) sobre taxonomia descreve que a informação será acessível quando estiver organizada de maneira a facilitar a busca desejada. A literatura apresenta diversas aplicações para a taxonomia. Uma destas aplicações para taxonomias é a navegacional, que destina-se a buscar a organização da informação conforme o seu público-alvo, isto é, uma forma flexível de utilização.

A proposta apresentada por Maia, Lima e Maculan (2017) descreve uma taxonomia dos tipos de relações semânticas para a organização e para a representação do conhecimento mediante uma revisão da literatura. Nesta investigação, as autoras descrevem os tipos de estruturas taxonômicas, como a corporativa e a facetada.

Para Lóscio, Burle e Calegari (2017), o enriquecimento semântico de dados é um conjunto de processos que podem ser usados para aprimorar, refinar ou melhorar os dados brutos ou previamente processados, contribuindo para tornar esses dados um ativo valioso para quase todos os negócios ou empresas. Uma proposta para o desenvolvimento de sistemas que apoia o processamento de dados brutos para a *web* é a taxonomia navegacional.

Portanto, conforme descrito por Aganette (2010), é necessário investigar as estruturas de taxonomias, tais como as taxonomias navegacionais. E por meio desses estudos identificar meios de utilizar esses instrumentos no processo de enriquecimento semântico. Segundo Sorrentino *et al.* (2013), a associação explícita entre um ou mais significados é uma ferramenta chave para o enriquecimento semântico. Utilizar o relacionamento estabelecido na taxonomia permitiria não só aumentar as possibilidades de pesquisa, mas facilitar a recuperação desses dados por sistemas automatizados.

O enriquecimento semântico destina-se a inserir conceitos a partes de um texto, de modo a criar uma estrutura para busca automática de seu significado (THAKKER *et al.*, 2012). Adotar taxonomias para ampliar o significado desses conceitos permitiria aos usuários interligar esses recursos por meio das relações existentes entre os mesmos e encontrar novos recursos. Contudo, as possibilidades de uso das taxonomias no processo de enriquecimento semântico não estão devidamente explicitadas na literatura da BCI.

Conclui-se, então, que, conforme descrito por Lopes, Aganette e Maculan (2020), “[...] a produção dos pesquisadores da área mostra-se muito dispersa e limitada a poucos trabalhos, [...]”. Assim, conforme as autoras, existe uma lacuna de pesquisas para uma área que busca consolidar-se como campo de pesquisa de instrumentos de representação da informação e do conhecimento.

3 Metodologia

A metodologia desenvolvida para a realização desta pesquisa priorizou o modo qualitativo para estabelecer o estado da arte da área da Ciência da Informação sobre o uso da taxonomia para o enriquecimento semântico de objetos. Para atingir este objetivo, optou-se pelo uso do Mapeamento Sistemático (PETERSEN *et al.*, 2007), ou seja, um levantamento de teorias, conceitos, modelos e métodos associado ao tema da pesquisa, conforme publicado na literatura.

Esse mapeamento exploratório possibilitou o aprofundamento em uma área do conhecimento e na temática de pesquisa, por meio de um processo constituído de cinco etapas, descritas a seguir.

3.1 Mapeamento Sistemático

O Mapeamento Sistemático (MS) foi realizado por meio de buscas a artigos científicos, livros, teses e dissertações que versam sobre o tema desta investigação, durante o período de janeiro a março de 2023. As etapas para o MS, conforme descrito por Petersen *et al.* (2007), são: 1 - estabelecer as questões de pesquisa; 2 - realizar as buscas em bases científicas; 3 - realizar a seleção do material (estudos) de interesse; 4 - efetuar o refinamento pelas palavras-chave e os resumos e 5 - extrair os dados e análises.

Na primeira etapa, são estabelecidas as questões de pesquisa, conforme Petersen *et al.* (2007), regularmente, deseja-se mapear as frequências de publicação ao longo do tempo para ver as tendências de uma investigação. Esses objetivos são refletidos nas questões de pesquisa, descritas na introdução, que foram estabelecidas com base no objetivo e possibilita delimitar o escopo do MS.

Ao realizar a segunda etapa, buscas em bases científicas, foram conduzidos estudos, que são pesquisas por fontes de dados científicos. Baseado na questão de pesquisa, é definido e executado um método para a consulta e recuperação destas fontes de informações.

As consultas foram realizadas em bases nacionais, internacionais e interdisciplinares nos últimos cinco anos. As bases estabelecidas para atingirem o objetivo estabelecido são: Repositório Institucional da Universidade Federal de Minas Gerais (RI/UFMG) [2], Springer Link [3], SciELO [4] e Wiley Online Library [5]. Esta consulta foi baseada em termos de busca. Uma boa forma de criar termos de busca (*strings*) é com o uso da *population, intervention, comparison, outcome and context* (PICOC) (KITCHENHAM; CHARTERS, 2007; PETTICREW; ROBERTS, 2005).

Para a seleção do material de interesse, terceira etapa, foram estabelecidos os critérios de inclusão e exclusão (PETERSEN *et al.*, 2007). Estes são usados para excluir estudos que não são relevantes para responder às questões de pesquisa. Na quarta etapa, uma análise por palavras-chave foi estabelecida pela aplicação das

strings e pelas leituras dos resumos, introduções, resultados e conclusões para a classificação dos materiais.

Para finalizar, a quinta etapa se destina a realizar a análise dos estudos identificados como fonte de dados para a investigação. Este levantamento possibilitou uma classificação do material obtido pelas etapas anteriores e é um método sistemático para selecionar apenas os artigos relevantes para o estudo em questão.

Após estabelecimento da questão de pesquisa, descrita como objetivo na seção 1, foram realizadas as buscas conforme os critérios descritos em 3.1.1 e os critérios de inclusão e exclusão em 3.1.2.

3.1.1 Critérios de busca

A estratégia de busca de pesquisas relacionadas iniciou com o estabelecimento das palavras-chave. Estas foram utilizadas de forma combinada ao utilizar operadores lógicos para uma pesquisa experimental com base no objetivo traçado. Os operadores utilizados foram o “AND” para o agrupamento dos conceitos e o “OR” para pesquisas combinadas com o mesmo conceito. O primeiro passo realizado foi estabelecer as palavras-chave para a língua inglesa (*Taxonomy*, *Semantic enrichment*) e portuguesa (Taxonomia, Enriquecimento semântico).

No segundo passo, para cada palavra-chave foi formado um conceito principal de busca para os artigos científicos e outros estudos relacionados. Os conceitos estabelecidos para a língua inglesa são: *Taxonomy*: *taxonomy*; *Semantic enrichment*: *semantic*, *semantic enrichment*. Já para a língua portuguesa, temos os seguintes conceitos: Taxonomia: *taxonomia*; Enriquecimento semântico: *semântico*, *semântica*, *enriquecimento semântico*.

Além disso, nos casos solicitados foi estabelecida a área de conhecimento das buscas como "Ciências Sociais Aplicadas" para a língua portuguesa e "*Applied Social Sciences*" para a língua inglesa. Na base de dados do RI/UFMG, foi estabelecida, também, a busca para as investigações desenvolvidas para programas de pós-graduação, como, por exemplo, o Programa de Pós-Graduação em Gestão e Organização do Conhecimento, isto devido à ausência da opção de seleção da área do conhecimento.

Como exemplo deste critério de busca é descrita, a seguir, a combinação dos conceitos das palavras-chave em uma *string* de busca para as bases de pesquisa *Springer Link* e a *SciELO*: '(*taxonomy AND (semantic enrichment OR semantic)*) OR (*taxonomia AND (enriquecimento semântico OR semântico OR semântica)*)'. Outro exemplo, para a base de pesquisa *Wiley Online Library*, é a seguinte *string*: '*for "taxonomy" in Abstract and "semantic" in Abstract and "taxonomy" in Keywords and "semantic" in Keywords*'.

3.1.2 Critérios de inclusão e exclusão

Os resultados obtidos pela busca foram filtrados na etapa três do MS. Nesta etapa, seleção do material (estudos) de interesse, foi realizada a seleção das pesquisas de interesse, atividade orientada por uma Doutora em Ciência da Informação e executada por dois alunos de doutorado. Esta seleção seguiu os critérios de inclusão e exclusão para o aceite dos estudos.

Foi considerado como critério de inclusão para as pesquisas de interesse os estudos estabelecidos com o objetivo de apresentar pesquisas sobre a taxonomia e/ou a semântica ou o enriquecimento semântico de objetos. Já os critérios de exclusão são: pesquisas que não possuem como assunto a taxonomia e/ou a semântica ou o enriquecimento semântico; pesquisas incompletas apresentadas como trabalhos com menos de quatro páginas; estudos que não estejam em português ou inglês e pesquisa já publicada e contabilizada em outro periódico ou conferência.

4 Resultados

As pesquisas selecionadas foram filtradas pelas palavras-chave, *string* de busca nas bases de dados, e, posteriormente, refinadas pelo conteúdo contido no estudo, etapa quatro do MS. Uma descrição quantitativa dos dados coletados, refinados e analisados nesta etapa, encontra-se na subseção 4.1. A subseção 4.2, revisão da literatura, apresenta a extração e a análise qualitativa dos dados, etapa cinco do MS.

4.1 Dados coletados

Após a coleta de estudos, etapa 3 do MS, que estavam alinhados com o objetivo desta pesquisa, realizou-se o refinamento pelas palavras-chave, resumos e leitura das introduções, resultados e conclusões (pela aplicação de três filtros). Os estudos foram quantificados e os valores obtidos encontram-se descritos na Tabela 1. Esta tabela apresenta nas linhas os nomes das bases de pesquisas utilizadas e nas colunas o identificador dos devidos "Filtros 1, 2 e 3". Sendo que, cada filtro apresenta a quantidade de estudos obtidos por base.

A coluna “Filtro 1” apresenta a quantidade de estudos obtidos pelas *strings* de busca nas bases de dados. Para cada uma dessas fontes de informações científicas foi realizada uma pesquisa com duas *strings* de buscas, conforme especificação das palavras-chave, conceitos de busca, e respeitando a forma de manipulação das quatro bases escolhidas para este estudo.

O quantitativo de estudos para cada uma das bases de dados, após a aplicação do filtro de leitura dos títulos e resumos, ao utilizar do critério de inclusão, encontra-se na coluna “Filtro 2” dessa Tabela 1. Por fim, a coluna “Filtro 3” apresenta o quantitativo final que foi obtido de cada base após a leitura da introdução, resultados e/ou conclusão de cada estudo.

Este último filtro considera o critério de exclusão estabelecido, descrevendo o quantitativo de estudos analisados na quinta etapa do MS. O baixo número de resultados de pesquisas, 67 estudos, já era esperado, isto devido ao período de pandemia da COVID-19 (hipótese identificada na análise dos dados a seguir) e, principalmente, a especificidade do objeto deste nosso estudo.

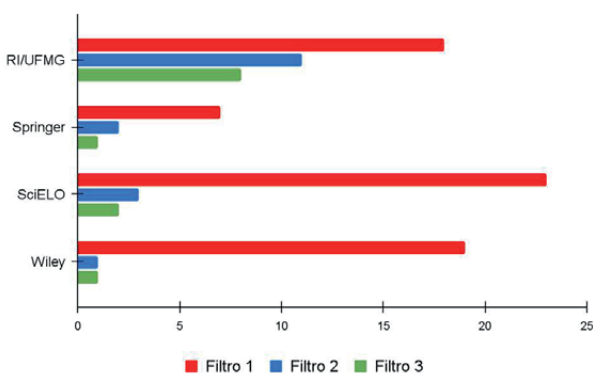
Tabela 1. *Quantitativos de estudos obtidos*

Bases	Filtro 1	Filtro 2	Filtro 3
RI/UFMG [2]	18	11	8
Springer [3]	7	2	1
SciELO [4]	23	3	2
Wiley [5]	19	1	1
Total	67	17	12

Fonte: Elaborada pelos autores (2023).

Pode-se verificar na Tabela 1 que dos 67 estudos, apenas 12 foram considerados alinhados à temática. Esse total de estudos selecionados no "Filtro 3" representa 17,9% das pesquisas obtidas no primeiro filtro. Ao analisar estes estudos selecionados por ano, após a aplicação dos três filtros, obteve-se os Gráficos 1 e 2. O Gráfico 1 apresenta a quantidade de estudos encontrados por base de dados em cada momento de aplicação desses três filtros. Pode-se constatar que, apesar da busca inicial com *strings* ter obtido um número de resultados com média simples de 16,8 estudos para as quatro bases, o aproveitamento das pesquisas ao considerar a temática foi pequeno, uma média simples de 3 estudos.

Gráfico 1. *Quantidades de estudos por filtro*



Fonte: Elaborado pelos autores (2023).

Ao analisar o Gráfico 1, conclui-se que, para a busca estabelecida, as bases de dados de pesquisas científicas com melhor aproveitamento percentual foram as do RI/UFMG e *Springer*. Na base do RI/UFMG, o aproveitamento percentual foi de 44,4% e para a *Springer*, ocorreu um aproveitamento de 14,3%. Para a

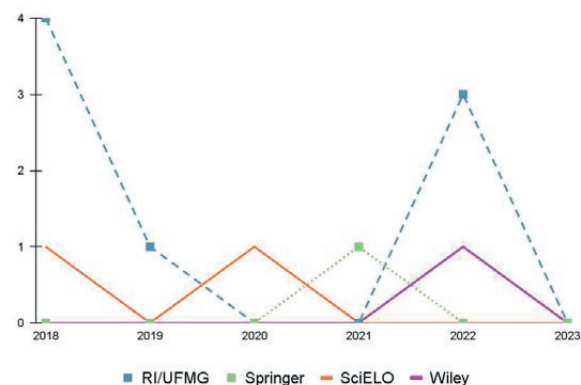
SciELO, ocorreu o aproveitamento percentual de 8,7% e na base *Wiley*, foi obtido o aproveitamento de 5,3%.

Já o Gráfico 2 apresenta a quantidade de publicações de estudos obtidas nas quatro bases de dados durante o período de 2018 a fevereiro de 2023. Percebeu-se o aproveitamento de mais estudos nas fontes da RI/UFMG, SciELO e *Wiley* em 2018 e 2022. Sendo que, como destaque, no ano de 2018, o processo de busca e a aplicação dos filtros resultou no total de 5 estudos e em 2022, este processo obteve 4 estudos.

Deve-se considerar, conforme Tabela 1, que em números absolutos as bases do RI/UFMG e SciELO representam o melhor retorno, com respectivamente 8 e 2 estudos, após a aplicação do "Filtro 3". Atribuímos este resultado expressivo no aproveitamento de estudos na base da RI/UFMG a elementos isolados da temática de pesquisa, como, por exemplo, trabalhos ligados somente ao SOC, pertencente ao objetivo de nosso estudo, a taxonomia.

Percebe-se também na análise do Gráfico 2 que a produção científica para a temática diminuiu para o período de 2019 a 2021, uma possível explicação seria a pandemia de COVID-19, pois esta ocorreu no mesmo período.

Gráfico 2. *Quantidades de estudos das bases de dados*



Fonte: Elaborado pelos autores (2023).

Além dos elementos quantitativos apresentados nesta subseção, a análise dos dados coletados proporcionou o Quadro 1 apresentado no Apêndice A. Esse quadro descreve os estudos obtidos pelo "Filtro 3" para suas respectivas temáticas de pesquisa e as interpretações sobre os objetivos traçados pelos autores. Deve-se observar que entre os resultados verificou-se que há artigos e trabalhos de pesquisas mais complexos e extensos, como dissertações e teses.

Estes estudos abordam de forma isolada a taxonomia, a representação semântica e o enriquecimento semântico, em suas devidas particularidades. Porém, nenhum dos estudos encontrados apresentou uma pesquisa em profundidade sobre a taxonomia e sua relação com o enriquecimento semântico de objetos. Estes fatos

quantitativos, descritos anteriormente, e a análise qualitativa, do Quadro 1, comprovam a necessidade de ampliar os estudos nessa temática de pesquisa estabelecida. Uma análise qualitativa inicial desse Quadro 1 encontra-se descrita na subseção 4.2 - Estado da arte.

4.2 Estado da arte

A aplicação do MS resultou em publicações cujas análises apontaram para o estado da arte na literatura para os últimos cinco anos, em quatro bases de dados de pesquisas científicas. Isto permitiu estabelecer um retrato da temática estabelecida e contemplar os principais conceitos que circundam o objeto deste projeto de investigação. Assim, os estudos contemporâneos descritos, a seguir, contextualizam uma análise primária para o objetivo estabelecido, o uso da taxonomia para o enriquecimento semântico de objetos.

A análise do estado da arte foi baseada nas pesquisas obtidas, que estão resumidas no Quadro 1, entre essas, como exemplos, descrevemos os estudos destinados aos SOC (LOPES; AGANETTE; MACULAN, 2020), a processos e à representação semântica (LISKA, 2018; OLIVEIRA, 2018). Além disto, como destaque, apresentamos os estudos da taxonomia (ANDERSON *et al.*, 2022; SANTOS, 2018; MENDES, 2022; VIANA, 2022) e do enriquecimento semântico (NAZÁRIO e DIAS, 2018).

Uma pesquisa que abordou a taxonomia foi a de Mendes (2022). Neste estudo, a taxonomia é apresentada como um

[...] processo de classificação e agrupamento categorizado, que define-se como organização, ou seja, partindo de um assunto cria-se classes gerais que se dividem em classes e subclasses, hierarquicamente, formando uma lista de assunto estruturada” (MENDES, 2022, p. 44).

Neste contexto, temos então a taxonomia corporativa e a facetada. Lopes, Aganette e Maculan (2020) descrevem que a taxonomia corporativa pode existir em distintos ambientes na organização, pois ela

[...] estrutura e organiza o conhecimento registrado em poder das organizações, organiza produtos e serviços, e se constitui como um mecanismo de apoio à gestão. Ela é construída de modo a atender a um público particular, servindo de interface sistema-usuário por meio da navegação (LOPES; AGANETTE; MACULAN, 2020, p. 164).

Ainda segundo Lopes, Aganette e Maculan (2020), ressalta-se a teoria da classificação facetada como aporte teórico para a construção de taxonomias:

A taxonomia facetada emprega os princípios da Teoria da Classificação Facetada, desenvolvida por Ranganathan (1967), que foi criada com o objetivo de auxiliar na organização de materiais bibliográficos (LOPES; AGANETTE; MACULAN, 2020, p. 164).

As autoras apontam também as vantagens do uso de classes e subclasses de assunto para a recuperação da informação:

Na construção de instrumentos de representação da informação e do conhecimento, a teoria fornece cinco categorias mais abstratas de mapeamento do domínio representado, formando-se múltiplas hierarquias de classes (facetadas) e subclasses (subfacetadas) que exprimem as diferentes dimensões desse dado domínio. Esse modo de classificação permite que o usuário encontre primeiro o assunto mais amplo, e, através dele, chegue à informação mais restrita. Essa navegação torna mais acessível e prática a recuperação da informação por parte do usuário (LOPES; AGANETTE; MACULAN, 2020, p. 164-165).

Duas pesquisas que também tratam com mais profundidade, de forma contemporânea, a taxonomia corporativa e a taxonomia facetada são as de Mendes (2022) e a de Viana (2022). A pesquisa de Mendes (2022) investiga a taxonomia corporativa e facetada. Este estudo concluiu que para iniciar uma pesquisa nesta temática é necessário verificar o foco nas taxonomias corporativas, como vocabulário controlado, para a representação da informação, quando se pesquisa um ambiente organizacional. Esta última autora também conclui que é necessário

[...] utilizar outro insumo tecnológico para o desenvolvimento e até mesmo para o aprimoramento do uso da taxonomia corporativa facetada, como, por exemplo, algum sistema informacional [...] (MENDES, 2022, p. 78).

A investigação de Viana (2022) propõe a taxonomia facetada em ambientes corporativos, considerando-a como instrumento na gestão e na organização da informação. Além disso, Viana (2022) discute os benefícios do uso da taxonomia e da classificação, possibilitando confirmar os resultados dos estudos de Aganette (2010) ao evidenciar que as taxonomias têm representado soluções para encontrar informações em diversos ambientes.

As pesquisas de Viana (2022) e Mendes (2022) indicam que a Teoria da Classificação Facetada, conforme o contexto de necessidades, é uma solução para a melhoria dos sistemas de classificação corporativos, considerando que existem várias possibilidades de classificar assuntos de documento. A citada pesquisa conclui também que, apesar da escassez de estudos, a literatura da Ciência da Informação indica a necessidade de utilizar os princípios classificatórios explícitos para o desenvolvimento de taxonomias.

Já sobre a *web* semântica e a interoperabilidade semântica, duas pesquisas contemporâneas com profundidade que foram encontradas são as de Firmino (2019) e Emygdio (2021). Contudo, apesar da importância das pesquisas para os SOC e do relacionamento com a temática estabelecida, estes estudos não abordam diretamente o enriquecimento semântico.

O enriquecimento semântico foi discutido na pesquisa de Nazário (2018) e em Nazário e Dias (2018), que apresentou uma proposta metodológica para avaliar o enriquecimento semântico de objetos publicados na *web* através do *Linked Data* [6]. Neste estudo, foi desenvolvida uma Matriz de Técnicas e recursos para o enriquecimento semântico de objetos (Matriz Treso), que foi utilizada na avaliação dos modelos de dados *Europeana Data Model* (EDM) [7] e *Bibliographic Framework Initiative* (BIBFRAME) [8]. A pesquisa também apontou a necessidade de aplicação da matriz a outros modelos de dados, bem como das possibilidades de uso da metodologia proposta para criar novos modelos de dados enriquecidos semanticamente.

Já o estudo de Maia (2018) apresentou um modelo de extração de relações semânticas para a representação do conhecimento de documentos acadêmicos no contexto do idioma português brasileiro, implementado em um sistema *web* chamado Semantizar, que realizou o enriquecimento semântico da amostra pesquisada.

Importante destacar que apesar da relevância do enriquecimento semântico para a recuperação e organização do conhecimento no contexto da BCI, foram poucas as pesquisas obtidas para o período de 2018 a fevereiro de 2023 nas bases selecionadas que abordam esse enriquecimento semântico. Os filtros de busca recuperaram, também, artigos sobre relacionamento semântico, semântica e taxonomia. Portanto, esta pesquisa é um estudo em forma de primeira análise sobre os dados coletados, e não foi possível contextualizar de forma completa, nem tão pouco descrever o estado da arte desejado, ao considerar o universo pesquisado em seu contexto.

5 Considerações finais

Este artigo apresenta a análise inicial dos dados que estão sendo coletados para duas pesquisas, que proporcionarão trabalhos com mais profundidades nas temáticas estabelecidas. Podemos concluir que, conforme dados quantitativos e afirmado por Lopes, Aganette e Maculan (2020), existe uma carência de pesquisas sobre os SOC e essa temática, fato comprovado pelo número de pesquisas encontradas.

Ressalta-se que tanto a temática do enriquecimento semântico quanto a de taxonomias e suas possibilidades de utilização para favorecer o processo de enriquecimento semântico demandam mais estudos, tendo em vista que na amostra analisada não foi encontrado nenhum resultado que abordasse esses temas de modo conjunto. Isto é, nenhum dos estudos encontrados apresentou uma pesquisa em profundidade sobre a taxonomia e sua relação com o enriquecimento semântico de objetos. A importância do tema pressupõe a necessidade de seu aprofundamento, sobretudo pelo fato cada vez mais constante da disseminação de conteúdos e informações sem qualquer controle na *web*.

Assim, estudos futuros precisam ser desenvolvidos com o objetivo de identificar necessidades, tecnologias e metodologias para o enriquecimento semântico de dados publicados na *web*. Além disso, pretende-se ampliar a busca realizada para o estado da arte da temática estabelecida para dez anos em mais bases de dados de pesquisas científicas.

Como trabalho futuro, pretende-se, também, ampliar essa pesquisa do estado da arte para outras temáticas relacionadas, como as taxonomias em sistemas educacionais e sua influência no desenvolvimento da competência informacional, propondo um modelo taxonômico para isso. Para o enriquecimento semântico de objetos, pretende-se avaliar se de fato o uso de taxonomias e ontologias pode contribuir para o seu objetivo e ainda propor um novo modelo para esse enriquecimento semântico de dados no contexto corporativo para publicação de dados em *linked data*.

Notas

- [1] ANSI/NISO Z39.19-2005: **Guidelines for the construction, format, and management of monolingual controlled vocabularies**. Bethesda: NISO, 172 p. 2005. Disponível em: <https://www.niso.org/publications/ansiniso-z3919-2005-r2010>. Acesso em: 28 mar. 2023.
- [2] Repositório Institucional da Universidade Federal de Minas Gerais (RI/UFMG) é um sistema que possibilita o acesso à base de dados de produções científicas da Universidade Federal de Minas Gerais. Disponível em: <https://repositorio.ufmg.br>. Acesso em: 22 mar. 2023.
- [3] *Springer Link* é uma editora mundial baseada na publicação de livros texto, livros de referência acadêmica, e periódicos de artigos, isto é, uma base de dados de estudos científicos. Disponível em: <https://link.springer.com>. Acesso em: 22 mar. 2023.
- [4] *Scientific Electronic Library Online* (SciELO) é uma biblioteca digital de livre acesso e modelo. Disponível em: <https://www.scielo.org/>. Acesso em: 22 mar. 2023.
- [5] *Wiley Online Library* concentra suas atividades em diferentes áreas da ciência, publica material para a comunidade científica, para pesquisadores, práticos (técnicos) e estudantes. Disponível em: <https://www.wiley.com/en-us>. Acesso em: 22 mar. 2023.
- [6] *Linked Data* são princípios e boas práticas para publicação de dados na *web*. Estes princípios denominados *Linked Data* têm como objetivo fundamental facilitar a integração de dados de diferentes fontes, de forma a torná-los compreensíveis também para as máquinas (BIZER; HEATH; BERNERS-LEE, 2009).
- [7] *Europeana Data Model* (EDM) é um modelo para dados proposto pela European - uma biblioteca virtual com a maior coleção *online* de arte. Sendo considerado a *Europeana Data Model primer* (2013) como referência. Disponível em: https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf. Acesso em: 27 mar. 2023.

[8] *Bibliographic Framework Initiative* (BIBFRAME) é um modelo de metadados que foi desenvolvido para publicar os registros bibliográficos na *web*, utilizando os princípios do *Linked Data*. Disponível em: <http://www.loc.gov/bibframe/>. Acesso em: 27 mar. 2023.

Referências

ANDERSON Elizabeth J.; MIDGLEY Katherine J.; HOLCOMB Phillip J.; RIÉS, Stephanie K. Taxonomic and thematic semantic relationships in picture naming as revealed by Laplacian-transformed event-related potentials. *Psychophysiology*, [S.l.], v. 59, n. 11, Nov. 2022.

AGANETTE, Elisângela Cristina. **Taxonomias corporativas**: um estudo sobre definições e etapas de construção fundamentado na literatura publicada. 2010. 104f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010. Disponível em: <https://repositorio.ufmg.br/handle/1843/ECID-87BJSR>. Acesso em: 19 jan. 2023.

AQUINO, Idalécio; CARLAN, Eliana; BRASCHER, Marisa. Princípios classificatórios para a construção de taxonomias. **PontodeAcesso**, Salvador, v. 3, n. 3, p. 196-215, dez. 2009. Disponível em: <https://www.periodicos.ufba.br/index.php/revistaici/article/view/3626/2744>. Acesso em: 20 jan. 2023.

AUSTIN, Patrik. Theory of language: a taxonomy. *SN Soc Sci*, [S.l.], v. 1, n. 78, p. 1-24, 2021. Disponível em: <https://doi.org/10.1007/s43545-021-00085-x>. Acesso em: 26 mar. 2023.

BARBOSA, Alice Príncipe. Classificações facetadas. **Ciência da Informação**, Brasília, DF, v. 1, n. 2, p. 73-81, jan. 1972. Disponível em: <http://revista.ibict.br/ciinf/article/view/10/10>. Acesso em: 07 fev. 2023.

BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked Data: The story so far. **Semantic services, interoperability and web applications**: emerging concepts. United States of America: 2009. p. 205-227. Disponível em: https://www.researchgate.net/publication/225070216_Linked_Data_The_Story_so_Far. Acesso em: 27 mar. 2023.

BRASCHER, Marisa; CARLAN, Eliana. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO, J.; BRASCHER, M. (org.). **Passeios pelos bosques da informação**: estudos sobre representação e organização da informação e do conhecimento. Brasília: IBICT, 2010. p. 147-176.

DIAS, Célia da Consolação; BARROSO, Juliana Martins de Castro; ROSADO, Keila Mara Lara; SANTOS, Sarah Rúbia de Oliveira. SOC e Interoperabilidade: desafios e tendências no âmbito da Conferência Internacional da ISKO. **FRC: Fronteiras da Representação do Conhecimento**, Belo Horizonte, v. 1, n. 2, p. 1-26, 2022.

EMYGDIO, Jeanne Louize. **Interoperabilidade semântica orientada por ontologia para a Ciência da Informação**: a metodologia Onto4All-Interoperability como resultado de estudo de caso no domínio de energia. 2021. 325f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2021. Disponível em:

<https://repositorio.ufmg.br/handle/1843/42050>. Acesso em: 19 fev. 2023.

FAGUNDES, Priscila Basto; FREUND, Gislaiane Parra; VITAL, Luciane Paula; BARROS, Camila Monteiro de; MACEDO, Douglas Dyllon Jeronimo de. Taxonomias, ontologias e tesaurus: possibilidades de contribuição para o processo de Engenharia de Requisitos. **Em Questão**, Porto Alegre, v. 26, n. 1, p. 237-254, jan. 2020.

FIRMINO, Geraldo José Rodrigues. **Reúso de recursos da web semântica para a construção de vocabulários controlados no contexto da ciência da informação**. 2019. 175f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2019. Disponível em: <https://repositorio.ufmg.br/handle/1843/32372>. Acesso em: 19 fev. 2023.

HJØRLAND, Birger. Theory and metatheory of information science: a new interpretation. *Journal of Documentation*, [S.l.], v. 54, n. 5, p. 606-621, dez. 1998. Disponível em: <http://www.emeraldinsight.com/doi/abs/10.1108/EUM0000000007183>. Acesso em: 19 mar. 2023.

HJØRLAND, Birger. Concept theory. *Journal of the American Society for Information Science and Technology*, [S.l.], v. 60, n. 8, p. 1519-1536, 2009. Disponível em: <https://dl.acm.org/doi/abs/10.5555/1572678.1572685>. Acesso em: 22 mar. 2023.

KITCHENHAM, Barbara; CHARTERS, Stuart. **Guidelines for performing systematic literature reviews in software engineering**. Technical Report EBSE-2007-01. School of Computer Science and Mathematics: Keele University. 2007. Disponível em: https://www.elsevier.com/_data/promis_misc/525444systematicreviewsguide.pdf. Acesso em: 18 fev. 2023.

LIMA, Flávia Danielle Santiago; GOMES NETO, José Mário Wanderley. Autocontenção à brasileira? Uma taxonomia dos argumentos jurídicos (e estratégias políticas?) explicativo(a)s do comportamento do STF nas relações com os poderes majoritários. **Revista de Investigações Constitucionais**, Curitiba, v. 5, n. 1, p. 221-247, jan./abr. 2018.

LIRA, Márcio Ângelo Bezerra de. **Uma abordagem para enriquecimento semântico de metadados para publicação de dados abertos**. 2014. 94f. Dissertação (Mestrado em Ciência da Computação) - Centro de Informática, Universidade Federal de Pernambuco, Recife, 2014. Disponível em: <https://repositorio.ufpe.br/handle/123456789/11570>. Acesso em: 24 mar. 2023.

LISKA, Geraldo José Rodrigues. **O estudo do léxico na sala de aula [manuscrito]**: investigação do ensino dos processos semânticos de formação de palavras sob a perspectiva da Semântica de Contextos e Cenários (SCC). 2018. 265f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: <http://hdl.handle.net/1843/LETR-AYKMP9>. Acesso em: 19 fev. 2023.

- LÓSCIO, Bernadette Farias; BURLE, Caroline; CALEGARI, Newton. **Data on the Web Best Practices**. [online]. W3C, 2017. Disponível em: <https://www.w3.org/TR/dwbp>. Acesso em: 27 mar. 2023.
- LOPES, Pâmela; AGANETTE, Elisângela; MACULAN, Benildes. Taxonomia corporativa e taxonomia facetada: usos e aplicações na ciência da informação no Brasil. **Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información**, [S.l.], v. 34, n. 82, p. 159-173. 2020. DOI <http://dx.doi.org/10.22201/iibi.24488321xe.2020.82.58062>. Acesso em: 22 mar. 2023.
- MAIA, Lucinéia Souza. **Extração e explicitação de relações semânticas para a representação do conhecimento de documentos acadêmicos: um estudo de caso a partir de uma estrutura classificatória**. 2018. 248f. Tese (Doutorado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: <https://repositorio.ufmg.br/handle/1843/VAFA-BB5JHW>. Acesso em: 19 fev. 2023.
- MACULAN, Benildes Coura Moreira dos Santos; LIMA, Gercina Ângela Borém de Oliveira. Taxonomia facetada navegacional: agregando valor às informações disponibilizadas em bibliotecas digitais de teses e dissertações. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 12., 2011, Brasília. **Anais [...]**. Brasília: ANCIB, 2011. p. 696-714. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/174991>. Acesso em: 20 fev. 2023.
- MAIA, Lucinéia Souza; LIMA, Gercina Ângela Borém de Oliveira; MACULAN, Benildes Coura Moreira dos Santos. Taxonomia dos tipos de relações semânticas para a organização e a representação do conhecimento: uma proposta a partir da literatura. **Tendências da Pesquisa Brasileira em Ciência da Informação**, João Pessoa, v. 10, n. 1, p. 1-22, jan./jun. 2017. Disponível em: <https://brapci.inf.br/index.php/res/v/104795>. Acesso em: 21 jan. 2023.
- MENDES, Paula Raphisa. **Taxonomia Corporativa Facetada no Apoio ao Processo de Avaliação para Autorização de Cursos de Instituições de Ensino Superior pelo INEP**. 2022. 99f. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2022. Disponível em: <http://hdl.handle.net/1843/48840>. Acesso em: 20 mar. 2023.
- NAZÁRIO, Claudiane Emanuele. **Proposta de metodologia para avaliar o enriquecimento semântico de objetos publicados na web através do Linked Data**. 2018. 140f. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: <https://repositorio.ufmg.br/handle/1843/ECIP-B3UJKW>. Acesso em: 19 fev. 2023.
- NAZÁRIO, Claudiane Emanuele; DIAS, Célia da Consolação. Proposta para avaliação das técnicas e dos recursos para o enriquecimento semântico de objetos publicados em Linked Data. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 11., 2018. Londrina. **Anais [...]**. Londrina: ANCIB, 2018.
- OLIVEIRA, Elaine Diamantino. **Proposta de um modelo de hipertexto com abordagem semântica para a representação do conhecimento no domínio temático da Intensificação Agropecuária**. 2018. 227f. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: <https://repositorio.ufmg.br/handle/1843/BUOS-B2JP2A>. Acesso em: 20 mar. 2023.
- PETERSEN, Kai; FELDT, Robert; MUJTABA, Shahid; MATTSSON, Michael. Systematic mapping studies in software engineering. In: INTERNATIONAL CONFERENCE ON EVALUATION AND ASSESSMENT IN SOFTWARE ENGINEERING, 12., 2007, Swindon. **Proceedings [...]**. Swindon: BCS, 2007. p. 68-77. Disponível em: <https://dl.acm.org/doi/10.5555/2227115.2227123>. Acesso em: 10 jan. 2023.
- PETTICREW, Mark; ROBERTS, Helen. **Systematic Reviews in the Social Sciences: A Practical Guide**, Blackwell Publishing, 2005. ISBN 1405121106. Disponível em: <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470754887>. Acesso em: 18 fev. 2023.
- RANGANATHAN, Shivali Ramamrita. **Colon classification, basic classification**. 6. ed. New York: Bombay, 1960.
- RANGANATHAN, Shiyali Ramamrita. **Prolegomena to Library Classification**. Bom- bay: A. P. House, Ed. 1967.
- SANTOS, Thais Helen do Nascimento. A taxonomia e a folksonomia na representação da informação de fotografias. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 23, n. 1, p. 89-10, 2018. Disponível em: <https://doi.org/10.1590/1981-5344/2395>. Acesso em: 10 mar. 2023.
- SORRENTINO, Serena; BERGAMASCHI, Sonia; FUSARI, Elisa; BENEVENTANO, Domenico. Semantic Annotation and Publication of Linked Open Data. In: INTERNATIONAL CONFERENCE – ICCSA, 13., 2013, **Proceedings [...]**. Ho Chi Minh City, Vietnam, 2013. p. 462-474. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-39640-3_34. Acesso em 25 mar. 2023.
- THAKKER, Dhavalkumar; DESPOTAKIS, Dimoklis; DIMITROVA, Vania; LAU, Lydia; BRNA, Paul. Taming Digital Traces for Informal Learning: A Semantic-driven Approach. In: EUROPEAN CONFERENCE ON TECHNOLOGY ENHANCED LEARNING, 7., 2012, Berlin. **Proceedings [...]** Heidelberg: Springer-Verlag, 2012. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-33263-0_27 Acesso em: 26 mar. 2023.
- VIANA, Jordânia Quintão. **Metodologia para a construção de taxonomia corporativa facetada**. 2022. 168f. Dissertação (Mestrado em Gestão e Organização do Conhecimento) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2022. Disponível em: <https://repositorio.ufmg.br/handle/1843/44638>. Acesso em: 19 fev. 2023.



**DIMENSÃO POLÍTICA E SOCIAL DA
ORGANIZAÇÃO DO CONHECIMENTO**

*Political and Social Dimension
of the Knowledge Organization*

Aspectos éticos da organização e representação do conhecimento no contexto das políticas de desenvolvimento em coleções de obras raras

Ethical aspects of knowledge organization and representation in the context of rare works collection development policies

Júlio César Silveira Tauil (1), Rosane Suely Alvares Lunardelli (2), Francisco Carlos Paletta (3)

(1) Universidade Estadual de Londrina – UEL, Rodovia Celso Garcia Cid, PR- 445, Km 380 – Campus Universitário - Londrina - Paraná – Brasil, jtauil86@gmail.com, (2) lunardelli@uel.br, (3) fcpaletta@usp.br.

Resumo

Introdução: As coleções de obras raras e especiais no Brasil estão espalhadas em inúmeras bibliotecas universitárias, públicas e especializadas, suas características específicas se diferem das coleções gerais. As coleções especiais são caracterizadas por várias tipologias, sendo as obras raras, a principal tipologia que compõem as coleções especiais. Não existe uma padronização universal definida a respeito dos critérios de raridade, há diversas definições internacionais que servem de modelo para as instituições brasileiras, nesse sentido as tipologias se diversificam nas bibliotecas analisadas. **Objetivo:** Investigar documentos públicos que evidenciam aspectos éticos nas políticas de acesso das coleções de obras raras e especiais. **Metodologia:** Análise exploratória e descritiva, a partir da seleção de vinte e sete bibliotecas analisadas no levantamento publicado em 2021 por Greenhalgh e Greenhalgh no artigo 'Coleções Especiais: uma análise da formação dos acervos nas bibliotecas brasileiras', buscou-se mapear os principais documentos disponibilizados com o intuito de analisar a relação das políticas de desenvolvimento e acesso das coleções de obras e especiais com aspectos éticos da representação e organização do conhecimento nos sítios institucionais das bibliotecas. **Resultados:** Não foram encontrados documentos públicos relacionados com a temática em questão. **Conclusão:** O posicionamento público das instituições no desenvolvimento ético das políticas de acesso das coleções de obras raras e especiais não está evidenciado, sendo necessárias discussões para a elaboração de diretrizes éticas que norteiem a representação e organização dos acervos, e consequentemente a publicação na internet em suas páginas oficiais, buscando assim, estabelecer um elo de maior transparência com a comunidade de usuários.

Palavras-chave: Coleções de Obras Raras. Políticas de acesso e desenvolvimento. Aspectos éticos. Organização e Representação do Conhecimento.

Abstract

Introduction: The collections of rare and special works in Brazil are spread across numerous university, public and specialized libraries, their specific characteristics differ from the general collections. Special collections are characterized by several typologies, being rare works the main typology that make up special collections. There is no universally defined standardization regarding rarity criteria, there are several international definitions that serve as models for the Brazilian institutions; in this sense, the typologies diversify in the libraries analyzed. **Objective:** To investigate public documents that highlight ethical aspects in the access policies of rare and special works collections. **Methodology:** Exploratory and descriptive analysis, based on the selection of twenty-seven libraries analyzed in the survey published in 2021 by Greenhalgh and Greenhalgh in the article 'Special Collections: an analysis of the formation of the collections in Brazilian libraries', the aim was to map the main documents made available in order to analyze the relationship of the development and access policies of the collections of works and special collections with ethical aspects of the representation and organization of knowledge in the libraries' institutional sites. **Results:** No public documents related to the topic in question were found. **Conclusion:** The public positioning of institutions in the ethical development of access policies for collections of rare and special works is not evidenced, requiring discussions for the development of ethical guidelines that guide the representation and organization of collections, and consequently the publication on the Internet in their official websites, thus seeking to establish a link of greater transparency with the user community.

Keywords: Rare Works Collections. Access and development policies. Ethical considerations. Knowledge Organization and Representation.

Resumen

Introducción: Las colecciones de obras raras y especiales en Brasil están diseminadas por numerosas bibliotecas universitarias, públicas y especializadas, sus características específicas difieren de las colecciones generales. Las colecciones especiales se caracterizan por varias tipologías, siendo las obras raras la principal tipología que compone las colecciones especiales. No existe una estandarización universalmente definida en cuanto a los criterios de rareza, existen varias definiciones internacionales que sirven de modelo para las instituciones brasileñas, en este sentido las tipologías se diversifican en las bibliotecas analizadas. **Objetivo:** Investigación de documentos públicos que ponen de relieve aspectos éticos en las políticas de acceso a las colecciones de obras raras y especiales. **Metodología:** Análisis exploratorio y descriptivo, a partir de la selección de veintisiete bibliotecas analizadas en la encuesta publicada en 2021 por Greenhalgh y Greenhalgh en el artículo "Coleções Especiais: uma análise da formação dos acervos nas bibliotecas brasileiras", se buscó mapear los principales documentos puestos a disposición con el objetivo de analizar la relación

de las políticas de desarrollo y acceso de las colecciones de obras y colecciones especiales con aspectos éticos de la representación y organización del conocimiento en los sitios institucionales de las bibliotecas. **Resultados:** No se encontraron documentos públicos relacionados con el tema en cuestión. **Conclusión:** No se evidencia el posicionamiento público de las instituciones en el desarrollo ético de las políticas de acceso de las colecciones de obras raras y especiales, siendo necesarias discusiones para la elaboración de directrices éticas que orienten la representación y organización de las colecciones, y consecuentemente la publicación en Internet en sus páginas oficiales, buscando de esta forma, establecer un vínculo de mayor transparencia con la comunidad de usuarios.

Palabras clave: Colecciones de obras raras. Políticas de acceso y desarrollo. Aspectos éticos. Organización y representación del conocimiento.

1 Introdução

Um fortuito ciclo de vida dos documentos nas unidades de informação está diretamente interligado ao acesso, uso e disseminação dos acervos e coleções. Por intermédio de consistentes aportes teóricos e metodológicos da Biblioteconomia, Museologia e Arquivologia, um dos principais procedimentos da Ciência da Informação (CI) reside na estruturação da informação, visando torná-la acessível para qualquer usuário que necessitar dela (LUNARDELLI et al., 2019). Dentro deste arcabouço informacional se incluem os processos designados às políticas de desenvolvimento e acesso, e conseqüentemente, os aspectos éticos da organização e representação do conhecimento no que se refere às Coleções de obras raras especiais (CORES).

CORES são partes integrantes do patrimônio histórico-cultural (REIFSCHNEIDER, 2008), o conceito de obra geralmente está mais relacionado ao livro, porém é possível incluir vários outros tipos de documentos impressos, como por exemplo, mapas, periódicos e cartões-postais (SANT'ANA, 2001). Coleções especiais geralmente são definidas por meio de variedades de funções exercidas pelos acervos que as adicionam, e são criadas pelas instituições baseadas numa motivação, ancorados num propósito (GREENHALGH; GREENHALGH, 2021).

Nas unidades de informação, como bibliotecas públicas ou universitárias, as CORES devem ser caracterizadas como um elemento peculiar que integra uma estrutura mais ampla (SANT'ANA, 2001), isto significa que estas coleções não fazem parte da coleção geral (FARIA; PERICÃO, 2008). De acordo com a *University of Glasgow* (2019) estes acervos são caracterizados por introduzir valores de pesquisa e cultura. As definições de raridade de uma coleção especial podem se relacionar com o valor histórico, artístico, científico, cultural, até mesmo a partir de sua interligação com instituições e personalidades (ASSOCIATION OF RESEARCH LIBRARIES, 2003).

O ponto nevrálgico envolvendo estudos relacionados aos aspectos éticos no contexto da organização e representação do conhecimento é verificada na atuação profissional na unidade de informação, se tornando uma questão fundamental na área de CI, cooperando

em grande medida para a preservação e disseminação da memória cultural. (MILANI; GUIMARÃES, 2011).

Desenvolver critérios éticos no gerenciamento da organização e da representação do conhecimento aplicados num acervo geral de biblioteca, por si só já é uma tarefa complexa, e a extensão desta complexidade poderá ser ainda mais específica e peculiar no que se refere ao universo das coleções especiais de obras raras. O primeiro desafio gira em torno da abrangência de critérios adotados pelas instituições e autores na definição de livro raro (BATISTA, 2012; NARDINO; CAREGNATO, 2005; ANDRADE; CANTALINO, 2003; PINHEIRO, 1989).

A falta de profissionais da informação especializados em trabalhar com este tipo de coleção consiste num outro importante desafio, um profissional especializado desenvolve técnicas muito específicas no funcionamento das atividades do acervo, além dos aspectos técnicos, este profissional irá estudar o seu acervo, a história do seu arranjo, dentre outras particularidades (ORDOVÁS; STEINDEL, 2017). O profissional da informação que desenvolve a curadoria de CORES precisa ter conhecimento do assunto e do acervo (SILVA, 2017).

Compreender a relação dos itens, artigos, livros e demais artefatos com seus proprietários originários, sejam eles bibliófilos ou as antigas instituições de guarda, também se torna providencial ao se traçar um paralelo do recorte histórico, político e social do arranjo do acervo com a organização e representação do conhecimento na atualidade. Dessa forma, de acordo com Guimarães e Pinho (2006), destacar discussões e reflexões que debatam a dimensão social e política permeadas por aspectos éticos na organização do conhecimento é uma prática de suma importância no escopo da CI.

2 Aspectos éticos envolvendo mediação cultural e organização e representação do conhecimento

Conforme sublinhado por Borko (1968), a CI estuda um grupo de conhecimentos que contemplam uma gama de aspectos, como: origem, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação. A organização do conhecimento é direcionada no sentido de construir modelos de mundo caracterizados

pelas abstrações da realidade (BRASCHER; CAFÉ, 2008).

A informação é passível de ser organizada, processada, e reorganizada por meio de uma gama de critérios, além de ser adicionada a outras informações, isto é, um sistema pode se tornar ainda mais organizado como resultado de mais informações introduzidas nele (ROBREDO, 2003).

A composição das CORES pode ser traçada pelo seu valor, o qual pode sofrer alterações no decorrer dos anos. Há sempre a possibilidade de as obras ganharem outros status no que se refere ao valor patrimonial, social, histórico e de memória. Dependendo do contexto, da corrente de pensamento e ideologia, os livros podem perder ou ganhar valores sociais simbólicos, independentemente das características físicas, da cronologia, ou mesmo da temática (BERGER, 2014).

Algumas diretrizes de políticas de indexação desenvolvidas por Dal'avedove e Fujita (2017) apontam para as diferentes realidades e características das mais diversas bibliotecas existentes, afinal cada instituição pode se deparar com situações cotidianas distintas.

Beghtol (2002 apud GUIMARÃES; PINHO, 2006) apontava para a necessidade dos sistemas de classificação enquanto ferramentas de representação do conhecimento, desenvolverem uma estrutura relativamente flexível no que se refere à inclusão de conceitos que reflitam a diversidade cultural.

[...] os problemas de natureza ética que destacam, a partir da literatura e da prática profissional, são os que se referem ao preconceito na análise ou na representação e, ainda, nos sistemas de classificação, cujas hierarquizações estabelecem desrespeito entre diferentes culturas, reforçando a ideia de preponderância ou revelando de certa maneira proselitismo. Assim, ao profissional da informação cabe o encargo de assegurar a diversidade de acesso às informações culturais e pessoais, onde atuam no sentido de projetar, avaliar, dar manutenção e revisar os sistemas de classificação, de forma que esses se ajustem aos princípios éticos. (GUIMARÃES et al, 2008, p.148).

Diante dessa perspectiva, espera-se que os profissionais que já atuam na Área, e estudantes de graduação em Arquivologia e Biblioteconomia, futuros profissionais da área, tenham em mente os aspectos éticos que envolvem atividades importantes da Organização do Conhecimento e de como a dimensão ética é intrínseca ao desenvolvimento dessas ações, considerando ainda o 'poder' de ofertar acesso (SILVA; EVANGELISTA; GUIMARÃES; 2019).

Em 2002 Garcia Gutiérrez propôs a Epistemografia Interativa, este paradigma busca substituir o termo profissional da informação pelo termo mediador a partir de uma visão transdisciplinar, transdiscursiva e transcultural, abarcando dessa forma, todas as

características possíveis de um conteúdo documental, sublinhando a influência do mediador nos processos de representação e organização do conhecimento, e na etapa de tomada de decisão. Não existe possibilidade de neutralidade em tais operações, pois estas são profundamente mediadas (MILANI; GUIMARÃES, 2011).

Os aspectos éticos da representação e organização podem ajudar a resolver os graves problemas de falta de inclusão existentes na biblioteca. Afinal, boa parte da população quando adentram nesses espaços não se sente incluída, acolhida.

Um problema persistente no acesso às instituições culturais e na própria definição do que é cultural está no fato do público em geral não se reconhecer nesses locais. Todos os locais representam uma ideologia, no entanto, as instituições que detêm informações de interesse do povo, geralmente as oferecem sob um olhar distinto da grande massa. (SILVA; SABBAG, 2019, p.259).

Os livros ainda representam uma forma de dominação cultural e social (MACEDO, 2004), e as maneiras como são organizados e representados os conhecimentos dentro das unidades de informação, podem reverberar em formas de exclusão por meio de diversos tipos de preconceitos. O surgimento da Biblioteca Pública no Brasil no século XIX é concomitante com o alto índice de analfabetismo, pois a maior parcela da população brasileira não sabia ler, e apesar da taxa de analfabetismo diminuir vertiginosamente no século XX, seu principal objetivo consistia em apoiar à educação das classes dominantes tidas como cultas, e esta imagem classista atualmente desemboca num considerável desinteresse de grande parte da população brasileira que não enxerga a utilidade das bibliotecas, seja pelo mito de conseguir encontrar todas as informações necessárias na internet, ou mesmo por não se identificar com a função cultural, social e informacional das bibliotecas. (SILVA; SABBAG, 2019).

Conforme destacado por Oliveira (2017) e Canclini (2003), atualmente há diversos debates acerca dos monumentos históricos presentes em importantes pontos das cidades. Um exemplo muito pertinente gira em torno dos monumentos dos bandeirantes. Dentro da perspectiva dessa discussão, caso estes símbolos fossem retirados dos locais públicos, muitos desses monumentos deveriam ser organizados em museus e demais unidades de informação que dispusesse em acondicionar os artefatos. Partindo desta perspectiva, os monumentos em questão não devem ser destruídos ou meramente descartados, em razão do notório valor artístico existente em algumas obras.

Por outro lado, caberia ao profissional mediador da informação contextualizar os monumentos fazendo um contraponto do antigo pensamento oficial vigente, a qual exaltava a figura do bandeirante, com o

pensamento presente, que faz um contrabalanço histórico acerca da relevância dos bandeirantes ao denunciar os inúmeros crimes cometidos, incluindo o genocídio, contra principalmente, populações de indígenas e de escravizados pretos. A lógica da ética para a representação, organização e curadoria de uma coleção de obras raras deve seguir essa linha de raciocínio.

Observa-se, desse modo, que a mediação constitui um papel central nos estudos em Ciência da Informação, pois além de conferir novas aplicabilidades e percepções conceituais para este Campo, contribui de forma efetiva para pensar a atuação dos centros de informação, assim como está vinculada a fatores diversos da Ciência da informação, como aspectos teórico-epistemológicos, tecnologia, gestão, serviços, estudos de usuários, entre outros assuntos da área. (SILVA, 2015, p. 105).

Discussões e reflexões acerca posicionamento ético e dimensão política que transpassam as práticas de organização do conhecimento são de extrema importância, pois tais atividades precisam voltar cada vez mais atenção para os contextos envolvendo as dimensões sociais, no qual o produto e o processo de representação do conhecimento trabalham de maneira adaptável e útil à sua comunidade de usuários. (SILVA; EVANGELISTA; GUIMARÃES, 2019).

3 Divulgação de questões éticas na política de desenvolvimento das coleções de obras raras

Segundo Cunha (2010) as bibliotecas universitárias são caracterizadas como organizações que desenvolvem tarefas complexas, com múltiplas funções e uma série de procedimentos, produtos e serviços. Já a biblioteca pública visa possibilitar o uso informacional em diferentes níveis sociais, atentando para as necessidades de sua comunidade de usuários, e assim proporcionar a inclusão social e a prática da cidadania (FREITAS; SILVA, 2014). O principal objetivo das bibliotecas especializadas gira em torno do processo de recuperação de informações específicas. A função de seus respectivos acervos consiste em suprir as necessidades da instituição (MIRANDA, 2007).

O presente estudo é de natureza exploratória e descritiva, a partir de pesquisas teóricas e análises bibliográficas foi possível verificar de forma mais detalhada, o posicionamento oficial das bibliotecas selecionadas por meio do mapeamento realizado em documentos disponibilizados em seus sítios oficiais relativos aos aspectos éticos envolvendo organização e representação do conhecimento em coleções de obras raras. Os resultados das fontes bibliográficas forneceram um arcabouço conceitual na contextualização teórica da pesquisa (GIL, 2002).

Inicialmente a principal intenção da pesquisa consistia em comparar os aspectos éticos das políticas de

desenvolvimento de coleções de obras raras de duas bibliotecas universitárias que tivessem disponibilizados documentos nos seus sítios, porém após investigação, nada específico no que concerne a temática foi encontrado.

Em 2021 os autores Raphael Diego Greenhalgh e Mariana Giuberti Guedes Greenhalgh, publicaram um importante artigo a respeito das coleções de obras raras e especiais em bibliotecas brasileiras: ‘Coleções especiais: uma análise da formação dos acervos nas bibliotecas brasileiras’. Os autores selecionaram as bibliotecas que detalhavam as informações das coleções especiais em suas páginas institucionais, e dessa forma, foram escolhidas 27 bibliotecas: 11 universitárias, 8 públicas e 8 especializadas. Os autores analisaram diversas coleções especiais, divididas nas seguintes tipologias: temática; multimeios; memória local/regional e institucional; particulares e obras raras. Os pesquisadores também constataram que as obras raras são as mais presentes nas coleções especiais, em exatamente 22 bibliotecas de 27, correspondendo a 81,5% das instituições pesquisadas.

Em razão da qualidade do levantamento e da análise realizada pelos pesquisadores Greenhalgh e Greenhalgh (2021), seja por conta do aspecto quantitativo (um grande número de bibliotecas analisadas por todo país, exceto a região norte), seja por causa do diagnóstico observado no que se refere às descrições empregadas acerca das tipologias do acervo, no qual buscaram detalhar da forma mais precisa possível, seja pelo fato de ser uma pesquisa contemporânea e atual, o presente estudo investigou as mesmas unidades de informação.

Nesse sentido, foram investigadas as seguintes vinte e sete bibliotecas: Biblioteca Central da Universidade de Brasília; Biblioteca Central da Universidade Federal de Santa Catarina; Biblioteca Central da Universidade Federal do Espírito Santo; Biblioteca Central da Universidade Federal do Estado do Rio de Janeiro; Biblioteca Central Irmão José Otão da Pontifícia Universidade Católica do Rio Grande do Sul; Biblioteca Comunitária da Universidade Federal de São Carlos; Biblioteca da Câmara dos Deputados; Biblioteca da Escola de Comunicações e Artes da Universidade de São Paulo; Biblioteca da Faculdade de Ciências e Letras da Unesp de Araraquara; Biblioteca da Universidade Federal de Minas Gerais; Biblioteca de História das Ciências e da Saúde (Casa de Oswaldo Cruz –COC), do Rio de Janeiro; Biblioteca do Senado Federal; Biblioteca do Tribunal de Justiça do Estado da Paraíba; Biblioteca Geral do Poder Judiciário de Alagoas; Biblioteca Mario de Andrade, de São Paulo; Biblioteca Ministro Victor Nunes Leal, do Supremo Tribunal Federal; Biblioteca Nacional de Brasília; Biblioteca Pública Benedito Leite, do Maranhão; Biblioteca Pública do Espírito Santo; Biblioteca Pública do Estado de Pernambuco; Biblioteca Pública

do Paraná; Biblioteca Pública Epiphânio Dória, do Sergipe; Biblioteca Pública Estadual de Minas Gerais; Biblioteca do Tribunal de Justiça do Estado do Rio de Janeiro, da Escola de Magistratura do Estado do Rio de Janeiro; Centro de Informação e Biblioteca em Educação do Ministério da Educação em Brasília; Biblioteca da Universidade de Caxias do Sul; e Biblioteca da Universidade Federal de Uberlândia.

Um percurso importante na aplicação das questões éticas que norteiam a organização e representação do conhecimento das coleções consiste na elaboração documental que oficializa a posição da biblioteca por meio do estabelecimento de diretrizes bem definidas. Foram analisados os mais variados tipos de documentos disponibilizados online, com destaque para as políticas de desenvolvimento, regimentos internos, e guias de usuário.

Todas as páginas das bibliotecas citadas foram visitadas, e em nenhuma página das vinte e sete instituições, foram encontradas referências nos documentos acerca da política de desenvolvimento e acesso relacionados com aspectos éticos no contexto da representação e organização das coleções de obras raras, das demais coleções especiais, ou mesmo envolvendo as coleções gerais. É possível afirmar categoricamente que a palavra ‘ética’ não é citada na maioria dos documentos, e nas pouquíssimas vezes que a palavra é citada, ocorre de forma bem breve e sucinta, sem nenhum aprofundamento detalhado.

É importante ressaltar que o estudo em questão não entrou em contato com a direção de nenhuma das vinte e sete bibliotecas analisadas, e é possível que exista, em algumas instituições, documentos não fornecidos ao público a respeito dos aspectos éticos da representação e organização no contexto das políticas de desenvolvimento e acesso das coleções.

Por outro lado, vale sublinhar o quanto é importante na relação entre instituição e comunidade de usuários o posicionamento das bibliotecas no que concerne a diversos temas, (como por exemplo, no caso do identitarismo e da representatividade de minorias sociais) e como são pautados na classificação e até mesmo no arranjo organizacional. Conforme destacado por Froelich (1994) os padrões culturais influenciam diretamente os padrões legais, impactando diretamente as políticas de desenvolvimento de coleções de uma biblioteca.

4 Considerações finais

Conforme salientado, não foi possível encontrar documentos que abordassem, de forma clara e explícita, diretrizes éticas do escopo da organização e representação do conhecimento envolvendo políticas de desenvolvimento e acesso das coleções nos respectivos sítios institucionais das bibliotecas. A

constatação realizada pelo mapeamento da presente pesquisa pode gerar importantes apontamentos futuros no que concerne a temática em questão, por outro lado este breve levantamento não tem a pretensão de determinar conclusões simplistas e/ou prematuras.

O mapeamento e análise da presente pesquisa foram realizados com intuito de localizar documentos que refletissem as problemáticas em forma de diretrizes, normas, regras etc. das instituições selecionadas, é bem possível e provável que algumas das bibliotecas analisadas guardem internamente os documentos relacionados com estas diretrizes. Também é completamente salutar a compreensão da inexistência de documentos no que diz respeito a temática em questão, afinal muito possivelmente boa parte dos bibliotecários podem considerar que os aspectos éticos já estão implícitos em outros documentos, como por exemplo, nos documentos que abordam questões de acessibilidade, ou em documentos que abordam igualdade racial ou de gênero. Não há dúvidas que documentos que abordem estes assuntos são de suma importância na relação entre bibliotecários e comunidades de usuários, porém para uma relação cada vez mais fortuita entre ambos, se faz necessário detalhar da forma mais minuciosa possível, todos os aspectos éticos que regem as diretrizes de dada unidade de informação, e conseqüentemente externalizar a relação deste acervo com seus usuários.

As bibliotecas, ao definirem seus critérios éticos, e posteriormente ao tornar público detalhando suas diversas diretrizes, dentre elas os contextos éticos envolvendo a classificação, a indexação, a catalogação e a curadoria do arranjo organizacional do acervo, formarão um elo transparente com sua comunidade de usuários ao se posicionar publicamente por meio de documentação oficial nas páginas institucionais de seus sítios.

De uma forma geral, as unidades de informação precisam debater internamente, com sua equipe de profissionais, e externamente com a sua comunidade de usuários e com sociedade civil como um todo, as inúmeras temáticas e problemáticas envolvendo a ética profissional, afinal este estudo diagnosticou um fluxo extremamente baixo da palavra ‘ética’. Em compensação um novo levantamento pode esclarecer as razões deste fenômeno, dentre eles se existe um norteamento implícito dos aspectos éticos, mesmo sem citá-lo diretamente.

Definir as diretrizes que norteiam as questões éticas das coleções de obras raras é uma tarefa ainda mais desafiadora se comparado à coleção geral, afinal, a natureza restrita dessas coleções delimitam em muitos sentidos as políticas de acesso. Outra situação complexa gira em torno dos aspectos curatoriais necessários ao se introduzir uma relação ética entre as características do acervo e o usuário, tal contexto deve

situar o usuário no espaço-tempo em que a coleção foi criada e desenvolvida, também se faz necessário explicações detalhadas e minuciosas a respeito da hierarquia desenvolvida na classificação. A complexidade envolvendo os aspectos éticos das coleções de obras raras pode gerar estudos com características bem peculiares, dentro do escopo da Organização e Representação do Conhecimento, se comparado às questões éticas que norteiam as coleções gerais.

Nesse sentido, cabe às instituições criar uma política de realização e divulgação de ações que valorizem os aspectos éticos da organização e representação das políticas de desenvolvimento e acesso dos seus acervos, dentre eles as coleções de obras raras. Uma representação eticamente consistente possivelmente irá conseguir atender de forma esclarecida, humanizada, e contextualizada os mais diversos tipos de usuários.

Referências

- AMERICAN LIBRARY ASSOCIATION. **Special collections**. 2021. Disponível em: <https://www.ala.org/tools/challengesupport/selectionpolicytoolkit/special>. Acesso em: 09 jan. 2023.
- ANDRADE, R. H. R. de; CANTALINO, M. das G. N. A raridade como questão epistemológica e política: um novo paradigma para os curadores de acervos especiais. **Anais da Biblioteca Nacional**, Rio de Janeiro, v.123, p.49-58, 2003.
- ASSOCIATION OF RESEARCH LIBRARIES. *Colege & Research Libraries [2003]*, Chicago, USA. Disponível em: <https://crl.acrl.org/index.php/crl/issue/view/987>. Acesso em: 16 abr. 2023.
- BATISTA, A. H. **Conceitos e critérios para a qualificação de Obras Raras da Biblioteca de Direito da Universidade Federal de Pelotas**. 2012. 106 p. Dissertação (Mestrado em Memória Social e Patrimônio Cultural) Universidade Federal de Pelotas, Pelotas, 2012. Disponível em: <http://repositorio.ufpel.edu.br/handle/123456789/1032>. Acesso em: 08 jan. 2023.
- BERGER, S. E. **Rare books and special collections**. Chicago: Neal-Schuman, 2014.
- BORKO, H. Information science: what is it?. **American documentation**, Washington v. 19, n. 1, p. 3-5, 1968.
- BRASCHER, M.; CAFE, L. Organização da informação ou organização do conhecimento? *In: Encontro Nacional de Pesquisa em Pós-graduação em Ciência da Informação*, 9, 2008, São Paulo. **Anais[...]** Brasília: ENANCIB, 2008.
- CANCLINI, N. G. **Culturas híbridas: estratégias para entrar e sair da modernidade**. 4.ed. São Paulo: Edusp, 2003.
- CUNHA, M. B. A biblioteca universitária na encruzilhada. **Revista de Ciência da Informação**, v. 11, n. 6, 2010. Disponível em: <https://repositorio.unb.br/handle/10482/14869>. Acesso em: 26 jan. 2023.
- DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Proposta de diretrizes para elaboração de política de indexação em bibliotecas universitárias: um estudo sociocultural com protocolo verbal. **Scire: Representación y Organización del Conocimiento**, v. 23, n. 1, p. 13-23, 2017. Disponível em: <https://www.ibersid.eu/ojs/index.php/scire/article/view/4352>. Acesso em: 27 mar. 2023.
- FARIA, M. I. R.; PERICÃO, M. G. **Dicionário do Livro: da escrita ao livro eletrônico**. São Paulo: Edusp, 2008.
- FREITAS, M. A.; SILVA, V. B. Bibliotecas públicas brasileiras: panorama e perspectivas. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 12, n. 1, p. 123-146, jan/abr. 2014.
- FROELICH, T. J. Ethical concerns of information professionals in an international context. *In: ALVAREZ-OSSORIO, J. R.; GOEDGEBUURE, B. G. New worlds in information and documentation*. Amsterdam: Elsevier, 1994. p. 459-470.
- GIL, A. C. **Como Elaborar Projetos de pesquisa**. 3. ed. Belo Horizonte: Atlas, 1991.
- GUIMARÃES, J. A. C. *et al.* Ética nas atividades informativas: aspectos teóricos. **Ponto de Acesso**, Salvador, v. 2, p. 138-153, 2008.
- GUIMARÃES, J. A. C.; PINHO, F. A. Aspectos éticos em organização e representação do conhecimento (ORC). *In: ENANCIB, 7., 2006, Marília. Anais [...]*Marília: ENANCIB, 2006.
- GREENHALGH, R. D.; GREENHALGH, M. G.G. Coleções Especiais: uma análise da formação dos acervos nas bibliotecas brasileiras. **BIBLOS, [S. l.]**, v. 35, n. 2, 2022. Disponível em: <https://periodicos.furg.br/biblos/article/view/13340>. Acesso em: 14 out. 2022.
- LUNARDELLI, R. A. *et al.* A análise da informação e seu ensino nos Cursos de Arquivologia e Biblioteconomia da Universidade Estadual de Londrina: relato de experiência. **Biblos**, v. 33, n. 2, p. 60-74, 2019. Disponível em: <https://periodicos.furg.br/biblos/article/view/9094>. Acesso em: 03 jan. 2021. 2023.
- MACEDO, E. A imagem da ciência: folheando um livro didático. **Educação & Sociedade**, v. 25, p. 103-129, 2004.
- MILANI, S. O.; GUIMARÃES, J. A. C. Problemas éticos em representação do conhecimento: uma abordagem teórica. **DataGramZero**, Rio de Janeiro, v. 12, p. artigo-04, 2011.
- MIRANDA, A. C. C. Formação e desenvolvimento de coleções em bibliotecas especializadas. **Inf. & Soc.: Est., João Pessoa**, v. 17, n. 1, p. 87-94, jan./abr. 2007.
- NARDINO, A. T. D.; CAREGNATO, S. E. O futuro dos livros do passado: a biblioteca digital contribuindo na preservação e acesso às obras raras. **EM QUESTÃO**, v. 11, n. 2, p. 381-407, 2005. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/126>. Acesso em: 04 ago. 2021.
- OLIVEIRA, R. Por que nos importamos com símbolos escravagistas dos EUA e ignoramos os do Brasil?. **El País**, São Paulo, 4 set. 2017. Disponível em:

https://brasil.elpais.com/brasil/2017/09/02/politica/1504310652_774711.html. Acesso em: 25 jan. 2023.

ORDOVÁS, Gleide Bitencourte José; STEINDEL, Gisela Eggert. Políticas de Preservação do Acervo de Obras Raras da BU/UFSC: um estudo. *In: Congresso Brasileiro de Biblioteconomia e Documentação*, 27, 2017. **Anais do CBBB**. São Paulo: FEBAB, 2017. Disponível em: <https://docplayer.com.br/106161525-Políticas-depreservacaodo-acervo-de-obras-raras-da-bu-ufsc-umestudo.html>. Acesso em 30 jan. 2023.

PINHO, F. A. **Aspectos éticos em representação do conhecimento: em busca do diálogo entre Antonio García Gutiérrez, Michèle Hudon e Clare Beghtol**. 2006. 123 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006. Disponível em: <http://hdl.handle.net/11449/93693>. Acesso em: 27 fev. 2023.

PINHEIRO, A. V. **Que é livro raro?: uma metodologia para o estabelecimento de critérios de raridade bibliográfica**. Rio de Janeiro: Presença, 1989.

REIFSCHNEIDER, O. D. B. A importância do acesso às obras raras. **Revista Iberoamericana de Ciência da Informação (RICI)**, v.1, n.1, p.67-76, jan./jun. 2008. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/910/792>. Acesso em: 15 jan. 2023.

ROBREDO, J. **Da Ciência da Informação revisitada aos sistemas humanos de informação**. Brasília: Thesaurus, 2003.

SILVA, A. P.; EVANGELISTA, I. V.; GUIMARÃES, J. A. C. Valores éticos em Organização do Conhecimento: uma análise a partir dos estudantes de Arquivologia e Biblioteconomia da UNESP Marília. *In: BARROS, Thiago*

Henrique Bragato; TOGNOLI, Natália Bolfarini. (Org.). Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas. Belém: **Ed. da UFPA**, 2019, v., p. 460-467.

SILVA, B. D. O.; SABBAG, D. Memes como diálogo cultural em bibliotecas públicas: uma análise sob a perspectiva do hibridismo cultural. *In: Seminário em Ciência da Informação*, 2019, Londrina. **Informação Social no Contexto da Ciência da Informação**. Londrina: UEL, 2019.

SILVA, J. L. C. Percepções conceituais sobre mediação da informação. **InCID: Revista de Ciência da Informação e Documentação**, [S. l.], v. 6, n. 1, p. 93-108, 2015. DOI: 10.11606/issn.2178-2075.v6i1p93-108. Disponível em: <https://www.revistas.usp.br/incid/article/view/89731>. Acesso em: 30 jan. 2023.

SILVA, P. W. J. **Curadoria de acervos raros e especiais: Bibliotecária Berenice Bacelar**. 2017. 55 f. Trabalho de Conclusão de Curso (Graduação) – Faculdade de Biblioteconomia, Instituto de Ciências Sociais Aplicadas, Universidade Federal do Pará, Belém, 2017. Disponível em: <http://bdm.ufpa.br/jspui/handle/prefix/100>. Acesso em: .

SANT'ANA, R. B. Critérios para definição de obras raras. **ETD. Educação Temática Digital, Campinas**, v. 2, n.3, p. 1-18, 2001. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/etd/article/view/577/592>. Acesso em: 30 ago. 2021.

UNIVERSITY OF GLASGOW. **What are Special Collections**. Glasgow, UK: University of Glasgow, Special Collections, [2019]. Disponível em: <https://web.archive.org/web/20190521171905/https://www.gla.ac.uk/myglasgow/specialcollections/whatarespecialcollections/>. Acesso em: 14 out. 2022.

As bases epistemológicas do álbum fotográfico e suas relações com a Organização do Conhecimento

The epistemological bases of the photographic album and its relations with the Knowledge Organization

Cristina Ribeiro dos Santos (1), Lucilene Aparecida Francisco (2) Marcos Antônio de Moraes (3)

Universidade Estadual de Londrina (UEL), Rodovia Celso Garcia Cid, PR-445, Km 380 - Campus Universitário, Londrina - PR, CEP 86057-970, (1) cristina.ribeiro@uel.br. (2) lafrancisco08@gmail.com. (3) marcosmoraes@uel.br.

Resumo

Objetivo: Este trabalho visa explorar a compreensão e a extensão do conceito de “álbum fotográfico” nas publicações que integram os sistemas de informação de teses e dissertações brasileiras. A busca pela compreensão da extensão de um conceito comum a diferentes ciências, como é o caso do álbum fotográfico, visa entender, ou mesmo, estabelecer suas características principais que interferem diretamente nos ambientes informacionais. Parte-se das bases epistemológicas de “álbum fotográfico” com o propósito de elucidar as relações destes com a Organização do Conhecimento. **Metodologia:** Para tal, realiza-se uma investigação de caráter bibliográfico e qualitativo, a partir das publicações indexadas na Biblioteca Digital de Teses e Dissertações (BDTD) e aplica-se a técnica de análise de categorial de conteúdo aos estudos levantados, a fim de identificar os enunciados do conceito de álbum trazido por estes trabalhos, bem como, a relevância destes para a Organização do Conhecimento. **Resultados:** O levantamento realizado em março de 2023 apontou 73 documentos abordando o tema em diferentes áreas do saber. A análise revelou que o álbum fotográfico tem sido caracterizado como um recurso informacional, uma narrativa visual e uma forma de arquivamento, que traz métodos peculiares de organização e, principalmente, uma abordagem específica acerca da temática representada. **Conclusão:** Em síntese, este texto traz em seus levantamentos e análises os contornos e propriedades do conceito de álbum fotográfico na perspectiva dos estudos desenvolvidos no âmbito da pós-graduação brasileira. Destacando que quando observado sob o prisma da Organização do Conhecimento este objeto traz consigo um desafio conceitual e epistemológico que contempla diferentes objetivos e finalidades como: forma de registro do conhecimento, recurso informacional, suporte para a preservação da memória individual ou coletiva, que traz o pensamento colecionista e de organização e está presente nos mais diversos espaços acadêmicos de forma digital ou físico.

Palavras-chave: Álbum Fotográfico. Teoria do Conceito. Organização do Conhecimento.

Abstract

Objective: This work aims to explore the understanding and extension of the concept of “photographic album” in publications that integrate the information systems of Brazilian theses and dissertations. The search for understanding the extension of a concept common to different sciences, as is the case of the photographic album, aims to understand, or even establish, its main characteristics that directly interfere in the informational environments. It starts with the epistemological bases of “photographic album” with the purpose of elucidating their relations with the Organization of Knowledge. **Methodology:** To this end, a bibliographic and qualitative investigation is carried out, based on publications indexed in the Digital Library of Theses and Dissertations (BDTD) and the technique of categorical analysis of content is applied to the surveyed studies, in order to identify the statements of the concept of album brought by these works, as well as their relevance for the Organization of Knowledge. **Results:** The survey carried out in March 2023 pointed to 73 documents addressing the topic in different areas of knowledge. The analysis revealed that the photo album has been characterized as an informational resource, a visual narrative and a form of archiving, which brings peculiar methods of organization and, mainly, a specific approach to the theme represented. **Conclusion:** In summary, this text brings in its surveys and analyzes the contours and properties of the concept of photo album from the perspective of studies developed within the scope of Brazilian postgraduate studies. Emphasizing that when observed from the perspective of the Organization of Knowledge, this object brings with it a conceptual and epistemological challenge that contemplates different objectives and purposes such as: way of registering knowledge, informational resource, support for the preservation of individual or collective memory, which brings the collectionist and organizational thinking and is present in the most diverse academic spaces in digital or physical form.

Keywords: Photo album. Concept Theory. Knowledge Organization.

Resumen

Objetivo: Este trabajo tiene como objetivo explorar la comprensión y extensión del concepto de “álbum fotográfico” en publicaciones que integran los sistemas de información de tesis y disertaciones brasileñas. La búsqueda por comprender la extensión de un concepto común a diferentes ciencias, como es el caso del álbum fotográfico, pretende comprender, o incluso establecer, sus principales características que interfieren directamente en los entornos informacionales. Se parte de las bases epistemológicas del

“álbum fotográfico” con el propósito de dilucidar sus relaciones con la Organización del Conocimiento. **Metodología:** Para ello, se realiza una investigación bibliográfica y cualitativa, a partir de publicaciones indexadas en la Biblioteca Digital de Tesis y Disertaciones (BDTD) y se aplica la técnica de análisis categórico de contenido a los estudios relevados, con el fin de identificar las declaraciones del concepto de álbum que traen estos trabajos, así como su relevancia para la Organización del Saber. **Resultados:** La encuesta realizada en marzo de 2023 apuntó a 73 documentos que abordan el tema en diferentes áreas del conocimiento. El análisis reveló que el álbum de fotos se ha caracterizado como un recurso informativo, una narrativa visual y una forma de archivo, lo que trae peculiares métodos de organización y, principalmente, un enfoque específico del tema representado. **Conclusión:** En resumen, este texto trae en sus encuestas y analiza los contornos y propiedades del concepto de álbum de fotos en la perspectiva de los estudios desarrollados en el ámbito de los estudios de posgrado brasileños. Destacando que cuando se observa desde la perspectiva de la Organización del Conocimiento, este objeto trae consigo un desafío conceptual y epistemológico que contempla diferentes objetivos y finalidades tales como: forma de registro del conocimiento, recurso informacional, apoyo a la preservación de la memoria individual o colectiva, la cual trae el pensamiento coleccionista y organizacional y está presente en los más diversos espacios académicos en forma digital o física.

Palabras clave: Album de fotos. Teoría del concepto. Organización del conocimiento.

1 Introdução

A Organização do Conhecimento (OC) dedica-se a atividades como indexação e classificação de documentos com vistas a promover sua recuperação. Nestas atribuições, os conceitos são fundamentais, pois formam um conjunto de significações que representam o conteúdo dos documentos a partir de um domínio de conhecimento e da linguagem pertencente a este domínio (HJORLAND, 2008). Nesta linha de pensamento, Abbagnano (2007, 164), afirma que o conceito abarca, “[...] todo processo que torne possível a descrição, a classificação e a previsão dos objetos cognoscíveis.” Carregando esta característica, o conceito se envolve com as ações da OC de forma íntima, pois esta tem entre os seus ofícios a atividade de organizar, representar, definir, simbolizar, expressar, dentre outros processos que se utilizam da linguagem para a estruturação de Sistemas de Organização do Conhecimento (SOC), como é o caso dos tesouros, esquemas de classificação, ontologias, entre outros.

Assim, OC possui ações interdisciplinares que estudam os processos de organização e estruturação do conhecimento, com o objetivo de torná-lo acessível e utilizável. Para isso, dispõe de ferramentas como: de vocabulários controlados, padrões e normas que primam pela representação semântica dos conceitos e suas relações. Trata-se de uma forma de representação ordenada, que demanda o emprego de métodos de descrição, formalização e classificação.

A emersão crescente de novas realidades, necessidades informacionais, novos formatos dos recursos informacionais, ou mesmo novas perspectivas de uso destes, trazem para a OC, a demanda de examinar transversalmente as apropriações, bem como as percepções a respeito destes recursos informacionais que se instalam neste universo informacional heterogêneo. Complexidades a parte, este interesse, incide na tríade: nomear/conceitualizar, representar e organizar. Que por sua vez traz à tona a importância do conceito e seus enunciados/características, bem como a sua extensão para a OC. A busca pela compreensão da

extensão de um conceito comum a diferentes ciências, como é o caso do álbum fotográfico, tem como base a visa entender, ou mesmo estabelecer as características principais que interferem diretamente nos ambientes informacionais.

Barros (2016, p. 27) afirma que “[...] apesar de não possuir uma existência real, o conceito é um instrumento imprescindível não apenas para o conhecimento científico, como para a própria vida comum.” Para este autor há uma essencialidade nos conceitos, pois se caracterizam como um instrumento, sem os quais os objetos e fenômenos não poderiam ser “[...] concebidos em termos de semelhança e diferença, com ajuda de conceitos a ciência e uma série de outras atividades humanas fundamentais simplesmente não seriam possíveis. [...]” (BARROS, 2016, p. 27).

Tal fato é perceptível no dicionário de Organización y representación del conocimiento de Barité (2015) que define conceito como:

Abstracción o noción que refiere a una unidad de conocimiento, independiente de su expresión lingüística, y comprende el conjunto de sus rasgos esenciales. El concepto, en tanto representación simbólica, está en la base de la Teoría de la Clasificación y de la Terminología, pues es el elemento indivisible que permite representar el conocimiento contenido en los documentos y organizar los enunciados correspondientes a la idea que se tiene de cualquier cosa (BARITÉ, 2015, p. 33).

Nesta reflexão, Barité (2015) reforça o entendimento de que o conceito ocupa uma posição nuclear nas preocupações da OC. Assim, o interesse de examinar as características/enunciados das produções científicas de diversas áreas presente neste estudo, além de permitir a visualização das bases interdisciplinares do compartilhamento de um mesmo recurso informacional, explicita as características e os novos contornos de entendimento acerca do conceito de álbum fotográfico na contemporaneidade.

Partindo desta relevância para as bases da OC, há uma tentativa de identificar as apropriações, bem como as percepções a respeito do álbum fotográfico nos ambientes informacionais acadêmicos. Esta

justificativa alinhava-se, à ideia de Barros (2016), que afirma que os conceitos estão presentes na dinâmica da produção do saber científico. Tal autor argumenta ainda que, todas as ciências se organizam a partir de conceitos (BARROS, 2016). Desta forma compreender as características que o álbum possui no âmbito acadêmico, bem como, mapear a sua extensão corrobora com a apreensão das novas relações que são constituídas e com as percepções que se instalam com o advento do mundo digital. Isso pressupõe que novas características foram incorporadas, demandando um olhar para a transversalidade deste conceito, para as mudanças de sua extensão e compreensão ao longo do tempo.

Como visto acima, abordar o conceito de álbum fotográfico é fundamental para que se possa reforçar a presença deste objeto dentro dos estudos não apenas da Organização do Conhecimento (OC) como também de outras ciências. Vincula-se a esta justificativa, o fato de que há uma estreita relação entre o êxito, a eficiência e a validade dos instrumentos de tratamento e recuperação da informação a partir da utilização de um conceito coeso e alinhado às várias características e extensões que lhe são atribuídos.

Para atingir tal intuito, parte-se para a contribuição estrutural: a construção do seu percurso histórico. Isso porque ao longo da história, o álbum fotográfico vem desempenhando um papel significativo na preservação e compartilhamento de memórias pessoais e culturais, a partir de sua forma particular de organização, classificação e exposição dos fatos, eventos e fenômenos, comprometido, sobretudo, com a construção das narrativas pessoais e coletivas.

Cumprir observar que o álbum fotográfico surge no século XIX, no bojo das profundas mudanças históricas e sociais que marcaram esse período, elevaram a produção intelectual e artística e por conseguinte a demanda social por imagens. Nesse contexto, o álbum fotográfico se consolida como uma coleção de imagens organizadas que visa comunicar uma narrativa específica. Dessa forma, uma das suas principais características é a maneira como as imagens são organizadas para contar uma história (FERRAZ, 2014).

Fabris (2008, p. 44-45) lembra que a fotografia manteve-se inicialmente como uma atividade dirigida ao público aristocrático e burguês, uma vez que embora tivesse um custo mais baixo que as pinturas, não estava ao alcance de toda a sociedade. Contudo, com o aperfeiçoamento das técnicas empregadas, passou a constituir-se como um instrumento de coesão social, que possibilita “[...] a construção de um repertório de imagens comuns, que permitem viajar no tempo e no espaço, estabelecer um museu imaginário ideal”. Com o passar dos anos, o álbum fotográfico ganhou popularidade, tornou-se uma forma de preservar memórias pessoais e familiares, um documento, um

testemunho dos fatos. Sem o qual muitos momentos da história se perderiam no tempo e no espaço.

Para Mauad (2004, p. 18) a:

[...] fotografia com toda a sua possibilidade de encenação, inventa uma memória para ser perenizada, eternizando-se na emulsão fotográfica uma vontade de ser, algumas vezes risível, mas na maior parte, crível. Uma imagem capaz de criar uma representação ideal para ser lembrada no futuro, nos álbuns de família, guardiões das tradições, inventadas ou não.

É necessário perceber que ao passo que a fotografia remete a ideia de registro de um fato ou momento, o álbum representa o colecionar, o organizar e classificar dos objetos revestidos de valor afetivo e simbólico. Tal registro, possibilita, assim, reunir fotografias de grandes eventos e personalidades, atribuindo-lhes um arranjo específico.

Ferraz (2014, p. 64) acrescenta que o álbum surge como “[...] uma necessidade da mentalidade classificatória do século XIX [...] a primeira grande máquina moderna a documentar o mundo e a amearhar suas imagens”. O álbum fotográfico adquire, assim, uma importância cultural e histórica, ao ponto de ser frequentemente usado como recurso informacional, uma fonte primária de pesquisa, que permite entender a história das pessoas, suas famílias e comunidades, ou mesmo, um instrumento de trabalho como são os álbuns fotográficos alimentares utilizados pela nutrição.

Com a popularidade das redes sociais e dos *smartphones*, o formato tradicional do álbum fotográfico pode estar se tornando menos comum. No entanto, muitas pessoas ainda o valorizam como uma forma tangível e significativa de preservar e compartilhar memórias.

Embora o álbum fotográfico tenha sua relevância cultural e histórica reconhecida, as apreensões acerca das suas principais características e abordagens nas diferentes áreas do conhecimento ainda são incipientes, especialmente, na Ciência da Informação. Fato que justifica a proposição deste estudo, que tem como questão norteadora: Qual a compreensão/extensão de “álbum fotográfico” apresentada pelos estudos desenvolvidos no âmbito pós-graduação brasileira e como esses se relacionam com a Organização do Conhecimento?

Para responder a tal questionamento, desenvolve-se uma pesquisa bibliográfica de cunho qualitativo a partir das publicações indexadas na Biblioteca Digital de Teses e Dissertações (BDTD) e aplica-se a técnica de análise de categorial de conteúdo de Bardin (2016) aos estudos levantados, a fim de identificar a compreensão e extensão do conceito de álbum fotográfico trazidas por estes trabalhos. Os resultados apontam que o álbum tem sido tratado como recurso informacional, uma

narrativa visual e uma forma de arquivamento, que traz métodos peculiares de classificação/organização e, principalmente, uma abordagem específica acerca da temática representada. A análise desenvolvida mostra ainda, que a compreensão e extensão do conceito de álbum tem se modificado com o passar dos tempos, principalmente com o desenvolvimento das tecnologias, mas permanece destacando-se como instrumento de organização do conhecimento e preservação da memória individual ou coletiva.

Para discussão e fundamentação dos resultados desenvolve-se dois tópicos de revisão da literatura, buscando elucidar as principais relações entre o conceito e a OC, bem como, refletir a respeito das duas dimensões a serem consideradas no conceito de álbum fotográfico na visão de Barros (2016), ou seja, a “extensão” e a “compreensão”, que possuem como sinônimos “conteúdo” ou de “intenção”

Do ponto de vista filosófico, todo o conceito possui duas dimensões a serem consideradas: a “extensão” e a “compreensão” (Às vezes também chamada de “conteúdo” ou de “intenção”). Chama-se extensão de um conceito precisamente ao grau de uma abrangência em relação a vários fenômenos e objetos e chama-se compreensão de um conceito ao esclarecimento das características que o constitui. À medida que um conceito adquire maior extensão perde em compreensão (BARROS, 2016, p. 72).

Partindo do contexto de análise da compreensão e extensão do termo álbum fotográfico, este estudo tem a intenção de contribuir para o mapeamento dos contornos e propriedades deste conceito, observando as modificações trazidas com o passar dos tempos, principalmente com o desenvolvimento das tecnologias.

2 O Conceito e a Organização do Conhecimento

Ao se referir ao conceito de algo, inevitavelmente se remete à Teoria do Conceito. Tal proposta de Ingetraut Dahlberg, criada na década de 1970, torna-se referência para se refletir acerca dos conceitos especialmente no âmbito da Ciência da Informação. Tal questão sob o prisma da Organização do conhecimento, envolve a análise, a busca por relacionamentos, hierarquias, ou seja, ações complexas que buscam mapear: os conceitos de um domínio; seus atributos, suas fronteiras, suas características; sua compreensão e definições visando à organização e recuperação da informação.

Cabe lembrar que as propriedades dos conceitos e suas definições são caras à Organização do Conhecimento, tal afirmação se estabelece nas perspectivas de diversos autores como Dahlberg (1978), Hjørland (2008), Kobashi e Francelin (2011) entre outros. Tal relevância se caracteriza pelo caráter estruturante, bem como pelo vínculo com a intelectualidade e a linguagem. A linguagem, o nomear dos objetos e fenômenos estão

por sua vez na base da OC. Em uma perspectiva complementar, Abbagnano (1998, p. 164) ao delimitar o sentido de “conceito” afirma que esse termo é “[...] em geral, todo processo que torne possível a descrição, a classificação e a previsão dos objetos cognoscíveis.”

Retomando as ideias de Abbagnano (2007) e sua percepção a respeito de conceito, destaca-se que um:

[...] termo tem significado generalíssimo e pode incluir qualquer espécie de sinal ou procedimento semântico, seja qual for o objeto a que se refere, abstrato ou concreto, próximo ou distante, universal ou individual, etc. Pode-se ter um C [conceito] de mesa tanto quanto do número 3, de homem tanto quanto de Deus, de gênero e espécie (os chamados universais [v.]) tanto quanto de uma realidade individual, [...] (ABBAGNANO, 1998, p. 164).

Abbagnano (2007, p.168), destaca que a função atribuída ao conceito é a “[...] interpretação como essência, já que, por essa interpretação, o C [conceito] não tem outra função senão exprimir ou revelar a substância das coisas. Desse ponto de vista, a função identifica-se com a própria natureza do conceito.”

Corroborando com esse pensamento, Hjørland (2008) acrescenta que a OC está preocupada com a natureza, qualidade e a adequação de tais processos vinculados à organização do conhecimento. Esta atenção, se estende aos “[...] sistemas de organização do conhecimento (KOS) usados para organizar documentos, representações de documentos, obras e conceitos.” (HJORLAND, 2008, p. 86).

Para Hjørland (2008, p. 86) a Organização do Conhecimento (OC):

[...] trata de atividades como descrição, indexação e classificação de documentos realizadas em bibliotecas, bases de dados bibliográficas, arquivos e outros tipos de “instituições de memória” por bibliotecários, arquivistas, especialistas em informação, especialistas no assunto, bem como por algoritmos de computador e leigos.

Neste domínio, os conceitos são compreendidos como resultado de uma construção humana que envolve a significação, a linguagem, ou seja, formas linguísticas e têm o intuito de representar o conteúdo de determinados documentos. Relaciona-se, portanto, com o conteúdo informacional de um determinado domínio do conhecimento, com a linguagem e com os conceitos pertencentes a este domínio. (HJORLAND, 2008). Tais elementos precedem de uma contextualização, que busque identificar as características constitutivas essenciais dos seus elementos base: o conceito.

As ações que envolvem a OC implicam em observar as transformações motivadas pelo significado ou mesmo a ampliação que o conceito recebe. Barros (2016, p. 53) sinaliza para uma exigência lógica que aparece quando se tem uma definição conceitual em discussão, a chamada conversibilidade, compreendida como uma regra

[...] simples que deve ser observada em toda formulação conceitual, uma definição deve valer para todos os sujeitos e objetos que incluem ou se pretende incluir no âmbito de aplicação da coisa definida (a extensão do conceito) e tão somente para estes sujeitos e objetos. Dito de outra forma, e utilizando a linguagem mais filosófica, a definição deve ser conversível ao definido. (BARROS, 2016, p. 53)

O autor acrescenta que os conceitos surgem no interior de uma história, ou seja, são históricos e, portanto, sofrem alterações. Partindo desta questão, este autor em sua busca por refletir a respeito da compreensão e a extensão de um conceito contribui para a concepção do termo álbum fotográfico, especialmente quando afirma que o conteúdo de um conceito é formado concomitantemente pelas características que o definem.

A partir destas concepções, compreender a relação de inter-relação entre a OC e as implicações que o conceito carrega consigo, pressupõe considerar a estrutura conceitual, o contexto e a identificação do conjunto de características que tem o potencial de representar um determinado registro do conhecimento em sua atualidade. Essas ações, no entanto, devem considerar o fato de que os conceitos podem apresentar algum nível de polissemia. A “[...] polissemia - ou a possibilidade de uma palavra ou conceito apresentar uma certa variedade de sentido bem aceitos - é de modo geral uma característica dos conceitos em todas as áreas” (BARROS, 2016, p. 55).

A prerrogativa de presença do conceito nos esquemas que organizam e representam o conhecimento humano, tais como as classificações, taxonomias, vocabulário controlado e as ontologias que buscam dentro das bibliotecas, das bases de dados, dos arquivos entre outros centros de informação, promover a organização e representação do conhecimento com vista ao seu acesso trazem à tona esta relação tão complexa. Na próxima seção busca-se explorar as relações entre o conceito de álbum fotográfico e a ORC.

3 O Álbum Fotográfico e a ORC

O álbum fotográfico insere-se no contexto de ORC, como um recurso informacional, que surge no século XIX, face ao crescimento exponencial das fotografias, que por sua vez também são registros produzidos pela sociedade. Este recurso informacional, tomado como um conceito, encontra harmonia com os constructos de Dahlberg (1978), especialmente quando se considera que foi a partir da linguagem que o sujeito desenvolveu a habilidade de relacionar-se com os vários objetos que o circundam, tornando-se capaz de “elaborar enunciados sobre os mesmos.” (DAHLBERG, 1978, p. 101). Destaca-se que para esta autora “[...] todo enunciado sobre objetos contém um elemento do respectivo conceito”(DAHLBERG, 1978, p. 102).

Este recurso informacional, que em sua origem carregava consigo a “extensão” e a “compreensão”, vinculada ao mundo físico, especialmente a sua materialidade, atualmente está imerso em uma combinação entre os mundos físico e digital. Tal mudança inclui, além do contexto de produção e uso, as construções teóricas atinentes a este recurso nas diferentes áreas dos conhecimentos na contemporaneidade. “Assim, qualquer alteração na compreensão de um conceito repercute imediatamente na extensão do mesmo.” (BARROS, 2016, p.72) Ou seja, ao se modificar a compreensão de álbum fotográfico, considerando o avanço tecnológico e social, incorpora-se à sua extensão, os álbuns digitais.

Historicamente, o álbum fotográfico tem na fotografia a sua base, e esta, por sua vez, tem na revolução industrial a sua origem, ou seja, meados do século XVIII e se estende pelo século XIX. Inicialmente com as pesquisas e experimentos que envolviam o conhecimento do princípio ótico, sensibilização de diversos suportes, para registro da luz, passando pelo próprio registro e seguindo com meios de popularizar desta invenção.

Borges (2005), ao desenvolver um cronograma referente à criação e desenvolvimento da fotografia, afirma que no ano de 1814 Joseph Nicéphore Niépce (1765 – 1833) iniciou suas pesquisas sobre a fixação de imagens da câmara escura. Manini (2008) relata que estas pesquisas culminaram na obtenção da primeira imagem a partir de 1825. Para Borges (2005, p. 116) foi em 1826 que as “[...] pesquisas de Nicéphore Niépce permitem-lhe captar a primeira imagem sobre a câmara escura: uma natureza morta. Para sua produção foram necessárias 14 horas”. Em 1829, Niépce se associa a Louis J. M. Daguerre, e os dois prosseguem com as pesquisas sobre a representação de imagens na câmara escura.

Com o desenvolvimento destas pesquisas, a fotografia expandiu-se pelo mundo, gerando e transmitindo conhecimentos. A invenção da fotografia facilitou o conhecimento de outros lugares, pessoas, enfim, outras realidades. O desenvolvimento da indústria gráfica favoreceu a multiplicação e a disseminação da imagem fotográfica. Segundo Kossoy (2001, p. 26).

[...] Com a descoberta da fotografia e, mais tarde, com o desenvolvimento da indústria gráfica, que possibilitou a multiplicação da imagem fotográfica em quantidades cada vez maiores através da via impressa, iniciou-se um processo de conhecimento do mundo, porém de um mundo em detalhe, posto que fragmentário em termos visuais e, portanto, contextuais.

Os processos iniciais da fotografia eram complexos e demorados. Benjamin (1994) destaca os clichés de Daguerre (equipamento de fazer fotografia da época) consistiam em placas de prata, iodadas e “[...] expostas na câmara obscura; elas precisavam ser manipuladas

em vários sentidos, até que se pudesse reconhecer, sob uma luz favorável, uma imagem cinza-pálida. Tão raras, que [...] eram guardadas em estojos, como jóias.” (BENJAMIN, 1994, p. 93). Gera-se assim um acúmulo de imagens fotográficas e obras de arte.

Para Benjamin (1994) a evolução da fotografia foi rápida. Em 1840 parte significativa dos pintores de miniaturas já tinham se transformado em fotógrafos, inicialmente de forma esporádica e pouco depois exclusivamente. De acordo com o autor:

Foi nessa época que começaram a surgir os álbuns fotográficos. Eles podiam ser encontrados nos lugares mais glaciais da casa, em consoles ou guéridons, na sala de visitas — grandes volumes encadernados em couro, com horríveis fechos de metal, e as páginas com margens douradas, com a espessura de um dedo, nas quais apareciam figuras grotescamente vestidas ou cobertas de rendas [...] (BENJAMIN, 1994, p. 97).

Na perspectiva de Benjamin (1994), o álbum fotográfico, tem como características: grandes volumes, encadernados em couro, com horríveis fechos de metal, e continha páginas, bem como margens douradas, além de ter a espessura de um dedo.

A radicalidade no uso e no consumo de imagens pelo florescer da modernidade fotográfica é explorada por Rouillé, no entendimento de que captar, apanhar, registrar e fixar uma imagem funciona como uma máquina de ver e renovar. O termo “arquivar”, torna-se então, a palavra-chave para o autor:

Uma das grandes funções da fotografia-documento terá sido a de erigir um novo inventário do real, sob a forma de álbuns e, em seguida, de arquivos. O álbum, enquanto mecanismo de reunir e tesarizar as imagens; a fotografia, enquanto mecanismo para ver (óptico) e para registrar e duplicar as aparências (químico). Assim, esse inventário fotográfico do real constituiu-se no cruzamento de dois procedimentos de tesarização: o das aparências, pela fotografia; e o das imagens, pelo álbum e pelo arquivo (ROUILLÉ, 2009, p. 97).

O álbum possui, na visão deste autor, o enunciado vinculado a um mecanismo de reunir, registrar e tesarizar as imagens. Observa-se nesta percepção, o álbum como um arquivo. Turazzi (2002, p. 23), ao pesquisar o acervo de Marc Ferrez, percebe o início da popularização da fotografia e seu acúmulo no álbum fotográfico, destaca a presença de Legendas, passe-partouts:

[...] álbuns completavam as imagens fotográficas com uma ornamentação bem característica: letras esculpidas, vinhetas douradas, guirlandas com motivos tropicais, encadernações em couro, guardas em tecido ou papel marmoreado, fechos metálicos cuidadosamente elaborados.

Percebe-se neste estudo, a presença de elementos trazidos de outros enunciados (encadernações em couro, ornamentação, fechos metálicos cuidadosamente

elaborados). E a presença de novos: letras esculpidas, vinhetas douradas, guirlandas com motivos tropicais, guardas em tecido ou papel marmorizado.

A característica dos álbuns fotográficos que se destaca e merece um aprofundamento é o acúmulo, uma necessidade que pode ser contextualizada no número crescente de imagens colecionadas. Bloom (2003) destaca que esta ação colecionadora surge na sociedade a partir do crescimento do espírito científico, característico do Renascimento (século XVI). Havia neste período “[...] uma grande quantidade de coleções que procuravam explorar e apresentar o mundo como ele parecia àquela altura”. (BLOOM, 2003, p. 35).

Bloom (2003, p. 35) questiona se “seria no século XVI que a Europa viveu seu primeiro surto de atividade colecionadora, na verdade a primeira atividade colecionadora que não se restringe a poucas pessoas desde os tempos de Roma?”. A resposta a tal indagação, está fundamentada em uma explicação mundana, vinculada a expansão do conhecimento ocorrida no século XVI, que exigia “[...] novas respostas, novas abordagens para os novos fenômenos. Juntamente com essa revolução mundana, ocorria uma mudança na maneira de perceber a morte e o mundo material” (BLOOM, 2003, p. 35).

Este acúmulo se constituiu em uma tarefa árdua, de organizar, selecionar imagens, legendas, adornos, que se alinhassem com o tema do álbum fotográfico. Esta característica particular, acrescida, por vezes, de recursos sonoros, permite a análise que sua composição promove uma conceitualização, que tem entre os seus enunciados ainda mais materialidade.

Silva (2008, p. 18) corrobora com essa caracterização, acrescentando que:

[...] o álbum conta histórias, mas não somente sobre fotos, pois a ele são acrescentados outros objetos: cartões, lembretes, recortes de jornal, relíquias e partes do corpo: umbigos de recém-nascidos, gotas de sangue, mechas de cabelo, unhas de mãos e marcas de pés. Em sentido literal, o álbum é um pedaço de nossos corpos.

A reprodução em grande escala fomenta o acúmulo e valorização das fotografias, das imagens e dos álbuns fotográficos. O acúmulo, dos registros do conhecimento, neste caso imagético, surge quando as máquinas fotográficas, e antes delas no século XV, Gutenberg com a sua prensa gráfica, que também incluía registros para além dos textuais, abrem o caminho para uma sociedade que vem de mudanças sociais e desejam registrar e acessar todo tipo de conhecimento, principalmente o imagético que permite ser visualizado e colecionado. Emanam dessa possibilidade álbuns temáticos com toda a sorte de conteúdo.

Dada a essa diversidade e popularidade, o álbum passou a ser observado a partir de várias áreas do

conhecimento, como Enfermagem Fundamental, Odontologia, História, educação, Artes visuais, Ciência da Informação, Literatura, Sociologia, Ciência da Computação, Geografia, Engenharia de Alimentos, Nutrição, Biologia entre outras onde passou a ser abordado, ora como objeto, ora como um instrumento de pesquisa.

Na Ciência da Informação, os estudos desse tema concentram-se especialmente nos domínios da OC, onde é compreendido como um recurso informacional com características peculiares que demanda de uma organização tal que permita não apenas o acesso físico, mas a compreensão de sua narrativa.

Buscando avançar no conhecimento acerca do conceito de álbum, este estudo visa identificar nas produções de pós-graduação cadastradas na Biblioteca Digital de Teses e Dissertações (BDTD) novas concepções e abordagens para o tema.

Há em destaque o fato que com o surgimento da fotografia, do álbum fotográfico no século XIX, bem como a sua reprodução em grande escala, tem-se a presença destes em centros de informação, arquivos, bibliotecas e museus. Um acúmulo e valorização das imagens, que fomentam a necessidade de tratamento visando o seu acesso, bem como, preocupações vinculadas ao contexto sociocultural de sua produção.

4 Metodologia

As análises empreendidas neste trabalho decorrem de uma pesquisa exploratória e descritiva que traz como delineamento metodológico o estudo bibliográfico de abordagem qualitativa, com o intuito de sintetizar resultados trazidos por outros estudos no intuito de ampliar a compreensão sobre o tema (CRESWELL, 2016).

O levantamento bibliográfico no portal da BDTD foi realizado em março de 2023, utilizando os termos: “álbum fotográfico” e “álbum de fotografias” nas variações singular e plural, sem delimitação de tempo ou mesmo área científica. Importa mencionar que esta base de dados visa reunir e “[...] disseminar os textos completos das teses e dissertações defendidas nas instituições brasileiras de ensino e pesquisa”. (BDTD, 2023 online) Trata-se de uma ferramenta que “proporciona a maior visibilidade à produção científica nacional e a difusão de informações de interesse científico e tecnológico para a sociedade em geral.” (BDTD, 2023 online).

Os dados foram analisados de acordo com a técnica de Análise Temática Categórica compreendida como uma variante da técnica de Análise de Conteúdo que “se propõe a descobrir os núcleos de sentido que compõem uma comunicação, cuja presença ou frequência seja significativa para o objetivo analítico visado”

(MINAYO, 2010, p. 316). A operacionalização desta técnica se dá pelo desmembramento do texto em unidades ou categorias, segundo reagrupamentos analógicos. A “análise temática é rápida e eficaz na condição de se aplicar a discursos diretos (significações manifestas) e simples” (BARDIN, 2016, p. 101). Dadas essas características, acredita-se que a metodologia apresentada seja pertinente aos objetivos propostos.

5 Apresentação dos Resultados

Partindo da questão inicialmente proposta - Qual a compreensão/extensão de “álbum fotográfico” apresentada pelos estudos desenvolvidos no âmbito da pós-graduação Brasileira e como este se relaciona com a OC? - e empregada a metodologia prevista, chegou-se ao resultado de 73 trabalhos abordando o tema. Após a primeira análise destes trabalhos, verificou-se que 23 eram repetidos e 2 não apresentaram o termo “álbum fotográfico” no resumo, portanto foram desconsiderados. Dessa forma, restaram 48 trabalhos que foram examinados a partir das informações trazidas em seus resumos. A tabela a seguir sintetiza o processo de seleção dos documentos para análise

Tabela 1. *Processo de seleção dos documentos para análise*

Documentos recuperados	73
Repetidos	23
Não possuem o termo “álbum” no resumo	2
<i>Total de documentos analisados</i>	<i>48</i>

Fonte: Elaborada pelos autores.

Inicialmente, este estudo traçou um percurso teórico no intuito de revelar as principais aproximações entre os conceitos e a atuação da OC, evidenciando as relações destes com as compreensões e extensões provenientes de álbum fotográfico.

Em seguida, foram identificados os enunciados relacionados às compreensões e extensões de álbum, a partir dos trabalhos selecionados, tomando como base a percepção de que

À medida que um conceito adquire maior extensão perde em compreensão. De maneira inversa à medida que em que enriquecemos a compreensão de um conceito-fornecendo elementos que melhoram o delimitam - a extensão diminui, pois o conceito deixa de se aplicar a muitos casos e passa a corresponder a um menor número de casos ou objetos que devem corresponder necessariamente a todos os elementos agregados a compreensão do conceito (BARROS, 2016, p.72).

Assim, os principais enunciados encontrados para designar e caracterizar o álbum fotográfico, encontram-se representados na figura 1, abaixo.

Figura 1. *Enunciados relacionadas a delimitação da compreensão e extensão acerca de álbum fotográfico*



Fonte: Elaborada pelos autores.

Observa-se a partir da figura a forte relação delimitação de álbum com a ideia de registro da informação e do conhecimento, que atua com um documento, um recurso visual que traz uma narrativa que contribui para preservação da cultura e memória de determinadas pessoas e localidades. É percebido ainda como um recurso informacional, que se dissemina a partir da ideia de colecionismo, e da iminente necessidade de classificação e organização do conhecimento com vistas a sua materialidade e circulação.

O álbum é associado ainda a um arquivo de imagens pessoais, com marcas particulares e de cunho pessoal, um diário íntimo, que diante das tecnologias de informação e comunicação (TICs) torna-se online, e com isso mais facilmente acessado e compartilhado.

Há certa necessidade de se obter definições corretas acerca dos conceitos, pois estes se atrelam ao desenvolvimento do conhecimento e da linguagem como visto acima. Barros (2016) corrobora com o tema afirmando que existem características que ajudam a delimitar com precisão um determinado conceito, atrelado a esta afirmação, tem-se a busca pela eficiência, contextualização e o respeito às apropriações sociais.

A importância das definições evidencia-se também quando se tem em vista a comunicação do conhecimento. Neste sentido, Barros (2016) afirma que que a constituição da compreensão de um conceito tem suas bases na busca por características apropriadas, pois com características insuficientes, inapropriadas o conceito não terá a capacidade de interagir entre si e com o todo evidencia-se, nesse sentido, o mérito desta busca pela compreensão e extensão dos conceitos, pois estes fazem parte das enciclopédias e dicionários, por exemplo, que são fontes relevantes de informação.

Devido à importância das definições reveladas acima, em somatória as ideias de Barros (2016) é necessário considerar também, a questão do “arejamento conceitual”, ou seja, o fato de que os conceitos, em alguns casos, necessitam de uma atualização, um ajuste teórico de suas características, visto que a versão disponível, já não corresponde à realidade. Estes ajustes, revisões podem ser percebidos em alguns momentos da análise realizada, especialmente em relação à fisicalidade do álbum fotográfico, uma vez que se observa uma transição para uma questão mais híbrida, físico-digital.

Outra questão bastante evidente é que as funções e finalidades do álbum na sociedade vão além da cultural, alcançando questões acadêmicas. A exemplo do que propõe a dissertação de Barbieri (2009), em que este registro informacional é utilizado para demonstrar estimativa do volume da porção alimentar de indivíduos com diabetes mellitus tipo 2.

A pesquisa de Azarias (2020) realizada na área de Nutrição e Saúde Pública também percebe o álbum como um recurso imagético para demonstrar o volume das porções de alimentos e dos utensílios utilizados durante o preparo. A principal diferença é que desta vez, é utilizado no formato digital.

A dissertação de Sales (2002), na área de Sociologia, numa perspectiva distinta, buscou mapear e analisar a gama das dimensões, significados e papéis sociais da fotografia em uma família de classe popular. Para atingir tais objetivos, utilizou-se de um álbum fotográfico como objeto de suas análises.

A tese de Arruda (2013) destaca um trecho de uma notícia publicada no jornal do Correio Paulistano de 30 agosto 1876, para ilustrar a importância cultural e social do álbum de Walter Bradley, alvo das análises de sua tese:

[...] teve a ótima ideia de tirar, em photographia, grande quantidade de vistas desta capital, em ponto pequeno, e com as quaes formou uma espécie de album a que intitolou — Lembranças de S. Paulo. Contem cada elegante livrinho desses, diversas vistas não só de ruas e praças da cidade como também de muitas igrejas e edificios importantes, sendo o trabalho photographico feito com todo o esmero artístico (ARRUDA, 2013, p. 126).

Observa-se neste fragmento da notícia de 1876, que o álbum é comparado a um livro, talvez pelo seu formato, e ainda que lhe é atribuído forte vínculo com a questão artística.

Cavalcanti (2013) em sua tese, na área da Ciência da Computação, desenvolveu um sistema destinado a dar suporte aos seguintes processos: (1) a seleção de fotografias de interesse e (2) a estruturação diagramada das fotografias de interesse em álbuns fotográficos, demonstrando a preocupação com a seleção, classificação e organização das fotografias, problemas

característicos da área de seu trabalho (CAVALCANTI, 2013, p. 1).

Já Santos (2017) em sua dissertação também na área da Ciência da Informação, propõe uma abordagem de álbum fotográfico como um recurso informacional considerando o formato digital ou físico. Para a autora, o álbum é:

[...] um documento que guarda as informações, representações, e conhecimento de uma instituição, de um grupo ou mesmo de um determinado período de tempo e espaço. Tal fonte de informação possui o pressuposto de ser projetado e alimentado pela mente humana (SANTOS, 2017, p. 9).

Com o intuito de aprofundar as análises acerca do conceito de álbum fotográfico, busca-se também observar as definições direcionadas ao álbum fotográfico elencadas em alguns dicionários e tesouros. Isso porque estes materiais desenvolvem listas, definições ou mesmo analisam termos diversos em determinados domínios do saber. Estas produções exercem especial contribuição na construção, compreensão e disseminação de conceitos, por isso são muito valiosas à OC.

Nesta perspectiva, Barité (2015, p. 13) destaca que

La Bibliotecología cuyo objeto de estudio es la organización de la información para transferirla y darla a conocer a la sociedad también analiza los términos y los conceptos asociados porque son la forma de vehicularlos contenidos documentales abreviados. [...] Los diccionarios especializados, tanto impresos como digitales, constituyen una de las grandes herramientas de consulta para el mediador lingüístico, ya sea bibliotecólogo, traductor, terminólogo o informático.

Figueiredo (1913) no Novo Dicionário da Língua Portuguesa define álbum fotográfico como um “livro, cujas fôlhas são molduras de cartão para guardar retratos. Livro, cujas fôlhas são destinadas a desenhos, versos, lembranças de família e de amigos.” (FIGUEIREDO, 1913, p. 70) Além de serem “Tábua branca, em que se publicaram os edictos do pretor romano. (*Lat. album*)” (FIGUEIREDO, 1913, p. 70). Observa-se novamente a aproximação dos enunciados de álbum ao conceito de livro.

Cunha e Cavalcanti (2008) desenvolveram um dicionário de Biblioteconomia e Arquivologia, com objetivo de definir de forma clara, sucinta e simples, “[...] os termos utilizados por bibliotecários, arquivistas e demais profissionais da ampla e multifacetada área de Ciência da Informação, facilitando a expansão de seus conhecimentos” (CUNHA; CAVALCANTI, 2008, p. 8).

Neste dicionário é apresentada uma sequência muito interessante para a definição de álbum fotográfico: “Caderno ou classificador pessoal, que se destina a receber desenhos, fotografias, autógrafos e coleções diversas, p.ex.: álbum de retratos, álbum de selos.”

(CUNHA; CAVALCANTI, 2008, p. 9). E sequencialmente “1.2 Coletânea impressa de ilustrações e documentos iconográficos, documentação iconográfica” (CUNHA; CAVALCANTI, 2008, p. 9).

Verifica-se assim, que álbum tem sido citado nas diferentes áreas do conhecimento, e na atualidade seu conceito tem absorvido novas perspectivas, trazidas pelos desenvolvimento tecnológico, a exemplo do que expõe Azarias (2020) ao perceber o álbum como um recurso imagético digital, ao qual pode ser atribuídas várias possibilidades de configurações, publicações, postagens, compartilhamentos e visualizações e interações especialmente, quando disponibilizado nas redes sociais.

Face ao exposto, observa-se que as características de álbum encontradas nos trabalhos e por fim, apresentadas nos dicionários compõem a extensão e compreensão do seu conceito. É válido lembrar que a compreensão de um conceito ou as suas características e extensão se estabelecem e se transformam ao longo do tempo. Assim, novas características aos poucos vão se infiltrando e modificando a compreensão e a extensão do conceito. Dessa forma, percebe-se que a compreensão de uma característica não é apenas constituída pelas suas características, mas também pelas relações entre as qualidades/propriedades que compõem o conceito (BARROS, 2016).

Isso significa dizer que o álbum, embora tenha se modificado com o passar dos tempos, principalmente com o desenvolvimento das tecnologias, segue se destacando como instrumento de organização do conhecimento e preservação da memória individual ou coletiva, que conta com uma forma singular de classificação, pautada, especialmente, na liberdade, criatividade e intencionalidade do editor e da narrativa que se pretende construir.

No que diz respeito à sua extensão observa-se um deslocamento de recurso informacional físico, com páginas e ornamentos, fechaduras e fotografias impressas para recurso informacional imagético, que pode ser encontrado tanto em suporte físico/analógico, quanto digital. Constituindo-se, para além destas questões de seu suporte, como uma rica fonte de pesquisa para diferentes áreas do saber.

6 Conclusão

Este estudo buscou explorar a compreensão e a extensão do conceito de “álbum fotográfico” nas publicações que integram os sistemas de informação de teses e dissertações brasileiras. A busca pela compreensão da extensão de um conceito comum a diferentes ciências, como é o caso do álbum fotográfico, visa entender, ou mesmo, estabelecer as características principais que interferem diretamente nos ambientes informacionais.

Os resultados evidenciam que os enunciados acerca do álbum apresentam forte relação com a ideia de registro da informação e do conhecimento, equivalente a um documento, um recurso visual que traz uma narrativa com vistas a contribuir para a preservação da cultura e da memória. É também um recurso informacional que traz consigo os princípios do colecionismo e a iminente necessidade de classificação e organização do conhecimento visando a sua materialidade e circulação.

O álbum é associado ainda a um arquivo de imagens pessoais, com marcas particulares e de cunho pessoal, um diário íntimo, que diante das tecnologias de informação e comunicação (TICs) torna-se online, logo, mais facilmente acessado e compartilhado.

Face a compreensão e a extensão do conceito de “álbum fotográfico” evidencia-se que as relações entre o álbum e a OC se estabelecem na medida em que o primeiro se infiltrou nos espaços informacionais, gerando necessidades de organização, que tem em sua base as definições, o nomear, o conceituar e a identificação das principais características que servirão de base para a representação e organização do conhecimento, com vistas a sua recuperação.

Sob a chave das características elencadas a respeito de álbum fotográfico, percebe-se que quanto mais próximos são do período de origem do álbum, século XIX, maior a ocorrência de características e compreensão vinculadas a materialidade, como o seu formato, ornamentações e materiais utilizados na sua produção, da mesma forma que quanto mais contemporâneos são as produções teóricas que destinam parte do seu texto a percepção do conceito álbum este se vincula ao digital, por exemplo, mas há sempre presente os enunciados vinculados a memória, a coleção de imagens fotográficas e a construção da narrativa visual.

O álbum fotográfico, desta forma, caracteriza-se como uma forma de registro do conhecimento e preservação da memória individual ou coletiva, que traz métodos peculiares de coleção/organização e, principalmente, uma narrativa específica acerca da temática representada, que conta com uma forma singular de organização pautada, especialmente, na liberdade, criatividade e intencionalidade do editor.

Referências

- ABBAGNANO, Nicola. **Dicionário de filosofia**. Cidade do México: FCE, 2004.
- ARRUDA, Rogério Pereira de. **Cidades-Capitais imaginadas pela fotografia**: La Plata (Argentina), Belo Horizonte (Brasil), 1880 – 1897. Fino Traço: Belo Horizonte, 2013.
- AZARIAS, Heloisa Gambarelli de Araújo. **Validade e reprodutibilidade do questionário de frequência alimentar com suporte de imagem da Coorte de Universidades Mineiras (projeto CUME)**. 2020.
- Dissertação (Mestrado em Nutrição e Saúde) -- Universidade Federal de Minas Gerais, Belo Horizonte, 2020. Disponível em: <http://hdl.handle.net/1843/35872>. Acesso em: 22 maio 2023.
- BARBIERI, Anelise Faloni Siman. **Estimativa da porção alimentar de indivíduos com diabetes mellitus tipo 2 utilizando o álbum fotográfico**. 2009, 130f. Dissertação. (Mestrado em Enfermagem) -- Escola de Enfermagem de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2009. Disponível em: <https://www.teses.usp.br/teses/disponiveis/22/22132/tde-18082009-130933/pt-br.php>. Acesso em: 22 maio 2023.
- BARDIN, Laurence. **Análise de conteúdo**. São Paulo: Edições 70, 2016.
- BARITÉ, Mario. **Diccionario de Organización del Conocimiento**: Clasificación, Indización, Terminología. 6. ed. corregida y aumentada. Montevideo: Ediciones Universitarias, 2015.
- BARROS, José D'Assunção. **Os conceitos**: Seus usos nas ciências humanas. Petrópolis, RJ: Vozes, 2016. 204.
- BENJAMIN, Walter. Pequena história da fotografia. (Trad. Sérgio P. Rouanet). In: BENJAMIN, Walter. **Magia e técnica, arte e política**: ensaios sobre literatura e história da cultura. São Paulo: Brasiliense, 1994.
- BDTD. **BIBLIOTECA BRASILEIRA DE TESES E DISSERTAÇÕES**. Disponível em: <https://bdtb.ibict.br/vufind/>. Acesso em: 05 mar. 2023.
- BLOOM, Philipp. **Ter e manter**: uma história íntima de colecionadores e coleções. Rio de Janeiro: Record, 2003.
- BORGES, Maria Eliza L. **História & Fotografia**. Belo Horizonte: Autêntica, 2005.
- CAVALCANTI, Cláudio Sebastião Vasconcelos da Cunha. **Análise e diagramação automática de álbuns fotográficos**. 2013. 270 f. Tese (Doutorado em Ciência da Computação) -- Centro de Engenharia Elétrica e Informática, Universidade Federal de Campina Grande, Campina Grande, 2013. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/6704>. Acesso em: 5 abr. 2023.
- CRESWELL, John W. **Projeto de pesquisa**: métodos qualitativos, quantitativos e mistos. 3. ed. Porto Alegre: Artmed, 2016.
- CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Briquet de Lemos, 2008.
- DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**. Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978.
- FABRIS, Annateresa. A invenção da fotografia: repercussões sociais. In: FABRIS, Annateresa. (org.) **Fotografia**: usos e funções no século XIX. São Paulo: Editora da Universidade de São Paulo, 2008, p. 11-37.
- FERRAZ, Rosane Carmanini. Memória e narrativa visual nos álbuns de fotografias oitocentistas das famílias Ferreira Lage e Cavalcanti. **Resgate**: Revista Interdisciplinar de Cultura, Campinas, v. 22, n. 2, p. 63-72, 2014. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/resgate/artic/e/view/8645780/13079>. Acesso em: 5 abr. 2023.

FIGUEIREDO, Cândido de. **Novo Dicionário da Língua Portuguesa**. Lisboa: Livraria Clássica Editora, 1913.

HJORLAND, Birger. What is a Knowledge Organization (KO)? **Knowledge Organization**, Frankfurt, Alemanha, v. 35, n. 2-3, p. 86-101, 2008.

KOBASHI, Nair Yumiko; FRANCELIN, Marivalde Moacir. Conceitos, categorias e organização do conhecimento. **Informação & Informação**, v. 16, n. 2, p. 1-24, 2011.

KOSSOY, Boris. **Realidades e Ficções na Trama** Fotográfica. Cotia/SP: Ateliê Editorial, 2001.

MANINI, Miriam Paula. A fotografia como registro e como documento de arquivo. In: BARTOLO; Linete; MORENO, Nádina Aparecida. (Orgs.). **Gestão em Arquivologia: abordagens múltiplas**. Londrina: EDUEL, 2008, v. 1, p. 102-161.

MAUAD, Ana Maria. Entre retratos e paisagens: modos de ver e representar no Brasil oitocentista. **Studium**, Campinas, n. 15, p. 3-9, 2004. Disponível em: <https://econtents.bc.unicamp.br/inpec/index.php/studium/article/view/11764/7706>. Acesso em: 5 abr. 2023.

MINAYO, Maria Cecília de Souza. **O desafio do conhecimento: pesquisa qualitativa em saúde**. 12 ed. São Paulo: Hucitec, 2010.

ROUILLE, André. **A fotografia: entre documento e arte contemporânea**. São Paulo: SENAC, 2009.

SALES, Bruno Sampaio. **Gritos silenciosos: trajetória e significados do álbum de família em um conjunto habitacional**. 2012. 127f. – Dissertação (Mestrado) – Universidade Federal do Ceará, Programa de Pós-graduação em Sociologia, Fortaleza (CE), 2012.

SANTOS, Cristina Ribeiro dos. **Gênese documental de álbuns fotográficos escolares: entre avanços e ausências**. 2017. 123f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual de Londrina, Londrina, 2017. Disponível em: <http://www.bibliotecadigital.uel.br/document/?code=vtls000214285>. Acesso em: 15 mai. 2023.

SILVA, Armando. **Álbum de família: a imagem de nós mesmos**. São Paulo: SENAC, 2008.

TURAZZI, Maria Inez. **Marc Ferrez**. Rio de Janeiro: Cosac & Naify, 2002.

O capacitismo e a ética na representação do conhecimento: um estudo quanto à ética na representação do conhecimento sobre a temática da neuroatipicidade

Ableism and ethics in the knowledge representation: a study regarding ethics in knowledge representation on the topic of neuroatypicality

Fernanda Felite Teixeira (1), Ilemar Christina Lansoni Wey Berti (2)

(1) Universidade Estadual de Londrina, Londrina, fernanda.felite@uel.br, (2) christinaberti@uel.br

Resumo

Contextualização: Propomo-nos neste artigo uma investigação em relação à influência dos processos na representação de assuntos, tendo o enfoque na perpetuação de estigmas sociais que tornam o capacitismo uma vivência comum entre pessoas com deficiência, com o recorte às pessoas neuroatípicas. Objetivo: Os objetivos neste estudo é compreender o que é o capacitismo e como ele atinge pessoas com deficiência; além de refletir como a ética na representação de assuntos pode colaborar no rompimento de questões capacitistas do processo social. Metodologia: A pesquisa trata de um estudo bibliográfico, exploratório, de natureza básica e abordagem qualitativa, a partir da análise de artigos das áreas do conhecimento: ciência da informação, educação, psiquiatria e psicologia, recuperados pelos termos: capacitismo E ética, ética E representação do conhecimento nas plataformas Scielo e BDTD. Nas buscas nas bases, recuperou-se um total de 211 documentos dos quais foram selecionados 4 artigos, considerando como critério os resumos que apresentavam o foco na descrição do capacitismo e a ética e ética e representação de assunto. Conclusão: Através dos resultados obtidos no processo de elaboração da pesquisa, foi possível detectar que a representação de assuntos possui o potencial de contribuir com o enfrentamento das condições capacitistas existentes na sociedade atual, ao se utilizar de seus recursos para estabelecer parâmetros para novas expressões a serem utilizadas para a realização de buscas por assuntos referentes ao grupo com neuroatipicidades.

Palavras-chave: Representação do conhecimento. Capacitismo. Ética. Neuroatipicidades.

Abstract

Context: The article proposes an investigation into the influence of processes in the representation of issues. With a focus on the perpetuation of social stigmas that make ableism a common experience among people with disabilities, with an emphasis on neurodiverse individuals. Objective: This study aims to understand what ableism is and how it affects people with disabilities also reflects on how ethics in the representation of issues can contribute to breaking ableist matter in the social process. Methodology: The research is an exploratory bibliographic study, of basal nature and qualitative approach, based on the analysis of articles in the knowledge areas of information science, education, psychiatry, and psychology, retrieved by the terms "ableism AND ethics," "ethics AND representation of knowledge" on the platforms Scielo and BDTD. In the search on the databases, a total of 211 retrieved documents, of which four selected articles were considered as a criterion for the abstracts that focused on the description of ableism and ethics, and ethics and subject representation. Conclusion: Based on the results obtained in the research process, it was possible to detect that the issue representation has the potential to contribute to facing the ableist conditions that exist in a society of the day by using its resources to establish parameters for new expressions to employ for searches related to the group of people with neurodiversity.

Keywords: Knowledge representation. Ethics. Neuroatypicalities. Ableism.

Resumen

Contexto: El artículo plantea una investigación sobre la influencia de los procesos en la representación del conocimiento, enfocándose en la perpetuación de los estigmas sociales que hacen del capacitismo una experiencia común entre las personas con discapacidad, con un enfoque en las personas neuroatípicas. Objetivo: El objetivo de este estudio es comprender qué es el capacitismo y cómo afecta a las personas con discapacidad, así como reflexionar sobre cómo la ética en la representación del conocimiento puede contribuir a romper cuestiones capacitivas en el proceso social. Metodología: La investigación se basa en un estudio bibliográfico, exploratorio, de naturaleza básica y enfoque cualitativo, a partir del análisis de artículos de áreas de conocimiento como ciencia de la información, educación, psiquiatria y psicología, recuperados por los términos "capacitismo y ética", "ética y representación del conocimiento" en las plataformas Scielo y BDTD. En las búsquedas en las bases de datos se recuperaron un total de 211 documentos, de los cuales se seleccionaron 4 artículos, considerando como criterio los resúmenes que presentaban el enfoque en la descripción del capacitismo, la ética y la representación del conocimiento. Conclusión: Los resultados obtenidos en el proceso de elaboración de la investigación revelaron que la representación del conocimiento tiene el potencial de contribuir al enfrentamiento de las condiciones capacitivas existentes en la sociedad actual, al utilizar sus recursos para establecer parámetros para nuevas expresiones que se utilizarán en la búsqueda de temas relacionados con el grupo de personas con neuroatipicidades.

Palabras clave: Representación del conocimiento. Ableísmo. Ética. Neuroatipicidades.

1 Introdução

A sociedade humana é repleta por uma série de conceitos, tabus, estereótipos e padrões estruturados e, em sua maioria, inalcançáveis e excludentes. Padrões que excluem todos que não se enquadram, seja por não possuírem corpos em tamanhos, cores ou etnias “adequadas” ou por possuírem alguma deficiência, independentemente de ser: física, intelectual ou sensorial.

Pode descrever-se tais padrões sociais através do uso de termos como a corponormatividade, advinda de *Compulsory able-bodiedness* que, de acordo com Mello (2014), seria uma tradução adequada para o termo em inglês utilizado com a proposta de descrever corpos capazes ou aptos ao serviço militar, por exemplo.

Outro termo presente nas discussões é o capacitismo, que entra em pauta para suprir a necessidade de uma maneira de representar as violências sofridas diariamente por pessoas com alguma deficiência. O sofrimento destas pessoas não se trata exclusivamente da violência física ou psicológica, mas permeia os ambientes de convívio público, de segregação pelo descrédito social, pela falta de recursos que possibilitem uma vivência igualitária, independente das características de cada ser humano.

Partindo desta perspectiva, pode dizer-se que atitudes capacitistas fogem aos padrões éticos definidos socialmente ao longo da trajetória histórica da humanidade, em relação à importância de cada ser humano no mundo. Desta maneira, para uma melhor explanação quanto ao que se faz ser ético, é necessário compreender do que se trata tal questão e como a ética influi nos processos de inclusão, exclusão e discriminação.

Busca-se o entendimento do que é a ética dentro deste contexto, ao qual compreende-se que o “ser ético” perpassa as diversas áreas da sociedade, ou seja, pode ser entendida como um parâmetro de consciencialização sobre o certo e errado na atividade humana.

Segundo Crisostomo *et al.* (2018), os primeiros filósofos diziam que a “Ética é a morada do homem” e esse seria então um local onde os habitantes teriam segurança e que se vivessem segundo as normas e leis existentes nas *pólis*; a sociedade seria melhor, tornando os indivíduos protegidos e confiantes.

Já no século XIX, Friedrich Nietzsche trabalhou sobre uma nova maneira de abordar a ética, separando-a da esfera religiosa e formulando um conceito científico, ao observar os preceitos da moralidade e da ética vigentes na sociedade que reprime a natureza humana e torna, assim, possível a convivência humana em sociedade.

A ética na contemporaneidade pode abranger os mais diversos contextos sociais, como no âmbito profissional.

Portanto, permanece sendo foco constante de pesquisadores, presentes em debates e pesquisas, sendo estudada e atualizada conceitualmente com o objetivo de contemplar a realidade atual de cada sociedade.

Neste artigo, a ética é descrita como um conjunto de regras e padrões estabelecidos socialmente, para que assim seja uma base de harmonia e segurança no convívio humano. Portanto, é importante ressaltar que ao serem desenvolvidos socialmente, os parâmetros quanto ao que é ou o que deixa de ser ético, podem sofrer alterações através da conjectura social vigente.

As normas éticas estão presentes em todos os aspectos sociais, inclusive nas áreas de atuação profissional, conforme o apresentado por Crisostomo *et al.* (2018) sobre os estabelecimentos expostos nos devidos códigos de ética profissional, dentre estas a área da ciência da informação. Com foco na organização e representação do conhecimento, este trabalho se refere a um conjunto de procedimentos com o fim de expressar o conteúdo de documentos científicos e facilitar a recuperação da informação sobre a ética na representação de assunto e os conceitos estruturantes capacitistas.

A atividade de representação do conhecimento, enquanto atividade profissional na área da informação, possibilita a união entre o conhecimento produzido e a geração de um novo conhecimento, em sua dimensão social, os quais possuem valores éticos que devem ser identificados e estudados.

A representação do conhecimento não é e nem se define enquanto uma atividade meramente técnica, mas sim, uma atividade intelectual que exige do profissional uma postura consciente e crítica aos valores humanistas. Ademais, é importante que se conheça os aspectos históricos e sociais que envolvem o conhecimento a ser representado, ou seja, o domínio ao qual o conhecimento está relacionado.

Sendo assim, tanto a atividade profissional como os instrumentos utilizados e criados por ele, podem ser afetados por uma visão parcial sobre determinado tema e, portanto, estar suscetíveis à presença de preconceitos e distorções. Considerando que por mais que seja idealizado o profissional neutro, imparcial, sem dogmas, preconceitos ou tradições, que reflitam no exercício de suas atividades profissionais e na construção de certos instrumentos de trabalho.

Considerando o potencial de influência já descrito anteriormente, pode dizer-se que o uso dos termos que reforçam discriminações sociais pode acarretar a manutenção dessas questões como: machismo, racismo, xenofobia, homofobia e o capacitismo, sendo esse último, abordado neste estudo.

Trata-se nesta perspectiva, da necessidade de mencionar a ação do “fazer” do bibliotecário, responsável pela representação de assuntos. De forma que há riscos quando à ausência reflexiva e demasiadamente

conservadora do bibliotecário que incorre no emprego de termos capacitistas na representação de assuntos, reforça estereótipos ou expõe barreiras já enfrentadas por pessoas com deficiência.

Visando responder sobre a relação do capacitismo e a representação de assunto, de maneira a refletir sobre os malefícios causados às pessoas com deficiência, especialmente nos casos das neuroatipicidades, os seguintes objetivos foram elaborados: Compreender o que é o capacitismo e como ele atinge pessoas com deficiência e refletir sobre a ética na representação de assuntos, de forma a colaborar com o rompimento de questões capacitistas.

A pesquisa se justifica pela contribuição reflexiva, conduzida pelo direcionamento à temática na área acadêmica, trazendo a discussão ao campo da Ciência da informação, visando à melhoria nas condições de convívio social para pessoas com deficiência.

Para abordar os conteúdos propostos, procedeu-se do ponto de vista metodológico: elaboração de uma pesquisa com base em estudo bibliográfico, exploratório, de natureza básica e qualitativa; num primeiro momento foram realizadas buscas pelos termos: capacitismo E ética, ética E representação do conhecimento nas plataformas, Scielo e BDTD.

Para o refinamento dos resultados obtidos realizou-se a leitura dos resumos e a partir deles foram selecionados 4 dos trabalhos que embasam este estudo.

Além de materiais da ciência da informação e educação, utilizaram-se artigos e publicações de áreas médicas ou psiquiátricas, visando à melhor compreensão quanto às neuroatipicidades e suas características, a fim de abordar o assunto do ponto de vista técnico-científico.

2 Desenvolvimento

Nesta seção são abordadas as questões que fundamentam os conceitos referentes à ética, ao profissional bibliotecário na ação de representação de assuntos e o capacitismo e neuroatipicidades. Com o objetivo de ampliar o referencial teórico dos assuntos abordados e possibilitar a elaboração de propostas, os assuntos foram mencionados no texto de forma intencionalmente reflexiva.

2.1 Ética na representação de assuntos

A ética enquanto tema pode fazer parte de discussões em todas as áreas do conhecimento. Compreender o que é a ética e como ela é necessária para o pleno exercício das diversas atividades profissionais, torna o tema plural e relevante, pois envolve aspectos epistemológicos do saber fazer, imbricado com as relações humanas. Contudo, em uma breve conceituação do termo e de seu papel em meio social, é possível considerar uma

complexidade de ordem filosófica, que não é aprofundada neste estudo exploratório.

De acordo com o exposto no dicionário UNESP do português contemporâneo (2004), a “Ética pode ser entendida como um conjunto [...] de conhecimentos, valores morais, princípios, ideias de conduta, bem como um modelo de comportamento a ser seguido pela sociedade e/ou no exercício de uma profissão” (BORBA, 2004).

Já de acordo com CHAÚÍ (2000, p. 434) “[...] a ética exprime a maneira como uma cultura e uma sociedade definem para si mesmas o que julgam ser o mal e o vício, a violência e o crime e, como contrapartida, o que consideram ser o bem e a virtude.” É ainda relevante enfatizar que essas questões devem estar submetidas à autorreflexão e ao questionamento constante, considerando que cada sociedade em sua época institui aquilo que lhe é considerado enquanto ético ou não.

Com base no apresentado, pode afirmar-se que, em um contexto profissional, há aspectos que se apresentam particulares e peculiares para cada especialização, ou seja, há um complexo de valores pertinentes a cada profissão. Sob essas perspectivas são apresentadas as deontologias, que se referem à área da Filosofia, a qual “trata dos princípios, fundamentos e sistemas da moral”, sendo esses conjuntos de deveres, porém vinculados às questões profissionais (BORBA, 2004).

Assim, a deontologia é um conjunto de regras que se expressam de maneira formal e explícita e aquele que a transgredir é passível de sanção. Nesse sentido, a deontologia não pode ser igualada à ética profissional que, por sua vez, está “além da manutenção da integridade e da coerência de uma prática” (SIROUX, 2007, p. 405)

Um exemplo de código ético e deontológico relacionado à área da informação é o aprovado pelo Conselho Federal de Biblioteconomia, que estabelece parâmetros éticos a serem seguidos por bibliotecários de todo o Brasil.

Para o fim de maior compreensão quanto ao que se propõe tal documento, pode citar-se que a resolução CFB nº 207/2018 apresenta 17 artigos, que fundamentam as normas orientadoras da conduta do exercício e das atividades profissionais do bibliotecário. Tendo por objetivo “fixar as normas orientadoras de conduta no exercício de suas atividades profissionais”. (CONSELHO FEDERAL DE BIBLIOTECONOMIA, 2018)

A natureza, fundamento e objeto de trabalho, assim descrito, abrange além da ética e diretrizes da profissão, os direitos humanos e os deveres de um bom profissional. O conselho imprime também no código os deveres e proibições, classificando e elencando o que são as infrações ético-disciplinares que podem ocorrer e as possíveis penalidades a serem tomadas pelo órgão regulador.

Destaca-se que o texto do Conselho Federal de Biblioteconomia se estrutura no artigo quinto da norma ao ressaltar os deveres e enfatizar ações com respeito à dignidade da pessoa humana, zelo, honestidade, cooperação, colaboração e o cumprimento das diretrizes determinadas pela deontologia descrita. Nesse sentido, o que se apresenta nesse contexto é o reflexo do entendimento da sociedade vigente, quanto ao que é ser um profissional ético.

Assim, os deveres de um bibliotecário, conforme o expresso pelo código, é o que a sociedade esperaria de um bom profissional, porém além de um texto que reflete meramente uma ética social em forma de resolução, a redação possui uma finalidade, fortemente enfatizada no combate às infrações éticas-disciplinares, através da aplicação de penalidades.

O recorrido ao longo dos artigos e parágrafos, da norma da profissão, fixa em especificar diversas problemáticas passíveis de serem exercidas por bibliotecários, tais como: a negligência, o descumprimento das normas, o desrespeito, a calúnia, o roubo e entre várias outras ações que podem ocorrer nas ações do indivíduo em questão. De forma socialmente entendida, esses atos antiéticos, presentes na resolução, poderiam ser reduzidos em palavras como injúria, corrupção, preconceito e discriminação.

Considerando as discussões elencadas, coloca-se em evidência o fato de um indivíduo, indiferentemente de sua qualificação, possuir de características individuais, cultura, preconceitos, dogmas e crenças particulares, que o acompanham em seus afazeres profissionais. Entende-se nesse aspecto que a existência de um documento que forneça a garantia do exercício da profissão de maneira condizente com a ética, pode balizar a relação do indivíduo com a profissão exercida e minimizar a ocorrência de desvios éticos ou tendenciosidades, em meio ao exercício oferecido à sociedade.

Entende-se como desvios, toda atitude que foge ao considerado ético ou recomendável para o pleno exercício profissional, não sendo guiado por posicionamentos particulares ou que se evite o quanto possível, compreendendo que o profissional não será neutro em sua atuação.

Dentre os exercícios realizados por profissionais da informação, pode destacar-se, para o fim do presente estudo, a representação do conhecimento, que, segundo Barité (2015, p. 136, tradução nossa) é:

O ramo da organização do conhecimento que compreende o conjunto dos processos de simbolização notacional ou conceitual do saber humano no âmbito de qualquer disciplina. Na representação do conhecimento se compreende a classificação, a indexação e o conjunto de aspectos informáticos e linguísticos, relacionados com a tradução simbólica do conhecimento.

Trata-se do uso e desenvolvimento de recursos e atividades que possuem o objetivo de representar as informações contidas em um documento informacional, a fim de tornar tais conteúdos acessíveis, além de possibilitar o desenvolvimento de ferramentas de organização de acervos. Algumas exemplificações destas representações da informação referenciadas são os resumos, palavras-chave e sistemas de classificação.

Percebe-se, ainda, a influência da globalização, bem como das tecnologias de informação e comunicação, a atuação profissional bibliotecária tem sido repensada no intuito de contribuir de forma positiva às necessidades da sociedade como um todo. Sob essa influência, tem-se incorporado, cada vez mais, os valores morais e as virtudes éticas, a fim de responder às exigências da sociedade ou de uma cultura.

Com base nas afirmações citadas quanto as representações de assuntos elaboradas pelos bibliotecários, vale destacar as questões relacionadas à ética do cuidado e o capacitismo no contexto dos assuntos descritos. Pode afirmar-se que a demanda por adequações e combate ao uso de instrumentos de representação das formas, que se contam sobre as deficiências, pode agravar tais questões, as quais serão destacadas no tópico “Capacitismo” apresentado a seguir.

2.2 Capacitismo

Em busca pela compreensão quanto ao significado do termo capacitismo, faz-se necessário percorrer o caminho pelo qual a ciência tem se empenhado ao longo dos anos.

Esse trajeto destaca que por muito tempo não houve um termo que amparasse as questões abarcadas pelo termo capacitismo. O próprio termo, relativamente novo, surge no ano de 2012, embasado pela literatura estrangeira.

Mello (2014) apresenta uma proposta elaborada por Robert McRuer, que seria a base para discussões referentes ao que seriam os corpos pertencentes ou não a um padrão socialmente estabelecido e com base neles, também seriam definidos quais corpos poderiam ou não ter acesso a direitos e privilégios.

De acordo com Mello (2014, p. 3266), o termo

compulsory able-bodiedness, cuja tradução em português de *able-bodied* é apto, faz referência à condição de um corpo apto e fisicamente capaz para o serviço militar, por exemplo. Em linhas gerais, dependendo do contexto das frases usadas com este termo, poder-se-ia traduzir como corpos são, hábeis, aptos, capazes ou sem deficiência. Entretanto, ao invés de “aptonormatividade”, considero corponormatividade uma tradução mais inteligível para o sentido de *able-bodiedness* em português. Assim, por convenção, adotarei corpos capazes, em vez de corpos aptos. Essa distinção etimológica é necessária para o acionamento da categoria capacitismo, materializada através de atitudes preconceituosas que hierarquizam

sujeitos em função da adequação de seus corpos a um ideal de beleza e capacidade funcional.

As relevâncias de tal temática em meio as diferentes áreas do conhecimento científico se evidenciam com base nas discussões anteriores e em maiores proporções, quando passam a ser nomeadas ou representadas por um termo que as defina. Alguns exemplos que demonstram a descrição de opressão e discriminação para com o corpo, podem ser citados o “[...] sexismo, no caso da discriminação por sexo; homofobia, no caso da discriminação pela orientação sexual; racismo, no caso da discriminação pela cor da pele ou etnia.” (MELLO, 2014, p. 3267).

O direito de não ser discriminado pelo corpo que se habita é assegurado pelo sistema constitucional, que veda toda discriminação, seja ela devido ao sexo ou raça.

Entretanto, no caso das agressões quanto aos corpos de pessoas com deficiência é relevante enfatizar que até a data presente do estudo, na língua portuguesa brasileira não há um termo que evidencie, demonstre ou questione tais atitudes que se relacionam à capacidade ou a falta dela e a dependência ou independência de tais pessoas.

Partindo das questões evidenciadas quanto à falta de uma representação dessas, Mello (2014) propõem que a partir do exemplo de Portugal,

passemos a adotar no Brasil a tradução de *ableism* para capacitismo na língua portuguesa, por duas razões principais: a primeira é a demanda de urgência para visibilizar uma forma peculiar de opressão contra as pessoas com deficiência e, por consequência, dar maior visibilidade social e política a este segmento; a segunda deriva do próprio postulado da Teoria Crip, ou seja, para desconstruir as fronteiras entre deficientes e não deficientes é necessário explorar os meandros da corponormatividade de nossa estrutura social ao dar nome a um tipo de discriminação que se materializa na forma de mecanismos de interdição e de controle biopolítico de corpos com base na premissa da (in)capacidade, ou seja, no que as pessoas com deficiência podem ou são capazes de ser e fazer. (MELLO, 2014, p. 3267)

Sendo assim o capacitismo pode ser entendido enquanto um neologismo utilizado com a potencialidade de descrever o distanciamento da capacidade em quadros de deficiência.

De acordo com essas concepções, torna-se possível o apontamento de que a deficiência se constitui como fruto da relação de um corpo fora dos padrões funcionais, seja esse intelectual ou fisicamente, e um espaço incapaz de atender as demandas necessárias para que estes sujeitos possam alcançar condições de igualdade, seja na inserção ou permanência, como também na participação nos mais diversos setores da vida humana.

Assim, dois exemplos seriam os citados por Ortega (2009) quando afirma que:

Andar de cadeira de rodas é um problema apenas por vivermos em um mundo cheio de escadas e consideramos deficientes indivíduos que não olham nos olhos quando se comunicam, como é o caso dos autistas, apenas por que nossa sociedade estabelece o contacto visual como um elemento básico da interação humana. Ortega (2009, p. 68)

Partindo de tais concepções, a sociedade, por vezes, desconsidera as potencialidades das pessoas com deficiência, reduzindo-as às suas lesões. O que se torna compreensível pelo fato que tais pessoas têm, “[...] desde sempre, convivido com a confusão entre o que realmente são, pessoas humanas e o que se pensa que elas são: deficientes” (SOUZA, 2008, p.23), havendo desse modo a anulação da pessoa enquanto humana e de suas características e particularidades em razão da deficiência.

Com relação à adoção do termo capacitismo, pode citar-se que se faz presente na relação de termos adotados e conceituados pelo glossário de acessibilidade na câmara dos deputados e, segundo tal documento, o Capacitismo se trata do

ato de discriminação, preconceito ou opressão contra pessoa com deficiência. É barreira atitudinal. Em geral, ocorre quando alguém considera uma pessoa incapaz, por conta de diferenças e impedimentos corporais. O capacitismo está focalizado nas supostas ‘capacidades das pessoas sem deficiência’ como referência para mostrar as supostas ‘limitações das pessoas com deficiência’. No capacitismo, a ênfase é colocada nas supostas ‘pessoas capazes’, as quais constituem a maioria da população e são supostamente consideradas ‘normais’. (CÂMARA DOS DEPUTADOS [202-])

A adoção do termo por parte do governo possui uma simbologia de conquista para os apoiadores e pertencentes à causa, já que com tal atitude o assunto passa a alcançar um maior destaque dentre a sociedade, além de ampliar o número de pesquisas quanto à temática, possibilitando assim o debate e a melhoria nas condições e na qualidade de vivência de pessoas amparadas e representadas pelo termo.

O termo em questão no atual contexto reflete a luta de diferentes grupos com suas características e necessidades particulares, as quais serão abordadas as relacionadas ao grupo de pessoas que enfrentam barreiras vinculadas às neuroatípicas ou deficiências intelectuais. A tensão provocada pelo uso do termo, possui a função de ampliar os conhecimentos quanto ao que seriam essas atipicidades e como as sociedades as tratam ao longo dos tempos.

2.3 Neuroatipicidades

A discussão sobre neuroatipicidades está amparada na contextualização histórica sobre as questões enfrentadas por pessoas com deficiência e modos como a sociedade vê e trata indivíduos que se enquadrem em tais condições, além de entender quais condições podem ser englobadas em tal terminologia.

Para isso, pode-se contextualizar que a representação das pessoas com deficiência, na sociedade contemporânea, tem tomado grande força nas últimas décadas, todavia o passado no qual esses indivíduos perpassaram possuía ainda mais preconceito e discriminação.

Conforme descrito por Figueirêdo, Delevati e Tavares (2014), havia um modo de diferenciação para essas pessoas, tendo suas vidas dificultadas, devido as suas condições, que implicavam em problemas sociais e familiares, o que resultou na busca por instituições capazes de atender as necessidades da população normativa para com seus “enfermos”.

Figueirêdo, Delevati e Tavares (2014) e Ramminger (2002) descrevem em seus trabalhos que o surgimento dos nomeados manicômios ou hospitais psiquiátricos, que tinham como função tratar pacientes com deficiências, resultou no que pode ser relacionado à perpetuação do rótulo de loucura, insanidade e ou incapacidade de pessoas com deficiência. A propagação, por gerações, de ideias pseudocientíficas, que nada agregariam ao entendimento, tampouco na melhora da qualidade de vida dos internados, tornou-se uma estrutura social de segregação. Estes indivíduos passavam boa parte de suas vidas sob efeito de sedações e terapias ineficientes, além de serem restringidos do acesso e luta por seus direitos.

Segundo Ramminger (2002), ainda no século XX, em 1934, foi promulgado a lei federal de Assistência aos Doentes Mentais, de decreto 24.559, onde determinava para essas pessoas o tratamento em hospitais psiquiátricos como a única solução ou possibilidade.

Somente na segunda metade do século passado, surge através do Movimento dos Trabalhadores de Saúde Mental, descrito como a Luta Antimanicomial. Contra a existência dessa instituição vigente, a história foi protagonizada inicialmente por Nise da Silveira, médica baiana que humanizou as ações de psiquiatria e iniciou um importante debate sobre a humanidade e a deficiência no Brasil.

Com a incidência de novos debates sociais sobre o assunto, em 2001 o decreto de 1934 foi substituído por uma nova forma de tratar pessoas, que se enquadravam como deficientes mentais perante a lei, surge assim a ampliação de direitos e tratamentos para pessoas com deficiência, possibilitando assim a maior participação delas em meio a sociedade.

Do início do século XXI até a atualidade, adicionaram-se novas formas de abordar o tema, estudar e debater os direitos, a vida e as causas médicas por de trás dessas pessoas, a contextualização cultural do termo neurodivergente que

Pode entender-se como pessoas neurodivergentes: autistas, pessoas com Transtorno de Déficit de Atenção e Hiperatividade (TDAH), pessoas com altas habilidades ou

superdotação (AH/ SD), dentre outras. Os demais indivíduos, ou seja, aqueles com o funcionamento cerebral considerado “típico” são denominados “neurotípicos”. Alencar, Barbosa e Gomes (2021, p. 2128).

Esta nova forma de pensar a respeito dos transtornos mentais, impulsionada pela cunhagem do termo “neurodiversidade”, pela socióloga australiana Judy Singer, a qual busca abranger condições neurológicas diferentes da normalidade social, tendo a finalidade de humanizar, dando voz ao movimento de igualdade e diversidade, sem a necessária busca de uma cura ou tratamento para o transtorno. Esse novo ponto de vista retira o debate da exclusividade clínica e passa a ter um caráter social e político mais proeminente.

A influência da mudança expressada anteriormente, passou a ter impacto na categorização da pessoa com deficiência, mas ainda insuficiente para resolver as tensões e evitar injustiças sociais. Conforme consta no Manual de Diagnóstico e Estatístico de Transtornos Mentais (DSM-5), as pessoas deficientes são categorizadas formalmente de modo clínico. Os transtornos e espectros, nestes casos, condicionam o indivíduo diagnosticado ao tratamento devido. Desta forma, ele será tratado com base em seus transtornos e espectros mentais catalogados como, por exemplo, os transtornos da Comunicação, Motores, do Espectro Autista, os de Déficit de Atenção e Hiperatividade, Deficiências Intelectuais e o Espectro da Esquizofrenia. Além da classificação sobre a gravidade, podendo ser definida como: leve, moderada, grave e profunda.

A deficiência intelectual (transtorno do desenvolvimento intelectual) caracteriza-se por déficits em capacidades mentais genéricas, como raciocínio, solução de problemas, planejamento, pensamento abstrato, juízo, aprendizagem acadêmica e aprendizagem pela experiência. Os déficits resultam em prejuízos no funcionamento adaptativo, de modo que o indivíduo não consegue atingir padrões de independência pessoal e responsabilidade social em um ou mais aspectos da vida diária, incluindo comunicação, participação social, funcionamento acadêmico ou profissional e independência pessoal em casa ou na comunidade. (AMERICAN PSYCHIATRIC ASSOCIATION, 2014, p. 31)

Pode averiguar-se, através dessa contextualização, de que o assunto a respeito da neurodiversidade ou das deficiências intelectuais é vasto e abrangente, podendo ter cunho político e social, porém muito se embasa principalmente na psiquiatria.

Para fins de debate é importante enfatizar que o tema se mantém em constante estudo e mudança, visando alcançar tratamentos, terminologias e condições cada vez mais acessíveis, respeitadas, inclusivas, igualitárias e seguras para pessoas neuroatípicas em meio a sociedade.

3 Metodologia

A metodologia da pesquisa é descrita como um estudo bibliográfico, que trata de um estudo geral sobre os principais trabalhos já realizados, revestidos de importância por serem capazes de fornecer dados atuais e relevantes relacionados com o tema (MARCONI; LAKATOS, 2022, p. 186). É exploratório de acordo com o afirmado por Marconi e Lakatos (2022, p. 218) são:

investigações de pesquisa empírica, cujo objetivo é a formulação de questões ou de um problema, com tripla finalidade: (1) desenvolver hipóteses; (2) aumentar a familiaridade do pesquisador com um ambiente, fato ou fenômeno, para a realização de uma pesquisa futura mais precisa; (3) modificar e clarificar conceitos. Empregam-se geralmente procedimentos sistemáticos ou para a obtenção de observações empíricas ou para as análises de dados (ou ambas, simultaneamente)

A pesquisa pode ser definida enquanto de natureza pura ou básica que, segundo Gil (2019, p. 25) “busca o progresso da ciência, procura desenvolver os conhecimentos científicos sem a preocupação direta com suas aplicações e consequências práticas.” e de abordagem qualitativa que, para Minayo (2015, p. 21) “responde a questões particulares”. Em Ciências Sociais, preocupa-se com “um nível de realidade que não pode ser quantificado”, ou seja, “ela trabalha com o universo de significados, dos motivos, das aspirações, das crenças, dos valores e das atitudes”.

No processo de busca na BDTD, utilizaram-se os termos ética e representação do conhecimento, sem recorte temporal devido à amplitude do debate do assunto, recuperando 209 teses ou dissertações, sendo selecionados dois trabalhos de mestrado e doutorado de um mesmo autor, os termos Capacitismo e ética foram aplicados à plataforma SCIELO onde foram recuperados dois artigos, dos quais selecionou-se um. O refinamento dos resultados para a definição dos textos a serem utilizados baseou-se no conteúdo selecionado pela leitura dos resumos. Os trabalhos que possuíam maior relevância para a elaboração do embasamento e fundamentação do conteúdo estudado, a fim de responder aos questionamentos propostos no estudo.

Por fim, estruturou-se o estudo com base nos materiais recuperados, que possibilitaram a argumentação teórica do estudo e embasamento para a elaboração de resultados que pudessem refletir sobre as questões indicadas na problemática proposta para o estudo.

4 Resultados e Conclusão

Visando ao entendimento social da área da ciência da informação, buscou-se retratar uma fração dos conhecimentos acadêmicos a respeito da elaboração da temática neuroatípica na representação de assunto e a

ética dos profissionais que atuam nesta área do conhecimento.

O conhecimento é de natureza multidisciplinar, visto que pessoas neuroatípicas podem necessitar de apoio em diferentes contextos sociais, como nas áreas de educação e ciências da saúde, além da necessidade de uma rede de apoio em um contexto particular, como familiares e amigos que contribuam com as atividades e demandas destes indivíduos no dia a dia.

Nesse aspecto, refletir sobre os aspectos culturais, os preconceitos e as tensões da área da representação de um assunto, envolve considerar as mudanças contemporâneas. Para tanto, é necessário que os profissionais deem voz às pessoas deficientes e à rede de apoio, de modo que possam opinar socialmente sobre os próprios saberes e serem assim, representados corretamente. Isso significa, permitir que colaborem ativamente no desenvolvimento de leis, estudos e demais atividades, ao abordarem suas vivências e necessidades, bem como questões de acesso e respeito aos seus direitos.

A adesão do termo “capacitismo” dá visibilidade aos preconceitos e discriminações historicamente construídos. A possibilidade de novas formas de representação de assuntos, que envolvem as pessoas com alguma deficiência, possibilita que as informações referentes às vivências e às violências enfrentadas por elas possam ser combatidas. Portanto, conclui-se que há necessidade de que sejam mais bem descritas e amparadas, ampliando o debate com o enfoque para questões inclusivas, assim como outras minorias que veem conquistando seus espaços de fala, socialmente estabelecidos em instituições acadêmicas.

Em relação à ética e ao contexto amplo da filosofia sobre a vida em sociedade, a qual contempla a ideia de esta ser parte do reflexo da sociedade vigente, é necessário autá-la. Ademais, uma das esferas do conhecimento que interage com os demais assuntos abordados, são os códigos de ética das mais variadas profissões, aos quais devem ser impressos o respeito a todos os seres humanos.

O conhecimento sobre as deficiências é construído com base na diversidade dos espectros e transtornos. Contudo, a dignidade humana é antológica à esfera psiquiátrica e deve combater os desvios sociais e culturais. Frequentemente, referindo-se a esta questão, os esforços e aprendizados muitas vezes recaem sobre o próprio paciente, beneficiando apenas a sociedade que opta por manter os “diferentes” afastados.

Sabe-se que a aliança entre o projeto de inclusão e o arcabouço teórico oferecido pelos modelos social e crítico de deficiência pode ser vista como potencialmente libertadora da opressão imposta por um modelo médico e de uma educação especial caracterizada como representante desse discurso, e que a perspectiva inclusiva impactou os

sistemas educacionais com a introdução de novas legislações e políticas educacionais. (Ullrich, 2022, p. 49).

Conclui-se que, ao representar o assunto, a carga semântica da palavra escolhida para expressar uma condição da pessoa humana, abriga um valor que poderá ou não refletir de maneira negativa no social. Dessa maneira, o que não soava preconceituoso em uma determinada época, poderá ser compreendido de outra forma, em um outro tempo histórico.

A partir de uma visão ética e científica, também se observa o avanço do discurso da neurodiversidade para acolher pessoas com funcionamento cerebral diferente do típico. Nessa conjuntura, a visão de “retardo mental”, “loucura” ou “neurose” para justificar o comportamento desses indivíduos está sendo contestada. Alencar, Barbosa e Gomes (2021, p. 2133).

Com base nos argumentos apresentados no texto, o uso de termos considerados excludentes ou capacitistas possuem a potencialidade de perpetuação de tabus e preconceitos, já estruturados em meio a sociedade. Trata-se da corponormatividade, mantendo assim as barreiras já enfrentadas e excludentes, vivenciadas por pessoas que não se enquadram nos padrões estabelecidos.

Em teoria, a existência de padrões éticos para a elaboração de produtos da representação de assuntos deveria garantir a padronização e evitar o uso de termos que possam excluir, prejudicar ou inferiorizar diferentes grupos sociais, inclusive acompanhando o debate social. Entretanto, o fator humano presente no processo profissional pode influenciar, seja por falta de conhecimento ou por questões particulares, cabendo aos órgãos reguladores da prática profissional encontrarem formas de evitar ações de propagação a preconceitos. Além da colaboração de instituições na formação de futuros profissionais, visando a uma maior concepção das temáticas.

Portanto, entende-se que são necessários novos estudos para propor caminhos aos processos de representação da informação, a fim de buscar soluções para o não uso de termos capacitistas. Esperamos que a leitura reflexiva deste artigo impulse uma questão a ser discutida na Ciência da Informação, enfatizando a relação dos termos sensíveis, para que sejam compreendidos a partir de estudos críticos. E que possam ser estabelecidas alternativas para que haja um início de mudança no contexto da representação de assuntos.

Referências

- ALENCAR, Helenira Fonseca de; BARBOSA, Heloisa Fonseca; GOMES, Robéria Vieira Barreto. Neurodiversidade: Aspectos históricos, conceituais e impactos na educação escolar. *In*: CASTRO, Paula Almeida de; SILVA, Gessika Cecília Carvalho da; SILVA, Alex Vieira da Givanildo; CAVALCANTI, Ricardo Jorge de Sousa (org.) **Escola em tempos de conexões**. Campina Grande: Realize editora, 2022.v.2, p.: 2125 – 2142. ISBN 978-65-86901-50-4. Disponível em: <http://10.46943/VII.CONEDU.2021.02.111>. Acesso em: 29 jan. 2023.
- AMERICAN PSYCHIATRIC ASSOCIATION. **Manual diagnóstico e estatístico de transtornos mentais: DSM-5**. 5 ed. Tradução: Maria Inês Corrêa Nascimento *et al.* Porto Alegre: Artmed, 2014. Disponível em: <http://www.institutopebioetica.com.br/documentos/manual-diagnostico-e-estatistico-de-transtornos-mentais-dsm-5.pdf>. Acesso em: 29 jan. 2023.
- BARITÉ, Mario. **Diccionario de Organización del Conocimiento**: Clasificación, Indización, Terminología. 6 ed. Montevideo: Universidad de la República, 2015.
- BORBA, F. S. (Org.) **Dicionário UNESP do português contemporâneo**. São Paulo: UNESP, 2004.
- CAMARA DOS DEPUTADOS. **Glossário de acessibilidade**. Brasília, [202-]. Disponível em: <https://www2.camara.leg.br/a-camara/estruturaadm/gestao-na-camara-dos-deputados/responsabilidade-social-e-ambiental/acessibilidade/glossarios/glossario.html>. Acesso em: 28 jan. 2023.
- CHAUÍ, Marilena. **Convite à filosofia**. São Paulo: Ática, 2000.
- CONSELHO FEDERAL DE BIBLIOTECONOMIA. Resolução CFB nº207/2018. Aprova o Código de Ética e Deontologia do Bibliotecário brasileiro, que fixa as normas orientadoras de conduta no exercício de suas atividades profissionais. Diário Oficial da União: seção 1, Brasília, DF, 2018. p. 155 e 156. Disponível em: <https://crb8.org.br/wp-content/uploads/2021/08/Co%CC%81digo-de-E%CC%81tica-e-Deontologia-do-CFB-1.pdf>. Acesso em: 29 jan. 2023.
- CRISOSTOMO, Alessandro Lombardi; VARANI, Gisele; PEREIRA, Priscila dos Santos; OST, Sheila Beatriz. **Ética**. Porto Alegre: SAGAH, 2018. *E-book* (216 p.) ISBN 9788595024557. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788595024557/>. Acesso em: 09 fev. 2023.
- FIGUEIRÊDO, Marianna Lima de Rolemberg; DELEVATI, Dalnei Minuzzi; TAVARES, Marcelo Góes. Entre loucos e manicômios: história da loucura e a reforma psiquiátrica no Brasil. **Cadernos de graduação ciências humanas e sociais**. Maceió, v.2, n.2, p. 121-136, nov. 2014. ISSN 2316-672x. Disponível em: <https://periodicos.set.edu.br/fitshumanas/article/view/1797/1067>. Acesso em: 29 jan.2023.
- GIL, Antônio Carlos. **Métodos e Técnicas de Pesquisa Social**. 7ed. São Paulo: Atlas, 2019. *E-book* (230 p.) ISBN 9788597020991. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788597020991/>. Acesso em: 10 mar. 2023.
- MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de Metodologia Científica**. São Paulo: Atlas, 2022. *E-book* (354 p.) ISBN 9788597026580. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788597026580/>. Acesso em: 09 fev. 2023.

MATIAS-PEREIRA, José. **Manual de Metodologia da Pesquisa Científica**. São Paulo: Atlas, 2016. *E-book (187 p.)*. ISBN 9788597008821. Disponível em: <https://app.minhabiblioteca.com.br/#/books/9788597008821/>. Acesso em: 09 fev. 2023.

MELLO, Anahi Guedes de. Deficiência, incapacidade e vulnerabilidade: do capacitismo ou a preeminência capacitista e biomédica do Comitê de Ética em Pesquisa da UFSC. **Ciência & Saúde Coletiva**, 21(10), p. 3265-3276, out. 2016. Disponível em: <https://www.scielo.br/j/csc/a/J959p5hgv5TYZgWbKvspRtF/abstract/?lang=pt#>. Acesso em: 02 fev. 2023.

MINAYO, Maria Cecília de Souza. O desafio da pesquisa social. *In*: MINAYO, Maria Cecília de Souza (org.). **Pesquisa social: teoria, método e criatividade**. 34 ed. Petrópolis: Vozes, 2015.

ORTEGA, Francisco. Deficiência, autismo e neurodiversidade. **Ciência & Saúde Coletiva**, Rio de Janeiro, 14(1), p. 67-77, jan. 2009. Disponível em: <https://www.scielo.br/j/csc/a/CPcMbsxyfF3CXSLwTprwC/abstract/?lang=pt#>. Acesso em: 03 fev. 2023.

PINHO, Fabio Assis. **aspectos éticos em representação do conhecimento**: em busca do diálogo entre Antonio García Gutiérrez, Michèle Hudon e Clare Beghtol. Orientador: José Augusto Chaves Guimarães. 2010. 157 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências da UNESP, Marília, 2006. Disponível em: <https://repositorio.unesp.br/handle/11449/93693>. Acesso em: 02 fev. 2023.

PINHO, Fabio Assis. **Aspectos éticos em representação do conhecimento em temáticas relativas à homossexualidade masculina**: uma análise da precisão em linguagens de indexação brasileiras. Orientador: José Augusto Chaves Guimarães. 2010. 157 f. Tese (Doutorado em Ciência da

Informação) - Faculdade de Filosofia e Ciências da UNESP, Marília, 2010. Disponível em: <https://repositorio.unesp.br/handle/11449/103379>. Acesso em: 02 fev. 2023.

RAMMINGER, Tatiana. A saúde mental do trabalhador em saúde mental: um estudo com trabalhadores de um hospital psiquiátrico. **Boletim da Saúde**, Rio Grande do Sul, v. 16, n. 1, jan./jun. 2002. Disponível em: <http://www.boletimdasaude.rs.gov.br/conteudo/1168/sumario>. Acesso em: 05 fev. 2023.

SILVA, Andrieli Pachú da. Código de ética ou de deontologia para o arquivista brasileiro? Questionamentos preliminares. *In*: congresso nacional de arquivologia 8., 2018, João Pessoa. **Anais eletrônicos...** Revista Analisando em Ciência da Informação - RACIn, João Pessoa, v. 6, n. especial, p. 483-497, out. 2018. Disponível em: http://arquivologiauepb.com.br/racin/edicoes/v6_nesp_racin_v6_nesp_TA_GT06_0483-0497.pdf. Acesso em: 02 fev. 2023.

SIROUX, Danièle. Deontologia. *In*: CANTO-SPERBER, M. (Org.) **Dicionário de Ética e Filosofia Moral**. Editora Unisinos, v.1, p.405-406, 2007.

SOUZA Olga Solange Herval. **Itinerários da inclusão escolar**: Múltiplos olhares, saberes e práticas. Porto Alegre: ULBRA. 2008.

ULLRICH, Wladimir Brasil. **Deficiência intelectual, educação e as possibilidades do reconhecimento**. Orientador: Carla Karnoppi Vasques. 2022. 140 f. Tese (Doutorado em Educação) - Faculdade de Educação da Universidade Federal do Rio Grande do Sul, Porto Alegre, 2022. Disponível em: <https://www.lume.ufrgs.br/handle/10183/237471>. Acesso em: 02 fev. 2023.

Chatbots em bibliotecas: aplicações, desafios e perspectivas

Chatbots in libraries: applications, challenges and perspectives

Maria Cristina Piumbato Innocentini Hayashi (1), Alexandre Masson Maroldi (2) Marcia Regina da Silva (3)

(1) Universidade Federal de São Carlos, dmch@ufscar.br, (2) Universidade Federal de Rondônia, alexandre@unir.br,

(3) Universidade de São Paulo, marciaregina@usp.br

Resumo

A inteligência artificial (IA) e suas aplicações já se tornaram realidade em diversas esferas da vida moderna, e não é diferente em ambientes informacionais. **Objetivo.** Investigar quais são as aplicações e perspectivas de uso dos *chatbots* em bibliotecas refletidas na literatura científica sobre o tema. **Metodologia.** As abordagens das análises bibliométricas e de conteúdo foram utilizadas para analisar a produção científica disponível no Portal de Periódicos Capes e Google Scholar. As seguintes variáveis foram investigadas: temporalidade dos documentos, canais de divulgação (periódicos, eventos e preprints), tipos de estudos (teóricos, estudos de caso, revisão de literatura) denominação dos *chatbots*, aplicações dos *chatbots* em bibliotecas. **Resultados.** Foram analisados artigos (n=41), trabalhos em eventos (n=5) e preprints (n=1) publicados em periódicos científicos (n=27) e do exterior entre 2002 e 2022. As aplicações dos *chatbots* é majoritária em serviços técnicos e suporte aos usuários, os tipos de pesquisa que deram origem à produção científica analisada referem-se a estudos de caso (n=21), estudos teóricos (n=19) e estudos de revisão (n=7). Os *chatbots* específicos para bibliotecas focalizados na produção científica analisada revelaram a predominância de nomes femininos, tais como Emma, Anna, Bia, Eva, Ivy, Paige, Pixel, Rasa. **Conclusão.** Dois aspectos chamaram a atenção na literatura analisada: a questão de gênero na denominação dos *chatbots*, e a necessidade de incluir as tecnologias de IA, como os *chatbots*, no ensino e educação de bibliotecários.

Palavras-chave: Bibliotecas. *Chatbots*. Inteligência Artificial. Produção Científica.

Abstract

Artificial Intelligence (AI) and its applications have already come true in various spheres of modern life, and is no different in informational environments. **Objective:** To investigate what are the applications and perspectives of using *chatbots* in libraries reflected in the scientific literature on the subject. **Methodology:** The approaches to bibliometric and content analysis were used to analyze the scientific production available on the Capes and Google Scholar journal portal. The following variables were investigated: temporality of documents, dissemination channels (periodicals, events and preprints), types of studies (theoretical, case studies, literature review) *chatbots* name, *chatbots* applications in libraries. **Results:** Articles (n = 41), event work (n = 5) and preprints (n = 1) published in scientific journals (n = 27) and from the abroad between 2002 and 2022. *chatbots* are majority in the events of scientific journals (n = 27). Technical services and support to users, the types of research that gave rise to the scientific production analyzed refer to case studies (n = 21), theoretical studies (n = 19) and review studies (n = 7). *Chatbots* specific to libraries focused on the scientific production analyzed revealed the predominance of female names such as Emma, Anna, Bia, Eva, Ivy, Paige, Pixel, Rasa. **Conclusion:** Two aspects caught the eye in the literature analyzed: the gender issue in the name of *chatbots*, and the need to include AI technologies, such as *chatbots*, in teaching and library education and education.

Keywords: Libraries. *Chatbots*. Artificial intelligence. Scientific production.

Resumen

La inteligencia artificial (IA) y sus aplicaciones ya se han hecho realidad en varias esferas de la vida moderna, y no es diferente en entornos informativos. **Objetivo:** investigar cuáles son las aplicaciones y perspectivas del uso de *chatbots* en bibliotecas reflejadas en la literatura científica sobre el tema. **Metodología:** Los enfoques para el análisis bibliométrico y de contenido se utilizaron para analizar la producción científica disponible en el portal de Capas y Google Scholar Journal. Se investigaron las siguientes variables: temporalidad de documentos, canales de difusión (publicaciones periódicas, eventos y preimpresiones), tipos de estudios (teóricos, estudios de casos, revisión de la literatura) Nombre de *chatbots*, aplicaciones de *chatbots* en bibliotecas. **Resultados:** Artículos (n = 41), trabajo de eventos (n = 5) y preimpresiones (n = 1) publicados en revistas científicas (n = 27) y desde el extranjero entre 2002 y 2022. Los *chatbots* son mayoritarios en los eventos de revistas científicas (n = 27). Los servicios técnicos y el apoyo a los usuarios, los tipos de investigación que dieron lugar a la producción científica analizada se refieren a estudios de casos (n = 21), estudios teóricos (n = 19) y estudios de revisión (n = 7). Los *chatbots* específicos de las bibliotecas centradas en la producción científica analizada revelaron el predominio de nombres femeninos como Emma, Anna, BIA, Eva, Ivy, Paige, Pixel, Rasa. **Conclusión:** Dos aspectos llamaron la atención en la literatura analizada: el tema de género en nombre de los *chatbots* y la necesidad de incluir tecnologías de IA, como *chatbots*, en la enseñanza y la educación en la biblioteca y la educación.

Palabras clave: Bibliotecas. *Chatbots*. Inteligencia artificial. Producción científica.

1 Introdução

Os profissionais da área da informação vivenciam um grande desafio alicerçados na prerrogativa da necessidade constante de inovação. As mudanças tecnológicas colocam esses profissionais constantemente em estado de reflexão, repensando o papel das bibliotecas neste cenário em que o acesso às informações não acadêmicas acontece mediante a um clique em aplicativos e mídias sociais. Para além dos diversos fatores, como a vinculação de notícias falsas ou fraudulentas, que implicam na validação desses meios de comunicação, o que mais tem afetado o comportamento de busca dos usuários é o imediatismo. O acesso rápido que acontece ao acessar o WhatsApp, por exemplo, torna os usuários de bibliotecas, muitas vezes, descontentes com o tempo despendido entre a formulação de estratégias de busca, resposta dos bancos de dados e filtragem dos resultados.

A busca pela informação científica requer competência em informação, ou seja, exige conhecimento e estudo. Para além do papel do bibliotecário na formação do usuário, existe uma necessidade latente e constante de propor novos serviços ou aprimorar os já existentes. Hoje temos à disposição diversas possibilidades de serviços digitais oferecidos por meio de plataformas. A adoção de processos e ferramentas digitais é uma tarefa complexa, pois envolve mudanças que afetam o produto informacional, o processo e os serviços da biblioteca tradicional. Hoje as bibliotecas podem criar repositórios de recursos digitais visando tanto a preservação como a própria disseminação na rede. (SINGH; ASIF, 2019).

As bibliotecas deixam de exercer seu papel principal de responsável pela custódia de recursos tradicionais de informação para provedora de recursos de informação digital, orientada para o oferecimento de serviços (CHANNAVEERAI; WILLIAM, 2008). Os dois anos de isolamento (2020-2021) decorrente da pandemia do COVID 19 aceleraram as mudanças relacionadas ao atendimento em vários setores, inclusive nas bibliotecas, que tiveram que adotar de forma rápida medidas para dar continuidade ao oferecimento de informações e capacitações (treinamentos e demonstrações) aos seus usuários. O uso de plataformas digitais, e-mails e atendimento telefônico foram as formas mais utilizadas no processo de mediação. No entanto, os avanços da inteligência artificial têm trazido novas possibilidades no campo informacional, denotando que podem se firmar também na simulação de interações humanas, o que implica diretamente no atendimento ao usuário de bibliotecas.

Os *chatbots* são sistemas de inteligência artificial (IA) que utilizam conceitos de Interação em Linguagem Natural (NLI) projetados para simular uma conversa. Esse sistema pode ser baseado em texto onde a interação é fornecida através de uma entrada e saída de texto e agentes de conversação incorporados, onde a interface é

representada por uma figura com um corpo e/ou rosto que interage com o usuário (ALLISON, 2012). Segundo Khanna *et al.* (2015), os *chatbots* respondem como uma entidade inteligente, por meio de texto ou voz, podendo interpretar uma ou mais línguas humanas por Processamento de Linguagem Natural (PNL).

Estudos recentes sobre chatbots apontam que estes precisam de melhorias, tais como abordagem avançada de reconhecimento de padrões, bases de conhecimento mais abrangentes, melhor organização e representação de conhecimento. (ARSOVSKI *et al.*, 2017).

Como referência para a IA, além do termo *chatbots*, encontramos na literatura o uso dos termos *smart bots*, agentes interativos, assistentes digitais ou entidades de conversação artificial (ADAMOPOULOU; MOUSSIADES, 2020). No entanto, o termo agente virtual é quase sempre usado em associação com o termo “*chatbot*” como sinônimo ou para sua definição (ASHFAQ, 2020).

O intenso uso desses robôs em serviços de atendimento ao público abriu um nicho no mercado de trabalho para criadores de conteúdo ou roteiristas para a conversa entre a máquina e o cliente da empresa. Hoje, tanto o mercado privado como o público recorrem aos *bots* para agilizar o atendimento aos clientes e usuários, tornando cada vez mais a experiência bem-sucedida. Além do *e-commerce*, os *chatbots* são aplicados na área educacional, recuperação da informação, negócios etc. O primeiro *chatbot* foi lançado em 1966 com o nome de Eliza, e agia como psicoterapeuta devolvendo as afirmações do usuário em forma de perguntas (WEIZENBAUM, 1983).

A *Gartner Hype Cycle* previu que em 2020 os *chatbots* seriam responsáveis por 85% do atendimento ao cliente. Já para 2022, Gartner destaca como tecnologia emergente a expansão de tecnologias imersivas, exemplificando com o gêmeo digital do cliente (Digital Twin of the Customer, DToC), que é uma representação virtual dinâmica de um cliente que simula e aprende a imitar e prever o comportamento. Ele pode ser usado para modificar e aprimorar a experiência do cliente (UX) e dar suporte a novas iniciativas de digitalização, produtos, serviços e oportunidades (GARTNER HYPE CYCLE, 2022).

O uso de chatbots foi considerado um potencial divisor de águas para instituições centradas nas pessoas já há algum tempo. Em sintonia com a tendência, várias bibliotecas e centros de recursos de conhecimento também adotaram chatbots com o objetivo de fornecer serviços aprimorados aos usuários. (BAGCHI, 2020). Na visão desse autor, as bibliotecas sempre estiveram na vanguarda do reconhecimento, absorção e efetivação da tecnologia em comparação com outras instituições acadêmicas.

Vinckze (2017) considera a adoção de entidades conversacionais artificiais para fins como entrevistas de referência, na tentativa de atrair novamente os usuários para utilizar seus serviços.

Na esfera das Bibliotecas, os *chatbots* ainda podem ser utilizados em tarefas rotineiras e repetitivas que podem contribuir para responder perguntas comuns, já que podem ser programados ou alimentados com informações necessárias. A IA consegue formatar essas informações e apresentá-las para atender às necessidades dos usuários. Por meio de um algoritmo de respostas os *chatbots* conseguem gerenciar pesquisas básicas e fazer o direcionamento de pesquisas mais complexas para um bibliotecário (ALLISON, 2012).

Outros estudos também mostram as vantagens do uso de *chatbots* em bibliotecas, tais como:

- a) uma alternativa mais conveniente, interativa e exclusiva ao atendimento aos usuários tradicional (CHUNG *et al*, 2020);
- b) aplicação nos serviços de referência virtuais para processamento automático de consultas e fontes de informação (BAGCHI, 2020)
- c) oferecer serviços de informação comunitária, disseminação automática de informações autorizadas e minimização de desinformação em caso de gerenciamento de risco durante pandemias, como recentemente ocorreu (pandemia de COVID-19) (BAGCHI, 2020);
- d) fornecer suporte 24 horas por dia, 7 dias por semana podendo garantir uma resposta imediata em tempo real (ZUMSTEIN; HUNDERTMARK, 2017; COX; PINFIELD; RUTTER, 2019);
- e) permitir a comunicação móvel do usuário e a compreensão de suas necessidades por serviços de *chat* (COX; PINFIELD; RUTTER, 2019).
- f) receber sugestões do serviço de biblioteca para mapear a intenção do usuário de acordo com a intenção de busca anterior (BAGCHI, 2020);
- g) o humor do *chatbot* permanece inalterado, demonstrando cuidado e gentileza com os usuários de forma constante (LUO *et al*, 2019);
- h) executar tarefas de tomada de decisão com uma taxa de erro minimizada, oferecendo soluções direcionadas que refletem os desejos do usuário (KAPLAN; HAENLEIN, 2019);
- i) capacidade de analisar o texto recebido por meio de uma abordagem de correspondência de padrões, identificando assim as palavras-chave, frases e construções de sentenças (RESE; GANSTER; BAIER, 2020; NGUYEN; SIDOROVA, 2018; KÖHLER *et al*, 2011);
- j) prever o comportamento do usuário, sentir as emoções e identificar preferências específicas de produtos (NGUYEN; SIDOROVA, 2018);
- k) capacidade de gravar um histórico de conversação ao qual podem se referir usando sua memória artificial, a fim de formular respostas mais satisfatórias e personalizadas. (NGUYEN; SIDOROVA, 2018).

Outros estudos, como o de Arsaovski et al (2017) argumentam que existem muitos projetos em ontologias, organização do conhecimento e representação do conhecimento visando aplicar essas funcionalidades de raciocínio em *chatbots*. Por sua vez, Bagchi (2020) comenta que trabalhos futuros sobre *chatbots* de biblioteca deveriam abordar a integração destes com sistemas de gerenciamento de conhecimento semântico, uma vez que esses sistemas continuam sendo a estrutura de serviço de informação ideal compatível com integração semântica de ponta a ponta, juntamente com tecnologias de IA.

Dessa perspectiva Bagchi (2020) defende que seria interessante desenvolver um sistema de gerenciamento de conhecimento semântico praticamente viável no contexto de bibliotecas e serviços de informação, que automatize o processamento e a resolução de necessidades de informação específicas de domínio, com foco no contexto, linguagem e diversidade. Ou seja, para o autor, um *chatbot* de biblioteca desenvolvido nas premissas técnicas de IA conversacional de código aberto, integrado com tecnologias semânticas como gráficos de conhecimento e mecanismos de inferência de raciocínio semântico no back-end continua sendo uma opção viável para tal estrutura. (BAGCHI, 2020).

Matthew (2020) analisou sistemas de organização de conhecimento em *chatbots* e arquiteturas de agentes para entender como esses sistemas podem contribuir para a experiência do usuário final. Para o autor, existem maneiras demonstráveis de que um serviço orientado a sistemas de organização do conhecimento pode não apenas representar o conhecimento de domínio com o qual um usuário está interagindo, mas também moldar e facilitar o diálogo em si.

Nesta perspectiva, evidencia-se que a IA e a utilização dos *chatbots* em ambientes informacionais constitui em um tema que merece ser mais bem explorado considerando a literatura científica produzida sobre esse assunto.

Sendo assim, a seguinte questão norteou a realização dessa pesquisa: Quais são as aplicações e perspectivas de uso dos *chatbots* em bibliotecas? O estudo realizado objetivou traçar um panorama do estado da arte sobre *chatbots* com base na literatura científica sobre esse tema.

2 Procedimentos metodológicos

Trata-se de pesquisa exploratória e descritiva que utilizou metodologia quanti-qualitativa realizada por meio da análise bibliométrica para a elaboração de indicadores de produção científica sobre *chatbots* em bibliotecas e análise de conteúdo para categorizar o corpus selecionado. Imperioso ressaltar que a combinação entre esses dois métodos amplia as discussões e os resultados de uma investigação do tipo estado da arte (SILVA; HAYASHI; HAYASHI, 2011; BARDIN, 2011)

Os dados foram coletados em novembro do ano de 2022, e as fontes de dados foram o *Google acadêmico* e o *Portal de Periódicos Capes*. Foram utilizados os seguintes termos de busca: *chatbot*; *digital assistant*; *intelligent assistant*; *virtual assistant*; *personal assistant*; *voice assistant* combinados com o termo *library*, nas formas singular e plural, sem filtro para temporalidade. A disponibilidade do texto completo e a aderência à temática pesquisada foram adotadas como critério de inclusão, sendo excluídos aqueles repetidos. Desse modo o corpus de análise (=47) constou de artigos científicos (n=41), trabalhos em eventos (n=5) e um preprints. Todos foram registrados em uma planilha eletrônica do Excel para organização e categorização das informações que abrangeram as seguintes variáveis: temporalidade dos documentos, canais de divulgação (periódicos, eventos e preprints), tipos de estudos (teóricos, estudos de caso, revisão de literatura) denominação dos *chatbots*, aplicações dos *chatbots* em bibliotecas. Os resultados obtidos foram analisados à luz do referencial teórico advindo da Biblioteconomia e da Ciência da Informação.

3 Resultados e Discussão

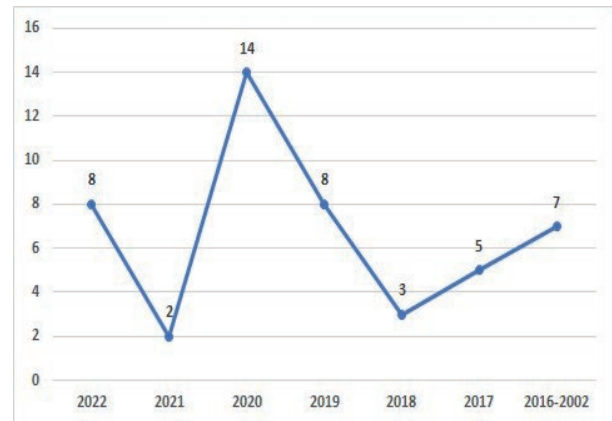
Esta seção está organizada em duas perspectivas: a primeira apresenta um panorama quantitativo da produção científica analisada; a segunda, expõe os resultados qualitativos obtidos mediante a análise de conteúdo da produção científica analisada.

3.1 Panorama bibliométrico da produção científica

O corpus analisado cobriu o período entre 2002 e 2022, com maior concentração entre 2017 e 2022 (n=42), e com distribuição irregular nos anos de 2002, 2004, 2005, 2010, 2012 e 2016 com apenas uma produção cada (Figura 1).

Essa distribuição revelou uma média de publicações (n=6) nos últimos seis anos. Nesse período houve um crescimento gradual de publicações sobre a temática dos *chatbots* em bibliotecas, embora com oscilações, sendo o maior índice (n=14) em 2020, e o menor (n=2) em 2021.

Figura 1. *Temporalidade da produção científica (2002-2022)*



Fonte: Autores.

Tais resultados demonstram que embora essa produção seja irregular do ponto de vista temporal, a área de Biblioteconomia e Ciência da Informação não deixou de abordar temáticas relacionadas à IA e *chatbots* nos ambientes informacionais.

Na Tabela 1 observa-se a distribuição da produção científica.

Tabela 1. *Distribuição da produção científica conforme o canal de publicação*

Periódicos / Eventos/ Repositório	Total
Journal of the Australian Library and Information Association	4
Library Hi Tech	4
Library Hi Tech News	3
Journal of Academic Librarianship	2
Journal of Electronic Resources in Medical Libraries	2
Journal of the Medical Library Association	2
Library Technology Reports	2
<i>Periódicos com um artigo cada:</i> Business Information Review; Computers in Library; Information Services & Uses; Information Technology and Libraries; Information Today; Interlending & Document Supply; International Journal of Innovative Science and Research Technology; International Research Journal of Multidisciplinary Technovation; Iranian Journal of Information Processing and Management; Journal of Southwest Jiaotong University; Journal of Librarianship and Informatics; Journal of Library & Information Technology; Journal of Library Administration; Journal of Management Information and Decision Sciences; Journal of Web Librarianship; Law Library Journal; Portal Libraries and the Academy; Public Library Quarterly; Public Services Quarterly; The Acquisitions Librarian; The Electronic Library	22
<i>Repositório de preprints – SSRN</i>	1

<i>Periódicos / Eventos/ Repositório</i>	<i>Total</i>
<i>Trabalhos em eventos: 27th SEBD - Italian Symposium on Advanced Database Systems; 12th International CALIBER-2019 (Library 2030: Moving Towards Smart Technologies, Services and Resources); 19th National Conference ACRL (Association of College & Research Libraries); ALISE 2019 Conference (Association for Library and Information Science Education; XX SNBU 2018 (Simpósio Nacional de Bibliotecas Universitárias)</i>	5
<i>Total</i>	47

Fonte: Autores.

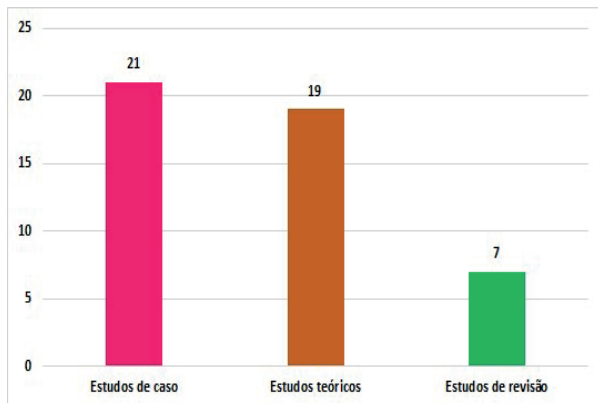
Observa-se na Tabela 1 que a maioria dessa produção científica (n=41) foi publicada em periódicos científicos, apresentada em eventos (n=5), e depositada em repositório de preprints (n=1) no exterior. Apenas um trabalho foi apresentado em evento tradicional da área de Biblioteconomia no Brasil.

Em relação ao gênero dos autores (n=95) os resultados apontaram que as mulheres são maioria (n=55) e os homens (n=40) representaram 42,1% do total. Prevaleram as publicações (n=28) em co-autorias e as autorias individuais (n=19) foram minoria.

3.2 Perfil qualitativo da produção científica sobre chatbots

A Figura 2 apresenta os resultados da análise de conteúdo da produção científica quanto ao tipo de pesquisa que foi realizada nos artigos (n=41), trabalhos em eventos (n=5) e preprints (n=1) publicados.

Figura 2. Tipos de pesquisa



Fonte: Autores.

Nota-se que em 44,7% do total prevaleceram os estudos de caso (n=21), seguidos dos estudos teóricos (n=19) e estudos de revisão (n=7).

Os estudos de caso (n=21) tiveram como objetivo investigar a opinião de bibliotecários e usuários de bibliotecas sobre os riscos e desafios da adoção de ferramentas de inteligência artificial, entre elas o *chatbots*, nas áreas de serviços das bibliotecas, bem como apresentar *chatbots* que foram construídos para

executar serviços técnicos de biblioteca. Esses estudos de caso foram realizados em bibliotecas acadêmicas dos seguintes países: Paquistão; Puglia Digital Library – Itália; University of Duisburg-Essen – Alemanha; University of Sydney-Austrália; PUC-Rio – Brasil; bibliotecas do Reino Unido e outros países; bibliotecas dos EUA (Nebraska-Lincoln, City of New York (CUNY); North Salem-NY; Southern California; Ohio; Boise State; University of California Irvine); University of Alberta – Canadá, bem como em outras bibliotecas de universidades canadenses, australianas e dos EUA sem especificar nomes e localidades.

Entre os estudos de caso relatados vale expor aqui o único que aborda uma experiência brasileira. Trata-se do desenvolvimento da BIA, a assistente virtual das Bibliotecas PUC-Rio (FERREIRA, 2018). O autor descreve os procedimentos adotados no desenvolvimento do personagem e apresenta detalhes do sistema planejado para controlar a inteligência artificial e o fluxo de respostas do mesmo. Além disso, aponta as medidas aplicadas para otimizar o funcionamento do sistema a partir do *feedback* dos utilizadores e conclui, apontando os aspectos vantajosos que devem ser considerados na implantação de um projeto dessa natureza.

Por sua vez, os estudos teóricos (n=19) abordaram os impactos e potencialidades das tecnologias de IA em bibliotecas por meio de relato de experiências com *chatbots* específicos para essas unidades de informação, bem como a utilização de outros assistentes de voz proprietários, tais como Alexa e Siri, e as implicações relacionadas à privacidade e segurança. Esses estudos também problematizaram questões como o gênero dos *chatbots* de bibliotecas, e o estímulo oportunizado pela utilização de IA, entre elas os *chatbots* para autoatendimento online durante a pandemia de COVID-19.

Entre os estudos teóricos vale comentar aquele realizado por Tait e Pierson (2022). As autoras abordam o uso de IA e robôs em biblioteconomia e ciência da informação sinalizando que este tema está chamando a atenção devido às primeiras aplicações e seu potencial de contribuir para a transformação digital das profissões da informação.

O referido artigo avalia os desafios e oportunidades para a educação LIS (Library and Information Science) por meio de um estudo que revisou o currículo de cinco cursos dessa área credenciados pela Australian Library and Information Association (ALIA). Para tal, as autoras revisaram as descrições de assuntos desses currículos e a documentação de conhecimento da fundação ALIA. Os resultados indicaram que apenas um assunto menciona a IA para posicionar o conteúdo do assunto e nenhum menciona a robótica, bem como que há múltiplas áreas para a inclusão desses tópicos dentro das cinco categorias de conhecimento da fundação ALIA.

Tait e Pearson (2022) argumentam que será necessária uma forma de integração desses tópicos na educação profissional de LIS. Essa visão é bastante relevante no atual cenário da educação de bibliotecários tendo em vista os currículos dos cursos que formam esses profissionais. Ou seja, é necessário que estes conheçam as potencialidades e uso das tecnologias de IA em seus futuros ambientes de trabalho para estarem aptos a atender às necessidades dos usuários.

Nessa direção, Queiroz e Valls (2022) analisaram currículos dos cursos presenciais de Biblioteconomia da cidade de São Paulo e verificaram que os estudantes recebem boas ferramentas para o desenvolvimento de competências para atuação como analistas de *chatbot*, apresentando uma nova possibilidade profissional para os bibliotecários. Nesta perspectiva, evidencia-se que a IA é uma temática de relevância para a área e precisa ser mais bem explorada.

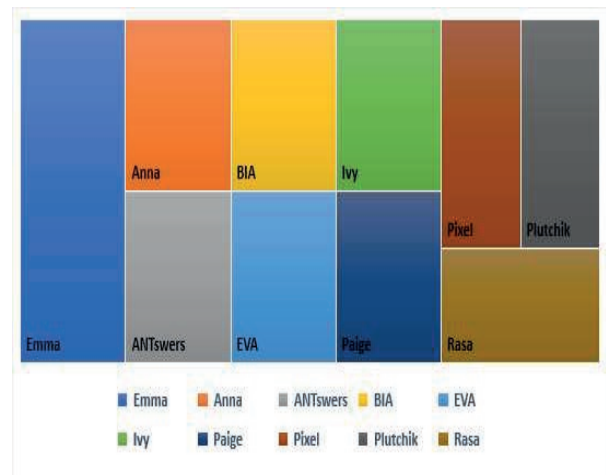
Os estudos de revisão (n=7) tiveram como foco analisar a literatura científica sobre;

- bibliotecas inteligentes sob a perspectiva de sistemas especialistas, IA, robôs, e Internet das Coisas (n=2);
- capacidades e aplicações de sistemas especialistas e *chatbots* para serviços de biblioteca;
- a aplicação de Inteligência Artificial e sistemas baseados em robótica em bibliotecas e centros de informação;
- aplicações integradas de *chatbots* de inteligência artificial em serviços de referência de bibliotecas acadêmicas para atender às demandas de usuários millenials;
- assistentes pessoais inteligentes (IPAs) como o Echo Dot Kids da Amazon e seu uso por crianças e;
- tipos de robôs utilizados nas bibliotecas, suas funções, e a prontidão dos bibliotecários para recebê-los.

A Figura 3 mostra as denominações dos *chatbots* específicos para bibliotecas focalizados na produção científica analisada. Pode-se perceber a predominância de nomes ou acrônimos (n=7) que remetem à feminilização dos *chatbots*, conforme comentado por Brown (2022).

Para essa autora, a maioria dos assistentes digitais, incluindo *chatbots*, recebeu nomes, vozes, representações visuais e até “personalidades” que são estereotipicamente femininas e refletem a ideologia patriarcal.

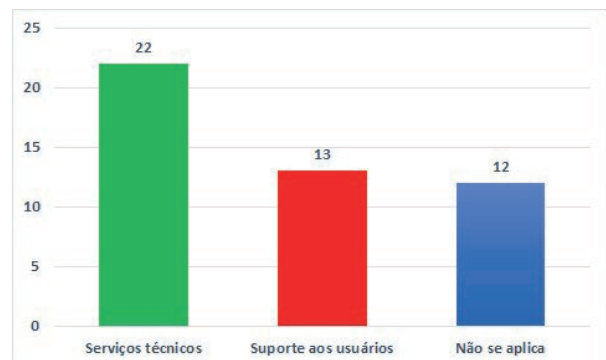
Figura 3. Denominações dos *chatbots* de bibliotecas



Fonte: Autores.

Em relação às aplicações dos *chatbots* nas bibliotecas (Figura 4) os resultados mostraram que prevaleceram aquelas direcionadas aos serviços técnicos, tais como o serviço de referência, seguido pelo suporte aos usuários.

Figura 4. Tipos de aplicações dos *chatbots* em bibliotecas



Fonte: Autores.

Observa-se que doze estudos não especificaram o tipo de aplicação, abordando o tema de forma genérica.

McNeal *et al.* (2013) levantaram alguns aspectos da utilização de *chatbots* no serviço de referência em Bibliotecas, tais como: ajudar a compensar déficit de pessoal; encaminhar perguntas complexas ao bibliotecário; proporcionar uma experiência interativa mais animada quando comparada a um mecanismo de busca; criar um vínculo mais forte com o público jovem; realizar conversas em aplicativos móveis no formato de bate-papo; permanecer paciente e educada durante toda a interação com o usuário; interagir com usuários mesmo quando as bibliotecas estiverem fechadas e; aumentar a visibilidade da biblioteca na comunidade de usuários.

4 Conclusão

O estudo realizado apontou as principais características da produção científica sobre os *chatbots* em bibliotecas, bem como as potencialidades e implicações do seu uso nesses ambientes informacionais, tanto para os profissionais bibliotecários, como também para os usuários. Nesse contexto, a incorporação de chatbots em bibliotecas agrega conhecimento para bibliotecários e faz com que entendam melhor a tecnologia de IA e seus potenciais de aplicações em diferentes setores das bibliotecas.

Apesar da literatura analisada se referir a um período de 30 anos que abrangeu os anos entre 2002 e 2022, a maioria das publicações, representada por 85% do total, são dos últimos seis anos. Além disso, apenas um artigo foi publicado em periódico brasileiro, denotando que a literatura científica sobre esse tema pode ser incrementada com outras pesquisas.

Dois aspectos chamaram a atenção na literatura analisada: a questão de gênero na denominação dos *chatbots*, e a necessidade de incluir as tecnologias de IA, como os *chatbots*, no ensino e educação de bibliotecários.

Sugere-se que futuros estudos ampliem as fontes de dados de modo a permitir comparações com os achados dessa pesquisa.

Referências

- ADAMOPOULOU, Eleni; MOUSSIADES, Lefteris. An overview of *chatbot* technology. *In: Artificial Intelligence Applications and Innovations: 16th IFIP WG 12.5 International Conference, AIAI 2020, Neos Marmaras, Greece, June 5–7, 2020, Proceedings, Part II 16*. Springer International Publishing, 2020. p. 373-383. Disponível em: https://link.springer.com/chapter/10.1007/978-3-030-49186-4_31. Acesso: 10 jan. 2023.
- ALLISON, DeeAnn. *Chatbots* in the library: is it time? **Library hi tech**, v. 30, n. 1, p. 95 - 107, 2012. Disponível em: <https://doi.org/10.1108/07378831211213238>. Acesso: 10 mar. 2023.
- ARSOVSKI, Sasa *et al.* (2017) Analysis of the chatbot open source languages AIML and Chatscript: a review. *In: DQM International Conference on life cycle engineering and management, 9 Proceedings...* Lille.
- ASHFAQ, Muhammad *et al.* I, *Chatbot: Modeling the determinants of users' satisfaction and continuance intention of AI-powered service agents. Telematics and Informatics*, v. 54, p. 101473, 2020. Disponível em: <https://doi.org/10.1016/j.tele.2020.101473>. Acesso: 10 mar. 2023.
- BAGCHI, Mayukh. Conceptualising a Library *Chatbot* using Open Source Conversational Artificial Intelligence. **DESIDOC Journal of Library & Information Technology**, v. 40, n. 6, 2020. Disponível em: <https://drdo.gov.in/sites/default/files/inline-files/document3.pdf>. Acesso: 10 mar. 2023.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Ed. 70, 2011.
- BROWN, Laila M. Gendered Artificial Intelligence in Libraries: Opportunities to Deconstruct Sexism and Gender Binarism. **Journal of Library Administration**, v. 62, n. 1, p. 19-30, 2022. Disponível em: <https://doi.org/10.1080/01930826.2021.2006979>. Acesso: 2 mar. 2023.
- CHANNAVEERAIHAH, Lalithamba; WILLIAMS, Theresa. **From automation to transformation: impact of ICT in LIS: major shifts & practices**. *In: International CALIBER, 6., University of Allahabad, Allahabad, 2008*. Disponível em: <http://ir.inflibnet.ac.in:8080/ir/bitstream/1944/1249/1/5.pdf>. Acesso: 2 mar. 2023.
- CHUNG, Minjee *et al.* *Chatbot e-service and customer satisfaction regarding luxury brands. Journal of Business Research*, v. 117, p. 587-595, 2020. Disponível em: <https://doi.org/10.1016/j.jbusres.2018.10.004>. Acesso: 2 mar. 2023.
- COX, Andrew M.; PINFIELD, Stephen; RUTTER, Sophie. The intelligent library: thought leaders' views on the likely impact of artificial intelligence on academic libraries. **Library Hi Tech**, v. 37, n. 3, p. 418-435, 2019. Disponível em: <https://doi.org/10.1108/LHT-08-2018-0105>. Acesso: 2 mar. 2023.
- FERREIRA, Giuliano. **BIA: um estudo sobre o desenvolvimento da assistente virtual das bibliotecas PUC-Rio**. *In: XX SNBU, 2020*.
- GARTNER HYPE CYCLE. 2022. Disponível em: <https://www.gartner.com.br/pt-br/artigos/novidades-nogartner-hype-cycle-for-emerging-technologies-de-2022>. Acesso: 12 mar. 2023.
- KAPLAN, Andreas; HAENLEIN, Michael. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. **Business horizons**, v. 62, n. 1, p. 15-25, 2019. Disponível em: <https://doi.org/10.1016/j.bushor.2018.08.004>. Acesso: 5 mar. 2023.
- KHANNA, Anirudh *et al.* A study of today's AI through *chatbots* and rediscovery of machine intelligence. **International Journal of u-and e-Service, Science and Technology**, v. 8, n. 7, p. 277-284, 2015. Disponível em:

<https://doi.org/10.14257/ijunesst.2015.8.7.28>. Acesso: 5 mar. 2023.

KÖHLER, Clemens F. *et al.* Return on interactivity: the impact of online agents on newcomer adjustment. **Journal of Marketing**, v. 75, n. 2, p. 93-108, 2011. Disponível em: <https://doi.org/10.1509/jm.75.2>. Acesso: 12 mar. 2023.

LUO, Xueming *et al.* Frontiers: Machines vs. humans: The impact of artificial intelligence *chatbot* disclosure on customer purchases. **Marketing Science**, v. 38, n. 6, p. 937-947, 2019. Disponível em: <https://doi.org/10.1287/mksc.2019.1192>. Acesso: 10 mar. 2023.

MATTHEW, Paul. (2020) Knowledge Organisation Systems for chatbots and conversational agents: a review of approaches and an evaluation of relative value-added for the user. *In: Proceedings of the Sixteenth International ISKO Conference, 2020 Aalborg, Proceeding...* Denmark, pp. 274-283.

MCNEAL, Michele L.; NEWYEAR, David. Introducing *chatbots* in libraries. **Library technology reports**, v. 49, n. 8, p. 5-10, 2013. Disponível em: <https://www.journals.ala.org/index.php/ltr/article/view/4504/5281>. Acesso em: 10 mar. 2023.

NGUYEN, Quynh N.; SIDOROVA, Anna. Understanding User Interactions with a *Chatbot*: a Self-determination Theory Approach. *In: Proceedings of the Twenty-Fourth Americas Conference on Information Systems, New Orleans, LA, USA. 2018.* p. 16-18. Disponível em: https://web.archive.org/web/20200322163359id_/https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1419&context=amcis2018. Acesso: 2 mar. 2023.

QUEIROZ, Thais dos Santos; VALLS, Valéria Martin. O bibliotecário analista de *chatbot*: as competências desenvolvidas nos cursos presenciais de bacharelado em biblioteconomia da cidade de São Paulo. **Revista Brasileira de Biblioteconomia e Documentação**, v. 18, 1-25, 2022. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/1534>.

RESE, Alexandra; GANSTER, Lena; BAIER, Daniel. *Chatbots* in retailers' customer communication: how to measure their acceptance? **Journal of Retailing and Consumer Services**, v. 56, p. 102176, 2020. Disponível em: <https://doi.org/10.1016/j.jretconser.2020.102176>. Acesso: 10 mar. 2023.

SILVA, Márcia Regina da; HAYASHI, Carlos Roberto Massao; HAYASHI, Maria Cristina Piumbato Innocentini. Análise bibliométrica e cientométrica: desafios para especialistas que atuam no campo. **InCID: revista de ciência da informação e documentação**, v. 2, n. 1, 2011. Disponível em: <https://revistas.ffclrp.usp.br/incid/article/viewArticle/52>. Acesso: 12 mar. 2023.

SINGH, K. K.; ASIF, Mohammad. Emerging trends and technologies for digital transformation of libraries. **IP Indian Journal of Library Science and Information Technology**, v. 4, n. 2, p. 41-43, 2019. Disponível em: https://www.researchgate.net/publication/337942956_Emerging_trends_and_technologies_for_digital_transformation_of_libraries/citation/download. Acesso: 12 mar. 2023.

TAIT, ELIZABETH; PIERSON, CAMERON M. Artificial intelligence and robots in libraries: opportunities in lis curriculum for preparing the librarians of tomorrow. **Journal of the Australian Library and Information Association**, v.71, n.3, p. 256-274, 2022. Disponível em: https://www.researchgate.net/publication/361421461_Artificial_Intelligence_and_Robots_in_Libraries_Opportunities_in_LIS_Curriculum_for_Preparing_the_Librarians_of_Tomorrow.

VINCZE, Joseph. (2017) Virtual reference librarians (chatbots). **Library Hi Tech**, v.34, n.3, pp. 5-8.

WEIZENBAUM, Joseph. Eliza—a computer program for the study of natural language communication between man and machine. **Communications of the ACM**, v. 26, n. 1, p. 23-28, 1983. Disponível em: <https://doi.org/10.1145/365153.365168>. Acesso: 2 mar. 2023.

ZUMSTEIN, Darius; HUNDERTMARK, Sophie. *Chatbots*--an interactive technology for personalized communication, transactions and services. **IADIS International Journal on WWW/Internet**, v. 15, n. 1, 2017. Disponível em: https://www.researchgate.net/publication/322855718_Chatbots_-_An_Interactive_Technology_for_Personalized_Communication_Transactions_and_Services Acesso: 2 mar. 2023.

O desafio da representação da etnoliteratura: o caso do Glossário Decolonial de Macunaíma

The challenge of representing ethnotexts: the case of Macunaíma's Decolonial Glossary

Míriam Gontijo de Moraes

Universidade Federal do Estado do Rio de Janeiro, Av. Pasteur, 458, sala 413, miriam.moraes@unirio.br

Resumo

Objetivo: Entre os desafios colocados para a organização do etnoconhecimento destaca-se o atendimento ao princípio teórico-metodológico da garantia literária. Esse trabalho tem como objetivo a construção de um instrumento terminológico que atenda tanto ao princípio da garantia literária, como à perspectiva decolonial. Encontramos na obra Macunaíma, de Mário de Andrade, a conexão com uma etnoliteratura que nos serviu como garantia literária para levantarmos um sistema nocional na perspectiva decolonial. **Problematiza:** A especificidade documental de um etnoconhecimento baseado na oralidade, nos coloca como uma 'área problema' na qual a garantia literária encontra obstáculos como a dispersão e déficit documental a exemplo dos espaços interdisciplinares. **Metodologia:** Foi composto um corpus baseado no glossário da obra produzido por M. Cavalcanti Proença que, por sua vez, faz sua pesquisa junto às fontes tupinológicas, compreendendo registros de uma etnoliteratura que só chegou ao nosso conhecimento por meio de publicações baseadas em narrativas da esquecida Língua Geral da Amazônia, o Nheengatu. Constam das fontes de Proença ainda o sulamericanista brasileiro Capistrano de Abreu, e o etnólogo alemão Theodor Koch-Grünberg. **Resultados:** Foram identificadas 2112 entradas referentes a uma compilação de termos e sinônimos, cada uma seguida de definição que leva em conta o aspecto semântico e o contexto da história, que no seu conjunto nos dão a dimensão da riqueza do vocabulário brasileiro em suas raízes. **Conclusões:** Mesmo enfrentando problemas de representação similar aos dos espaços interdisciplinares, O Glossário Decolonial de Macunaíma irá contribuir para a documentação do Inventário Nacional da Diversidade Linguística, instituído pelo Decreto nº 7.387, de 9 de dezembro de 2010.

Palavras-chave: Etnoconhecimento. Organização do Conhecimento. Decolonial. Macunaíma. Glossário.

Abstract

Objective: Among the challenges posed for the organization of ethnotexts, the fulfillment of the theoretical-methodological principle of literary guarantee stands out. This work aims to build a terminological instrument that meets both the principle of literary guarantee and the decolonial perspective. We found in the work Macunaíma, by Mário de Andrade, the connection with an ethnotext that served us as a literary guarantee to raise a notional system in the decolonial perspective. **Problem:** The documentary specificity of an ethnotext based on orality places us as a 'problem area' in which the literary guarantee encounters obstacles such as dispersion and documentary deficit, as in the case of interdisciplinary spaces. **Methodology:** A corpus was composed based on the glossary of the work produced by M. Cavalcanti Proença who, in turn, does his research with Tupinologist sources, comprising records of an ethnotext that only came to our knowledge through publications based on narratives of the forgotten General Language of the Amazon, the Nheengatu. Proença's sources also include the Brazilian South Americanist Capistrano de Abreu, and the German ethnologist Theodor Koch-Grünberg. **Results:** 2112 entries were identified referring to a compilation of terms and synonyms, each followed by a definition that takes into account the semantic aspect and the context of history, which together give us the dimension of the richness of Brazilian vocabulary in its roots. **Conclusions:** Even facing problems of representation similar to those of interdisciplinary spaces, The Decolonial Glossary of Macunaíma will contribute to the documentation of the National Inventory of Linguistic Diversity, instituted by Decree nº 7.387, of December 9, 2010.

Keywords: Ethnotexts. Knowledge Organization. Decolonial. Macunaíma. Glossary.

Resumen

Objetivo: Entre los desafíos que se plantean para la organización del etnoconocimiento, se destaca el cumplimiento del principio teórico-metodológico de garantía literaria. Este trabajo pretende construir un instrumento terminológico que responda tanto al principio de garantía literaria como a la perspectiva decolonial. Encontramos en la obra Macunaíma, de Mário de Andrade, la conexión con una etnoliteratura que nos sirvió como garantía literaria para plantear un sistema nocional en la perspectiva decolonial. **Problema:** La especificidad documental de un etnosaber basado en la oralidad nos sitúa como un 'ámbito problema' en el que la garantía literaria encuentra obstáculos como la dispersión y el déficit documental, como en el caso de los espacios interdisciplinarios. **Metodología:** Se compuso un corpus a partir del glosario de la obra producida por M. Cavalcanti Proença quien, a su vez, realiza su investigación con fuentes tupinológicas, comprendiendo registros de una etnoliteratura que sólo llegó a nuestro conocimiento a través de publicacio-

nes basadas en relatos de los olvidados Lengua General de la Amazonia, el Nheengatu. Las fuentes de Proença también incluyen al sudamericanista brasileño Capistrano de Abreu y al etnólogo alemán Theodor Koch-Grünberg. **Resultados:** Se identificaron 2112 entradas referentes a una compilación de términos y sinónimos, cada uno seguido de una definición que tiene en cuenta el aspecto semántico y el contexto de la historia, que juntos nos dan la dimensión de la riqueza del vocabulario brasileño en sus raíces. **Conclusiones:** Incluso enfrentando problemas de representación similares a los de los espacios interdisciplinarios, el Glosario Decolonial de Macunaíma contribuirá a la documentación del Inventario Nacional de Diversidad Lingüística, instituido por el Decreto n° 7.387, de 9 de diciembre de 2010.

Palabras clave: Etnoconocimiento. Organización del conocimiento. Decolonial. Macunaíma. Glosario.

1 Introdução

Apesar da origem tecnicista e positivista, há espaço, no campo da Ciência da Informação, para problematizar e refletir sobre a natureza crítica no campo da organização e representação da informação e do conhecimento, para além de sua função imediata de dar respostas operacionais para os problemas exclusivamente de natureza técnica. As dimensões ideológicas envolvidas nos processos de representação da informação e do conhecimento também estão na agenda de pesquisa desde o estudo de Berman (1993) sobre a Lista de Cabeçalho de Assuntos da Library of Congress, que identificou uma série de manifestações de etnocentrismo, machismo, racismo na maneira como se busca prescrever formas de representar e classificar documentos.

Neste sentido, a construção de obras de referência, a exemplo de glossários e outros trabalhos terminológicos também se dá mediante uma postura crítica em relação ao eurocentrismo na representação do conhecimento em toda a América Latina.

A diversidade cultural e linguística da região é muito pouco conhecida sobretudo quando se tem em mente o universo das línguas e culturas ameríndias e africanas sobreviventes ao silenciamento imposto pelos Estado, meios de comunicação e sistema educacional.

Como obra de referência, o Glossário Decolonial de Macunaíma é um projeto de elaboração de um instrumento terminológico para a recuperação de informação, conforme a perspectiva decolonial, no marco do movimento modernista brasileiro e se vê diante da necessidade de romper com esse cenário de silenciamento e subjogação de povos, etnias, classes e, ou, identidades de grupo. Entre os muitos desafios colocados pelo projeto, destaca-se o atendimento ao princípio teórico-metodológico cunhado em 1911 pelo bibliotecário britânico E. Wyndham Hulme, como garantia literária. Segundo Barité et al (2010), em sua concepção original, a garantia literária demarca que a literatura de um domínio de conhecimento deve ser a fonte para extração e validação da terminologia a ser incorporada em um sistema de classificação, ou em qualquer outro sistema de organização do conhecimento. Ainda de acordo com esse autor, diante desse princípio, a documentação opera como instrumento que sintetiza o conhecimento científico especializado de forma que classificacionistas, por meio de metodologias de construção

previamente estabelecidas em normas, padrões ou mesmo acordos locais, processam a reapresentação desse conhecimento em sistemas de organização do conhecimento de diversas naturezas: sistemas de classificação, terminologias, tesouros, listas de cabeçalhos de assunto, taxonomias, ontologias temáticas etc.

Na discussão levantada por Barité para uma revisão crítica ao princípio da garantia literária, é destacado que o atendimento a esse princípio não parte de uma organização do conhecimento apriorística ou baseada em aspectos formais da teoria da classificação, “uma vez que é a documentação que efetivamente atua como fonte de validação dos termos a serem incluídos em um sistema de organização do conhecimento, pois o que se classifica e se indexa são temas presentes nos documentos”. Encontramos na obra Macunaíma e na figura de Mário de Andrade a conexão para uma proposta de construção de um instrumento terminológico que atendia tanto ao princípio da garantia literária, como à perspectiva decolonial a ser impressa no levantamento dos termos. Mário de Andrade desenvolveu um trabalho de documentação sobre a história, o povo, a cultura e especialmente a música do interior do Brasil, tanto em São Paulo quanto no Nordeste brasileiro como também a Amazônia. “O Turista Aprendiz” é o título do diário de viagem do escritor modernista pela Amazônia no ano de 1927. Tal viagem foi emblemática na consolidação do seu pensamento vinculado à temática nacional. A partir daí ele publicaria clássicos como “Clã do Jabuti” (novembro de 1927) e “Macunaíma” (1928), nos quais as referências à Amazônia são evidentes. Tendo como referencial empírico o pensamento de um grupo modernista brasileiro, notadamente registrado na obra Macunaíma de Mario de Andrade, definiu-se um recorte de seu aspecto em conformidade à visão decolonial na sua leitura de mundo e ressignificação da nossa cultura. Para evidenciar essa faceta do movimento modernista inaugurado em 1922, e, a partir da estratégia de análise de domínio, buscou-se identificar o campo semântico modernista brasileiro tendo como objeto essa obra síntese do movimento, que se caracteriza por ser um trabalho de documentação trabalhosa de lendas, superstições, frases feitas, provérbios, modismos de linguagem da paisagem brasileira e conter uma narrativa de projeção continental que questiona as fronteiras das rígidas nacionalidades modernas impostas pelo colonialismo europeu.

2 A composição do Corpus para a análise terminológica

Atendendo ao princípio da garantia literária, foi composto um corpus baseado no glossário da obra produzido por M. Cavalcanti Proença que, por sua vez, faz sua pesquisa junto às fontes tupinólogas, compreendendo registros de uma etnoliteratura que só chegou ao nosso conhecimento por meio de publicações baseadas em narrativas da esquecida Língua Geral da Amazônia, o Nheengatu. Constam das fontes de Proença ainda o sul americanista brasileiro Capistrano de Abreu, e o etnólogo alemão Theodor Koch-Grünberg, fonte primária de Mário de Andrade. Também buscamos compor o corpus com a consulta às 520 notas e a compilação de 706 verbetes, que contribuem para a compreensão de uma das obras mais instigantes da literatura brasileira, e que compõem a edição de Macunaíma da Editora FTD, que identifica em notas elaboradas pela professora de Literatura Brasileira, Noemi Jaffe, muitos dos termos encontrados no glossário de Proença e acrescidos de outros. M. Cavalcanti Proença, o autor do Roteiro de Macunaíma, cita que os livros guias de Macunaíma de Mário de Andrade, além da obra do etnógrafo alemão Koch-Grünberg, foram os registros de Couto de Magalhães e o seu *O Selvagem*; de Capistrano de Abreu, no caso do livro *a Língua dos Kaxinawás*, e também Gustavo Barroso, Basílio de Magalhães, e Silvio Romero, entre outros.

Segundo Lúcia Sá, autora do livro *Literaturas da Floresta- textos amazônicos e cultura latino-americana*, em *Roteiro de Macunaíma* Cavalcanti Proença mapeou as fontes da obra, configurando um admirável estudo que identifica a complicada rede de textos e subtextos que compõem a rapsódia Mario Andradina, as narrativas indígenas que fornece ao autor seus principais personagens e a maioria das tramas. Ainda conforme Lúcia Sá, grande parte do enredo foi inspirada de Koch-Grünberg, embora algumas histórias tenham vindo também de Capistrano de Abreu, Barbosa Rodrigues, Brandão de Amorim, e Couto de Magalhães.

Em uma releitura esclarecedora sobre as fontes indígenas de Macunaíma, Lucia Sá comenta que os modernistas se encantaram com a coleção de lendas indígenas amazônicas recolhidas por Brandão Amorim e publicadas, postumamente, em 1928, em prosa brasileira. Na época, comenta a autora, não existia nada similar no que diz respeito à reprodução de uma sensibilidade e uma linguagem popular, no caso, amazônica. Mário de Andrade, Oswald de Andrade, Cassiano Ricardo, Menotti del Picchia leram as 33 lendas recolhidas por Amorim, publicadas em português e nheengatu. Na avaliação de Sá, tal coleção de lendas, escrita na língua falada no norte do país, não deveria ser considerada apenas um texto-fonte de que diferentes autores modernistas se valeram para elaborar suas obras, mas sim um dos trabalhos modernistas dos mais significativos.

2.1 Tupinólogos e sul americanistas: a etnoliteratura inspiradora de Mário de Andrade

O discurso colonialista desqualificou a poesia e os mitos indígenas e a historiografia jogou no esquecimento a existência da Língua Geral da Amazônia e suas características de meio de comunicação interétnica, durante 200 anos, o Nheengatu.

Conforme o prof. Jose Ribamar Bessa Freire, do programa de Memória Social da Unirio e aposentado na UERJ (2021), que coordenou durante 30 anos o Programa de Estudos dos Povos Indígenas (PROINDIO), no final do século 20, olhando para esse conjunto de estudiosos interessados na literatura indígena, encontramos os chamados tupinólogos, porque a maioria deles, com exceção do Theodor Koch-Grünberg e do Capistrano de Abreu, era de estudiosos do Nheengatu. Segundo Bessa, eles surgiram e se interessaram pela língua geral através do contato com uma etnoliteratura, ou seja, narrativas que em uma cultura oralizada são as enciclopédias, as histórias, as obras de referência de toda a sabedoria contida nessa narrativa.

Foram intelectuais que coletaram narrativas indígenas na segunda metade do sec. XIX até o final do sec. XX, registrando por escrito histórias narradas oralmente quase todas em Nheengatu e as traduziram ao português. Segundo o prof. Bessa, são eles: Couto de Magalhães (1837-1898), nascido na fazenda de gado de seu avô, em Diamantina (MG), onde passou sua infância embalado por “lendas tocantes e poéticas, metade cristãs, metade indígenas” contadas pelos vaqueiros. Formado em direito, procurou “as cores do país” não só em arquivos e bibliotecas, mas também em dez viagens por grotões do Brasil profundo, quando ouviu “lendas tupi” que transcreveu em – *O Selvagem* – editado em várias línguas: francês, inglês, alemão e italiano. Seu interesse cresceu, quando foi nomeado presidente da Província do Pará pelo Imperador Pedro II.

O geólogo canadense Charles Hartt (1840-1878), que veio na missão científica do naturalista Louis Agassiz, inventariar as riquezas da Amazônia. Na parada do navio em Óbidos (PA) ouviu uma velha senhora cantando histórias em Nheengatu na calçada de sua casa, cercada por crianças e jovens. Quando traduziram as histórias, fascinado, aprendeu a língua e coletou mitos amazônicos com o jaboti como personagem central. Descobriu que essa era a riqueza mais importante da região: sua literatura repleta de sabedoria e a língua usada para fazê-la circular.

O conde italiano Ermano Stradelli (1852-1926), nascido num castelo em Borgo Val di Taro, uma comuna na Itália. Sua intenção era visitar a Amazônia durante algumas semanas. De noite, numa maloca no Rio Negro, deitado em sua rede, ouviu os índios contarem histórias, o que o fez modificar seu plano ao se inteirar da tradução. Passou o resto da vida no Amazonas até a sua

morte em Manaus. Aprendeu a língua, fez um dicionário Nheengatu-Português-Nheengatu, registrou as tradições, entre elas o Mito do Jurupari e as “lendas dos Tariana”, entre uma e outra cuia de caxiri.

Outro que não resistiu aos encantos das narrativas orais foi o botânico mineiro João Barbosa Rodrigues (1842-1909), filho de um comerciante português. Professor do Colégio Pedro II, no Rio de Janeiro, mudou para Manaus em 1872, contratado para criar o Museu Botânico. Aprendeu o Nheengatu e nessa língua coletou cantigas e contos avaliados por ele como “flores da imaginação de um povo” publicadas no *Poranduba Amazonense*. Quando perguntava o nome de uma planta dele desconhecida, respondiam com uma história na qual a planta era descrita. Percebeu que em sociedades orais, histórias constituíam as enciclopédias populares.

Temos ainda Brandão de Amorim (1865-1926) nascido em Manaus, filho de um comerciante português criador da companhia que fez a navegação direta de Liverpool a Manaus. Ele publicou 33 narrativas do alto Rio Negro em edição bilingue, sem mencionar que haviam sido recolhidas por Maximiano José Roberto, índio descendente dos Manaú e dos Tariana do rio Uaupés. Em seu livro, Lúcia Sá nos chama a refletir o fato de que essa etnoliteratura, ou mesmo literatura, que foi sempre considerada como simples matéria-prima etnográfica, só passa a fazer parte da categoria de arte quando trabalhada por intelectuais não indígenas.

entre as inúmeras fontes em que Mário de Andrade se documentou para a construção de *Macunaíma*, destaca-se o livro de Capistrano de Abreu, “Língua dos Caxinauás” resultado dos seus estudos sobre as línguas indígenas brasileiras (Kaxinawás e Bacaeris), em conformidade com o movimento sul-americanista, que reuniu um grupo de especialistas sobre a cultura e línguas sul-americanas de 1890 a 1929 (época das viagens dos etnógrafos alemães à região Amazônica) e o contexto do domínio europeu no conhecimento etnográfico e de linguística dos povos americanos. Capistrano de Abreu como sul-americanista foi interlocutor da etnografia alemã da qual se destacava Theodor Koch-Grünberg, importante fonte de Mário de Andrade. A obra de Capistrano fornece o tema central do capítulo 4 (Boiuna Luna) e do capítulo 13 (A Piolhenta do Jigüê) de *Macunaíma*.

Beatriz Protti Christino, Professora Associada II do Departamento de Letras Vernáculas (Setor de Língua Portuguesa) da Universidade Federal do Rio de Janeiro (UFRJ), investigou o circuito de produção e recepção das pesquisas acerca de línguas sul-americanas de 1890 a 1929 e, em especial, a obra de Capistrano de Abreu sobre a língua Kaxinawá, em seu trabalho de doutorado intitulado *A rede de Capistrano de Abreu (1853-1927): uma análise historiográfica do rã-txa hu-ni-ku-~i em face da Sul-americanista dos anos 1890-1929*.

Segundo Christino, essa comunidade de especialistas que, na verdade, se autodenominavam americanistas (como os estudiosos da América Central e de povos da América do Norte também se denominavam americanistas), mas ela preferiu destacar como sul americanistas, porque o diálogo se dava especialmente entre os especialistas de povos de um mesmo continente, tinha a característica de estar deixando de ser eurocêntrica, nesse período entre 1890 e 1929.

Nesse período, pesquisadores da América do Sul também passam a ser parte dessa rede, informa Christino. A rede propriamente dessa comunidade de especialistas que contava também com chilenos, como o antropólogo Robert Lehmann-Nitsche, venezuelanos como Luís Ramón Oramas. Nesse cenário, vários pesquisadores de povos da América do Sul passam a trabalhar, inclusive, na América do Sul e em diálogo direto e bastante horizontal com os pesquisadores radicados na Europa, que vinham fazer viagens de pesquisa ou que depois organizavam dados em pesquisas de gabinete.

Segundo Christino, essa geração de 1890 é a primeira que tem nas culturas e nas línguas dos vários povos o principal objeto de estudo, porque os pesquisadores anteriores eram pesquisadores de ciências naturais, por assim dizer, que nas suas viagens de pesquisa também recolhiam informações sobre povos e línguas. Um estudioso paradigmático da geração anterior, destaca a professora, é Carl Friedrich Philipp Von Martius, que compôs os *Glossaria Linguarum Brasiliensium*, reunindo todas as fontes que havia até aquele momento no tempo do império sobre línguas brasileiras e também coletou nas suas viagens algumas palavras, alguns vocabulários com os povos com quem teve contato na viagem como o zoólogo Johann Baptist von Spix, que era, sobretudo, um botânico e fez vários volumes de um catálogo sobre a flora brasileira. A geração seguinte a essa, que a pesquisadora focaliza, são os Sul-americanistas dos quais Capistrano de Abreu faz parte, como também Theodor Koch-Grünberg, cuja obra *Vom Roraima zum Orinoco*, vol. II (1917) se baseou Mário de Andrade.

Esses especialistas, de forma geral, se viam na urgência de fazer registros etnográficos e linguísticos, porque estavam assistindo a um processo de genocídio muito intenso; e acreditavam, naquele início de século XX, que podiam estar vendo os últimos representantes de muitos povos. Nesse sentido, Capistrano de Abreu acreditava que o registro da cultura dos Kaxinawá e da cultura dos Bakairi, que ele pretendia fazer, podia vir a ser o registro de povos que deixariam de existir por conta de genocídio e etnocídio; A conjuntura do momento, então, especialmente na região tradicional dos Kaxinawá, o Alto Juruá, podia ser caracterizada historicamente pelo fato de que ela passou, em muito pouco tempo, de uma região que era completamente à parte do jogo do capitalismo internacional para ser fornece-

dora da indústria de ponta daquele momento, que é a indústria automobilística. Por conta dessa grande demanda por borracha, as terras dos Kaxinawá e de outros povos da região foram invadidas, eles foram escravizados, foram afugentados. A pesquisadora destaca que era uma realidade muito ameaçadora e os sul-americanistas se viam com a premência de fazer esses registros, como que na última oportunidade, no último momento (ainda) possível.

A historiadora Rebeca Gontijo, professora associada do Departamento de História e do Programa de Pós-Graduação em História da Universidade Federal Rural do Rio de Janeiro, apresentou a trajetória e o legado de Capistrano de Abreu, um ícone da historiografia nacional, considerado “o pai fundador” de uma história moderna e científica no Brasil no livro *O velho vaqueano. Capistrano de Abreu: historiografia, memória e escrita de si* (7Letras, 2013). Segundo a historiadora, Capistrano de Abreu dedicou-se ao estudo da história do Brasil (séculos XVI e XVII) e das línguas indígena, desde o final do século XIX. Seu primeiro texto sobre os Bakairis é de 1895, e ao longo de toda a primeira década do século XX, ele vai se dedicar a esses estudos. Supostamente, o estudo das línguas indígenas faz parte de um projeto de estudo sobre a história brasileira, “e não me parece que eles sejam um desvio como alguns intérpretes pensaram” por que a grande questão do Capistrano é entender o processo de povoamento do Brasil, de povoamento do interior, dos sertões. Esse é o tema que vai ocupar as suas investigações e, para entender esse processo, esse autor considerava necessário compreender as relações estabelecidas com os povos indígenas, sobretudo nos dois grandes centros de povoamento dos sertões, Pernambuco e São Paulo. Ele supunha que nessas regiões alguma coisa de diferente aconteceu, e, no seu entendimento, as relações com os povos indígenas favoreceram a ocupação dos sertões. A ausência de estudos sobre os indígenas deixou uma lacuna na história brasileira. Capistrano inseriu os indígenas na história, não como uma curiosidade erudita ou como um elemento exótico da formação nacional, mas como agente necessário dessa história, cujas ações eram constitutivas da própria nacionalidade. Para Capistrano, sem entender a relação com os indígenas, a história do Brasil ficaria incompleta, observa Rebeca Gontijo.

2.2 O desafio da garantia literária do etnoconhecimento de povos originários

Segundo Barité et al (2010) o fato de basear a garantia literária no que estiver efetivamente documentado reforça a concepção de que o universo da documentação possui leis ou, pelo menos, manifestações diferentes das do universo das classificações do conhecimento, uma vez que, a documentação - notadamente - rege-se por lógicas relacionadas com o modo pelo qual os autores propõem, definem, privilegiam, inter-

cambiam, acordam, associam e analisam criticamente os tópicos que estudam, em seus conceitos e denominações. Na avaliação de Barité, na documentação são expressos tanto os tópicos cujo estudo não perde vigência ao longo do tempo como os que já são considerados obsoletos, mas também aqueles que recebem um tratamento intenso em função de situações da realidade ou de necessidades concretas. E isso ocorre em qualquer área, em virtude de seu desenvolvimento ou especialização, que gere documentos (temáticos, regulamentares, interpretativos, de divulgação etc.) em uma quantidade significativa.

Ainda segundo Barité et al (2010), o propósito de Hulme ao instituir a documentação como suporte para um princípio validador não consistia em negar o valor ou mesmo a transcendência das classificações científicas nem sua condição de estatutos de referência para sistemas de organização do conhecimento, mas sim em propor um novo enfoque, que se referencia na busca e extração de terminologia a partir da análise da documentação, assumindo não apenas um status empírico como, também, de representação do conhecimento. Neste sentido, a garantia literária encontra-se estreitamente relacionada aos aspectos semânticos das formas de representação próprias à classificação e à indexação (descritores, cabeçalhos de assunto, notações classificatórias).

Segundo Guedes e Moura (2016), a concepção de linguagem como o meio em que o conhecimento se manifesta, como veículo do pensamento, e a forma de administrar o conhecimento — pois é através da representação do conhecimento, mediado por um sistema de signos, que se torna possível organizá-lo, introduz para a OC questões de ordem linguístico-filosófica. Desde então o preceito de garantia semântica como um princípio de validação de uma linguagem documentária tem sido aplicado ao longo dos tempos na conjuntura de práticas da ciência da informação, uma vez que a representação do conhecimento ocorre a partir de um ponto de vista, tomando lugar num contexto sociocultural que é influenciado por condicionantes históricos. Dessa maneira, a garantia seria o compromisso ontológico que asseguraria que as unidades terminológicas incorporadas a um SOC carreguem o significado adequado para a finalidade e utilidade de tal dispositivo, sendo que a aceitação e confiabilidade dos SOCs vêm do alinhamento semântico estabelecido entre o sistema e à audiência em foco.

A garantia semântica implica que um instrumento destinado a representar conhecimentos por meio de uma estrutura de conceitos somente é eficaz caso seu contingente de termos consiga retratar tanto o campo semântico do domínio representado como a realidade contextual daqueles que o utiliza.

Indo ao encontro desta perspectiva, Bessa Freire (2014) destaca que uma língua contém todo o território onde é

falada, na medida em que classifica, nomeia, descreve, avalia, hierarquiza e dá sentido a tudo que nele existe: flora, fauna, acidentes geográficos, seres encantados e desencantados que o povoam, além de crenças e conhecimentos que revelam a relação entre eles. As línguas, segundo Bessa Freire (2014) não só comunicam informações, mas realizam toda uma série de práticas sociais, construindo discursos que estabelecem vínculos social, ritualizam, contam histórias, cantam, brigam, amam e contribuem para criar comunidades que se formam a partir dessas afinidades. Servem, enfim, como critério para o reconhecimento da identidade de seus falantes e as suas fronteiras. Neste caso, a extensão e a natureza do território são avaliadas pelo grau de intercomunicação entre os que convivem em um dado espaço geográfico, identificadas como fronteiras linguísticas.

Ainda segundo Freire (2014 p365), no século XVI no território que é hoje o Brasil, Pindorama, eram faladas mais de 1.300 línguas de diferentes famílias e troncos linguísticos, todas elas portadoras de narrativas orais, de conhecimentos e de memória. No entanto, longe de serem consideradas como patrimônio a ser cuidado e preservado, eram tratadas como obstáculo ao desenvolvimento da colônia, e mais tarde, à unidade nacional e talvez por isso, no período colonial não foi realizado um inventário delas, de modo que muitas das quais já extintas, foram registradas em documentação fragmentada e dispersa, produzida especialmente por missionários e menos escala por cronistas e viajantes. Um documento importante foi a tentativa de identificação dados territórios onde eram faladas feita por Curt Nimuendaju, no seu mapa Etno-histórico, elaborado entre 1942, resultando em uma cartografia linguística. Também existem documentos criados por missionários que criavam catecismos, gramáticas, dicionários, vocabulários, produzindo um metadiscorso sobre as línguas usadas nas suas áreas de atuação, no entanto, a documentação existente é dispersa e rara. Diante desse quadro de déficit documental de um etnoconhecimento baseado na oralidade, voltamos à questão também colocada por Barité et al (2010) da existência de uma 'área problema' na qual a garantia literária: os espaços interdisciplinares ou disciplinas. Segundo Zipp (1999) como uma consequência natural da expansão de um espaço interdisciplinar que se consolida, surgem as bibliotecas e centros de documentação que dão cobertura às disciplinas, e que encontram dificuldades sérias para representar o conhecimento contido nos documentos das novas áreas. Como no caso da representação do etnoconhecimento de povos originários, aponta Barité (2010) várias ordens de problemas de representação do conhecimento em torno dos espaços interdisciplinares, ao menos dos que começam a ser tratados com a contribuição de certas formas da garantia literária. Entre elas, o autor menciona as dificuldades para estabelecer um corpus documental representa-

tivo e equilibrado, considerando-se o nível de dispersão disciplinar da produção especializada de uma interdisciplina, o que, além disso, expressa-se na dispersão de sua documentação em bibliotecas especializadas em distintas disciplinas.

3 Conclusão

Enquanto obra de referência, O Glossário Decolonial de Macunaíma pretende contribuir para a documentação do Inventário Nacional da Diversidade Linguística, como instrumento de identificação, documentação, reconhecimento e valorização das línguas portadoras de referência à identidade, à ação e à memória dos diferentes grupos formadores da sociedade brasileira, inventário este instituído pelo Decreto nº 7.387, de 9 de Dezembro de 2010, tendo em vista a necessidade de ações de fortalecimento de línguas identitárias e de subsídio aos estudos nas áreas afins, dada a sua natureza epistemológica transversal, enriquecendo as políticas educacionais. Uma das maiores demandas dos grupos de falantes de línguas minoritárias está relacionada ao direito de acesso a serviços públicos na sua língua de referência e de implementação de projetos de apoio à produção literária e audiovisual, que buscam favorecer a preservação e a transmissão intergeracional.

Foram identificadas 2112 entradas referentes a uma compilação de termos e sinônimas, cada uma seguida de definição que leva em conta o aspecto semântico e o contexto da história, que no seu conjunto nos dão a dimensão da riqueza do vocabulário brasileiro em suas raízes.

As 2112 entradas que compõem o glossário foram categorizadas se utilizando da análise facetada do matemático e bibliotecário indiano Shiyali Ramamrita Ranganathan que é uma das estratégias para a representação de um determinado corpus. Ranganathan estabeleceu cinco categorias fundamentais para o seu sistema de classificação facetada: Personalidade, Matéria, Energia, Espaço, Tempo, conhecidas pela sigla PMEST. Essas categorias são mutuamente excludentes, ou seja, o que é personalidade não é matéria, nem energia, nem espaço, nem tempo. No entanto, uma palavra de um léxico pode ser categorizada em qualquer uma delas, dependendo do seu sentido e contexto.

A categoria Personalidade, para efeito de organização do léxico de Macunaíma, agrupou todos os termos referentes a fauna, flora, entidades, personagens, instituições. A categoria Energia abrange ações, reações, atividades, operações, processos, técnicas, tratamentos, problemas etc. A categoria Matéria consiste em todos os tipos de materiais e substâncias de que são feitas as coisas. A categoria Espaço corresponde às divisões geográficas e a categoria Tempo, às divisões cronológicas. Com essa categorização fizemos o controle semântico das palavras que aparecem na obra Macu-

naíma, incluindo as sinonímias encontradas no glossário de Roteiro de Macunaíma e buscadas em outras referências. Neste Glossário, utilizamos a obra de Mário de Andrade como a documentação necessária para a garantia literária de uma etnoliteratura cuja documentação é deficitária, e implica uma intertextualidade específica, uma vez que muitas das narrativas de povos originários estão ali presentes.

Referencias

BARITÉ, Mario; FERNÁNDEZ-MOLINA, Juan Carlos; GUIMARÃES, José Augusto Chaves; MORAES, João Batista Ernesto de. Garantia literária: elementos para uma revisão crítica após um século. **Transinformação**, v.22, n.2. Disponível em: <https://periodicos.puc-campinas.edu.br/transinfo/article/view/6201>. Acesso em 24 maio 2023.

AMORIM, A. B. **Lendas em Nhegatu e em Português**. Revista do Instituto Historico e Geographico Brasileiro, Rio de Janeiro: Imprensa Nacional, 1928. p. 9-475

FREIRE, J. R. B. Índio falou, tá falado. **Revista de História da Biblioteca Nacional**, v.9, p. 44-46, 2014.

COUTO DE MAGALHÃES, GENERAL. **O Selvagem**. 4.ed. São Paulo: Comp. Edit. Nac., 1940.

CHRISTINO, B. P. Os vaivéns da rede (internacional) de Capistrano de Abreu. **Revista do Instituto de Estudos Brasileiros**, v. 45, p. 37-62, 2007.

GUEDES, R. M.; MOURA, M. A. O princípio da garantia semântica e os estudos da linguagem. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 9, n. 2, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/119649>. Acesso em: 12 abr. 2023.

HART, C. F. **Contribuição para a Etnologia do Vale do Amazonas**. Archivos do Museu Nacional do Rio de Janeiro v. VI (Consagrado a Exposição Anthropologica Brasileira), p. 1-174. Rio de Janeiro: Typ. e Lith. Economica, de Machado & C. Disponível em: <http://www.etnolinguistica.org/biblio/hartt-1885-contribuicoes>. Acesso em: 23 abr. 2023.

HULME, E.W. Principles of book classification: chapter III – on the definition of class headings, and the natural limit to the extension of book classification. **Library Association Record**, n.13, p. 444-449, 1911.

SÁ, L. **Literaturas da floresta: textos amazônicos e cultura latino-americana**. SciELO-EDUERJ, 2012. [Ebook]

ZIPP, L. Core serial titles in an interdisciplinary field: the case of environmental geology. **Library Resources & Technical Services**, v.43, n.1, p.28-36, 1999.

Desafios da Organização do Conhecimento para a representação temática dos saberes originários em sistemas de informação: reflexões preliminares

Challenges of Knowledge Organization for the thematic representation of original knowledge in information systems: preliminary reflections

Luciana de Souza Gracioso, Luzia Sigoli Fernandes Costa, Omar Lopes da Silva

Universidade Federal de São Carlos (UFSCar), Rod. Washington Luís, s/n - Monjolinho, São Carlos - SP, 13565-905

Resumo

Introdução: A participação indígena nas Universidades Brasileiras tem se tornado realidade sendo este movimento, uma das principais conquistas decorrentes das lutas dos povos originários pelos seus direitos. Um dos desdobramentos desse movimento, é a produção de pesquisas, trabalhos de conclusão de curso, dissertações de mestrado e teses de doutorado que passam agora, a serem desenvolvidos pelos povos originários, ora sobre eles, ora sobre o modo como suas concepções e representações de mundo podem entender e analisar as problematizações científicas de toda ordem. Esse movimento nos coloca questões importantes frente as nossas práticas de organização e representação desses saberes, em nossos sistemas de informação. **Objetivo:** discorrer sobre alguns dos desafios que se apresentam em torno das ações de representação dos saberes plurais produzidos no escopo da produção científica acadêmica, uma vez que estas, cada vez mais, estão sendo registradas e autoarquivadas em ambientes institucionais (Repositórios Institucionais - RIs). **Metodologia:** base exploratória a partir da qual são tecidas considerações que buscam contribuir para a reflexão e posterior proposição de ações direcionadas as práticas de indexação autoral em sistemas de informação acadêmicos. **Resultados e conclusões:** é feito o indicativo de que, somadas as ações já propostas em estudos anteriores para o estabelecimento de políticas de indexação de RIs, sejam viabilizadas e estimuladas a indicação de palavras-chave nas línguas originárias das pessoas produtoras das pesquisas, e que seus sentidos e significados sejam também indicados, pelas pessoas autoras/indexadoras, neste metadado de assunto.

Palavras-chave: Indígenas. Ações afirmativas. Repositórios Institucionais.

Abstract

Introduction: The indigenous participation in Brazilian universities has become a reality and this movement is one of the main achievements resulting from the struggles of native peoples for their rights. One of the outcomes of this movement is the production of research, course completion papers, master's theses and doctoral dissertations that are now being developed by native peoples, sometimes about them, sometimes about how their conceptions and representations of the world can understand and analyze the scientific problems of all kinds. This movement raises important questions about our practices of organizing and representing this knowledge in our information systems. **Objective:** To discuss some of the challenges that are presented around the representation actions of plural knowledge produced in the scope of academic scientific production, since these are increasingly being registered and self-archived in institutional environments (Institutional Repositories - IRs). **Methodology:** exploratory basis from which considerations are made that seek to contribute to the reflection and subsequent proposition of actions directed towards authorial indexing practices in academic information systems. **Results and conclusions:** it is indicated that, added to the actions already proposed in previous studies for the establishment of policies of indexing of IRs, the indication of keywords in the original languages of the people who produce the research should be made viable and stimulated, and that their senses and meanings should also be indicated, by the authors/indexers, in this subject metadata.

Keywords: Indigenous. Affirmative Action. Institutional Repositories.

Resumen

Introducción: La participación indígena en las universidades brasileñas se ha convertido en una realidad y este movimiento es uno de los principales logros resultantes de las luchas de los pueblos nativos por sus derechos. Uno de los resultados de este movimiento es la producción de investigaciones, trabajos de fin de curso, tesis de maestría y disertaciones doctorales que ahora están siendo desarrolladas por los pueblos nativos, a veces sobre ellos, a veces sobre la forma en que sus concepciones y representaciones del mundo pueden comprender y analizar los problemas científicos de todo tipo. Este movimiento nos plantea importantes cuestiones sobre nuestras prácticas de organización y representación de estos conocimientos en nuestros sistemas de información. **Objetivo:** discutir algunos de los desafíos que se presentan en torno a las acciones de representación de los conocimientos plurales producidos en el ámbito de la producción científica académica, ya que éstos, cada vez más, están siendo registrados y autoarchivados en

ambientes institucionales (Repositorios Institucionales - RIs). **Metodología:** base exploratoria a partir de la cual se tejen consideraciones que buscan contribuir a la reflexión y posterior proposición de acciones dirigidas a las prácticas de indización autoral en los sistemas de información académica. **Resultados y conclusiones:** se indica que, sumado a las acciones ya propuestas en estudios anteriores para el establecimiento de políticas de indización de RIs, la indicación de palabras clave en los idiomas originales de las personas que producen las investigaciones, y que sus sentidos y significados también sean indicados, por los autores/indexadores, en estos metadatos temáticos.

Palabras clave: Indígena. Indización. Ações afirmativas. Repositórios Institucionais.

1 Introdução

Diferentes abordagens sobre os estudos em Organização do Conhecimento (OC) tornam-se, irrefutavelmente, cada vez mais necessárias, frente as atuais configurações e dinâmicas sociais que relacionam modos de vidas e intermediações tecnológicas. Tal configuração tem demandado a este campo, avançar para além das ações de representarão da informação, para atuar também em uma esfera de mediações simbólicas (como sugere Muniz Sodré (2014) no que diz respeito as ações de comunicação), uma vez que, cada vez mais, se intensificam os arranjos sociotécnicos que se articulam na sociedade, o que tem evidenciado o quanto somos pessoas sujeitas de ações comunicantes, que antes se serem sintáticas ou semânticas, são *transverbais*.

Esse preambulo se faz necessário para justificar que o estudo proposto se estabelecerá em um plano de reflexão e proposição ainda em construção, embora arrisque a apresentar ponderações em caráter conclusivo. No imediato, o que se almeja, de modo muito experimental, é transitar e especular sobre alguns dos fenômenos e eventos que tem emergido, especialmente nos ambientes universitários brasileiros, muito em função da diversidade e pluralidade de vozes, que atualmente ecoam nos campus acadêmicos, em decorrência da implementação de um conjunto de políticas de Ações Afirmativas, tanto na graduação, como na Pós-graduação, que trazem para consigo - principalmente pelos estudantes indígenas - uma diversidade cultural e de possibilidades de entendimento da realidade que reverberam nos modos de produção da ciência.

Nesse contexto é objetivo deste texto ponderar sobre algumas questões que se apresentam em torno das ações de representação da produção científica acadêmica em ambientes institucionais (Repositórios Institucionais - RIs), uma vez que estas, cada vez mais, estão sendo produzidas, registradas e autoarquivadas por seus atores/produtores cada vez mais plurais.

2 A linguagem cotidiana como ponto de partida

Pinho e Milani (2020) em caráter conclusivo ao estudo desenvolvido sobre ética em Organização do Conhecimento, nos dizem que:

As contribuições advindas da Linguística, Terminologia e das Teorias críticas são imprescindíveis para reforçar que a neutralidade não existe na organização do conhecimento e que, para abrigarmos as diferenças, devemos explorar os limites dos sistemas de organização do conhecimento e justificar as nossas decisões. (PINHO, MILANI, 2020, p. 84).

Assim, mobilizados por essas constatações, o presente estudo busca compreender, de modo introdutório e pela via interdisciplinar, sobre qual seria o lugar da linguagem comum ou cotidiana (ou ainda, originária), nas ações de representação da informação científica.

Em estudo recente Makoni e Severo (2020), nos sinalizam para a necessidade e atentarmos para a valorização das línguas locais e comuns, especialmente por conter, nelas, uma via para a construção de um comum global, que seja mais equitativo e inclusivo. Tais proposições se dão a partir de uma perspectiva linguística de integração decolonial e que está alinhada a defesa de que todas as línguas teriam igual valor. Em função disto, todas as línguas deveriam ter seus espaços e suas esferas de manifestação garantidas e respeitadas. Considerando este ponto de partida, é que o diálogo intercultural seria possível, e outras compreensões sobre o desenvolvimento social, poderiam ser também ponderados na conjuntura global. Os autores, se apoiam aos estudos pioneiros de Taylor (1997, 1981) para essa defesa, cuja pauta está cada vez mais presente nos espaços acadêmicos em função das novas configurações sociais e científicas e dos usos e mediações simbólicas promovidas pelas tecnologias da informação.

Nos anos de 2020 e 2023, por exemplo, a UFSCar (Universidade Federal de São Carlos) promoveu o SIELIPOP (Seminário Internacional de Estudos em Linguística Popular) (UFSCAR, 2023). No escopo deste evento, seus articuladores categorizam a *Folk linguística* como sendo a linha de trabalho dedicada a linguagem que defende que a mesma merece ser estudada independente dos grupos sociais que as produzem e acionam. Não haveria, portanto, uma lógica predominante de uma estrutura de linguagem, sobre outra. Alicerçados em autores como N. Niedzielski e D. Preston, fazem a defesa de que os saberes populares são substanciais para a construção do discurso científico. A *Folk linguistics* designaria então, todo o trabalho sobre linguagem, isto é, os saberes espontaneamente construídos pelos mais diversos

atores sociais, que não estão necessariamente fundamentados em uma lógica de uma teoria da linguagem. (SIELIPOP, 2023). Essas são também as defesas de Baronas, principal articulador do referido evento e que se apoia substancialmente nas obras de Marie-Anne Paveau, conferencista nas duas edições do simpósio, e que faz a defesa de que as abordagens científica e popular, da linguagem, não são excludentes. Assim, linguística popular e linguística acadêmica, poderiam se enriquecer mutuamente.

Esses apontamentos interdisciplinares gerais se dão em função da necessidade de argumentarmos a favor de que esforços continuem sendo direcionados, nos estudos informacionais, para compreensão e valorização, não só de saberes locais, mas como de linguagens locais, quando da proposição e aperfeiçoamento de sistemas de organização do conhecimento, estejam estes no plano das ações comunicativas em sistemas abertos de informação, ou no âmbito dos sistemas controlados de registro e acesso institucionais à informação científica, por exemplo. Essa defesa se localiza no bojo dos desdobramentos das políticas de ações afirmativas brasileiras, implementadas no ensino superior, e que tem oportunizado, ainda que de modo não suficiente, o acesso mais plural e diverso à Universidade, como é o caso do acesso de estudantes indígenas às Instituições Públicas de Ensino Superior Brasileiro que, mesmo recente, cresce exponencialmente.

3 Estudantes indígenas: em defesa de seus saberes, linguagens e sentidos na academia.

Dados levantados pelo Censo da Educação Superior do Inep (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira) indicam que, em 2018, 57.706 indígenas estavam matriculados no ensino superior, tendo havido um crescimento de 695% em relação a 2010, anos iniciais de implementação das políticas de ações afirmativas e vestibular indígena no Brasil. (LUCIANO, AMARAL, 2021).

É sempre necessário mencionar que muitas destas conquistas resultam das contribuições das lideranças do movimento negro em Instituições de ensino, como a Profa. Dra. Petronilha Beatriz Gonçalves e Silva, que foi conselheira da Câmara de Educação Superior do Conselho Nacional de Educação e relatora do Parecer CNE/CP 3/2004 que estabelece as Diretrizes Curriculares Nacionais para a Educação das Relações Étnico-Raciais e para o Ensino de História e Cultura Afro-brasileira e Africana, além de ter sido a coordenadora do Grupo Gestor do Programa de Ações Afirmativas da UFSCar, de 2007 a 2011.

Especificamente, no caso da UFSCar, reconhecida pelo seu pioneirismo na implementação das políticas e ações afirmativas, especialmente do vestibular

indígena, temos que, desde a primeira edição deste vestibular ocorrida em 2008, já passaram pela Universidade, mais de 43 etnias distintas. Em 2023, houve o maior número de inscrições neste vestibular (agora feito em parceria com a UNICAMP – Universidade Estadual de Campinas). Foram 3.480 inscrições para as 130 vagas da UFSCar e as 130 vagas da Unicamp. Atualmente, na UFSCar, constam mais de 400 estudantes indígenas com matrículas ativas. Mais recentemente em 2016 foi aprovada, nesta instituição, a Política de Ações Afirmativas orientadas para a Pós-graduação [1], sendo esta, resultante de um coletivo trabalho coordenado pela Secretaria de Ações Afirmativas, Diversidade e Equidade (SAADE). Em função dessa ainda recente implementação, a presença de estudantes indígenas na pós-graduação, é pequena, mas, em crescimento. Assim, é possível afirmar que, a partir das políticas constantemente implementadas e com o aumento sistemático de procura e ingresso de estudantes indígenas na graduação e na pós-graduação brasileira, a produção científica pelos povos originários, só irá aumentar.

São outras diferenciais desta instituição:

- 2009 - Construção identitária do Centro de Culturas Indígenas (CCI) – sendo este o primeiro centro desta natureza em Universidade Pública no país;
- 2010 – Criação dos Programas de Educação Tutorial: Pet Saberes indígenas e Pet Ações e Saúde;
- 2012 – Encontro Nacional dos Estudantes Indígenas (ENEI), sendo este o primeiro encontro desta natureza em Universidade Pública no país – atualmente o evento tem alcance nacional e está em sua nona edição.

As pessoas autoras indígenas, Gonçalves, Waurá, Marcondy e Batista (2021, p.58) complementam ainda que:

As conquistas do movimento indígena no contexto da UFSCar atingem instâncias nacionais e internacionais, gerando participações em eventos científicos de alta relevância como, por exemplo, o da Sociedade Brasileira para o Progresso da Ciência (SBPC), em cujo espaço foi demarcada a produção cultural, científica e de saberes tradicionais dos povos indígenas - SBPC Indígena. Assim como a criação de projetos de cooperação internacional, como o Projeto Abdias do Nascimento (Brasil-Espanha) junto à Universidade de Córdoba, que proporcionou aos estudantes indígenas a experiência de adentrar o território e cultura europeia, mediado por um olhar crítico acerca de suas relações com tal universo. (GONÇALVES, WAURÁ, MARCONDY, BATISTA, 2021, p. 58).

Assim, é possível constatar esforços contundentes para garantir o acesso, a visibilidade e a permanência estudantil indígena, nas esferas acadêmicas, mas ainda há muito a ser feito. A professora e ativista indígena,

Célia Nunes Correa, do povo Xakriabá em Minas Gerais, disse, em sua dissertação de mestrado, “(...) que há duas maneiras de matar o povo indígena coletivamente: quando nos negam o território e quando reproduzem o epistemicídio.” (CORREA Xakriabá, 2018, p. 102). Nesta mesma perspectiva de alerta, temos o indicativo feito no livro "Acadêmicos indígenas em Roraima e a construção da interculturalidade indígena na Universidade", publicado recentemente em 2022, no qual seu autor, João Francisco Kleba Lisboa argumenta que a Universidade deve ser um local de construção intercultural, ao invés de impor novas formas de colonização sobre os conhecimentos originários, e, para tanto, afirma:

Se, por um lado, os indígenas mantêm uma contínua construção de alianças com os mais diversos sujeitos não-índios para alcançar esse objetivo, também parece claro a eles que é necessária uma precaução constante quanto aos valores ambíguos e vacilantes da sociedade ocidental que os circunda. (LISBOA, 2022, p.235).

Com propósitos comuns, o trabalho de Lillian DePaula “Como/quando pensar em língua indígena? Resistências e conformidades”, problematiza a ação de tradução com recurso utilizado na interação entre pessoas, alertando para os riscos que se corre, nesse processo de uma forma de racionalizar o pensamento, se sobrepondo, propositalmente, a outra. A autora apresenta elementos para que seja possível reconhecermos os silenciamentos que se dão, nesse processo:

A língua indígena, por exemplo, quando considerada de segunda importância, define a relação que o falante terá com a língua, o engajamento que fará para melhor usufruir de sua língua-mãe. Desconsiderar o dialeto de um grupo social, tratando-o como inferior e de menor importância, reflete na maneira com a qual aquele ser lidará com sua própria identidade. (DEPAULA, 2008, p. 390).

Ainda, como agravante, temos que, “Cerca de 97% da população mundial fala somente 4% dessas línguas, e somente 3% das pessoas do mundo falam 96% de todas as línguas existentes.” (UNESCO, 2022). Em função deste estado crítico, a Assembleia Geral das Nações Unidas (Resolução A/RES/74/135) declarou que entre 2022 e 2032 será considerada a Década Internacional das Línguas Indígenas (*International Decade of Indigenous Languages – IDIL 2022-2032*), com o propósito de sensibilizar e otimizar a produção de ações pela sociedade que minimizem o risco de desaparecimento de línguas originárias, e junto a elas, todo um modo de conceber, representar e conhecer o mundo. (UNESCO, 2022).

Nessa direção, desde 2018, o Programa de Educação Tutorial (PET) “Conexões de saberes indígenas” da UFSCar vem envidando esforços para mapear territorialmente e caracterizar a diversidade de línguas

faladas e escritas pelos estudantes indígenas da UFSCar.

Os dados levantados estão sendo analisados, territorialmente, e relacionados à diversidade étnica para a geração de um “catálogo” das línguas ativas, entre os estudantes da UFSCar.

Figura 1. *Famílias Linguísticas Indígenas presentes na UFSCar*

Famílias Linguísticas presentes na UFSCar				
Aruak	Iranxe	Krenak	Pano	Tukano
Bororo	Jê	Maxakali	Rikbaktsá	Tupi-Guarani
Guaikuru	Karib	Mondé	Tikuna	

Fonte: Coordenadoria de Acompanhamento Acadêmico e Pedagógico para Estudantes (CAAPE) (UFSCAR, 2020).

Esta iniciativa visa contribuir para que os pesquisadores de diferentes áreas de conhecimento e para as próprias políticas institucionais tenham melhor representatividade em relação à produção acadêmica, por meio de proposta de inserção de palavras-chave e resumos em línguas indígenas em publicações, que tratam de temáticas indígenas, depositadas no Repositório Institucional da UFSCar. (WAURA, COSTA, SILVA E PALOMINO, no prelo).

Outros programas institucionais na Universidade, também tem se dedicado ao tema como o Laboratório LEETRA, coordenado pela Profa. Maria Sílvia Cintra Martins, responsável pela revista LEETRA Indígena, periódico diferenciado que viabiliza a publicação de poéticas indígenas e que é em geral publicada também nas línguas originárias.

Essas proposições tem sido construídas com a participação direta de pessoas pesquisadoras indígenas, que, no lugar de produtores da ciência, reconhecem as lacunas ainda existentes nos sistemas de informação para promoverem a devida visibilidade aos saberes produzidos. E a partir destes estudos, é que seguimos com uma sinalização pontual, relacionada a incrementação de alguns metadados de representação de conteúdos, no Ris.

4 Desafios da organização do conhecimento para a representação dos saberes originários em sistemas de informação: reflexões preliminares

Partimos do lugar da Biblioteconomia e da Ciência da Informação, para analisar como a área tem se voltado a desenvolver produtos e serviços relacionados a organização do conhecimento científico, agora produzidos dentro das especificidades dos saberes cosmológicos produzidos pelas pessoas autoras indígenas. Nos direcionamos a refletir, pontualmente, sobre o Repositório Institucional da UFSCar, apenas em caráter de exemplo e não propositivo, no intuito de

ponderar sobre como o sistema institucionalizado para registro e promoção do acesso aberto da produção científica da Universidade, poderia, para além da função de registrar, preservar, valorizar e dar visibilidade a produção científica indígena, também garantir a sobrevivência das línguas originárias e consequentemente das outras possibilidades de concepção de mundo, que se dão, a partir delas.

Mas, é válido reiterar que temos notadamente, um conjunto de estudos dedicados a problematizar e a sugerir proposições que atenuem os silenciamentos, as omissões e repressões, por vezes reproduzidas e reforçadas pelos sistemas de Organização do conhecimento e Representação da informação. Diferentes estudos relacionados a implicações éticas na Organização do Conhecimento têm sido produzidos em âmbito nacional e tem nos alicerçado para o desenvolvimento do presente estudo, pois nos permitem conjecturar de modo mais adequado qualquer proposição relacionada às questões relativas aos povos originários brasileiros. Assim, as pesquisas de José Augusto Guimarães, Fábio Assis Pinho, Suelen Milani, Vinícios Menezes, dentre outras pessoas pesquisadoras, tem dado o suporte teórico e conceitual, para seguirmos no percurso das proposições que serão ponderadas, no bojo deste estudo. Ao mesmo tempo, pesquisas recentes diretamente relacionadas ao desenvolvimento de métodos e políticas de indexação em Repositórios Institucionais tem subsidiado as reflexões e potenciais proposições futuras, relacionados ao assunto, e que trataremos, a seguir.

4.1 Repositório institucional: espaço de registro e emancipação dos saberes originários

Os estudos dedicados aos Repositórios Institucionais (em suas diferentes configurações), tem recebido cada vez mais atenção da área de Ciência da Informação. Tal como constatado em tese defendida recentemente por Brazilo (2022), a CI é a área que mais desenvolve e publica pesquisas sobre o assunto, tanto em âmbito nacional, como internacional. Especificamente, é possível também constatar, tal como indicado por Fujita *et al.* (2023) que no campo da Organização do Conhecimento, especificamente relacionadas aos temas contemplados no escopo dos Sistemas de Classificação da Literatura de Organização do Conhecimento (CSKOL), há uma produção de pesquisas importante, em curso. Em outros estudos, Fujita, Tartarotti, Panuto (2023), tem evidenciado a relevância e necessidade de atentarmos, em nossas atividades de pesquisa e profissionais, para o modo como as representações das informações estão sendo construídas nestes ambientes, uma vez que esses processos impactam diretamente no modo como seus conteúdos serão acessados (ou não), pela sociedade.

Considerando a predominância de práticas de autoarquivamento nos Repositórios, Silva (2022) faz uma contunde revisão sobre os conceitos e as práticas que envolvem esse processo. Em seu trabalho sobre “Atribuição de palavras-chave na prática do autoarquivamento: estudo das diretrizes de repositórios de dados de pesquisa ibero-americanos”, a autora conclui que:

(...) tem-se um cenário preocupante sobre a atribuição de palavras-chave na prática do autoarquivamento em repositórios de dados de pesquisa ibero-americanos, sendo oportunos estudos complementares que proponham orientações referentes ao metadado assunto nesses sistemas de informação.” (SILVA, 2022, p. 07).

Nesta mesma linha, o trabalho de Freitas, Dal’Evedove e Tartarotti (2021), confirmam a ainda insipiente prática de oferta de orientações, pelos Ris, para pessoas autoras, no que diz respeito a indicação de metadados de assunto durante o autoarquivamento de suas produções científicas. Neste contexto, Freitas (2019) indica, em caráter conclusivo de sua dissertação, em que analisou pontualmente o RI da UFSCar, que:

(...) mostra-se oportuna a construção de um vocabulário controlado próprio para o RI/UFSCar com a **participação do usuário enquanto elemento fundamental no processo de construção colaborativa na indexação**. Sendo o autor da publicação científica um especialista em sua área de atuação, a construção de um instrumento desta natureza contribuirá para o fortalecimento das relações semânticas entre os termos nos índices de assuntos dos repositórios institucionais.” (FREITAS, 2019, p. 79, grifo nosso).

Após ampla agenda de pesquisa, ancorada em métodos e proposição de modelos de avaliação dos processos de indexação de Ris, Fujita (2022) sintetiza uma proposta de sistematização e aplicação das metodologias de avaliação do controle de vocabulário em repositório, que, mesmo direcionado ao Repositório Institucional Unesp, podem ser adaptadas para os demais. Em caráter conclusivo deste estudo, diz a autora que se faz necessário a:

(...) elaboração de uma política de indexação adequada para repositórios que considere todos os atores e fatores em prol do controle de vocabulário e **da riqueza terminológica de especialidade da linguagem natural**. (FUJITA, 2022, p. 19, grifo nosso).

Assim, diante do conjunto de estudos analisados e considerando especialmente as conclusões alcançadas pelos mesmos, o presente texto, em caráter conclusivo, almeja apenas apresentar mais um conjunto de sinalizações sobre uma prática possível, relacionada a incrementação do metadado de assunto, no processo de indicação de palavra-chave, durante o processo de autoarquivamento de trabalhos acadêmicos.

5 Resultados

Visando contribuir para que os saberes científicos produzidos pelas pessoas estudantes indígenas não fiquem inviabilizados nas estruturas e sistemas de informação que controlam seus registros e fluxos, as pessoas autoras do presente estudo fazem o indicativo de que, somadas as ações já propostas em estudos anteriores, sejam viabilizadas e estimuladas a indicação de palavras-chave, não só nas línguas originárias das pessoas produtoras das pesquisas, mas que os seus sentidos e significados sejam também indicados neste metadado. Essa contextualização pragmática poderá, em certa circunstância, oportunizar que o mesmo saber, estruturado e representado nos sistemas e lógicas científicas, possam manter suas subjetividades e identidades, coletivas. Para além disto, sugere-se que sejam ampliados os metadados de modo que resumos e palavra-chave na linguagem originárias, possam ser incluídos. De modo complementar, em caráter descritivo, faz-se o indicativo de que seja viabilizado o metadado para autorias coletivas, indicação de gênero e etnia indígena.

Em síntese, as proposições feitas, até o presente momento e que já foram apresentadas a equipe bibliotecária responsável pelo Ris da UFSCar, são:

- Inclusão de metadado de assunto que possibilite a indicação do significado da palavra-chave, sugerida pela pessoa autora, no ato do autoarquivamento [2];
- Inclusão de metadado para resumo em língua originária [3];
- Inclusão de metadado para palavra-chave em língua originária;
- Inclusão de metadado para autorias coletivas,
- Inclusão de metadados para indicação de etnia indígena na autoria.

A proposição feita, tem como objetivo oportunizar a pessoa autora/ indexadora, que as palavras-chave que consideram representativas dos saberes apresentados em suas respectivas pesquisas, possam ser mantidas e validadas como um potencial descritor nos sistemas de representação da informação, tanto para a indexação de conteúdo, como para a recuperação a informação. Ao mesmo tempo, esta prática pode subsidiar pessoas bibliotecárias em suas ações de validação dos metadados de assuntos atribuídos pelas pessoas autoras depositantes de seus produtos científicos nos Ris e estabelecimento de relações semânticas, que validem os sentidos e significados atribuídos pelos estudantes indígenas quando da produção de linguagens controladas.

5 Conclusões

Respalgadas em referencial teórico oriundos da filosofia pragmática da linguagem, que indica ser, na explicação do uso da palavra que se daria a apresentação de seu sentido mais genuíno e ancoradas em orientações interdisciplinares da *Fok linguística* que sinalizam para a necessidade de validarmos os pontos de partida epistêmicos das vozes diversas que compõe, cada vez mais, a produção de saberes científicos, seguimos reiterando que estudos dedicados a estas questões tem se mostrado potentes para garantir o registro e divulgação das línguas indígenas, especialmente aquelas utilizadas nos ambientes universitários, na produção científica. Conclui-se que, com as de Políticas Ações Afirmativas implantadas nas Universidades, torna-se cada vez mais necessário que as Bibliotecas estejam integradas a esses movimentos e aos seus desdobramentos, intensificando seu compromisso de preservar as Línguas dos Estudantes Indígenas dentro da Universidade e nesse sentido, os Sistemas de informação e as políticas de Organização do Conhecimento, tem um lugar central, na defesa desta causa.

Assim, garantir que títulos, resumos e palavras-chave da produção acadêmica indígena possam ser recuperados também em sua versão originária, parece ser uma frente importante a ser assumida pelas estruturas dos Ris, especialmente o da UFSCar, em função da diversidade étnica e linguística de seus estudantes. Além disso, o conhecimento das línguas, para além de facilitar a comunicação e a interação entre as pessoas no ambiente científico, poderá minimizar preconceitos e outras formas de racismos e, como indicado na introdução deste texto, poderá viabilizar, comunicações *transverbais* e diálogos interculturais de modo que outras compreensões sobre o desenvolvimento social, possam também, ser consideradas.

Reconhecer, respeitar e preservar linguagens, são formas de construção de sociedades mais justas, uma vez que é, via linguagem, que se constituem todas as demais ações e relações que estabelecemos socialmente.

Notas

- [1] Portaria Normativa Nº 13, de 11 de maio de 2016, do Ministério da Educação (MEC), que dispõe sobre a indução de Políticas de Ações Afirmativas voltadas para negros, indígenas e pessoas com deficiência na Pós-Graduação.
- [2] Em caráter complementar a defesa da inclusão de explicação do uso de palavra-chave, indicamos a leitura de Gracioso (2008).
- [3] Em caráter complementar a defesa da inclusão de resumo, indicamos a leitura de Costa (2008).

Referências

- BARONAS, R., LOURENÇO COSTA, J., & BONANI CONTI, T. C. (2021). Resignificação discursiva em diferentes contextos: linguística popular e ludolinguistas. **Revista Da Anpoll**, 52, n.1, p.15–33. <https://doi.org/10.18309/ranpoll.v52i1.1530>
- BAZILIO, A. P. M. **Repositório institucionais a serviço da Ciência, Tecnologia e Sociedade: uma análise RiUFF e LUME**. Tese (Doutorado em Ciência, Tecnologia e Sociedade). Universidade Federal de São Carlos, 2022. 222f.
- CORREA Xakriabá, C. N. **O barro, o genipapo e o giz no fazer epistemológico de autoria xakriabá: reativação da memória por uma educação territorializada**. Dissertação de mestrado MESPT. UnB. 2018.
- COSTA, L. S. F. **Uma contribuição da teoria literária para a análise de conteúdo de imagens publicitárias do fim do século XIX e primeira metade do século XX, contemplando aspectos da natureza brasileira** / Luzia Sigoli Fernandes Costa. -- 2008. 261 f.: il. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista UNESP, Marília, 2008.
- DEPAULA Lillian. Como/Quando pensar em Língua Indígena? Resistências e conformidades. *In*: CONGRESSO DE ESTUDOS LITERÁRIOS, 8., 2008. Vitória: PPGEL/MEL, 2008.
- FREITAS, M. P. **Autoarquivamento e representação de assunto: estudo analítico de teses e dissertações do Repositório Institucional da UFSCar**. Dissertação de Mestrado (Programa de Pós-Graduação em Ciência da Informação), 2019. Universidade Federal de São Carlos, UFSCar, 2019.
- FREITAS, M. P. Dal'Evedove, P. R. TARTAROTTI, R.C.D.E. Políticas de autoarquivamento em repositórios institucionais brasileiros: estudo analítico do metadado assunto. **Páginas a&b**: arquivos e bibliotecas, 2021. 169-175.
- FUJITA, M. S. L.; DALLEVEDOVE, R. C. P. R.; PANUTO, J. C. . The indexing policy in the practices of Brazilian institutional repositories: a diagnostic study from the perception of managers and indexers. **Journal of Librarianship and Information Science**, v. 55, p. 1-19, 2023.
- FUJITA, M. S.L.; AGUSTIN-LACRUZ, C.; TOLARE, J. B.; TERRA, A. L.; BUENO DE LA FUENTE, GEMA . Institutional repositories and knowledge organization: a bibliographic study from Library and Information Science. **Education for Information**, v. 39, p. 51-66, 2023.
- FUJITA, M. S.L. Sistematização de modelo de avaliação do controle de vocabulários em repositórios: relato de pesquisa com o Repositório Institucional Unesp. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**. Campinas, v. 20, p e022013, 2022. Disponível em: <https://www.scielo.br/j/rdbci/a/7zdxRD8ss7fVvMk8KZ5FJDDy>. Acesso em: 28 maio 2023.
- GONÇALVES, G. S. WAUIRÁ, K. MARCONDY, M, BATISTA. Guia de saberes: busca de caminhos acadêmicos, éticos, estéticos e político-administrativos na formação indígena universitária. *In*: Silva, et al. **PET conexões saberes indígenas, UFSCar**: das origens aos horizontes. São Carlos : CPOI/UFSCar, 2021.
- GRACIOSO, L. S. **Filosofia da linguagem e ciência da informação: jogos de linguagem e ação comunicativa no contexto das ações de informações em tecnologias virtuais**. Tese (Doutorado em Ciência da Informação)- Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal Fluminense, Rio de Janeiro, 2008.
- LISBOA, J. F.K. **Acadêmicos indígenas em Roraima e a construção da interculturalidade indígena na universidade**. São Paulo : Bookerfield, 2022.
- LUCIANO, G. Y AMARAL, W. (2021). Povos indígenas e educação superior no Brasil e no Paraná: desafios e perspectivas. *Integración y Conocimiento*, v.10, n.2, p.13–37.
- MAKONI, S., SEVERO, C., Southern perspectives of language and the construction of the common, **Language & Communication**, <https://doi.org/10.1016/j.langcom.2022.06.003>.
- NIEDZIELSKI, N. A., PRESTON, D. R. **Folk Linguistics** Walter de Gruyter, 2000 – 375p.
- PINHO, F., MILANI, S., ética em organização do conhecimento: categorização de termos fronteiriços em relação a gênero e sexualidade. **LOGEION: Filosofia da Informação**, Rio de Janeiro, v. 6, n. 2, p. 84-103, mar./ago. 2020.
- SILVA, W. K. P. **Atribuição de palavras-chave na prática do autoarquivamento: estudo das diretrizes de repositórios de dados de pesquisa ibero-americanos**. Dissertação de Mestrado (Programa de Pós-Graduação em Ciência da informação), 2022. Universidade Federal de São Carlos, UFSCar, 2022.
- SODRÉ, M. **A ciência do comum** – Notas para o método comunicacional. Petrópolis, RJ, Vozes: 2014. 323p
- TAYLOR, T. **Enculturating language**. Theorizing Language. Pergamon Press, NY. 1997.
- TAYLOR, T.J. **Linguistic Theory and Structural Stylistics**. Pergamon, Oxford. 1981.
- UNIVERSIDADE FEDERAL DE SÃO CARLOS - UFSCAR. **Seminário Internacional de Estudos em Linguística Popular (SIELIPOP)**. São Carlos: UFSCAR, 2023. Disponível em: <http://www.sielipop.ufscar.br/index.php/SIELIPop/>. Acesso em: 27 maio 2023.
- UNIVERSIDADE FEDERAL DE SÃO CARLOS - UFSCAR. Coordenadoria de Acompanhamento Acadêmico e Pedagógico para Estudantes (CAAPE). São Carlos: UFSCAR, 2020. Disponível em: <https://www.acessibilidade.ufscar.br/apresentacao/caape>. Acesso em: 27 maio 2023.
- UNESCO. Lançamento oficial da Década Internacional das Línguas Indígenas no Brasil. 2022. Disponível em: <https://pt.unesco.org/news/lancamento-oficial-da-decada-internacional-das-linguas-indigenas-no-brasil>. Acesso em 20 de abril de 2023.
- WAURA, K. COSTA, L., SILVA, O. PALOMINO, T. Línguas e etnias de estudantes indígenas da UFSCar: um estudo preliminar. XI ENCONTRO NACIONAL DE ESTUDANTES INDÍGENAS, 2018. ENEI. **Anais...** [no prelo].

Em busca de uma terminologia: o circo no Brasil de 1820 a 2000

Searching for a terminology: circus arts in Brazil from 1820 to 2000

Ciro Ítalo Tertulino (1), Marcos Luiz Cavalcante de Miranda (2)

(1) Universidade Federal do Estado do Rio de Janeiro, Av. Pasteur, 458, Urca, Rio de Janeiro, 22290-240, Brasil, ciro.italo@gmail.com (2) mlmiranda@unirio.br

Resumo

Objetivo: Identificar os termos e o desenvolvimento/variação no tempo dos saberes e práticas do domínio circo presentes nas biografias dos circenses disponibilizadas no *site* Circodata: dicionário do circo brasileiro. **Metodologia:** O presente estudo se configura como uma pesquisa aplicada de abordagem quantitativa, exploratória em sua finalidade e baseado na Análise de Domínio (HJØRLAND, 2017) e na Análise de Conteúdo (BARDIN, 2015) como procedimento técnico. **Resultados:** O *corpus* de análise deste estudo se constituiu de 1180 biografias e resultou em um total de 555 termos relacionados a diferentes contextos do domínio circo. Possibilitando construir uma representação visual do conhecimento através de uma linha do tempo que abrange um período de 180 anos, de 1820 até 2000. Efetuou-se também, uma ordenação alfabética sistemática desses saberes e práticas. **Conclusão:** O texto biográfico se mostrou significativo para o mapeamento do universo terminológico do circo, refletindo o que vem sendo discutido na literatura especializada da área, inclusive, no que se refere à falta de uma terminologia padronizada para esse domínio.

Palavras-chave: Circo no Brasil. Conhecimento Circense. Organização do Conhecimento. Arte. Estudos Biográficos.

Abstract

Aim: to identify circus domain terms present in the circuses biographies available on the *website* “Circodata: dicionário do circo brasileiro” and to recognize the development/variation of these knowledges and practices over time. **Methodology:** the present study is characterized as an applied research following a quantitative approach with an exploratory finality. As technical procedure it is based both in domain (HJØRLAND, 2017) and content analyses (BARDIN, 2015). **Results:** A total of 1180 biographies and 556 terms related to variety of contexts in the circus domain were identified in this corpus analysis. Following these was developed a visual representation of the knowledge in a timeline of 180 years (from 1820 to 2000), and an systematic alphabetical ordering of these knowledge and practices. **Conclusion:** The biographical text demonstrated to be a significant tool for mapping the terminological circus universe. Corroborating what has been discussed in the specialized circus literature, moreover with regard to the lack of a standardized terminology for this domain.

Keywords: Circus Arts in Brazil. Circus Arts Knowledge. Knowledge Organization. Art. Biographical Studies.

Resumen

Objetivo: Realizar una distribución temporal de términos relativos a conocimientos y prácticas del dominio circense, presentes en las biografias de artistas de circo disponibles en el sitio web Circodata: dicionário do circo brasileiro y verificar la evolución/variación de estos términos. **Metodología:** Este estudio se configura como una investigación aplicada de enfoque cuantitativo, exploratoria en su propósito, basada en el Análisis de Dominio (HJØRLAND, 2017) y Análisis de Contenido (BARDIN, 2015) como procedimiento técnico. **Resultados:** El corpus de análisis de este estudio consistió en 1180 biografias y dio como resultado un total de 555 términos relacionados con diferentes contextos del dominio del circo. Esto permitió construir una representación visual del conocimiento a través de una línea de tiempo que abarca un período de 180 años, de 1820 a 2000. También se llevó a cabo una ordenación alfabética sistemática de estos conocimientos y prácticas. **Conclusión:** El texto biográfico resultó significativo para el mapeo del universo de la terminología circense, reflejando lo que ha sido discutido en la literatura especializada en el área, incluso, con relación a la falta de una terminología estandarizada para este dominio.

Palabras clave: Circo en Brasil. Conocimiento Circense. Organización del Conocimiento. Arte. Estudios Biográficos.

1 Introdução

A cultura circense se caracteriza, principalmente, pela transmissão dos seus conhecimentos por meio da oralidade, o que pode acarretar perdas parciais ou totais

de informações sobre os seus saberes, bem como em dificuldades para acessar informações e realizar pesquisas nesse contexto. Dito isso, o presente estudo objetiva identificar os saberes e práticas nas biografias dos circenses disponibilizados no *Circodata: dicionário*

do circo brasileiro, e verificar o desenvolvimento/variação desses saberes e práticas ao longo tempo.

Partindo do pressuposto de que os métodos biográficos tem como característica priorizar os elementos da oralidade para compreender situações e fenômenos vivenciados pelos sujeitos pesquisados, buscou-se investigar de que forma as biografias dos circenses poderiam contribuir para a identificação de termos e conceitos relacionados ao conhecimento do circo no Brasil. Assim, este estudo se caracteriza como uma pesquisa aplicada de abordagem quantitativa, exploratória em sua finalidade, com base na Análise de Domínio (HJØRLAND, 2017) e na Análise de Conteúdo (BARDIN, 2015) como procedimento técnico. Os resultados deste estudo fornecem dados relevantes sobre o conhecimento circense no Brasil desde meados do século XIX até o final do século XX.

2 O Circo no Brasil e a Produção do Conhecimento

Os primeiros registros da presença circense no Brasil datam do início do século XIX, mais especificamente no Rio de Janeiro no ano de 1818, a partir de uma nota na Gazeta do Rio de Janeiro sobre a chegada da Companhia Inglesa de Cavalinhos dirigida por Guilherme Southby, que se apresentou em uma arena no Campo de Santana (atual Praça da República) construída em função da aclamação de D. João VI e do casamento de D. Pedro I (LOPES; SILVA, 2022), como observa-se na Figura 1.

Figura 1. Arena em que Guilherme Southby realizou suas apresentações em 1818, retratada por Franz Josef Fruhbeck



Junto com a presença do circo no Brasil começa a se produzir também informações sobre esses grupos nos jornais (LOPES; SILVA, 2022), na pintura (MARAMBIO, 2022), na literatura (SILVA, 1996, 2009) e nas universidades (ROCHA, 2010). A junção dessas informações, presentes tanto na oralidade dos artistas circenses quanto em qualquer outro tipo de suporte, configura o universo de conhecimento do domínio Circo.

É neste ponto onde a Biblioteconomia se aproxima das Artes e, mais especificamente do Circo, ao ser capaz de investigar esse domínio a partir da Organização do Conhecimento. Uma disciplina ou campo de pesquisa dentro da Ciência da Informação que atenta para o modo como o conhecimento de determinado domínio é organizado, e fornece ferramentas capazes de “descrever, representar, arquivar e organizar documentos, assuntos e conceitos tanto por humanos quanto por programas de computador” (HJØRLAND, 2016, p. 475, *tradução nossa*), e que tem como um dos principais objetivos a Recuperação da Informação (RI).

3 A Organização do Conhecimento Circense no Brasil

No Brasil, iniciativas de sistematização e agrupamento dos conhecimentos produzidos pelos e sobre os circenses se fazem notar desde o início do século XIX. Como é o caso do *Pequeno Tratado de Acrobacia e Gymnástica* de 1933 (LOPES, 2019), *Monografia de Artes Circenses* produzido entre 1976 e 1977 (LOPES, 2019), o site *Pindorama Circus: notícias do circo brasileiro* em 2002, renomeado *Circonteúdo: o portal da diversidade circense* a partir de 2009 (CIRCONTEÚDO, c2022), *Encircopédia: dicionário crítico ilustrado do circo no Brasil* (MAVRUDIS, 2016), o site *Circodata: dicionário do circo brasileiro* em 2017 (CIRCODATA, c2017), *Minidicionário de Teatro e Circo* (MICHEL, 2018) e, mais recentemente, o *IBRARTES: indexador brasileiro das artes* em 2023, um repositório digital voltado para o agrupamento da produção científica brasileira relacionada às áreas de música, dança, circo, teatro, artes visuais, memória e artes integradas (FUNDAÇÃO NACIONAL DE ARTES, 2020). Além de uma série de estudos acadêmicos sobre diversos aspectos do circo em diferentes áreas do conhecimento (ROCHA, 2010; NEUMANN; TERTULINO, 2021).

Contudo, esse conhecimento circense se encontra disperso em diferentes suportes e fontes, de forma não relacionada e fazendo uso de diferentes terminologias. O que já se apresenta como um obstáculo nas pesquisas sobre essa temática, como identificado pelos pesquisadores ao dizerem que:

[...] um dos problemas mais relevantes observados na atualidade, é a falta de uma linguagem e também de uma terminologia comum (estandar, padronizada) neste âmbito artístico, fato que complica severamente o intercâmbio de informação e a sistematização dos conhecimentos, sejam eles patrimônios pessoais de artistas e mestres, ou mesmo de novos estudos realizados do ponto de vista científico (BORTOLIETO; CALÇA, 2007, p. 346).

Cabe refletir um pouco mais sobre essa questão e analisar a necessidade e até mesmo a viabilidade de uma terminologia comum para o circo, no sentido de determinar quais os termos e os conceitos que deverão

ser utilizados, legitimando assim apenas um conjunto de termos. Os tesouros, por exemplo, são ferramentas de controle terminológico capazes de controlar dispersões de natureza lexical, sintática e simbólica (eliminar ambiguidades, controlar sinônimos ou equivalentes), bem como estabelecer relacionamentos semânticos entre os conceitos (CARLAN; BRÄSCHER, 2015; DODEBEI, 2014). Reconhecendo assim, a legitimidade das diferentes terminologias presentes nos diversos contextos do circo, agregando e relacionando-as para que diferentes sujeitos, utilizando diferentes termos, possam encontrar a mesma informação.

A construção de uma ferramenta como o tesouro poderá contribuir para diminuir as dificuldades encontradas atualmente na pesquisa em circo e também como uma proposta de sistematização do conhecimento circense. A partir de um estudo que abarque tanto o conhecimento registrado em qualquer tipo de suporte quanto o conhecimento presente na oralidade e na história de vida dos circenses. Pois, devido à incipiência de publicações na área, a experiência de vida dos circenses e suas informações biográficas se caracterizam como fontes de informação importantes para mapear e identificar o universo de conhecimento do circo (BORTOLETO, 2008; BORTOLETO; CALÇA, 2007; ROCHA, 2010).

4 A Importância dos Estudos Biográficos para o Domínio do Circo e o Circodata.

Além da itinerância, uma outra característica muito ressaltada da comunidade circense é o uso da oralidade para transmissão dos seus saberes e práticas (SILVA, 1996; SILVA; ABREU, 2009). No entanto, por não ser registrado, esses saberes e práticas estão suscetíveis a perdas parciais ou totais. Por isso, registrar através de um meio que favoreça a durabilidade de tais conhecimentos para gerações futuras e que possibilite a recuperação dessas informações se faz necessário.

De acordo com Baptista (2015, p. 21) o registro textual possui relevância tanto como meio de organização e representação da informação quanto “[...] como portador de significados úteis à disseminação de conteúdos e à construção do conhecimento”. O pensamento da autora se reflete nos diários, livros de memórias e cartas produzidos pelos circenses que sobreviveram à passagem do tempo e, atualmente, são fontes de conhecimento. Contudo, essa ainda é uma prática pouco frequente no meio circense e o que se observa na atualidade, além da crescente produção de estudos acadêmicos, são esforços para registrar a existência e os conhecimentos desse grupo através de textos biográficos, um dos gêneros textuais mais comuns no domínio da arte (ØROM, 2003).

Partindo do princípio de que os métodos biográficos consideram elementos da comunicação oral como “[...] meio para a compreensão de fenômenos e situações a

partir da narração de uma trama de acontecimentos vivenciados no passado da pessoa pesquisada” (SANTOS; DAVEL, 2021, p. 431), optou-se por analisar as biografias dos artistas presentes no *site* Circodata: dicionário do circo brasileiro como “ponto de partida para descoberta, geração e disseminação do conhecimento” (BAPTISTA, 2015, p. 22).

O Circodata: dicionário do circo brasileiro foi idealizado pela pesquisadora Cristina Bend e lançado em 2017, o *site* agrupa informações sobre o circo brasileiro de meados do século XIX até o final do XX (CIRCODATA, c2017), objetiva contribuir com a bibliografia existente sobre a temática e tem como público alvo artistas, pesquisadores, professores, alunos e demais interessados sobre o circo brasileiro. Em sua estrutura, possui seções dedicadas aos circos, artistas, grupos e famílias circenses. Tais informações foram coletadas através de entrevistas, bem como de outras fontes, e estão organizadas em forma de verbetes que buscam responder questões relacionadas à trajetória pessoal, profissional e sobre a produção artística do biografado.

Desse modo, o presente estudo objetiva identificar e realizar uma distribuição temporal de termos relacionados ao domínio do circo presentes nessas biografias, contribuindo assim, para uma melhor compreensão dos aspectos culturais, sociais, costumes, tradições e muitos outros fatores refletidos na experiência de vida dos indivíduos desse grupo.

5 Metodologia

O presente estudo é de natureza aplicada, uma vez que busca contribuir para soluções de problemas concretos (CERVO; BERVIAN; SILVA, 2007), e utiliza-se de uma abordagem quantitativa, que se caracteriza pela quantificação tanto na coleta quanto no tratamento dos dados, buscando por padrões ou relações entre os mesmos (MARCONI; LAKATOS, 2022; MATIAS-PEREIRA, 2019). Em relação ao objetivo, caracteriza-se como exploratória, por investigar a presença de variáveis em um contexto determinado (KOCHE, 2015). Como procedimento técnico faz uso de abordagens da Análise de Domínio (HJØRLAND, 2017) e da Análise de Conteúdo (BARDIN, 2015).

Como objeto de estudo foi selecionado o *site* Circodata: dicionário do circo brasileiro, por este se configurar como uma fonte especializada no circo brasileiro, e por conter uma quantidade significativa de informações sobre os saberes e práticas circenses, identificadas durante as etapas de pré-análise, definição do *corpus* e leitura flutuante (BARDIN, 2015). Em relação ao *Corpus*, este se constitui de 1.180 verbetes presentes na aba Artistas do *site* no momento desta pesquisa.

Um segundo momento foi definir as abordagens da Análise de Domínio mais adequada para o estudo dessas

informações, sendo estas, *Estudos Bibliométricos e Estudos Históricos* (HJØRLAND, 2017). As quais, permitiriam realizar a distribuição espaço temporal de termos relacionados ao circo presentes nas biografias, a fim de identificar padrões e relações entre esses termos em determinado período de tempo.

Em seguida, foram coletados os dados referentes ao nome do artista, data de nascimento e morte, período de atuação ou qualquer outra data disponível na fonte, bem como atividade desenvolvida pelo artista. A coleta realizou-se de forma manual, entre 11 de dezembro de 2022 e 2 de março de 2023. Esses dados foram adicionados a planilhas no software *Excel* e organizadas em ordem alfabética de acordo com o nome dos biografados. Na fase de referência dos documentos (BARDIN, 2015) utilizou-se como código a primeira letra do nome do biografado precedida do número que representa a posição do mesmo na planilha, por exemplo, para Abdizio Silva o código é A1, ou seja, este é o primeiro registro da planilha A. Como unidade de registro (BARDIN, 2015) foi selecionado o termo (atividade desenvolvida pelo biografado), a este termo foi adicionado o código de referência para facilitar a recuperação da unidade de contexto (a frase que identifica a atividade desenvolvida pelo artista e que consta na biografia), como a seguir, Aramista (A1). Esta etapa é a que Bardin (2015) denomina de Codificação.

Por fim, todos os termos foram adicionados a uma única planilha e agrupados por década, a fim de verificar a dispersão dos mesmos ao longo do tempo em uma representação visual em formato de linha do tempo.

6 Resultados e Discussão

A análise das 1.180 biografias presentes no Circodata resultou em um total de 555 termos, relacionados a diferentes contextos do domínio circo. Possibilitando identificar termos do universo conceitual dessa área (Esquema 1., Apêndice B)

A partir das biografias foi possível construir a linha do tempo iniciando na década de 1820 até o ano 2000. Permitindo mensurar o desenvolvimento/variação de interesses e práticas do circo no Brasil (SOUZA, 1996), sendo o período entre 1910 e 1960 o de maior diversidade de termos (Figura 2., Apêndice A).

Observa-se também o uso de termos em outras línguas, como por exemplo, Partner, Clown, Doble Trapézio, Double Trapézio e Duo Acrobático, bem como referência a outras culturas em alguns dos termos, como ocorre com Corda Indiana, Balança Russa e Argola Chinesa. O que denota características singulares na terminologia da área, Silva (1996, p. 94) esclarece que “da fusão dos vários grupos étnicos, o circo acaba por criar uma linguagem e um vocabulário próprio”. Resultando em palavras herdadas de diferentes dialetos

e línguas como o inglês, francês, italiano, romeno, de dialetos dos ciganos, entre outros.

Um outro aspecto percebido nos dados é a natureza híbrida da construção do espetáculo ao verificar algumas funções dos biografados, como por exemplo, Músico, Musicista, Ator, Bailarina, Bailarino, Cantor, Cantora, Instrumentista, Jôquei, Ciclista, Ginasta e Dupla Caipira, caracterizando o circo como um espaço que promove:

[...] circularidades culturais entre expressões culturais populares, eruditas e de massa, e entre manifestações culturais distintas, como cinema, teatro, dança, ópera, esporte etc. (ROCHA, 2010, p. 56)

Contudo, mesmo que atualmente o circo venha ganhando espaço e sendo legitimado como arte, não é raro que ainda seja apenas enquadrado no aspecto da cultura popular. Reproduzindo assim, o discurso que o coloca como contraponto à cultura erudita e nega o seu caráter de hibridação (CANCLINI, 1998), nas suas diversas mesclas interculturais dentro do processo histórico e social.

Portanto, é imprescindível que qualquer iniciativa de organização do conhecimento circense considere a hibridação dos saberes e práticas envolvidos nos diversos contextos do circo, dentre eles, as modalidades, as técnicas, as funções desenvolvidas, expressões, instrumentos, entre outros, evitando refletir preconceitos, vieses e tendenciosidades (BORTOLETO, 2008; MILANI; GUIMARÃES, 2014).

7 Conclusão

O presente estudo alcançou o seu objetivo ao identificar termos referentes ao domínio circo presentes nas biografias dos circenses, bem como mensurar o desenvolvimento/variação de interesses e práticas do circo no Brasil.

O texto biográfico se mostrou significativo no mapeamento do universo conceitual do circo, refletindo o que vem sendo discutido na literatura especializada da área, inclusive, no que se refere a falta de uma terminologia padronizada nesse domínio.

Desta forma, identificar a terminologia utilizada no circo é um primeiro passo para definir os relacionamentos hierárquicos, semânticos, controlar a sinonímia entre os termos e, consequentemente, desenvolver linguagens documentárias para esse domínio do conhecimento. E, também, para investigar sobre o que se tem escrito ou deixado de escrever na literatura especializada.

Cabe destacar a dificuldade encontrada em identificar de forma específica alguns termos relacionados às modalidades circenses, uma vez que estes se apresentavam nas biografias de forma mais geral,

Trapézio por exemplo, não sendo possível identificar a que tipo de Trapézio (fixo, de voos, em balanço, de um ponto etc.) o texto se referia.

Como sugestão para pesquisas futuras, recomenda-se um estudo mais aprofundado dos conceitos relacionados a cada termo, a fim de identificar de forma mais científica suas semelhanças e diferenças.

Referências

- BAPTISTA, D. M. A Relevância do Texto na Organização e Representação da Informação. *In*: BAPTISTA, D. M.; ARAÚJO JÚNIOR, H. (org.). **Organização da Informação**: abordagens e práticas. Brasília: Thesaurus, 2015. p. 21-158.
- BARDIN, L. **Análise de Conteúdo**. São Paulo: Edições 70, 2015.
- BORTOLETO, M. A. C. (org.). **Introdução à pedagogia das atividades circense**. Jundiaí: Fontoura, 2008.
- BORTOLETO, M. A. C.; CALÇA, D. H. Circo e Educação Física: compendium das modalidades aéreas. **Movimento & Percepção**, Espírito Santo do Pinhal, v. 8, n. 11, p. 345-360, jul./dez. 2007.
- CANCLINI, N. G. **Culturas Híbridas**: estratégias para entrar e sair da modernidade. 2. ed. São Paulo: Edusp, 1998.
- CARLAN, E.; BRÄSCHER, M. Fundamentos Teóricos para Elaboração de Sistemas de Organização do Conhecimento. *In*: BAPTISTA, D. M.; ARAÚJO JÚNIOR, H. (org.). **Organização da Informação**: abordagens e práticas. Brasília: Thesaurus, 2015. p. 135-158.
- CERVO, A. L.; BERVIAN, P. A.; SILVA, R. Técnicas de Coleta de Dados. *In*: CERVO, A. L.; BERVIAN, P. A.; SILVA, R. **Metodologia Científica**. 6. ed. São Paulo: Pearson Prentice Hall, 2007. p. 57-70. *E-book*.
- CIRCODATA: dicionário do circo brasileiro. [S.l.], c2017. *Site*. Disponível em: <http://circodata.com.br/index.php?c=inicio&>. Acesso em: 09 mar. 2023.
- CIRCONTEUDO. O portal da diversidade circense. [S.l.], c2022. *Site*. Disponível em: <https://www.circonteudo.com>. Acesso em: 22 mar. 2023.
- DODEBEI, V. L. D. **Tesauro**: linguagem de representação da memória documentária. Niteroi: Intertexto, 2014.
- FUNDAÇÃO NACIONAL DE ARTES. **Ibrartes**: indexador brasileiro das artes. [S.l.], 2020. *Site*. Disponível em: <https://sistema.funarte.gov.br/tainacan/colecoes/ibrartes/>. Acesso em: 22 mar. 2023.
- HJØRLAND, B. Knowledge organization. **Knowledge Organization**, *Würzburg*, v. 43, n. 6, p. 475-484, 2016.
- HJØRLAND, B. Domain Analysis. **Knowledge Organization**, *Würzburg*, v. 44, n. 6, p. 436-464, 2017.
- KOCHE, J. C. O Fluxograma da Pesquisa Científica. *In*: KOCHE, J. C. **Fundamentos de Metodologia Científica**: teoria da ciência e iniciação à pesquisa. 34. ed. Petrópolis: Vozes, 2015. p. 121-136.
- LOPES, D. C. **Os circenses e seus Saberes sobre o Corpo, suas Artes e sua Educação**: encontros e desencontros históricos entre circo e ginástica. 2019. Tese (Doutorado em Educação) - Faculdade de Educação, Universidade de São Paulo, São Paulo, 2019.
- LOPES, D.; SILVA, E. **Um Brasil de Circos**: a produção da linguagem circense do século XIX aos anos 1930. Campinas: Circonteúdo/Prêmio Funarte de Estímulo ao Circo (2019), 2022.
- MARAMBIO, R. E. M. **O Circo e a Memória do Circense**: narrativas a partir da História da Arte. 2022. Dissertação (Mestrado em História da Arte) - Escola de Filosofia, Letras e Ciências Humanas, Universidade Federal de São Paulo, Guarulhos, 2022.
- MARCONI, M. A.; LAKATOS, E. M. Metodologia Qualitativa e quantitativa. *In*: MARCONI, M. A.; LAKATOS, E. M. **Metodologia Científica**. 8.ed. São Paulo: Atlas, 2022. p. 295-347. *E-book*.
- MATIAS-PEREIRA, J. **Manual de Metodologia da Pesquisa Científica**. 4.ed. São Paulo: Atlas, 2019. p. 19-40. *E-book*.
- MAVRUDIS, S. K. **Encicropédia**: dicionário crítico ilustrado do circo no Brasil. 2. ed. Belo Horizonte: Mútua Comunicação, 2016.
- MICHEL, V. C. **Minidicionário de Teatro e Circo**. Porto Alegre: AGE, 2018.
- MILANI, S. O.; GUIMARÃES, J. A. C. Problemas Relacionados a Biases em Sistemas de Organização do Conhecimento: perspectivas para a representação de assunto. **IRIS – Revista de Informação, Memória e Tecnologia**, Recife, v. 3, n. Especial, p.72-92, 2014.
- NEUMANN, S. E.; TERTULINO, C. I. Análise bibliométrica do site Circonteúdo: informação como subsídio para a formação circense. **BiblioCanto, Natal**, v. 7, n. 1, p. 62–84, 2021. DOI: 10.21680/2447-7842.2021v7n1ID25995.
- ØROM, A. Knowledge Organization in the domain of Art Studies: history, transition and conceptual changes. **Knowledge Organization**, *Würzburg*, v.30, n. 3-4, p. 128-143. 2003.
- ROCHA, G. O Circo no Brasil: estado da arte. **Revista Brasileira de Informação bibliográfica em Ciências Sociais - BIB**, São Paulo, n. 70, 2 semestre de 2010. p. 51-70.
- SANTOS, F. P.; DAVEL, E. P. B. Métodos Biográficos para a Pesquisa em Administração: princípios, potencialidades, práticas e desafios. **Revista Eletrônica de Administração**, Porto Alegre, v. 27, n. 2, p. 430-461, maio/ago. 2021.
- SILVA, E. **O circo** - sua arte e seus saberes: o circo no Brasil do final do século XIX a meados do XX. 1996. Dissertação (Mestrado em História) - Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 1996.
- SILVA, E.; ABREU, L. A. **Respeitável público... O circo em cena**. Rio de Janeiro: Funarte, 2009.
- SOUZA, R F. Padrões de Comunicação em Ciência: O caso da física da matéria condensada no Brasil, no período de 1950 a 1980. **Informare – Cadernos do Programa de Pós-Graduação em Ciência da Informação**, Rio de Janeiro, v.2, n.1, p. 14-24, jan./jun. 1996.

Apêndice B – Esquema 1. Ordem Alfabética Sistemática dos Termos

CIRCO

Acrobacia

Acrobacia Coletiva
Adágio Acrobático
Bandeira
Duo Acrobático
Dupla Acrobática (ver
Duo acrobático)
Estatuário
Pirâmide
 Antipodismo
Icários
 Contorcionismo (*ver* Con-
 torção)
 Balança Russa
 Barra (*ver* Barra Fixa)
 Barra Fixa
 Báscula
 Cama Elástica
 Contorção
Contorção da Boneca
Homem Sapo

Dândis

Deslocação (*ver* Contor-
 ção)

Despenhadeiro

Solo (saltos)

Charivari

Trampolim

Acrobacia Aérea

Argolas

Bambu

Corda

Corda Indiana

Corda Marinha

Escada Giratória

Faixa

Giro da Morte

Giro de Pé

Lira

Passeio Aéreo

Quadrante

Rede

Tecido

Trapézio

Doble (ver Doble Tra-
 pézio)

Doble Trapézio

Double Trapézio (ver

Doble Trapézio)

Trapézio de Cabeça

Trapézio de Voos

Voos (ver Trapézio de

Voos)

Trapézio Doble (ver

Doble Trapézio)

Trapézio Duplo (ver

Doble Trapézio)

Trapézio em Balanço

Trapézio Fixo

Trapézio Simples (ver

Trapézio)

Turbilhão

Adestramento (*ver* Doma)

Amestramento (*ver* Doma)

Aqualoco

Arquitetura

Arquitetura Nômade

Circo

circo de lona

circo de pau fincado

coberto de pano

Pavilhão

Picadeiros Simultâneos

Balé Aéreo (*ver* **Acrobacia**

Aérea)

Bordado

Capatazia

Circo

Circo Itinerante

Circo de Tourada

Circo Família

Circos de Tiro

Circos Mambembes

Comicidade

Comediante

Comediante Físico

Cômico(a)

Cômico de Entradas

Cômico Musical

Dupla Cômica

Excêntrico

Humorista

Humorista Caipira

Humorista Circense

Imitação

Palhaçaria

Clonisa

Clown

Clown dos Palha-

ços

Táxi Maluco

Toni

Dança

Bailados

Dança com Boleadeiras

Dança com Cobras

Declamação

Doma

de Cachorros

de Cães (*ver* Cachorros)

de Camelos

de Cavalos

de Chimpanzés

de Elefantes

de Feras

de Hipopótamo

de Leões

de Lhamas

de Macacos

de Pôneis

de Ursos

Equilibrismo

Antipodismo

Tranca

Arame

Arame Alto

Arame Bambo

Arame com Maromba

Arame Esticado

Bola

Ciclismo

Bicicleta

Cesto

Equilíbrio de Cadeira

Motociclismo

Globo da Morte

Equilíbrio de Objetos

Equilíbrio de Garrafas

Equilíbrio de Pratos

Equilíbrio de Xícaras

Escada

Escada de Garrafa

Escada de Livre

Escada de Ombro

Escada de Pé

Escada Pendurada

Escada Rolante

Escada Sete

Escada Simples

Escada Três

Homem Foca

Monociclo

Parada

Parada de Cabeça

Parada de Cadeira

Parada de Mãos

Patinação

Patinação sobre Ta-
blado

Pêndulo

Percha

<i>Percha de Testa</i>	barrista	de hipopótamo
Perna de Pau	<i>adestrador(a)</i>	de leões
Pratos Bailables (<i>ver</i>	(<i>ver domador</i>)	de lhamas
Equilíbrio de Objetos)	de animais	de pôneis
Rola (<i>ver</i> Rola-Rola)	de cachorros	de ursos
Rola-Rola	de cavalos	<i>encantador de</i>
Rola Rola (<i>ver</i> Rola-Rola)	de leões	<i>serpentes</i>
Trono	de lhamas	<i>engolidor de es-</i>
Equipamentos	de macacos	<i>padas</i>
Expressões	de pôneis	<i>equestre</i>
Cachê	<i>amestrador(a)</i>	<i>equilibrista</i>
Fazer Cachê	(<i>ver domador</i>)	de escada
Número	de cachorros	de garrafas
Picadeiro	de cães (<i>ver</i>	de objetos
Portô	cachorros)	de pêndulo
Rodar no Globo	de camelos	de pratos
Volante	de cavalos	de pratos gi-
Fonte Luminosa	de chimpan-	ratórios
Funções Organizacionais	zês	de rola-rola
Funções Administrativas	de elefantes	<i>equitador</i>
<i>Administrador(a)</i>	de leões	<i>evasionista</i>
<i>Capataz</i>	de macacos	<i>excêntrico(a)</i>
<i>Chefe da Contrarre-</i>	de pôneis	excêntrico
<i>gragem</i>	<i>antipodista</i>	musical
<i>Chefe de Transporte</i>	<i>aramista</i>	<i>faquir</i>
<i>Chefe de Vendas</i>	<i>artista aéreo(a)</i>	<i>faquiresa</i>
<i>Diretora de Marketing</i>	argolista	ginasta
<i>Empresário</i>	trapezista	globista
<i>de circo</i>	<i>trapezista</i>	<i>homem bala</i>
<i>circense</i>	<i>volante</i>	<i>homem dos sete</i>
<i>de parque</i>	<i>artista de circo</i>	<i>instrumentos</i>
<i>de parque de diver-</i>	(<i>ver Artista Cir-</i>	(<i>ver excêntrico</i>
<i>são</i>	<i>cense</i>)	musical)
<i>Encarregado</i>	<i>artista genérico</i>	<i>homem foca</i>
<i>de compras</i>	<i>atiradora</i>	<i>homem sapo</i>
<i>de marcenaria</i>	<i>atirador(a) de</i>	(<i>ver contorcio-</i>
<i>Gerente</i>	<i>facas</i>	nista)
<i>Produtor(a)</i>	<i>caubói</i>	<i>icarista</i>
<i>Proprietário</i>	<i>ciclista</i>	<i>ilusionista</i>
<i>de circo</i>	acrobático	<i>imitador</i>
<i>de pavilhão</i>	de cesto	<i>instrumentista</i>
<i>de parque de diver-</i>	de bicicleta	<i>jejuador</i>
<i>são</i>	<i>clonisa</i>	<i>jóquei</i>
<i>Secretário</i>	<i>clown</i>	<i>lutadora</i>
<i>de circo</i>	dos palhaços	<i>mágico(a)</i>
<i>de frente</i>	<i>comediante</i>	ilusionista
Funções Artísticas	comediante	<i>malabarista</i>
<i>Adaptador(a)</i>	físico	<i>malabarista</i>
<i>de comédias</i>	<i>cômico(a)</i>	<i>de objetos</i>
<i>de dramas</i>	<i>contorcionista</i>	<i>manipulador</i>
<i>de peças estrangei-</i>	<i>deslocador(a)</i>	de argola
<i>ras</i>	(<i>ver contorcio-</i>	chinesa
<i>de peças nacionais</i>	nista)	de baldes
<i>Aderecista</i>	<i>domador(a)</i>	de baralho
<i>Apontador de Dramas</i>	de cachorros	de bolas
<i>Apresentador(a)</i>	de cavalos	de cachim-
<i>Artista Circense</i>	de elefantes	bos
<i>acrobata</i>	de feras	

de cartas	<i>de teatro</i>	<i>Iluminador</i>
de cartas de baralho	<i>Dramaturgo</i>	<i>Montador</i>
de cigarros	<i>Ensaaiador(a)</i>	<i>de trapézio</i>
de dedais	<i>Escritor</i>	<i>de globo</i>
de moedas	<i>roteirista</i>	<i>de mesa</i>
de objetos	<i>Figurinista</i>	<i>de monociclo</i>
de pombos	<i>Instrumentista</i>	<i>Motorista</i>
de velas	<i>sax</i>	<i>Pintor</i>
<i>mímico</i>	<i>piano</i>	<i>de cartazes</i>
<i>monociclista</i>	<i>guitarra</i>	<i>Responsável</i>
<i>multiartista</i>	<i>Instrutor</i>	<i>pelo guarda-roupa</i>
<i>palhaço(a)</i>	<i>circense</i>	<i>por estábulos</i>
excêntrico	<i>de oficinas</i>	<i>por jaulas</i>
musical	<i>de circo (ver Instrutor circense)</i>	<i>Sonoplasta</i>
palhaço musical	<i>Locutor(a)</i>	<i>Técnico</i>
palhacinho	<i>Mestre de Pista</i>	<i>de montagem</i>
<i>paradista</i>	<i>Multinstrumentista</i>	<i>de som</i>
paradista de cabeça	<i>Musicista</i>	<i>de teatro</i>
<i>patinador(a)</i>	<i>Músico</i>	<i>Tratador de Animais</i>
<i>perchista</i>	<i>clarinetista</i>	<i>Vendedor</i>
<i>pirofagista</i>	<i>de orquestra</i>	<i>baleiro</i>
<i>saltador(a)</i>	<i>pandeirista</i>	<i>vendedor de doces</i>
<i>toureiro(a)</i>	<i>saxofonista</i>	<i>vendedor de pipoca</i>
<i>ventriloquo(a)</i>	<i>violinista</i>	Ilusionismo
<i>volteador(a)</i>	<i>Partner</i>	Evasionismo
<i>Ator</i>	<i>Professor</i>	Levitação
<i>cômico</i>	<i>de acrobacia</i>	Magia (ver Ilusionismo)
<i>de circo-teatro</i>	<i>de aéreos</i>	Mágica
<i>de teatro</i>	<i>de artes circenses</i>	<i>Mala Moscovita</i>
<i>de televisão</i>	<i>de circo</i>	<i>Caixa de Espadas</i>
<i>dramático</i>	<i>de comicidade</i>	Manipulação de Objetos
<i>figurante</i>	<i>de escolas de circo</i>	Manipulação com
<i>galã</i>	<i>no circo</i>	<i>Argolas Chinesas</i>
<i>Atriz</i>	Funções Técnicas (operações)	<i>Baldes</i>
<i>cômica</i>	<i>Ajudante</i>	<i>Baralhos</i>
<i>de circo-teatro</i>	<i>de animais</i>	<i>Bolas</i>
<i>de musicais</i>	<i>de tratador de camelos</i>	<i>Cachimbos</i>
<i>dramática</i>	<i>de tratador de cavalos</i>	<i>Cartas</i>
<i>caricata</i>	<i>de tratador de ursos</i>	<i>Cartas De Baralho</i>
<i>Autor</i>	<i>geral de montagem</i>	<i>Cigarros</i>
<i>Bailarino(a)</i>	<i>Caminhoneiro</i>	<i>Dedais</i>
<i>clássica</i>	<i>Bilheteiro(a)</i>	<i>Moedas</i>
<i>sapateadora</i>	<i>Cenotécnico</i>	<i>Velas</i>
<i>rumbeira</i>	<i>Cuidador</i>	Malabares
<i>Camareira</i>	<i>de animais</i>	<i>com Clavas</i>
<i>Cantor(a)</i>	<i>de elefantes</i>	<i>com Bolas de Ferro</i>
<i>cançonetista</i>	<i>Eletricista</i>	<i>de Contato</i>
<i>Cenógrafo</i>	<i>Engenheiro</i>	Música
<i>Compositor(a)</i>	<i>Fabricante</i>	Duplas Caipiras
<i>Coreógrafo(a)</i>	<i>de lona</i>	Trios Caipiras
<i>Dançarino(a)</i>	<i>de aparelhos de magia</i>	Instrumentos
<i>Declamadora</i>	<i>Fotógrafo</i>	Números com Animais
<i>Diretor(a)</i>		Cachorro Calculista
<i>artístico(a)</i>		com Cachorrinhos
<i>de espetáculos</i>		com Cavalos Adestrados
		com Chimpanzês
		com Cobras
		com Elefantes

Encantador de Serpentes
 Equitação
 com Pombos
 com Tigres
 Volteio Equestre
Volteio (ver Volteio Equestre)
Número de Força Física
 Equilíbrio de Objetos sobre a testa
 Lutas
Número de Resistência Física
 Canhão Humano (ver Homem-Bala)
 Faquirismo

Espadas
Estiletas
Facas
Objetos Cortantes
cacos de vidro
cama de prego
 Força Capilar
 Força Dental
 Homem-bala
 Número de Canhão (ver Homem-Bala)
Números de Variedades
Pirofagia
Teatro
 Circo-Teatro
 Peças de Teatro

Comédias
pastelão
entradas cômicas
escada cômica
esquetes cômicos
Dramas
Pantomimas
pantomimas aquáticas
mímica
Reprises
 Teatro de Revista
Telepatia
Ventriloquismo

Espoliação de artefatos africanos: o caso dos Bronzes de Benin e suas representações no British Museum

Dispossession of african artifacts: the case of the Benin Bronzes and their representations at the British Museum

Gláucia Aparecida Vaz

Universidade Federal do Rio Grande do Sul/Universidade Estadual de Londrina, glaucia-vaz@hotmail.com

Resumo

O objetivo deste trabalho é analisar um conjunto de peças dos Bronzes de Benin, no catálogo do British Museum disponibilizado na página do projeto Digital Benin. Os Bronzes de Benin, somam um total de 5.246 peças, que foram roubadas do Benin, atual Nigéria, em uma expedição militar da Inglaterra em 1897. Essa análise busca pontuar como se estrutura a representação da informação em torno de objetos espoliados da cultura africana. A trabalho utiliza da revisão de literatura para levantar a história das peças, bem como compreender a complexa relação colonizador e colonizado nas exposições museológicas. Os resultados apontam a dificuldade de identificação de vários contextos das peças, como produtor/artista e estilo/escola. Aponta, portanto para a necessidade das instituições desenvolverem outro olhar, que não apenas o eurocêntrico no processo de construção de catálogos.

Palavras-chave: Bronzes de Benin, Digital Benin, British Museum, representação da informação.

Abstract

The objective of this paper is to analyze a set of pieces from the Benin Bronzes, in the British Museum catalog available on the page of the Digital Benin project. The Benin Bronzes, add up to a total of 5,246 pieces, which were stolen from Benin, present-day Nigeria, in a military expedition from England in 1897. This analysis seeks to point out how the representation of information is structured around objects plundered from African culture. The work uses the literature review to raise the history of the pieces, as well as understand the complex relationship colonizer and colonized in museum exhibitions. The results point to the difficulty of identifying the various contexts of the pieces, such as producer/artist and style/school. It points, therefore, to the need for institutions to develop another look, which is not just the Eurocentric one in the process of building catalogs.

Keywords: Benin Bronze, Digital Benin, British Museum, Information Representation.

Resumen

El objetivo de este trabajo es analizar un conjunto de piezas de los Bronces de Benin, en el catálogo del Museo Británico disponible en la página del proyecto Digital Benin. Los Bronces de Benín, suman un total de 5.246 piezas, que fueron sustraídas en Benín, actual Nigeria, en una expedición militar procedente de Inglaterra en 1897. Este análisis busca señalar cómo se estructura la representación de la información en torno a los objetos espoliados de cultura africana. El trabajo utiliza la revisión bibliográfica para plantear la historia de las piezas, así como comprender la compleja relación colonizador y colonizado en las exposiciones museísticas. Los resultados apuntan a la dificultad de identificar los diversos contextos de las piezas, como productor/artista y estilo/escuela. Señala, por tanto, la necesidad de que las instituciones desarrollen otra mirada, que no sea solo la eurocéntrica en el proceso de construcción de catálogos.

Palabras clave: Bronces de Benin, Digital Benin, British Museum, representación de la información.

1 Introdução

O estudo das formas simbólicas que circulam e que reproduzem as relações de dominação e os discursos que surgem nas artes e na mídia, sustentam as relações de desigualdade racial e a manutenção do eurocentrismo.

No Brasil a arte desenvolvida, ligada às instituições europeias, herdou parte dessa forma de representação, pois apesar das imagens que confirmavam a suposta superioridade racial da população branca, associando

negros/as a situações degradantes, outros artistas, inclusive europeus, produziram imagens que questionavam a escravidão e valorizavam os elementos da cultura de matriz africana que sobreviviam e alteravam as características da sociedade brasileira em formação.

Menezes e Alvarez (2019), trazem em seu trabalho uma discussão em torno do termo "primitivo", que evoca, em nosso imaginário, vários sentidos. Nesse caso, em particular, as autoras apontam as conotações raciais e colonialistas quando se fala em "culturas

primitivas", 'já que essa expressão foi empregada para descrever aquelas culturas que, do ponto de vista ocidental, careciam de desenvolvimento econômico e social, de elaboração estética e mental. De fato, esse termo foi usado por acadêmicos ocidentais para falar sobre todas as culturas "exóticas" que estavam em contato com colonizadores e exploradores europeus.' (MENEZES; ALVAREZ, 2019)

O pesquisador Paul Wood, estudioso dos bronzes de Benin, destaca uma afirmação do ano de 1857, do estudioso John Ruskin, onde afirma que não havia "arte antiga pura e preciosa" na Ásia, África ou América, centralizando assim o termo "arte" na tradição europeia. Isso não quer dizer que Ruskin pensasse que as civilizações islâmica, chinesa e indiana careciam de arte como tal (na verdade, ele observa em outra parte da mesma palestra que "quase todas as grandes e intelectuais raças do mundo" produziram uma arte específica). Mas o coloca em uma tradição que depreciava sistematicamente os produtos de tais culturas em relação à Europa; e, de fato, no caso de outros lugares como Oceania, América do Norte e África, ele realmente os considerava desprovidos de 'arte'. (WOOD, 2019, p.3, tradução nossa)

Menezes e Álvarez (2019) recorrem ao termo 'matriz de classificação social', crido pelo sociólogo peruano Anibal Quijano, baseada na ideia de divisão da humanidade em grupos diversos, no que tange a biologia e o intelecto: as raças humanas (Arthur de Gobineau etcetera). A ideia de raça nasce, com a chegada dos europeus à América e com o tráfico de escravos para o continente. A partir desse evento, as sociedades europeias tornam-se obcecadas pela ideia de classificar os grupos humanos, com o objetivo de explicar e relacionar diferenças culturais a diferenças biológicas. (MENEZES; ALVAREZ, 2019)

Dessa forma, para reconstituir a identidade do colonizado e desalienar seu corpo, conferindo-lhes de fato, humanidade. A resposta a essa digressão está na reconversão do corpo objetivado (e sua extensão, a arte) em corpo-sujeito. Em outras palavras, o corpo dos "primitivos", transformado em objeto de arte para satisfazer a curiosidade europeia e sua vontade de dominação, o corpo-objeto, deve ser convertido em corpo sujeito. A restituição de obras de arte para as sociedades que as exigem é um caminho importante nessa conversão, pois permite que os povos que as produziram possam contar e contextualizar o real significado de suas produções.

Sendo o processo de representação da informação, também um fenômeno social, todo esse processo de colonização, a classificação por raças, criadas pelo próprio colonizador e imagem de inferioridade construída em torno dos povos colonizados, interferem nos processos de catalogação e indexação. De acordo com Menezes e Alvarez:

Assistimos, assim, à criação do mito fundador da modernidade: "a ideia de um estado originário de natureza na história das espécies e uma escala de desenvolvimento histórico que parte do primitivo (...) ao mais civilizado (...)". O fenômeno da colonização será, porém, suficientemente longo e profundo para gravar uma divisão fundamental entre corpo (natureza) e mente (cultura), graças ao pensamento cartesiano, responsável pela formação da visão/sentir do mundo das civilizações modernas. Essa divisão será a base de uma operação simbólica de separação entre aqueles que estariam próximos da natureza (primitivos, negros, nativos, indígenas) e aqueles cujo desenvolvimento cultural seria mais elevado (civilizado, branco, europeu). (MENEZES, ALVAREZ, 2019, p. 5)

Observa-se que o processo de classificação das raças, imposta pelos colonizadores, interfere na representação da informação de peças da cultura Africana nos museus.

Dessa forma o objetivo desse estudo é analisar um conjunto de obras da famosa coleção conhecida mundialmente como 'Bronzes de Benin', expostas atualmente no British Museum, por meio de análise de conteúdo, dos catálogos online disponibilizados pelo projeto Digital Benin. A técnica de Análise de conteúdo de Bardin (2004, 2010, 2011). A Análise de Conteúdo objetiva analisar o que foi dito em meio a uma investigação, construindo e apresentando concepções em torno de um objeto de estudo. A análise do material coletado segue um processo rigoroso frente às fases definidas por Bardin (2011), como: Pré-análise; Exploração do material e Tratamento dos resultados.

2 A história da fundição dos Bronzes de Benin

O pesquisador Esse Vivian Odiahi, da Benson Idahosa University, situada na Nigéria, realizou um estudo sobre a história da fundição dos Bronzes de Benin, perpassando desde sua criação, guerra, espoliação e reconstrução.

Odiahi (2017), nos conta que a arte da fundição de bronze floresceu sob o sistema de guilda estabelecido por Oba Oguola em 1280 d.C. A guilda dos fundidores de bronze do reino de Benin desempenhava funções indispensáveis ao reino como o registro dos eventos que aconteciam no reino em bronze retratando assim as memórias e conquistas dos reis anteriores e coisas que aconteceram no passado, serviam também para a ritualística necessidades do reino fundindo objetos de bronze que foram colocados no santuário, objetos de bronze foram usados para embelezar e decorar o palácio. Entre todas as guildas artesanais do reino de Benin, a guilda dos fundidores de bronze ocupa a posição mais alta por dois motivos; serviram como historiadores locais do reino e, em segundo lugar, pela sua capacidade de transformar areia em dinheiro que

fazia com que fosse atribuído grande valor aos seus produtos.

O florescimento da arte da fundição de bronze sob a guilda, entretanto, dependia da estabilidade do reino. A arte continuou a florescer até 1897, quando uma expedição punitiva foi realizada em Benin e a maioria das obras de arte do reino de Benin, incluindo obras de bronze, foram levadas para longe do palácio, o que resultou em um colapso temporário da guilda até 1914, quando uma nova monarquia foi restaurado no reino de Benin com Eweka II, o filho mais velho de Ovoramwen como o Oba que encorajou a guilda de fundidores de bronze a fundir réplicas e substituições para as obras de bronze que foram retirados especialmente do santuário real. A guilda mudou-se formalmente para a rua Igun, onde ainda residem até hoje. (ODAH, tradução nossa)

A guilda dos fundidores de bronze do reino de Benin, a guilda Igun Eromwon é chamada de Brass and Bronze, que é chamada localmente de Oze, é considerada e vista como um tesouro (Eromwon em Bini) e como os membros da guilda dos fundidores de bronze residiam em uma rua chamada Igun, que se tornou sua morada permanente foram retirados especialmente do santuário real. A guilda mudou-se formalmente para a rua Igun (INNEH, 2007), onde ainda residem até hoje.

Sobre o material utilizado nas peças, o pesquisador relata que apesar de eram chamados de Bronzes de Benin, o material utilizado para fundição seria uma mistura de zinco e cobre. Existe, portanto, uma diferença entre um trabalho de latão e um trabalho de bronze após a fundição que é desconhecida para o leigo. No entanto, tanto o latão quanto o bronze ainda estão em uso hoje, mas o latão é usado principalmente porque é mais barato. Este estudo usa ambos os conceitos, “latão e bronze” de forma intercambiável. Oba Oguola desejava introduzir a fundição de latão no Benin para produzir obras de arte semelhantes às que lhe foram enviadas de Ife. Ele, portanto, enviou ao Ooni de Ife um ferreiro de bronze e Igueghae foi enviado a ele.

O latão é uma liga de zinco e cobre, enquanto o bronze compreende em grande parte uma liga de estanho e cobre, mas a quantidade de cobre no bronze é muito maior do que no latão, o bronze e o latão parecem ouro na cor, mas o latão brilha mais que o bronze, o bronze é mais duro que o latão e mais durável devido à grande quantidade de cobre que contém, o bronze é mais caro que o latão. (Entrevista com Kingsley Inneh e Ehigie) Embora a guilda Igun eromwon se refira a si mesma como a guilda dos fundidores de bronze, suas obras de arte, especialmente antes do século 15, eram/são feitas principalmente de latão e não de materiais de bronze. (ODAH, 2017, p. 179, tradução nossa)

A maioria dos estudiosos e historiadores que estudam a história do Benin, incluindo membros da guilda Igun Eromwon, traçam a origem da guilda até 1280, durante o reinado de Oba Oguola, que dizem ter formado a guilda. Esta afirmação é, no entanto, contrária à afirmação do primeiro historiador local do Benin, Jacob U. Egharevba, de que Oba Oguola introduziu a arte da fundição de latão no Benin. No entanto, Akenzua (2011, p. 173) afirmou expressamente que a guilda de fundidores de bronze do reino de Benin nasceu em 1280 d.C. durante o reinado de Oba Oguola, enquanto a introdução da arte de fundição de bronze em Benin é anterior à era de Obaship.

Com o tempo, com muito estudo e investigação acadêmica, a afirmação de Egharevba foi contestada por escritores da história do Benin e estudiosos como Eyo Ita, Eweka Basimi, Erediauwa, etc. não é de origem Ife. A Arte foi introduzida em Benin por um fundidor de bronze nascido em Benin, Igueghae, que já era um fundidor de bronze bem conhecido em Benin antes de ir para Ife para fundir a cabeça de Oranmiyan; um príncipe de origem beninense que foi criado e viveu em Ife porque já estava acostumado com a cultura do povo, é provavelmente com base nisso que Igueghae foi erroneamente considerado um fundidor de bronze de Ife, pois Oranmiyan também foi erroneamente considerado um príncipe de Ife origem. Em segundo lugar, diz-se que a arte da fundição do bronze é um presente do Deus supremo ou superior ao Binis, que é o primeiro fundidor de bronze indígena. Igueghae não aprendeu a arte de lugar nenhum, ele nasceu com ela. Em terceiro lugar, Igueghae não é um nome Yoruba, mas um nome Bini que significa que eu (o rei de Benin) compartilhei com ele (o Ooni de Ife) (INNEH, 1998) Inne Em quarto lugar, de acordo com a tradição oral de Bini, a fundição de bronze em Benin é tão antiga quanto o próprio reino, pois afirma-se que as filhas de Ogiso usavam pulseiras fundidas em bronze. Isso foi apoiado por escavações arqueológicas feitas por Graham Connah, que revelaram objetos de pulseiras feitas de bronze. (ODAH, 2017, p. 180 tradução nossa)

Ele, no entanto, afirmou que esses objetos de pulseira foram feitos de forja de bronze e não de fundição de bronze (CONNAH, 1975, p. 142). Outro arqueólogo, Gerard, refutou a afirmação de Connah de que esses primeiros objetos de bronze foram produzidos por forjamento, eles foram produzidos com a técnica de cera perdida de fundição como um cadinho encontrado durante a escavação de Connah confirmou (INNEH, 2007, p. 105).

A sugestão de Connah de que essas primeiras obras de bronze foram feitas por ferragem surgiu da crença errônea inicial de que a técnica de fundição por cera perdida foi introduzida por não africanos, seguindo um argumento de que a tecnologia envolvida era

sofisticada demais para ser de origem africana. No entanto, na visão de William Fagg (um arqueólogo), nenhuma das primeiras peças de bronze do Benin mostra qualquer evidência de contato com os europeus (OMORUYI, 1981, p. 46).

É sob essa luz que a técnica de fundição de bronze em cera perdida no Benin pode ser considerada única e nativa do reino. “A mais maravilhosa coleção de arte e artesanato em bronze pode ser encontrada em Benin do que em qualquer outro lugar da África. A adesão à guilda é organizada de acordo com as linhas hereditárias da família da guilda, ou seja, todos os membros da guilda são parentes de sangue e com os chefes titulares da guilda, estranhos não são permitidos na guilda, no entanto, há uma exceção para pessoas que tem permissão real para se juntar à guilda. (Entrevista com Igbinosun) Essas pessoas são consideradas membros da família da guilda. Um exemplo é a família Omodamwen, que também é membro da guilda Igun eromwon.

Omodamwen, um Omada (portador do bastão) recebeu permissão real do Oba na década de 1950 para aprender e praticar a arte da fundição de bronze, enquanto ele foi ensinado residiu em Igun e depois de aprender a arte mudou de residência para Oloton (não muito longe da rua Igun) onde praticou a arte, seus descendentes praticam até hoje. É pertinente notar que não há guilda em Oloton, eles são membros da guilda Iguneromwon e se veem como tal, e participam das atividades da guilda em Igun, como as reuniões da guilda realizadas.

Os membros da guilda Igun eromwon são divididos em três graus de idade: O primeiro é o Ekhaemwen, o grupo de Oba que é o grau mais alto; compreendia chefes de fundição de bronze nomeados pelo Oba de Benin, os títulos de chefia na guilda são hereditários. O segundo ou médio grau é o Urhonigbe ou grupo de anciãos; eles são muito habilidosos, os membros deste grupo podem ser nomeados como chefes do reino de Benin pelo Oba, o maior favor do Oba em troca de seus bens e serviços.

O último e mais baixo grau é o grupo Iroghae; é composta por jovens artesãos entre 12 e 30 anos. No entanto, é pertinente saber que apenas os membros masculinos da família da guilda Igun eromwon podem aprender e praticar a arte, as mulheres das famílias Igun não podem conhecer a arte para duas razões; a arte requer tremenda força e mão de obra, em segundo lugar e mais importante, para evitar que a arte seja transferida para a família do marido. Uma forte estrutura organizacional existia dentro da guilda; os doze chefes da guilda em ordem de antiguidade.

A arte da fundição de bronze no reino de Benin floresceu no século 13 sob a guilda e durante o reinado de Oba Oguola, que teria dado a Igueghae o título de

Inneh (Igueghae foi o primeiro Inneh) e que é creditado por ampliar as categorias de objetos de bronze além da fundição de cabeças memoriais, usando-os para representar eventos importantes no reino, os fundidores de bronze tornaram-se registradores oficiais de eventos no reino (INNEH, 2007, pp. 105–107). Sempre que havia eventos no palácio, a guilda estava lá para testemunhar os eventos, eles observaram cuidadosamente e tiraram fotos mentais dos eventos, após os quais os fundiram coletivamente em bronze. Ao contrário de outros ofícios e guildas, a guilda dos fundidores de bronze foi proibida de fundir bronze e vender para qualquer pessoa fora do palácio (EYO, 1977, pp. 32 & 34). De acordo com Osarenren Kingsley Inneh, a arte era tão sagrada que nada (materiais ou obras) podia ser levado para fora do palácio.

Os membros das guildas de artesãos, incluindo a guilda dos fundidores de bronze, por medo de suas vidas e especialmente porque seu monarca e patrono não estava mais no reino, fugiram para o mato e aldeias vizinhas, no entanto, alguns membros da guilda Igun eromwon, principalmente os chefes, ficaram para trás em o Reino. Alguns dos fundidores de bronze que fugiram para o mato estabeleceram ou formaram novas aldeias e também estabeleceram a arte lá, por exemplo, foram estes fundidores de bronze de Benin que fugiram para Ora e Igueben em Esan que estabeleceram a arte, mas não havia guilda formada nessas aldeias e a prática da arte não era tão forte. A arte de fundir bronze tornou-se uma forma de ganhar a vida, enquanto os membros da guilda que não continuaram a arte em suas novas casas adotaram a agricultura como profissão. É pertinente saber que os membros da guilda Igun eromwon que permaneceram no reino de Benin ainda eram leais a seu rei enquanto estavam exilados em Calabar; durante este período, as mensagens escritas foram interceptadas e rastreadas antes de viver no reino e, portanto, os espiões do reino ou guardiões do palácio recorreram ao uso dos rodízios de bronze para inscrever mensagens como símbolos e imagens em forma de bronze que foram enviadas para Oba Ovoramwen no exílio como apenas ele conseguia decifrar os símbolos e imagens inscritos (NWACHUKWU, 2012).

Esta foi a origem do estabelecimento da arte de fundir o bronze em Calabar precisamente Oron (formalmente em Calabar mas agora em Akwa-Ibom) A guilda de fundidores de bronze do Reino de Benin foi iniciada novamente em 1914 pelos poucos fundidores de bronze que permaneceram no reino durante a invasão (os chefes da guilda e alguns outros membros).

Apesar dessas mudanças trazidas pela colonização do reino, a guilda ainda mantinha alguns aspectos de seu passado pré-colonial. Eles ainda prestam homenagem ao santuário de Igueghae, o fundador da guilda dos

fundidores de bronze do reino de Benin, eles ainda respeitam a posição de Inneh e o veem como seu chefe, eles ainda retêm e mantêm a estrutura administrativa da guilda, a única notável. A mudança foi a introdução do cargo de secretário e conselho de jovens. A posição do secretário, ao contrário de outras posições da guilda que é hereditária, é nomeada entre os membros da guilda. A qualificação para o cargo é a capacidade de ler e escrever (OSAWARU; EGHAFONA, 2004, p. 90). A posição tornou-se necessária ou importante quando os assuntos da guilda, especialmente sua reunião a cada duas semanas, precisavam ser documentados.

A posição posterior do conselho da juventude tornou-se necessária quando a guilda precisou afastar intrusos que não fazem parte da família original da guilda de praticar a arte da fundição de bronze. A guilda ainda manteve suas funções com exceção de registrar os eventos do palácio em bronze; essa função de registrar os eventos do palácio em bronze foi assumida por câmeras de vídeo. A estrutura de membros do sistema de idade e grau ainda é mantida, os filhos do sexo masculino dos membros da guilda aprendem a arte desde cedo, enquanto as mulheres da guilda ainda não têm permissão para praticar a arte.

3 Os museus e a representação da informação

A representação da informação se fundamenta na definição clássica de Greimas e Courtés:

Representação é um conceito da filosofia clássica que, utilizado em semiótica, insinua – de maneira mais ou menos explícita – que a linguagem teria por função estar no lugar de outra coisa, de representar uma ‘realidade’ diferente. Está aí, como se vê, a origem da concepção das línguas enquanto denotação: as palavras não são então nada mais do que signos, representações das coisas do mundo (GREIMAS; COURTÉS, 2008, p. 419, grifo dos autores).

Diante dessa ideia de estar no lugar de outra coisa, é importante atentar-se para aquilo que se destaca e para aquilo que se oculta no processo de representação. Sendo este processo, portanto, também um lugar de tensionalidade política e social. (ALBUQUERQUE, MARTINS; MOTA, 2019)

A escolha dos conceitos para representação de um objeto, é um processo complexo, que não envolve apenas conhecimento técnico, exige também o conhecimento do contexto desse objeto.

Dessa forma, Albuquerque, Martins e Mota (2019) definem a representação da informação como:

[...] uma prática de enunciação das propriedades de um objeto informacional (unidade de informação organizável) ou das relações desse objeto com outros que o identificam, através de operações que visam a sua descrição. (ALBUQUERQUE, MARTINS; MOTA, 2019, p.18)

A organização e representação da informação e do conhecimento [ORIC], reconhecido como processo intelectual, além de organizar e dividir os agrupamentos de conhecimento, também carrega em si a representação da informação e as relações e representações sociais. Sobre essas representações, Berger e Luckmann, 1962, apontam que:

Sendo a realidade socialmente construída, seus significados são compartilhados pela linguagem, instrumento fundamental para a objetivação das experiências individuais e sociais. Por meio da linguagem são construídos campos semânticos de significados, representações simbólicas e esquemas classificatórios sobre o esquema social de conhecimento. A construção de significados é um processo dialético e relacional, com inter-relação dos âmbitos individuais e coletivo, permeada por símbolos, representações sociais e (re)produzidos em diferentes instâncias da sociedade. (BERGER; LUCKMANN, 1962 *apud* ALBUQUERQUE; MARTINS; MOTA, 2019, p.162).

No campo da Museologia, Ceravolo e Tálamo, trazem a conceituação sobre a pesquisa museográfica, que registra, investiga, cria fichas de catalogação e elabora o tratamento documental sobre a massa de dados, produzindo instrumento de busca e acesso à informação (CERAVOLO; TÁLAMO, 2000).

Após a Segunda Guerra Mundial, nota-se uma consolidação de conceitos e práticas a fim de normalizar o processo de organização e representação da informação e conhecimento nos museus. O conceitos de musealização e fato museológico, deslocam o significado da instituição, não estando mais ligado apenas a um lugar. Com isso, a própria natureza do objeto de estudo da museologia também se amplia.

A partir da criação do ICOM, Conselho Internacional de Museus, vinculado à Unesco, em 1946, sua primeira ação foi a criação do CIDOC, Comitê Internacional de Documentação, que ajudou a estruturar um esquema de classificação museológica para bibliotecas e centros especializados em Museus. (CERAVOLO; TÁLAMO, 2000).

Considerando o cerne desta pesquisa, destacaremos aqui o papel desempenhado pelos catálogos nos museus. Os catálogos podem ser definidos como:

Geralmente estruturados a partir de listas, os catálogos apresentam dados completos das obras expostas e seus autores, contendo título da obra, suporte, datação, proveniência, entre outras informações. Com o passar do tempo, as publicações também passaram a conter textos introdutórios enxutos relacionados à temática da exposição e uma seleção bastante concisa de imagens em preto e branco. Ainda assim, os catálogos conservavam a simplicidade no modo de registrar a exposição. Com a ascensão do curador de exposições de arte no contexto internacional a partir da década de 1970, os textos passaram a ocupar espaços maiores nas

publicações vinculadas às exposições, provavelmente como forma de explicar mais detalhadamente os discursos curatoriais. Os catálogos são (ou ao menos deveriam ser) obras que possibilitam a compreensão do conceito proposto pelo curador, isto é, possibilitando que a linguagem empregada pelo autor seja compreendida pelo maior número de pessoas possível, sem a necessidade de decodificar termos e elementos textuais. Caso contrário, a publicação será um meio de comunicação e registro ineficiente, permanecendo restrita às prateleiras de especialistas e bibliotecas acadêmicas.

Analisando catálogos de exposições, nota-se que existem dois modelos de publicações: os mais acadêmicos, que apresentam textos extensos cheios de informações históricas e poucas imagens; enquanto outros são mais visuais, recheados de imagens e textos enxutos. Há também catálogos que mesclam os dois modelos, combinando de forma equilibrada e agradável elementos textuais e visuais na tentativa de produzir um registro bem fundamentado da exposição. Desde as últimas décadas do século XX, conforme apontam Béatrice Joyeux Prunel e Olivier Marcel (2016), os catálogos de exposições de arte passaram por transformações tanto físicas quanto de conteúdo, afastando-se lentamente da sua função primordial de apenas listar as obras exibidas em uma exposição. Os formatos e as proporções das publicações se diversificaram e, com o desenvolvimento tecnológico, as imagens passaram a ter melhor qualidade. (DIETZ, 2021, p. 27)

Silveira (2004) defende que o catálogo é definido por sua função. É a função que define o catálogo como tal. Ele indica, arrola, registra, classifica, ilustra, explica etc. O formato é secundário, podendo ser mesmo um prospecto de algumas páginas. Mas em geral se aceita a designação quando voltada ao menos para um pequeno livreto, mesmo que de pouquíssimas páginas. A lógica interna é a da organização de informações referentes a algo passível de ser classificado ou demonstrado como inserido num conjunto maior. Pode existir, acompanhando uma exposição ou um acervo de qualquer espécie, na forma de um guia, uma revista especializada, um cartaz diagramado para isso, um mapa explicativo, ou um impresso outro, mais elaborado. Se a sua identidade permanecer explícita, importa pouco a sua forma. Mas, por princípio, ele é um livro comum no aspecto externo, porém especializado e funcional. (SILVEIRA, 2004, p. 2)

O catálogo, portanto, reúne os objetos do museu, de maneira que possam ser organizados, classificados e principalmente possam ser divulgados ao público da instituição. A maneira como as peças são descritas nos catálogos, reflete o discurso e a intencionalidade da exposição, por isso carrega um caráter técnico, e também simbólico.

Dentre os tipos e formas de catálogos, destaca-se o Catálogo Raisonné, também chamado de catálogo Racional, que é uma compilação de a produção

conhecida de determinado artista, ou, um recorte bem definido da mesma, como um período ou técnica específica (pinturas, esculturas, desenhos, etc.).

Esse tipo de catálogo é caracterizado por um trabalho extenso e complexo, pois busca determinar a autoria de obras analisadas, baseando-se no fazer do artista estudado. Dessa maneira, o seu contexto de produção é fundamental, pois permite a preservação do legado de um artista e de em alguns contextos de seu povo.

Voltando à questão dos artefatos espoliados do continente africano, no violento processo de colonização, especialmente no caso aqui em questão, das peças da vasta coleção dos Bronzes de Benin; não é possível hoje afirmar que a preservação de seu legado está garantida, pelo simples fato de estarem expostas nos principais museus europeus. Mesmo que essas instituições ofereçam tecnologia avançada em preservação de patrimônio, elas são colocadas de maneira descontextualizada.

Considerando toda a complexidade do processo de fundição dessas peças, não é possível mais determinar a autoria dessas peças, nem quais fatos históricos ou cotidianos da cultura do antigo Benin foram retratadas. A espoliação levou consigo as peças e apagou boa parte da história e memória de um povo.

Em novembro de 2022, foi inaugurado o projeto 'Digital Benin'. O Digital Benin reúne todos os objetos, fotografias históricas e rico material de documentação de coleções em todo o mundo para fornecer uma visão geral há muito solicitada dos artefatos reais do Reino de Benin saqueados no final do século XIX. Os objetos históricos do Benin são uma expressão das artes, cultura e história do Benin, e foram originalmente usados como artes de representação real, para retratar eventos históricos, comunicar, adorar e realizar rituais. A plataforma digital apresenta estudos que conectam a documentação digital sobre os objetos espoliados a histórias orais, pesquisa de objetos, contexto histórico, um catálogo fundamental da linguagem Edo, locais de proveniência, um mapa do Reino de Benin e coleções de museus em todo o mundo. (DIGITAL BENIN, 2022)

O Digital Benin conecta dados de 5.246 objetos em 131 instituições em 20 países. O escopo do Digital Benin se concentra em objetos saqueados pelas forças britânicas do Reino de Benin (atual Estado de Edo, Nigéria) em fevereiro de 1897 e distribuídos logo após. Juntos, esses eventos e processos levaram à translocação mundial dos objetos mostrados nesta plataforma. Um pequeno conjunto de objetos é incluído no catálogo para representar o contexto mais amplo em que se situa a produção artística das guildas do Benin: marfins bini-portugueses, produzidos e circulados fora da África Ocidental no século XVI, objetos produzidos em regiões vizinhas do reino e uma seleção de obras

produzidas por artistas renomados após 1930, que fazem parte dos acervos dos museus. esses eventos e processos levaram à translocação mundial dos objetos mostrados nesta plataforma. Um pequeno conjunto de objetos é incluído no catálogo para representar o contexto mais amplo em que se situa a produção artística das guildas do Benim: marfins bini-portugueses, produzidos e circulados fora da África Ocidental no século XVI, objetos produzidos em regiões vizinhas do reino e uma seleção de obras produzidas por artistas renomados após 1930, que fazem parte dos acervos dos museus. esses eventos e processos levaram à translocação mundial dos objetos mostrados nesta plataforma. Um pequeno conjunto de objetos é incluído no catálogo para representar o contexto mais amplo em que se situa a produção artística das guildas do Benim: marfins bini-portugueses, produzidos e circulados fora da África Ocidental no século XVI, objetos produzidos em regiões vizinhas do reino e uma seleção de obras produzidas por artistas renomados após 1930, que fazem parte dos acervos dos museus. (DIGITAL BENIN, 2022)

É interessante notar que na página inicial do projeto na internet, foi colocado um alerta para ‘conteúdo sensível’. Visto que os catálogos apresentados, são os mesmos das instituições que possuem a guarda das peças catalogadas, grande parte situada na Europa, apresentando dessa forma, uma representação colonialista. O site reconhece que essa visão europeia é danosa para toda a população africana e sua diáspora.

As instituições com maior número de peças dos Bronzes de Benin, são o British Museum, na Inglaterra; o Ethnologisches Museum, na Alemanha; e o Field Museum, nos Estados Unidos. O British Museum abriga 944 objetos, das 5.246 catalogadas.

4 Definição da amostra e análise

Os objetos escolhidos para análise estão disponíveis no Catálogo online disponibilizado pelo projeto Digital Benin. Este catálogo apenas reproduz as fichas de registro de cada instituição, sem nenhuma alteração ou revisão.

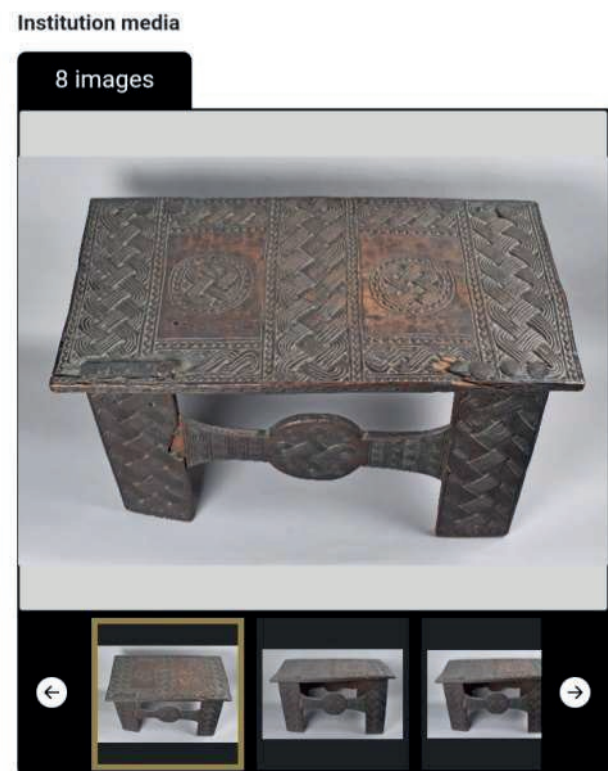
As obras selecionada fazem parte da coleção atualmente exposta no British Museum. Na fase de pré-análise foram selecionados para o universo dessa pesquisa, os objetos pertencentes à etnia Edo, e designação Agba. A etnia Edo possui 87 designações no conjunto de suas obras. A designação Agba se refere à objetos cerimoniais utilizados pelos mais velhos (ancião) do Reinado.

No total, são 10 objetos tridimensionais, feitos de madeira, ferro e vidro, esculpidos em alto relevo.

A metodologia utilizada será análise de conteúdo, fazendo leitura das fichas de registro de cada objeto.

Na fase de pré-análise observou-se que, as fichas de registro do catálogo apresentam um cabeçalho principal constando a designação do objeto, que neste caso, são todos da etnia Edo; a designação inglesa para o objeto, classificados como bancos ou partes de bancos; a categoria do objeto, sendo objeto cerimonial; número de registro ; proveniência, indicando o nome da instituição ou pessoa proprietária da objeto; e localização, que no conjunto dos objetos escolhidos, estão todos localizados em Londres.

Figura 1. Foto de um banco cerimonial disponibilizada na ficha de registro.



Fonte: Digital Benin, 2022.

Durante a fase de exploração do material, foi possível verificar que a ficha de registro interna, que é a reprodução do catálogo das instituições de origem, possui campos para descrição física, material, dimensões, data de produção, campo para designação do produtor/ artista, técnica utilizada, local de produção, meios de aquisição, local coletado, etnia de produção, estilo/escola, notas para informações de referência bibliográfica e informações adicionais sobre os objetos.

Figura 2. Ficha de registro dos objetos cadastrados no Digital Benin.

Internal Data (eng)	
<p>British Museum The data was transferred directly by the institution. Last Updated: 2022-08-04 Language: English</p>	
Registration Number	Af1957,11.3
Object Name - Comment	stool
Object Name - Object Name	stool
Title	
Title translation	
Title type	
Object History Note / Acquisition Note	
Physical description	Stool; made of wood. Rectangular in form; legs fixed to the seat by iron nails.
Material	wood-iron
Material - Comment	

Fonte: Digital Benin, 2022.

Após a análise verificou-se no tratamento dos resultados que, quanto à descrição física, os bancos cerimoniais são de madeira, ferro e vidro. A data de produção das obras não possui uma data precisa, sendo descrita em todos os objetos sendo possivelmente entre os séculos XVI e XIX.

Apesar da existência do campo produtor/artista, não consta essa informação em nenhuma das fichas. O mesmo se repete no campo da técnica utilizada, que se apresenta de forma bastante reducionista, indicando apenas que foram esculpidos, sem detalhar a complexidade ou exclusividade da técnica utilizada em Benin.

A exclusividade da técnica utilizada, foi o principal motivo da espoliação pelos militares ingleses, visto que se utilizavam de uma tecnologia até o momento amais encontrada. A curiosidade maior em torno das peças se dá ainda, pelo fato de o colonizador acreditar que a cultura do Benin, assim como todas as outras representações africanas, eram culturas primitivas, não no sentido de serem as primeiras, mas de serem uma subcultura.

A dificuldade em determinar um produtor/artista, é explicada pela própria história dos bronzes de Benin, no seu contexto de origem, onde eram produzidos para registrar momentos especiais do Reinado, ou até mesmo acontecimentos mais rotineiros do cotidiano local. O processo de aquisição, feita por espólio, fruto de ação

militar violenta, impossibilita qualquer tipo de identificação do seu contexto de produção.

Nos campos disponíveis para informação sobre os meios de aquisição, está registrado em 9 fichas apenas o nome do proprietário atual, se referindo a algum colecionador ou instituição localizada na cidade de Londres. Apenas uma ficha traz a informação de que o objeto foi adquirido por meio da campanha militar colonial britânica no Benin, em fevereiro de 1897. Este mesmo objeto também possui no campo 'notas', informações sobre o contexto histórico da peça. Neste caso, trata-se do banco de Overami, fazendo referência ao Rei de mesmo nome que o utilizou durante o seu reinado. Mesmo com essas informações adicionais, o registro não possui identificação do produtor/artista.

Na maioria das fichas, a indicação de local de coleta, menciona o Benin ou a Nigéria. Em apenas uma ficha, houve a indicação 'Africa', refletindo uma linha de pensamento ocidental e europeu, que considera o continente africano, uma massa única, e não a sua diversidade étnica distribuída em seus 54 países.

A exposição dessas peças, é realizada fora de seu contexto de produção, exibidos em caixas de vidro, com uma conotação minimalista, que pende do etnográfico, apenas para o estético, na maioria das vezes, descritos como artesanato. O pesquisador Paul Wood (2012), classifica essas características como parte de uma estrutura de 'exibição ocidental'.

5 Considerações finais

A história dos Bronzes de Benin é apenas uma amostra da violência física e simbólica sofrida por diversos povos do continente africano durante o processo de colonização europeu, que atuou sob a justificativa de trazer civilidade a esses povos, classificando-os como uma sub-raça, produtores de uma subcultura. O espólio dessas peças interferiu de maneira direta no processo de registro da história e da memória do Benin, atual Nigéria. Mesmo, mais de um século após a expedição militar da Inglaterra, o próprio país ainda não conseguiu resgatar ou reconstruir o processo de fabricação das peças como era anteriormente. Sendo este um dano histórico irreparável, que apenas corrobora para a manutenção de uma hegemonia eurocêntrica.

Existe hoje um movimento de devolução das peças para a Nigéria, o British Museum em seu site fala a respeito do assunto, demonstrando sua intenção em colaborar com este projeto, mas cai no discurso comum de colonizador, impondo condições para devolução e considerando que os objetos só estarão bem cuidados se estiverem em situação semelhante aos cuidados feitos pela instituição, abstando-se do reconhecimento do fato de que a posse desses objetos perpetua a

violência iniciada em 1897. Além da perpetuação de violência física e simbólica, continuam um movimento de desconfiar ou não creditar capacidade técnica e intelectual aos profissionais e instituições nigerianos.

O museu como espaço de preservação da história e da memória, também deve ser o espaço para desconstrução de estereótipos e instrumento para reparação e justiça histórica. Deve também aprender a incorporar na cultura institucional outras perspectivas, que não apenas europeia.

A análise das peças da designação Edo, conjunto Agba, é apenas uma amostra dos sérios problemas a serem superados no processo de catalogação de peças pertencentes a cultura negra, especialmente àquelas adquiridas por meio de espólio. É imprescindível que as instituições passem a contar no seu quadro profissional, pessoas capazes de rever a estrutura desses catálogos, até que o processo de devolução possa ser realizado.

Além da designação Edo, o catálogo do British Museum, possui mais 86 designações. Portanto, reconhece-se a necessidade de realizar um estudo maior, para ampliar a discussão sobre o tema e apontar outros desdobramentos possíveis.

Referências

- ALBUQUERQUE, M. E. B.C; MARTINS, G. K.; MOTA, D.A.R. (org.). **Organização da informação e do conhecimento**: intersecções teórico-sociais. João Pessoa: Editora UFPB, 2019. 208 p.
- BENIN DIGITAL. **Explore Digital Benin**, 2022. Disponível em: <https://digitalbenin.org/>. Acesso em: 22 maio 2023.
- AKENZUA, G. I. **The glorious reign of Omo N' Oba Erediauwa**. Lagos: Inter Press Limited, 2011.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2011.
- BARDIN, L. **Análise de conteúdo**. Lisboa: Edições 70, 2010.
- BARDIN, L. **Análise de conteúdo**. Lisboa: Edições 70, 2004.
- CERAVOLO, S. M.; TÁLAMO, M. F. Os museus e a representação do conhecimento: uma retrospectiva sobre a documentação em museus e o processamento da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 8, 2007, Salvador. **Anais [...]** Salvador: Universidade Federal da Bahia, 2007.
- CONNAH, G. **The archaeology of Benin**. Oxford: Clarendon Press, 1975.
- GREIMAS, A. J.; COURTÉS, J. **Dicionário de semiótica**. São Paulo: Contexto, 2008.
- INNEH, D. **The guilds working for the palace**. In: PLANKEINSTERNER, B. (ed.) *Benin kings and rituals, courts arts from Nigeria*. Chicago: Snoeck publishers, 2007.
- MENEZES, P. S.; ÁLVAREZ, E. P. A descolonização dos museus e a restituição de obras artes africanas; o debate atual na França. **CSONline** – Revista Eletrônica de Ciências Sociais, Juiz de Fora, n. 29, 2019.
- NWACHUKWU, M. P. A paper on Benin bronze casting. The story of power and royalty: an insight into the great bronze casting works of Benin kingdom, v. 3, 2012.
- ODIAHI, E. V. The Origin and Development of the Guild of Bronze Casters of Benin Kingdom up to 1914. **AFRREV IJAH: An International Journal of Arts and Humanities**, Bahir-Dar-Ethiopia, v. 6, n. 1, p. 176-187, 2017.
- OMORUYI, A. **Benin anthology**. Benin City: Benin Cultural Publications, 1981.
- OSAWARU, I. R.; EGHAFONA, K.A. An assessment of the Benin guild of bronze casters for tourism potentials in Nigeria. **South-South Journal of Culture and Development**. v.6, n.2, 2004.
- RUSKIN, J., *The Political Economy of Art*. Londres: Smith Elder, 1857.
- WOOD, P. Display, restitution and world art history: the case of the 'Benin Bronzes'. **Visual Culture in Britain**, v. 13, n. 1, p. 115-137, 2012.

O fenômeno do *Juridiquês* e seu impacto na representação do conhecimento

Legal slangs: a discredit to the law terminology, a threaten to democracy and hindrance to knowledge representation

José Augusto Chaves Guimarães (1); João Carlos Gardini Santos (2)

(1) Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências - Marília, SP – Brasil; (2) Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Campus Avançado Tupã IFSP - Tupã, SP - Brasil

Resumo

Problema: O uso de uma linguagem rebuscada e metafórica, semelhante à de textos literários – o *juridiquês* - constitui um problema reiterado na praxe processual brasileira, fenômeno que atua em desprestígio da terminologia jurídica definida em lei e consolidada na doutrina, gerando uma linguagem artificial e excludente que, por sua vez, dificulta o processo comunicativo e, como consequência, afeta a necessária transparência dos atos inerentes à democracia. **Objetivo:** Nesse sentido, analisa-se o *juridiquês* em seu aspecto conceitual e em suas características intrínsecas visando a tecer uma reflexão sobre suas danosas consequências, que podem atingir o processo e os instrumentos de representação do conhecimento. **Metodologia:** Analisa-se a incidência do *juridiquês* em acórdãos (ementas e votos) oriundos de exemplos trabalhados em cursos de Elaboração de ementas jurisprudenciais desenvolvidos em Tribunais do país ao longo dos últimos dez anos bem como em exemplos fornecidos por Guimarães (2004). **Resultados e Conclusões:** Evidencia-se que o necessário uso de linguagem técnica em peças processuais não é pressuposto para – e, ao contrário, se opõe a - o emprego do *juridiquês*, pois a linguagem técnica pauta-se pela busca por univocidade e desambiguação, valendo-se de recursos de clareza, objetividade e principalmente de precisão. Conclui-se que o *juridiquês* torna menos visível a terminologia jurídica e, em decorrência, traz danos à democracia pois fere o princípio de transparência. Em termos de representação do conhecimento, destacam-se as consequências extremamente danosas do *juridiquês* pois pode desvirtuar o sentido original do texto.

Palavras-chave: Juridiquês. Terminologia jurídica. Representação do conhecimento

Abstract

Problem: The use of a metaphorical language is becoming a serious problem in Court Cases, the so-called “legal slang” is a recurrent problem in the Brazilian judicial praxis. Such phenomenon brings serious damages to the Law terminology and creates an artificial and exclusionary language that makes the communication process more and more difficult and, as a consequence, affects the needed transparency of the democratic acts. **Objective:** This paper analyses the phenomenon – and practice – of legal slang in terms of its concept and characteristics in order to discuss its consequences. **Methodology:** court cases retrieved from training courses on Jurisprudential abstracts offered in the last decade and examples from Guimarães (2004) were analyzed in terms of their metaphorical dimension. **Results and Conclusions:** The use of technical language in court cases can be considered mandatory because They promote univocity, clarity, objectivity and precision to the information retrieval. In terms of knowledge representation, the presence of legal slangs can be extremely harmful because they can betray the original sense of the text.

Keywords: Legal slang; Legal terminology. Knowledge representation

Resumen

La utilización de un lenguaje barroco y metafórico em fallos judiciales – la jerga judicial – constituye un serio problema em la praxis procesal brasileña, Este fenómeno actúa em desprestigio a la terminología jurídica y dificulta el proceso comunicativo em dicha área afectando la transparencia de los actos democráticos. **Objetivo:** Este artículo analiza el fenómeno de las jergas judiciales a partir de su concepto y de sus características de manera a discutir sus consecuencias dañosas. **Metodología:** Se analiza la dimensión metafórica de jergas judiciales presentes em fallos judiciales recuperados em cursos de Elaboração de resúmenes judiciales impartidos em distintos tribunales brasileños em la última década, así como los ejemplos planteados por Guimarães (2004). **Resultados y Conclusiones:** Evidenciase la necesidad de la utilización de un lenguaje técnico em fallos judiciales para que se pueda efectivamente promocionar univocidad, desambiguación, claridad, objetividad y precisión. Em términos de representación del conocimiento, la presencia de jergas judiciales puede ser extremamente dañosa porque puede traicionar el sentido original del texto.

Palabras-clave: Jerga judicial. Terminología jurídica. Representación del conocimiento.

1. Introdução

A simplicidade constitui o último grau da sofisticação[1]

Observa-se, no âmbito jurídico, notadamente na atuação do Judiciário, seja em peças processuais, por parte dos advogados, seja em votos e decisões, por parte de magistrados, muitas peças permeadas por uma linguagem rebuscada e metafórica, que mais se assemelham a textos literários, fazendo uso de termos arcaicos, estrangeirismos, construções frasais barrocas e, em alguns casos, criando neologismos esdrúxulos. Tal prática denomina-se *juridiquês*;

Muitos defendem o uso do *juridiquês* como um elemento natural e caracterizador da “língua jurídica”, como se todos os que operam nessa seara fossem identificados pela natureza rocambolosa de sua linguagem. Tais defensores, por sua vez, se esquecem do fato de que o funcionamento da justiça é algo constitucionalmente previsto como um serviço ao cidadão dentro de um Estado democrático e que a informação daí gerada reveste-se de um caráter eminentemente técnico – que, por sua vez, se opõe ao literário – em que os documentos devem seguir uma estrutura textual pré-definida (como por exemplo a sentença, que nos termos do artigo 458 do Código de Processo Civil, deve apresentar uma sequência canônica: Relatório, Fundamentação e Dispositivo) valendo-se de termos técnicos que integram oficialmente a língua de especialidade do Direito e são definidos e caracterizados por lei ou pela doutrina jurídica.

A vista disso, o presente artigo discute o *juridiquês* procurando delimitar seu conceito e suas características e apresenta uma reflexão acerca do perigo de tal prática, seja por desprestigiar a terminologia jurídica efetivamente construída e sedimentada, seja por trazer um entrave ao exercício da democracia uma vez que esta pressupõe que os serviços de natureza pública estabeleçam uma comunicação clara, objetiva e precisa com a sociedade.

Para tanto, analisa-se a incidência do *juridiquês* em acórdãos (ementas e votos) oriundos de exemplos trabalhados em cursos de Elaboração de ementas jurisprudenciais desenvolvidos em Tribunais do país ao longo dos últimos dez anos bem como em exemplos fornecidos por Guimarães (2004).

2. *Juridiquês*: conceito e caracterização

O termo *juridiquês* constitui um neologismo brasileiro que surgiu como conotação pejorativa para designar uma linguagem utilizada em textos jurídicos, linguagem essa de natureza hermética e excludente (privativa apenas dos integrantes daquele meio, como se fosse um novo dialeto), permeada

pelo uso de jargões e de termos rebuscados e metafóricos, e prejudicando substancialmente a compreensão do texto de uma forma mais abrangente. Em geral, o termo *juridiquês* é utilizado em tom “depreciativo, irônico, referindo-se a um “conjunto de termos e expressões próprias da área jurídica, geralmente de difícil compreensão geral”. Não se confunde, pois, com a terminologia jurídica propriamente dita, pois esta decorre dos termos que são definidos em lei ou em doutrina como figuras ou institutos do Direito, como é o caso, por exemplo, do *Habeas Corpus*, previsto no artigo 5º., LXVIII da Constituição Federal de 1988. Bulhões (2008) contrapõe a terminologia jurídica, composta pelo vocabulário técnico, específico da ciência do Direito, ao *juridiquês*, este permeado pela ornamentação e pelo rebuscamento, com função antes decorativa que informativa, e caracterizado por Belém (2013, p. 316) como um “desvio da linguagem jurídica”.

Em uma caracterização geral, o *juridiquês* vale-se de uma linguagem evasiva e prolixa, com orações e parágrafos muito extensos, forte uso de adjetivos e advérbios (que, por sua vez, retiram sua objetividade), apostos e longas orações em ordem inversa (que interrompem a fluidez do raciocínio em uma oração), expressões ambíguas, e uso de termos estrangeiros (principalmente em latim), de neologismos e de metáforas, em total desprestígio, como já mencionado, à terminologia jurídica consagrada.

Tal prática, por sua vez, contraria três requisitos fundamentais de todo e qualquer documento técnico (e sentenças e acórdãos o são, por natureza): a clareza (de modo que a compreensão seja possível com o menor tempo e com uma única leitura), a concisão (de modo que se garanta o maior número de informações com a menor extensão textual) e a precisão (uso de termos específicos e inequívocos, que tenham caráter oficial) (GUIMARÃES; SANTOS, 2016).

Há, inclusive, situações extremas em que petições, contestações, recursos, sentenças, votos e pareceres chegam a ser redigidos muitas vezes de forma poética, descaracterizando completamente o objetivo precípuo daquele documento técnico pois “a utilização de metáforas e prosopopeias constitui em fator de imprecisão, prejudicando a inteligibilidade e a objetividade do texto” (GUIMARÃES, 2004, p. 89).

Tem-se, assim, um floreio excessivo da língua com o recurso desnecessário a termos pouco conhecidos do grande público, em um rebuscamento que, por sua vez, contribui para a morosidade da justiça pois traz entraves à franca comunicação. Ademais, o *juridiquês*, ao utilizar-se de expressões de difícil compreensão “a fim de elitizar a linguagem jurídica e segregar poder”, pode “trazer consequências

irreversíveis à justiça e à sociedade” (OLIVEIRA, 2021, p. 30983).

No âmbito do exercício profissional, destaca Moisés Junior (2015) que a simplificação do linguajar jurídico constitui dever do advogado, uma vez que o art. 45 do Código de Ética da OAB refere-se à singeleza, aspecto que se aplica não apenas ao comportamento, mas, sobretudo, à linguagem. No âmbito do Judiciário, a Associação dos Magistrados Brasileiros - AMB desenvolve, há quase duas décadas, a Campanha pela Simplificação da Linguagem Jurídica (AMB, 2005) que evidencia uma crescente preocupação com clareza e objetividade entre os próprios magistrados.

Tais iniciativas visam a coibir uma prática que vem se alastrando no mundo jurídico e que, ao impedir uma compreensão mais ampla dos conteúdos veiculados, por conta de construções barrocas e de um vocabulário empolado, acabam por dificultar, e até inviabilizar, o processo comunicativo e trazer hermetismo às atividades jurídicas, deixando a compreensão efetiva de seus atos a apenas um grupo quando, na verdade, tais atividades têm uma função social. Vale dizer que o recurso a termos incompreensíveis, ainda que integrantes da língua portuguesa, e a expressões em línguas estrangeiras, principalmente o latim, acaba sendo como um símbolo de poder por alguns, que disso se valem para demonstrar conhecimento, erudição e autoridade, sem levar em conta o efetivo destinatário do cumprimento de um serviço de natureza pública.

Em tempos atuais, quando a tecnologia imprime rapidez ao processo comunicativo e em que se avolumam cada vez mais as atividades do Judiciário, necessária se faz a adoção da tônica do simples, de modo a garantir rápida compreensão. Não se defende, aqui, a tradução de toda a terminologia efetivamente jurídica (reitera-se: aquele conjunto de institutos jurídicos nomeados, definidos e caracterizados em lei ou em doutrina), mas o abandono de uma “elitização da linguagem” permeada pela falta de concisão e de objetividade.

A tal respeito, o Projeto de Lei 3.326/2021, de autoria do deputado federal Paulo Bengston pretende alterar o Código de Processo Civil [2] para tornar as decisões judiciais plenamente compreensíveis por qualquer pessoa do povo, valendo-se de uma linguagem simples e clara (que não se confunde com a linguagem coloquial), também conhecida *plain language* [3].

Vale destacar que o discurso jurídico vem sendo tradicionalmente acusado, em distintas partes do mundo, de ser desnecessariamente difícil, tornando-se inacessível ao público leigo por conta daquilo que os países anglófonos denominam como *legalese*, ou seja, uma linguagem confusa, prolixa, indireta a que pode ainda incluir termos ou frases desnecessários (OATES; ENQUIST, 2009, p. 127).

Tal aspecto já havia sido objeto de projeto anterior - PL7.448/06, apresentado pela ex-deputada federal Maria do Rosário, em uma iniciativa direta contra o *juridiquês* para romper com a cultura de segregação por meio da linguagem, defendendo para todos o direito de compreender e transformando a comunicação jurídica em algo empático e inclusivo. Como justificativa, o referido projeto menciona que (grifos nossos):

[...] o Estado tem o compromisso político de dirigir-se diretamente ao cidadão que o procura para a solução de uma lide. Nesse passo, deve-se considerar que o Direito, de forma corriqueira, utiliza-se de linguagem normalmente inacessível ao comum da população, apresentando, no mais das vezes, um texto hermético e incompreensível. Assim, de pouco ou nada adianta às partes a mera leitura da sentença em seu texto técnico. Desse modo, a tradução para o vernáculo comum do texto técnico da sentença judicial impõe-se como imperativo democrático, especialmente nos processos que, por sua natureza, versem interesses peculiares às camadas mais humildes da sociedade, como as ações previdenciárias e relacionadas ao direito do consumidor.

Há de se convir que o uso de uma linguagem desnecessariamente rebuscada pode afastar o cidadão da Justiça, não sendo demais recordar que a democratização da linguagem propiciará uma maior aproximação entre a população e o Poder Judiciário.

Sem que haja um fundamento claro para tal, há quem defenda a manutenção do rebuscado *juridiquês*. Para esse grupo, a simplificação da linguagem, tornando-a mais clara e acessível pode afetar a autoridade, o prestígio e o respeito ao Poder Judiciário por parte da sociedade. Tal argumentação carece de coerência pois quanto mais compreensíveis forem as peças processuais e as decisões judiciais, maior aproximação e empatia haverá com a sociedade, grande destinatária dos serviços judiciais, visto que o objetivo da comunicação só é atingido quando o receptor compreende devidamente a mensagem.

Assim, e desmistificando uma falsa e elitista ideia de que escrever bem é escrever difícil, não se pode esquecer que a qualidade de um texto passa, necessariamente, pela sua clareza, objetividade, coerência, coesão, correção e precisão vocabular. Indo além, pode-se dizer que escrever de forma rebuscada, barroca e de difícil compreensão é fácil, pois quanto mais enviesado e rocambolesco for o discurso mais aparência de sofisticação ele terá [4]. No entanto, escrever de forma fácil, com clareza e objetividade, é muito mais difícil pois não haverá mais recursos retóricos que possam ocultar eventuais problemas de concordância e de regência, por exemplo, pois as orações estarão em sua forma mais pura, em que qualquer erro se nota de pronto. Às vezes, a preocupação com uma linguagem mais *sofisticada* acaba por atingir mesmo a coerência do texto, como na seguinte ementa de acórdão trazida como exemplo por Campestrini (1994, p. 39):

Procede culposamente o motorista que, em via pública, dotado de iluminação, no perímetro urbano, estaciona o veículo **para** em seguida, passar sobre a vítima embriagada e caída na rua, causando-lhe lesões gravíssimas.

Como ressalta o referido autor, a proposição *para* dá ideia de finalidade. Desse modo, ter-se-ia alguém que estaciona um veículo *com a finalidade de*, em seguida, passar sobre a vítima..., o que caracterizaria *dolo* e não *culpa* como afirma a ementa (pois se essa era a finalidade, havia clara intenção de fazê-lo). Essa incoerência não haveria ocorrido se o autor do texto houvesse optado pelo aditivo *e* que, embora aparentemente mais simples, não traz nexo finalístico ou causal (GUIMARÃES, 2004, p. 97).

Com relação ao uso de latinismos o problema é ainda mais sério pois desde 1961 (ou seja, há mais de seis décadas) o ensino do latim não é mais obrigatório nos níveis fundamental e médio no Brasil de tal modo que a geração que hoje se encontra às voltas com a redação de tais documentos não teve a oportunidade de estudar a gramática dessa língua clássica e acaba simplesmente repetindo expressões em latim lidas ou ouvidas aqui e acolá, sem, no entanto, analisar se elas estão gramaticalmente corretas e, como consequência, perpetuando erros. Não é raro, nessa tônica, encontrarem-se expressões redundantes como “um plus a mais”.

3. O *juridiquês* como um desprestígio à terminologia jurídica, uma ameaça à democracia e um entrave à representação do conhecimento na área jurídica

Como já mencionado, em muitas ocasiões os defensores do *juridiquês* argumentam tratar-se de “terminologia jurídica”. Cabe, pois, distinguir terminologia de jargão.

A terminologia pode ser entendida, sob o aspecto teórico, como uma área de estudo dos termos que compõem um campo conceitual (ou nocional) ou, sob o aspecto prático ou concreto, como o conjunto efetivo de termos que integra, caracteriza e representa os conceitos relacionados a uma língua de especialidade, na medida em que cada grupo social (uma profissão, por exemplo) utiliza um vocabulário específico, composto por termos constituídos de uma ou mais palavras para designar univocamente um conceito dentro de uma língua de especialidade (como o Direito, por exemplo). Desse modo, a terminologia jurídica pode ser entendida tanto como o estudo científico dos termos empregados no universo do Direito quanto pelo conjunto desses termos em si mesmos. Nessa última acepção, tem-se por objetivo facilitar e agilizar a comunicação tanto entre os especialistas em seu meio quanto entre os especialistas e o público em geral, tornando-a compreensível e sem ambiguidades (LARA, 2005; FOGAÇA; CERVANTES, 2021). Como destaca

Cabré (1995) a terminologia visa, em última análise, a fixar e sedimentar termos como formas normalizadas para o discurso de um dado campo de especialidade, para que estes possam servir elemento de referência e, como consequência, para que se possa descartar as demais variantes daquele conceito.

Cabe destacar que os termos técnicos são caracterizados pela sua definição precisa no campo a que pertencem e são formalmente reconhecidos, documentados e disseminados, fazendo parte, inclusive, dos processos educacionais. No caso do Direito, a fixação terminológica ocorre por meio do estabelecimento e definição de um termo como instituto, seja por previsão legal, seja por conceituação doutrinária, sendo objeto de estudo nos cursos de Direito e por meio desse processo de padronização garante-se uma comunicação inequívoca.

O jargão, por sua vez, embora próprio de um dado entorno profissional, caracteriza-se por uma linguagem hermética, viciada e excludente, muitas vezes com a “a intencionalidade de que a conversa não seja compreendida por quem não pertença ao círculo” (JARGÃO, [2023]), como no *juridiquês*.

A título de exemplo veja-se, no Quadro 1, em Apêndice, uma comparação entre elementos que efetivamente integram a terminologia jurídica e elementos do *juridiquês*. Em cada linha do referido quadro, apresenta-se um termo tal como previsto oficialmente na terminologia jurídica, seu fundamento legal ou doutrinário, seus correspondentes no *juridiquês* e uma análise das distorções que os permeiam.

Observe-se, nos exemplos apresentados, o uso metafórico de vergastar (chicotear) e objugar (censurar) como sinônimo de recorrer. Igualmente tem-se o uso inadequado do termo Carta Magna para designar a Constituição Federal brasileira constitui, além de um jargão metafórico, um erro histórico pois a Carta Magna, ou Magna Carta, constitui um documento assinado em 1215 pelo rei inglês João Sem Terra, que limitou os poderes dos monarcas ingleses e constitui formalmente o início do constitucionalismo (SILVEIRA, 2021).

No mesmo sentido o uso do termo *peça ancilar* para petição inicial ou de *autarquia ancilar* para o INSS é totalmente inadequado, pois deriva de ancila (servil), significando algo que é subsidiário, auxiliar, acessório, suplementar. Ter-se-ia, então, uma *autarquia suplementar*? Suplementar a que? Também o uso do termo *Corte* para designar tribunal não se justifica pois, em língua portuguesa, significa a residência de um monarca, incluindo a nobreza que a frequenta, podendo designar também o paço. No caso, tem-se um anglicismo derivado de *Court* (tribunal).

Vejam-se ainda termos de conotação fortemente emocional e subjetiva, como *flagrante* (aquilo que está diante dos nossos olhos, observável no momento exato de sua ocorrência, tornando-se incontestável) e *sobejamente* (excessivamente, demasiadamente).

A publicidade é um princípio inerente à democracia, e mais que uma ampla divulgação, pressupõe que tudo que se torna público deva ser suficientemente claro e compreensível, sem o que não cumpre o seu objetivo de promover um processo comunicativo. No caso do Judiciário, é a publicidade de seus atos, prevista no artigo 93, IX da Constituição Federal, que estabelece a ponte entre o juiz (representando o Estado) e o cidadão (WLASSAK, 2002; SALGADO, 2017).

Se para o texto legislativo existem normas específicas e precisas quanto à forma de estruturação de seu conteúdo para garantir a inteligibilidade no contexto social (cf. Lei Complementar 95, de 26 de fevereiro de 1998), tal não ocorre relativamente às informações geradas no âmbito do processo judicial, que ficam à mercê do *estilo pessoal* de cada advogado, membro do Ministério Público ou magistrado, criando seara fértil para o *juridiquês*.

Na tradição judicial portuguesa, observa-se um linguajar que tem conseguido atingir um nível de clareza e de objetividade bastante superior à tradição brasileira, como no caso dos acórdãos portugueses que denominam acertadamente de *resumo* aquilo que a tradição brasileira insiste em denominar com o já arcaico e polissêmico termo *ementa*.

Nesse sentido, destaca Carmo (2013, p. 148) que a relevância do uso da linguagem pelo Judiciário reside exatamente na construção de uma relação democrática entre o sistema de justiça e os cidadãos”, aspecto que “muito depende de melhor em mais acessível informação e da capacidade de compreensão pelos cidadãos acerca da justiça que é administrada em seu nome pelos tribunais”. E, indo além, reporta-se às recomendações do Conselho da Europa sobre a “utilização de uma linguagem compreensível para os intervenientes nos processos judiciais como uma exigência democrática e uma das condições de garantia do pleno acesso à justiça”. O referido autor segue em sua crítica ao afirmar que:

O encriptamento da linguagem judiciária assumida como jargão, o mimetismo face à linguagem jurídica científica, a confusão entre erudição e discurso dificilmente compreensível, a substituição da demonstração pelo pretense argumento de autoridade da citação e a prolixidade fastidiosa não são tão raros como deveriam ser (CARMO, 2013, p. 149).

À vista do exposto e partindo das premissas de que: a) a publicidade é inerente ao Estado democrático; b) essa publicidade pressupõe um processo comunicativo claro que atinja a sociedade; e c) há informações processuais elaboradas em uma linguagem hermética e de difícil compreensão, pode-

se concluir que o *juridiquês* constitui um efetivo entrave à democracia.

Cumprido destacar que o *juridiquês*, assim como muitos outros jargões de áreas de especialidade, quando presente em documentos técnicos, como é o caso do acórdão – que deverá ser objeto de análise e de representação temática para fins de pesquisa e de recuperação da informação –, constitui sério entrave pois pode levar a representações literais que enviesam a compreensão do texto.

Constituindo uma matéria de natureza simbólica que vem preocupando o mundo da documentação desde a sua origem, a representação do conhecimento pode ser abordada em duas dimensões: uma representação primária, voltada ao registro do conhecimento em um suporte documental e uma representação secundária, ocorrida em sistemas de documentação e de informação, em que se identificam os conceitos predominantes em um dado documento e os reapresentam, por meio de pontos de acesso normalizados que subsidiarão a recuperação da informação, também denominados por substitutos do conhecimento (*surrogates of knowledge*). Na representação secundária, os conceitos constantes dos documentos constituem insumos para a sua representação e devem ser identificados e descritos de modo a resgatar a temática do documento e também levando em consideração eventuais necessidades dos usuários potenciais. Deve-se recordar, outrossim, que ambas as dimensões da representação visam a propiciar o processo de comunicação humana. (VICKERY, 1986; OLSON, 2002; ALVARENGA, 2003).

Essa representação, por sua vez, vale-se de instrumentos denominados sistemas de organização do conhecimento – SOC (sistemas de classificação, cabeçalhos de assunto, arquivos de autoridade, redes semânticas, ontologias, dicionários, glossários, taxonomias e tesouros) que se caracterizam como sistemas de signos que visam à uniformização da linguagem utilizada em um Sistema de Informação, proporcionando uma representação padronizada do conteúdo informacional subsidiariamente. São sistemas funcionais que tratam o conteúdo dos documentos, para fins de armazenamento e recuperação com vistas à sua gestão, (HODGE, 2000; SOERGEL, 2001; SALES, 2007; BOCCATO, 2008; HJORLAND, 2008; LARA, 2015; MAZZOCHI, 2018).

O termo SOC (em inglês: Knowledge Organization Systems – KOS) foi inicialmente utilizado por Hodge (2000) para abarcar todos os tipos de esquemas de organização da informação e gerenciamento do conhecimento, desde sistemas de classificação, de categorização, cabeçalhos de assunto, listas de autoridades, tesouros, dicionários, glossários e redes semânticas e ontologias (HODGE, 2000; HJORLAND, 2008). O termo também foi

empregado por Soergel (2001), quando analisou as diversas normas dos instrumentos de organização do conhecimento (LARA, 2015, p. 92)

A vista disso, e considerando que: a) a representação do conhecimento pressupõe a apreensão dos temas mais importantes do documento; b) essa representação visa a propiciar a recuperação da informação a um público que, no caso da jurisprudência, é constituído não apenas por especialistas da área quanto pelas partes do processo, que são, via de regra, do público leigo; c) o profissional da informação responsável pela representação do conhecimento, ainda que tenha formação jurídica, não estará familiarizado com o *juridiquês* visto que esse é inerente àqueles que atuam especificamente no âmbito processual (advogados e magistrados), d) o profissional da informação vale-se de SOC para efetivar a representação, sendo que esses instrumentos se baseiam efetivamente na terminologia jurídica, pode-se deparar com situações extremas em que o profissional não dará conta de perceber os institutos jurídicos que se encontram encobertos pelas metáforas do *juridiquês* e, assim, não representar adequadamente o conteúdo dos documentos.

Desse modo, é fundamental que se tenha claro que as linguagens de indexação da área, em especial as que buscam a representação de conteúdos oriundos de decisões judiciais, onde o *juridiquês* é mais disseminado, valham-se tão somente das fontes terminológicas efetivas da área jurídica (dicionários e enciclopédias especializados, textos legais e textos doutrinários) pois estes visam a sedimentar os conceitos da área e a estabelecer a forma técnica pela qual devem ser oficialmente representados.

4 Conclusão

O discurso do Judiciário possui natureza eminentemente argumentativa no intuito de convencer, demonstrar, provar – seja na petição inicial e na contestação para convencimento do juiz singular, seja na decisão monocrática para o convencimento das partes, e seja ainda nas razões e contrarrazões de um recurso para convencimento do juiz relator, ou mesmo no voto do acórdão em que o relator visa a convencer os seus pares, as partes e as instâncias recursais superiores. Essa natureza eminentemente argumentativa se revela no fato de haver decisões sobre uma mesma situação que chegam a entendimento opostos em virtude de diferentes argumentações que, por sua vez, sustentam-se em distintas premissas metodológicas (VIGNAUX, 1979).

Nesse contexto, o necessário uso de linguagem técnica em peças processuais não é pressuposto para – e, ao contrário, se opõe a – o emprego do *juridiquês* pois a linguagem técnica pauta-se, como já afirmado,

pela busca pela univocidade, pela desambiguação, valendo-se de recursos de clareza, objetividade e principalmente de precisão. A essa precisão terminológica aliam-se a coesão e a coerência textuais bem como a correção gramatical. Por meio da coesão textual garante-se a concetenação, o encadeamento das palavras e frases, dando-lhe fluidez, textura e por meio da coerência garante-se uma conexão conceitual que leve à estruturação do sentido (FÁVERO, 1991, p. 12-13, 59). Já a correção gramatical leva àquilo que Nalini (1994, p. 125) denomina como “o bom uso da palavra e a correta interpretação”.

Se, como argumentado, essa linguagem possui característica altamente especializada, com uma terminologia inequivocamente sedimentada a partir de cuidadosas definições legais e doutrinárias, não faz sentido o uso de um jargão emocional, hermético, ambíguo[5], permeado de estrangeirismos e neologismos, e usado de forma prolixa que põe em risco a coesão e a coerência textuais e mesmo a correção gramatical, truncando o processo comunicativo e dificultando a consecução da justiça e da publicidade em um Estado democrático.

Como destacam Guimarães e Santos (2016, p. 40) :

Idos são os tempos em que o Direito era uma seara de latinismos, construções barrocas e textos herméticos em um *juridiquês* artificial e excludente pois, mormente em tempos de amplo acesso à informação, necessário se torna garantir que essa informação não apenas esteja disponível, mas, e principalmente, que seja compreensível ao cidadão, pois a atuação do judiciário, por ser pública, é mantida pela sociedade e a ela se destina.

Desse modo, cabe a todos combater a figura esdrúxula do *juridiquês* em prol de a linguagem jurídica clara, correta, concisa e precisa (ANDRADE, 2021). E esse combate incia-se nas instâncias de formação jurídica (faculdades de Direito, cursos preparatórios para concurso etc), bem como pela produção e disseminação de compêndios de terminologia jurídica para que se possa garantir a univocidade e a precisão almejadas.

Ao bibliotecário da área, notadamente em seu papel educativo, cabe estar atento às eventuais variantes de *juridiquês* de um termo jurídico, muitas vezes alertando, nos SOC que eventualmente organizar, para a possibilidade de ocorrência de tais termos não autorizados para fins de pesquisa, sempre conduzindo o pesquisador a uma busca que considere e prestigie a efetiva terminologia da área.

Nunca é demais recordar que a formação jurídica suplementar e a busca pelo contato com peças processuais onde possa ocorrer o *juridiquês* é sempre útil ao bibliotecário como subsídio a uma maior precisão no processo de representação do conteúdo dos documentos.

Notas

[1] Citação cuja autoria é controvertida: Leonardo da Vinci? Clare Boothe Luce? Leonard Thiessen? Elizabeth Hillyer? William Gaddis? Eleanor All? Apple Computer Company? Anonymous? (<https://quoteinvestigator.com/2015/04/02/simple/>)

[2] Ementa: Determina a reprodução / tradução da sentença em linguagem coloquial para compreensão da parte interessada que integrar processo judicial, alterando a redação do artigo 458 do Código de Processo Civil, que dispõe sobre as decisões judiciais. No texto do projeto, tem-se, no inciso IV, que “a reprodução do dispositivo da sentença em linguagem coloquial, sem a utilização de termos exclusivos da linguagem técnico-jurídica e acrescida das considerações que a autoridade Judicial entender necessárias, de modo que a prestação jurisdicional possa ser plenamente compreendida por qualquer pessoa do povo”. E melhor esclarecendo, dispõe o § 1º que “A utilização de expressões ou textos em língua estrangeira deve ser sempre acompanhada da respectiva tradução em língua portuguesa, dispensada apenas quando se trate de texto ou expressão já integrados à técnica jurídica”.

[3] Essa expressão deriva-se do termo *plain English* como destaca Garner (200, p. xiv): “Actually, plain English is typically quite interesting to read. It’s robust and direct—the opposite of gaudy, pretentious language. You achieve plain English when you use the simplest, most straightforward way of expressing an idea. You can still choose interesting words. But you’ll avoid fancy ones that have everyday replacements meaning precisely the same thing”. Colares (2010, p. 10), por sua vez, ressalta a existência, desde a década de 1970 do século XX, do *Plain language movement*, mais especialmente nos Estados Unidos, na Inglaterra, na Suécia, na Alemanha e em Israel, que passou a “questionar o uso da linguagem em contextos institucionais e o abuso no uso da linguagem pelos detentores do poder, no exercício de suas atividades profissionais” como forma de preservar “o direito do cidadão de entender e ser entendido de modo [...] visando a defender “o direito do cidadão comum de entender e ser entendido”.

[4] Pereira Junior (2020, p. 1), de forma bastante firme e direta, chega a afirmar: “Na verdade, toda essa parafernália verbal é sintomática e se resume a uma questão: carência de leitura qualificada. Os profissionais estão se atendo, quando muito, à leitura técnica, achando que isso é suficiente para o melhor domínio do vernáculo e, na verdade, há engano nisso, pois a leitura técnica é árida, não cria estilo próprio e nem expande o leque verbal do leitor, pois circunscrito que é àquelas formas de expressão que evidenciam um único sotaque”.

[5] “[...] teóricos da filosofia e semanticistas também se preocuparam em estudar as disfunções da linguagem e da lógica do discurso humano relacionados à ambiguidade, à vagueza e outras imprecisões” (BARROS, 2016, p. 146).

Referências

ABREU, Antônio Suárez. **A arte de argumentar:** gerenciando razão e emoção. 3.ed. Cotia: Ateliê Editorial, 2001.

ALVARENGA, Lídia. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli**, v. 15, p.18-40, 2003.

AMB. Associação dos Magistrados Brasileiros. **AMB lança campanha para simplificar linguagem jurídica.** Rio de Janeiro, 11 de agosto de 2005. Disponível em: <https://www.amb.com.br/amb-lanca-campanha-para-simplificar-linguagem-juridica/>. Acesso em: 23 maio 2023.

ANDRADE, Valdeciliana da Silva Ramos. **O juridiquês e a linguagem jurídica:** o certo e o errado no discurso. Direito legal, 24 de fevereiro de 2021. Disponível em: <https://direitolegal.org/artigos/o-juridiquês-e-a-linguagem-juridica>. Acesso em: 23 maio 2023.

BARROS, Lucivaldo. **Teoria e prática da pesquisa em jurisprudência:** da procura e uso da informação para sustentar teses e estudos jurídicos. Belo Horizonte: Forum, 2016.

BELÉM, Mariana. A simplificação da linguagem jurídica como meio de aproximação do cidadão à justiça. **Revista Jurídica da Seção Judiciária de Pernambuco**, Recife, n. 6, p. 313–320, 2013.

BOCCATO, Vera Regina Casari. A linguagem documentária como instrumento de organização e recuperação da informação. In: HOFFMANN, W. A. M., FURNIVAL, A. C. (Orgs.) **Olhar:** ciência, tecnologia e sociedade. São Paulo: Pedro e João Ed., CECH-UFSCar, p.269-78, 2008.

BULHÕES, Eliane Simões Pereira. O tradicionalismo na linguagem jurídica. **Signo**, Santa Cruz do Sul, RS, v. 33, n. 55, p. 66–77, jul/dez. 2008. Disponível em: <https://online.unisc.br/seer/index.php/signo/article/view/543>. Acesso em: 23 maio 2023.

CABRÉ, Maria Teresa. La terminología hoy: concepciones, tendencias y aplicaciones. **Ciência da Informação**, v. 24, n. 3, 1995. DOI: 10.18225/ci.inf.v24i3.567. Acesso em: 10 jun. 2022.

CAMPESTRINI, Hildebrando. **Como redigir ementas.** São Paulo: Saraiva, 1994.

CARMO, Rui do. Os tribunais administram a justiça em nome do povo. In: Simões, Rita Basílio; Camponez, Carlos; Perixinho, Ana Teresa (Org.). **Justiça e comunicação:** o diálogo (im)possível. Coimbra: Imprensa da Universidade de Coimbra, 2013. p. 147-155.

COLARES, V. (Org). **Linguagem & direito.** Recife: Editora Universitária UFPE, 2010.

FAVERO, Leonor Lopes. **Coesão e coerência textuais.** São Paulo: Ática, 1991.

FOGAÇA, M. G.; CERVANTES, B. M. N.; CERVANTES, B. M. N. Estudo comparativo de terminologias da área jurídica em sistemas de classificação bibliográfica à luz do código de processo civil brasileiro. **Revista Brasileira de Educação em Ciência da Informação**, v. 8, 2021. DOI: 10.24208/rebecin.v8i.276 Acesso em: 10 jun. 2022.

GARNER, Bryan. **Legal writing in plain english.** Chicago: University of Chicago, 2001.

- GUIMARÃES, José Augusto Chaves. **Elaboração de ementas jurisprudenciais: elementos teórico-metodológicos**. Brasília: Conselho da Justiça Federal, 2004.
- GUIMARÃES, José Augusto Chaves; SANTOS, João Carlos Gardini. Ementa jurisprudencial como resumo informativo em um domínio especializado: aspectos estruturais. **Brazilian Journal of Information Science: research trends**. v.10, n.3, p.32-43, 2016.
- HJORLAND, Birger. Knowledge Organization Systems (KOS). In: HJORLAND, B. **Lifeboat of Knowledge Organization**. 2008. Disponível em: https://arkiv.iva.ku.dk/kolifeboat/CONCEPTS/knowledg_e_organization_systems.htm
- Hjørland, Birger. Domain analysis in information science: eleven approaches-traditional as well as innovative. **Journal of Documentation**, v. 58, n. 4, p. 422-462. 2002.
- HODGE, Gail. **Systems of knowledge organization for digital libraries: beyond traditional authority files**. Washington, D.C.: The Digital Library Federation Council on Library Information Resources, 2000. Disponível em: <http://old.diglib.org/pubs/dlf090/dlf090.pdfgt.>,
- LARA, Marilda Lopes Ginez. Propostas de tipologias de KOS: uma análise das referências de formas dominantes de organização do conhecimento. **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf, Florianópolis**, v. 20, n. esp. 1, Fev. 2015. Disponível em: <https://repositorio.usp.br/item/002732707>. Acesso em: 23 maio 2023.
- MAZZOCCHI, F. Knowledge organization systems (KOS): an introductory critical account. **Knowledge Organization**, v. 45, n. 1, p. 54-78, 2018. Disponível em: <https://www.isko.org/cyclo/kos> . Acesso em: 23 maio 2023.
- MOISÉS JUNIOR, Jorge. A simplificação da linguagem jurídica e o artigo 45 do Código de Ética a OAB. In: URBANO, Alexandre Figueira de Andrade *et al.* (Org.) **Advocacia, ética e o exercício da profissão**. Belo Horizonte: Del Rey, 2015. p.264
- NALINI, José Renato. A formação do juiz. In: NALINI, José Renato. (Org.) **Formação jurídica**. São Paulo: Revista dos Tribunais, 1994. P.6-148.
- OATES, Laurel Currie; ENQUIST, Anne. **Just writing: grammar, punctuation, and style for the legal writer**, New York: Aspen Publishers. 2009.
- OLIVEIRA, Rodrigo Rios Farias de. A educação e o conhecer das normas: consequências da linguagem do jurídiquês. **Brazilian Journal of Development**, Curitiba, v.7, n.3, p. 30980-30990 mar 2021.
- PEREIRA JUNIOR, João. **Quando o "juridiquês" se torna brega. Aqui acontece**. Penedo (AL): Organização Hélio Lopes, 22 de setembro de 2010. Disponível em: <http://www.aquiacontece.com.br/artigo/joao-pereira-junior/22/09/2010/quando-o-juridiques-se-torna-brega/84>. Acesso em: 23 maio 2023.
- RODRIGUES, Maria da Conceição Campinha. **Linguagem, discurso e direito: algumas questões de linguística jurídica**. Revista do Ministério Público, n.111, p.5-36, jul.-set. 2007.
- SALES, R. Suportes teóricos para pensar linguagens Documentárias. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.5, n. 1, p 95-114, jul/dez. 2007.
- SALGADO, Eneida Desirée. Princípio da publicidade. In: **Enciclopédia jurídica da PUCSP**. Abril de 2017. Disponível em: <https://enciclopediajuridica.pucsp.br/verbete/37/edicao-1/principio-da-publicidade>. Acesso em: 23 maio 2023.
- JARGÃO. In: Significados. [Sl:sn], [2023]. Disponível em: <https://www.significados.com.br/jargao/>
- SILVEIRA, Gustavo Siqueira. Carta Magna não é sinônimo de Constituição: uma análise do conceito no Brasil e uma breve história do documento medieval. **Revista direito e práxis**, 2021 (Ahead of print). DOI: 10.1590/2179-8966/2021/59938. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/revistaceaju/article/view/59938>. Acesso em: 23 maio 2023.
- SOERGEL, Dagobert. **The representation of Knowledge Organization Structure (KOS) data: a multiplicity of standards**. Roanoke: JCDL, 2001. Disponível em: <http://www.dsoergel.com/cv/B75.pdf>. Acesso em: 23 maio 2023.
- VICKERY, Brian. C. Knowledge representation: a brief review. **Journal of documentation**, v.42, n.3, sept. 1986. p. 145-59.
- VIGNAUX, Georges. Argumentation et discours de la norme. **Langages**, v.12, n.53, p.67-85, 1979.
- WLASSAK, Thomas. O princípio da publicidade.: Considerações sobre forma e conteúdo. **Revista Jus Navigandi**, v.7, n.60, 1 nov. 2002. Disponível em: <https://jus.com.br/artigos/3425>. Acesso em 13 jun. 2022.

Apêndice A - Quadro 1

Quadro 1. *Comparação entre a terminologia jurídica e o juridiquês*

<i>Terminologia jurídica</i>	<i>Fundamento</i>	<i>Jargão (juridiquês)</i>	<i>Explicação</i>
Decisão recorrida	Art. 458 e 994, CPC	<i>Decisum vergastado</i> <i>Decisum fustigado</i> , <i>Decisum humilhado</i> <i>Decisum pisoteado</i> <i>Decisum objugado</i>	Latinismo e Termos metafóricos (Ex: vergastar: chicotear)
Constituição Federal ou Constituição da República Federativa do Brasil	Cf. disposto em seu Preâmbulo	<i>Carta Política Democrática</i> <i>Lei Maior</i> <i>Carta Magna</i>	Termos metafóricos
Réu revel	Art. 344 do CPC	<i>Réu silente</i>	Termo metafórico
Petição inicial	Art. 319, CPC	<i>Peça: ancilar, atrial, autoral, de arranque, de ingresso, de intróito, dilucular, ovo, exordial, gênese, inaugural, incoativa, introdutória, pórtico, preambular, prefacial, preludial, primeva, primígena, prodrômica, proemial, prologal, umbilical, vestibular, madrugadora, dilucular, prodrômica</i>	Termos metafóricos
Recurso extraordinário	Art. 102, III da CF	<i>Apelo extremo</i>	Termo metafórico
Supremo Tribunal Federal	Art. 101 a 103, C.F.	<i>Excelso sodalício</i>	Termo metafórico
Inquérito policial	Art. 4º a 23 do CPP	<i>Caderno indiciário</i>	Termo metafórico
Código de Processo Civil ou Código de Processo Penal	Lei 13.105/ 2015 Decreto-lei 3.689/ 1941	<i>Repositório adjetivo</i>	Termo metafórico
Consolidação das Leis do Trabalho ou C.L.T.	Decreto-lei 5.452/1943	<i>Digesto obreiro</i> <i>Estatuto consolidado</i>	Termo metafórico
Instituto Nacional de Seguro Social (INSS)	Decreto 99.350/1990	<i>Autarquia Ancilar</i>	Termo metafórico
Reclamação trabalhista	Arret. 840 da C.L.T.	<i>Apelo obreiro</i>	Termo metafórico
Contrato de trabalho	Art. 442 da CLT	<i>Pacto laboral</i>	Termo metafórico
Perito	Art 156 a 158 do CPC	<i>Vistor</i>	Termo metafórico
Mandado de Segurança	Art. 5º., LXIX da CF	<i>Writ</i>	Estrangeirismo
Habeas Corpus	Art. 5º., LVXIII da C.F.	<i>Remédio heroico</i>	Termo metafórico
Tribunal	Art. 101 a 126, C.F.	<i>Corte</i>	Termo metafórico Anglicismo
Denúncia	Art. 24, C.P.P.	<i>Exordial increpatória</i>	Termo metafórico
Ônus da prova	Art. 373, C.P.C.	<i>Onus probandi</i>	Latinismo
Litigante de má-fé	Art. 79 a 81 do C.P.C.	<i>Improbis litigator</i>	Latinismo
Ministério Público	Art. 127 a 130 da C.F.	<i>Parquet</i>	Estrangeirismo
Desprezível		<i>Despiciendo</i>	Termo arcaico
Caracteriza carência de direito		Caracteriza <i>flagrante</i> carência de direito	Uso de adjetivo com tom emocional e subjetivo
Bastante comprovado		<i>sobejamente</i> comprovado	Uso de advérbio com tom emocional e subjetivo
Entendimento da turma		Entendimento <i>turmário</i>	Neologismo
Advogado	Art. 133 da C.F.	Paracleto	Termo arcaico
Juízo de primeiro grau Juízo de primeira instância	Primeira Instância ou Primeiro Grau: porta de entrada do Poder Judiciário, por onde se inicia a maior parte dos processos que são analisados e julgados por um juiz.	juízo <i>planicial</i>	
Sentença de prisão / Sentença penal condenatória	Art. 387 C.P.P.	decreto <i>ergastulário</i>	Termo metafórico

Fonte: Elaborado pelos autores.

La garantía indígena: aportes para su aplicación al patrimonio cultural de los pueblos aborígenes latinoamericanos

The indigenous warrant: contributions to its application to the cultural heritage of latin american aboriginal peoples

Mario Barité (1), Sonia Matos Moutinho (2)

(1) Universidad de la República, Uruguay, mario.barite@fic.edu.uy, (2) Universidade Estadual Paulista (UNESP), Marília-SP, Brasil, sonia.matos-moutinho@unesp.br.

Resumo

A situação de muitas culturas indígenas na Austrália, América do Norte, Central ou do Sul pode ser descrita como marginalizada, subjugada ou minorizada. Dentro da Ciência da Informação, a questão indígena é cada vez mais levantada em textos que identificam elementos críticos como os vieses às concepções colonialistas implantados nos esquemas dos sistemas classificatórios. **Objetivo:** contribuir para a sistematização do crescente corpo de literatura sobre garantia indígena na Organização do Conhecimento, a partir de uma perspectiva crítica, a fim de promover uma abordagem latino-americana. **Procedimentos metodológicos:** Revisão da literatura sobre garantia indígena. Acesso a documentos originais. Leitura e interpretação crítica do corpus a partir de quatro categorias de análise. **Resultados:** As ideias mais significativas advindas da leitura crítica do corpus distribuíram-se nas seguintes categorias de análise: critérios de organização do conhecimento das culturas indígenas; caracterização da garantia indígena como forma de garantia cultural; processos de institucionalização das culturas indígenas; Lista de métodos úteis para a aplicação da garantia indígena. **Conclusão:** Confirma-se que existe uma bateria consistente de metodologias para orientar a aplicação da garantia indígena. Ressalta-se a urgência de promover a produção acadêmica sobre o tema da Organização do Conhecimento na América Latina, considerando o conjunto -até agora extremamente reduzido- de precedentes, todos eles brasileiros, que começam a liderar o caminho.

Palavras-chave: Culturas indígenas. Garantia cultural. Garantia indígena.

Abstract

The situation of many indigenous cultures in Australia, North, Central or South America could be described as marginalized, subjugated or minorised. Within Information Science, the indigenous question is increasingly raised in texts that identify critical elements such as biases towards colonialist conceptions implanted in the schemes of classification systems. **Objective:** to contribute to the systematization of the growing body of literature on indigenous warrant in Knowledge Organization, from a critical perspective, in order to promote a Latin American approach. **Methodological procedures:** Literature review on indigenous warrant. Access to original documents. Reading and critical interpretation of the corpus based on four categories of analysis. **Results:** The most significant ideas arising from the critical reading of the corpus were distributed in the following categories of analysis: criteria for the organization of knowledge of indigenous cultures; characterization of the indigenous warrant as a form of cultural warrant; institutionalization processes of indigenous cultures; List of useful methods for the application of the indigenous warrant. **Conclusion:** It is confirmed that there is a consistent battery of methodologies to guide the application of the indigenous warrant. The urgency of promoting academic production on the subject from the Knowledge Organization in Latin America is highlighted, considering the set -until now extremely reduced- of precedents, all of them Brazilian, which have begun to lead the way.

Keywords: Indigenous cultures. Cultural warrant. Indigenous warrant.

Resumen

La situación de muchas culturas indígenas en Australia, América del Norte, Central o del Sur, podría calificarse como marginalizada, sojuzgada o minorizada. Dentro de la Ciencia de la Información, la cuestión indígena se plantea en forma creciente en textos que identifican elementos críticos como los **sesgos** hacia concepciones colonialistas implantadas en los esquemas de sistemas de clasificación. **Objetivo:** contribuir a la sistematización del cuerpo creciente de literatura sobre garantía indígena en Organización del Conocimiento, desde una perspectiva crítica, a efectos de promover un enfoque latinoamericano. **Procedimientos metodológicos:** Revisión de literatura sobre garantía indígena. Acceso a los documentos originales. Lectura e interpretación crítica del corpus en base a cuatro categorías de análisis. **Resultados:** Las ideas más significativas surgidas de la lectura crítica del corpus, se distribuyeron en las siguientes categorías de análisis: criterios para la organización del conocimiento de las culturas indígenas; caracterización de la garantía indígena como forma de la garantía cultural; procesos de institucionalización de las culturas indígenas; Relación de métodos útiles para la aplicación de la garantía indígena. **Conclusión:** Se confirma que existe una batería consistente de metodologías para orientar la aplicación de la garantía indígena. Se remarca la urgencia de promover la producción académica en la materia desde la Organización del Conocimiento en América Latina, considerando el conjunto -hasta ahora extremadamente reducido- de antecedentes, todos ellos brasileños, que han comenzado a marcar el camino.

Palabras clave: Culturas indígenas. Garantía cultural. Garantía indígena.

1 Introducción

La situación de muchas culturas indígenas tanto en América del Norte, como en América Central o del Sur, pero también en Australia u otras regiones, una vez superados los lamentables episodios de la conquista por parte de los imperios colonialistas europeos o sus herederos criollos, podría calificarse como penosa. Siguen siendo todas ellas culturas sojuzgadas o marginadas, a las que incluso se les niega o retacea el derecho a conservar sus propios territorios, de los cuales muchas veces son expulsados. Sus integrantes viven así una especie de existencia paralela a la de los descendientes de los colonizadores, con escasas posibilidades de integración real.

Sin embargo, la mayor fortaleza de estas comunidades aborígenes ha sido su notable capacidad de resistencia, pues aún confinadas en el corazón mismo de una cultura dominante, han conseguido mantener su identidad, su lengua, sus valores y creencias y sus interpretaciones del mundo.

En las universidades, el tratamiento tradicional de la cuestión indígena (una visión construida desde las bases clásicas del pensamiento occidental, que más o menos puede entenderse como la visión de la cultura académica hegemónica), ha sido contestado por un conjunto de intelectuales dispuestos a instalar nuevos posicionamientos. Entre los autores más respetados, destacan el antropólogo y educador Darcy Ribeiro, quien dividió su vida entre la política, la educación y el desarrollo pionero de los estudios sobre los indígenas brasileños, y el del ghano-canadiense Georges Dei, quien ha promovido con firmeza marcos teóricos que parten de la necesidad de decolonizar el pensamiento académico tradicional, considerando que el conocimiento indígena está basado fundamentalmente en la experiencia, y en miradas locales, holísticas y relacionadas con la idea de resistencia (DEI, 2000; DEI; ASGHARZADEH, 2001).

Dentro de la Ciencia de la Información (CI), y más específicamente en la Organización del Conocimiento (OC), la cuestión indígena se viene planteando de forma creciente desde los años '70 del siglo pasado, particularmente en textos que identifican elementos críticos (YEH; FROSIO, 1971), en especial los sesgos hacia concepciones colonialistas implantadas en los esquemas de algunos sistemas de clasificación (PACEY, 1989).

La atención a los temas indígenas se incrementó notoriamente en lo que va del siglo veintiuno, y la producción más significativa proviene de Canadá, Estados Unidos y Australia. En Canadá destacan las contribuciones de la investigadora Ann. M Doyle en trabajos individuales o en coautoría (DOYLE, 2006, 2013; DOYLE; LAWSON; DUPONT, 2015; BURNS; DOYLE; JOSEPH; KREBS, 2009; WEBSTER;

DOYLE, 2008), y de Sharon Farnel (FARNEL *et al.*, 2016; FARNEL, 2021).

En Estados Unidos prevalecen los trabajos críticos sobre los esquemas de clasificación correspondientes a culturas aborígenes en diversos sistemas tales como la lista Library of Congress Subject Headings, o la Library of Congress Classification (YEH; FROSIO, 1971; TOMREN, 2004; LEE, 2011, CARRON, 2019). También se registran buenos estudios conceptuales (OLSON, 1999; AGRAWAL, 2002; LITTLETREE; BELARDE-LEWIS; DUARTE, 2020). Destaca la publicación de un número temático de la revista *Cataloging & Classification Quarterly* (el número 5/6 del volumen 53, correspondiente al año 2015), coordinado por Ann M. Doyle y Cheryl Metoyer, dedicado a la organización del conocimiento indígena.

Australia es otro país con una producción regular en torno a la organización del conocimiento indígena (NAKATA; LANGTON, 2005; EAST, 2008; NAKATA *et al.*, 2008; GODBOLD, 2009; THORPE; GALASSI, 2014).

En América Latina, pese a contar con un extraordinario legado de las culturas originarias (particularmente la inca, la maya, la azteca y la guaraní), desde la CI y particularmente desde la OC existen escasos antecedentes de producción académica sobre la organización de los documentos, los objetos y la información que constituyen su historia. Cabe mencionar antes que nada, las contribuciones pioneras de França y Silveira (2004), de Maimone y Matos (2019), y de Albuquerque y Moraes (2023).

En el primero de los documentos, se realiza una búsqueda exploratoria en catálogos en línea de acceso público sobre conceptos antropológicos e dados sobre grupos indígenas do Brasil, con el objeto de identificar elementos relevantes para una representación bibliográfica “digna” de las obras literarias indígenas do Brasil (FRANCA; SILVEIRA, 2004). El segundo se ocupa de identificar las causas y las consecuencias de la escasa presencia de la temática indígena en la literatura de la CI, y en sus conclusiones los autores establecen que:

Ainda que seja visível uma melhora significativa no âmbito da documentação, biblioteconomia e museologia, é ingênuo fechar os olhos para problemas políticos que ainda existem, como a demarcação de terras, direitos sociais e questões ambientais [...] Quando em contato com a questão indígena, o bibliotecário, documentalista ou museólogo deve se afastar do etnocentrismo, desse ímpeto colonizador que enquadra o outro como exótico, não civilizado (MAIMONE; MATOS, 2019, p. 53).

El tercer trabajo analiza la producción científica publicada en revistas brasileñas de CI sobre las culturas indígenas, y cuestiona su tratamiento desde la perspectiva de la OC. En las conclusiones, sus autores establecen que, aunque el corpus reunido no trata

directamente de OC, demuestra potencial para caracterizar aspectos de comprensión de las culturas indígenas, lo que puede auxiliar en la elaboración de sistemas de organización del conocimiento, en adelante, SOC (ALBUQUERQUE; MORAES, 2023).

Por último, corresponde destacar la producción de Mello, quien se ha dedicado a poner en valor los archivos sobre culturas indígenas brasileñas, como una forma de garantizar la visibilidad y la supervivencia de las mismas (v. por ejemplo, MELLO, 2014).

Salvas estas destacadas excepciones, ha faltado en los países latinoamericanos una consideración dedicada a las cuestiones propias de las culturas indígenas desde la CI. En países como Estados Unidos, Canadá y Australia se verifican en cambio, intentos serios por lograr una aproximación al fenómeno indígena como algo integral, a partir de algunas de sus peculiaridades: la tradición oral por sobre la escrita, la diversidad de objetos que pueden ser objetos de archivo o de intervención museística, los condicionamientos provocados por la influencia de la cultura dominante y de las tradiciones académicas funcionales a ella, y la necesidad de establecer formas participativas de gestión y curaduría de bibliotecas, archivos y museos.

La representación temática de contenidos, esa área delicada que tiene el “poder de nominar” y que al decir de Olson es un “acto político” en entornos culturales (OLSON, 2001) constituye también una responsabilidad para con las culturas minoritarias y marginales. En este ámbito el concepto crucial es el de garantía, entendida como el criterio de justificación de las expresiones destinadas a la clasificación, la indización y la recuperación temática de documentos e información.

En el siguiente apartado, se revisará la noción de garantía cultural como pauta metodológica genérica necesaria para la representación temática de las culturas originarias.

3. Objetivos

El objetivo general de este trabajo es contribuir a la sistematización del cuerpo creciente de literatura sobre garantía indígena en OC, desde una perspectiva crítica, a efectos de promover un enfoque latinoamericano.

Se establecen asimismo tres objetivos específicos:

- i) Identificar los presupuestos y las cuestiones principales que deben tenerse en cuenta para establecer el alcance del concepto de garantía indígena;
- ii) Favorecer la integración de los conocimientos acumulados sobre garantía cultural y garantía indígena, de modo de tender en forma progresiva, a una teoría general de las garantías;

- iii) Aportar sustrato teórico para el desarrollo y la organización temática de bibliotecas, archivos, museos y otros servicios de información especializados en culturas aborígenes latinoamericanas.

2 La garantía cultural

Con una formulación que ha sido citada en forma recurrente hasta hoy, informa Beghtol que

the warrant of a classification system can be thought of as the authority a classificacionist invokes first to justify and subsequently to verify decisions about what classes/concepts to include in the system, in what order classes/concepts should appear in the schedules, what unit classes/concepts are divided into, how far subdivision should proceed, how much and where synthesis is available, whether citation orders are static or variable and similar questions (BEGHTOL, 1986, p. 110-111).

Sistemas universales ampliamente utilizados como Library of Congress Classification o la Dewey Decimal Classification, y otros especializados como el de Medical Subject Headings, utilizan la garantía literaria como criterio principal para organizar su terminología desde hace décadas. Este criterio se basa en un dato empírico: si se cuenta con una cantidad suficiente de obras publicadas sobre un asunto, entonces ese asunto cuenta con validez suficiente para ser incorporado a los esquemas de un SOC. En el caso de la Dewey Decimal Classification, ese número ‘suficiente’ es veinte (CUSTER, 1958; BEALL, 2003). Para paliar las insuficiencias de la garantía literaria, particularmente por su carencia de base semántica, más de una veintena de garantías han sido propuestas hasta el presente (MARTÍNEZ ÁVILA; BUDD, 2017; BARITÉ, 2017; ZAMBONI, 2018; LEE, 2021). Lo que distingue a una garantía de otra es el criterio utilizado como pauta metodológica, más allá de los fundamentos teóricos que los diferentes autores le han atribuido a cada una.

Todas las garantías procuran asegurar consistencia y homogeneidad terminológica a los esquemas de clasificación, para mejorar la comunicación con los usuarios que realizan búsquedas temáticas. La utilización consistente de una garantía para la selección de los términos y las relaciones entre términos que se organizarán en la estructura conceptual del SOC provee un determinado sentido, una lógica nocional, a la representación temática de un dominio. Para Beghtol, la garantía cultural es uno de los cuatro tipos de garantía semántica, junto con la literaria, la científico-filosófica, la educacional y la cultural y expresa:

cultural warrant posits that every classification system is based on the assumptions and preoccupations of a certain culture, whether the culture is that of a country, or of some smaller or larger social unit (e.g. ethnic group, academic discipline, arts domain, political party, religion and/or language) (BEGHTOL, 2002a, p. 45).

Más allá de reconocer la acuñación y los antecedentes del término en Hulme, Austin y Joel Lee, será Beghtol quien más se preocupará por afinar el concepto, ofrecer ejemplos, y profundizar en sus postulados teóricos. En su trabajo canónico de 1986 señala que la garantía cultural es un ‘concepto paraguas’ que permite explicar los cambios que se producen en las concepciones y usos de las garantías literaria, científica-filosófica y educacional (BEGHTOL, 1986, p. 121). En 2002 Beghtol vuelve sobre la garantía cultural en una ponencia presentada ante la 7° Conferencia internacional de la International Society for Knowledge Organization (ISKO). En la ponencia Beghtol, además de ofrecer una definición precisa del concepto, transcrita más arriba, establece dos premisas importantes:

- i) que todo SOC es más probable que pueda ser útil y apropiado para los miembros de una cultura que para quienes comparten una cultura diferente; y, ii) que por derivación, los SOC pueden no reconocer o excluir elementos de esa cultura diferente, lo que conspira contra la integración del conocimiento a través de las fronteras geográficas, culturales y lingüísticas (BEGHTOL, 2002a, p. 45).

Hay tres elementos cruciales que cruzan transversalmente estas premisas. El primero es que en cualquier sociedad hay una cultura hegemónica, dominante y eventualmente mayoritaria, así como otras minoritarias, que tanto pueden ser autónomas, mantener hasta cierto punto su integridad, o estar sofocadas o acosadas en mayor o menor medida.

Las formas de relación que se establecen entre una cultura dominante y el resto de las culturas que conviven en un mismo espacio geográfico o social pueden darse a través de distintas modalidades: la convivencia pacífica, la indiferencia, ciertas formas de control social o informativo, el sojuzgamiento o incluso el intento de erradicación o aniquilación. Esto quiere significar que una cultura dominante suele tolerar tácitamente la existencia de una serie de culturas minoritarias (por ejemplo, no molesta a los centros de poder que existan los cultores del ajedrez, aunque también el ajedrez ha sido protagonista de varios episodios de la guerra fría y de culturas enfrentadas). Por otra parte, la cultura dominante puede impulsar asimismo políticas activas de integración, aculturación o destrucción. Lo que transforma a una cultura minoritaria en una cultura minorizada no es ningún dato cuantitativo, sino la calidad del movimiento de péndulo hacia la convivencia mutuamente aceptada o hacia el conflicto. Ello tiene consecuencias en la mayor o menor vulnerabilidad de los integrantes de una cultura minoritaria, en términos políticos, económicos y sociales.

El segundo elemento, nada menor en sus repercusiones, es que cada persona que vive en una sociedad regida por una cultura dominante, a su vez forma parte del devenir de una cantidad de culturas minoritarias o minorizadas

(profesionales, deportivas, religiosas, étnicas, etc.). La intensidad del sentimiento de pertenencia o de identificación externa con cada una de esas culturas minoritarias o minorizadas, así como la reacción de la cultura dominante en cada caso, es lo que determina y puede condicionar la situación social de una persona.

El tercer elemento es que para definir estrategias que aporten soluciones para contribuir a la integridad de culturas específicas desde las bibliotecas, los museos y los archivos, deben definirse políticas sustentables que respondan del modo más abierto y libre a los intereses, los fundamentos y necesidades de la cultura minoritaria, y que no transformen el espacio intelectual y material atribuido en una especie de ghetto. La garantía cultural constituye sólo uno de los ámbitos donde se exigen esas políticas y estrategias: su jurisdicción se establece en relación a la representación temática de la documentación y la información de una cultura minoritaria.

En bibliotecas que frecuentan personas con diferentes culturas, el sentido cultural específico de una clasificación o de un sistema de indización puede provocar problemas de mediación y recuperación de información, dificultades de comprensión, invisibilización de contenidos culturales, e incluso quejas por la utilización de lenguaje discriminatorio o peyorativo. Lo expuesto es particularmente importante en las bibliotecas públicas, por tratarse de instituciones

de cohesión social, que pueden actuar como factor integrador de minorías e instrumento difusor de una cultura democratizadora, en tanto que se use como un espacio favorable para el desarrollo personal y colectivo; en esta práctica cobra sentido la afirmación de que la biblioteca pública se constituye en un compromiso de voluntad política (KOLENAS 2008, p. 111).

Y esto es así, porque al decir de Puente Hernández (2013), la biblioteca debe “integrar a las distintas culturas en procesos de interculturalidad y no desintegrarlas con prácticas discriminatorias” (PUENTE HERNÁNDEZ, 2013, p. 150).

Tampoco debería establecerse una biblioteca, un museo o un archivo especializado en la cultura de una comunidad minoritaria, organizando sus documentos desde las perspectivas de la cultura dominante. No es posible resolver esos problemas si no se cuenta con orientaciones éticas que, preferentemente, puedan acordarse con los interesados (GUIMARÃES, 2006; GUIMARÃES; FERNÁNDEZ-MOLINA; PINHO; MILANI, 2008).

La garantía cultural debe apuntar entonces, a asumir las perspectivas culturales de los usuarios de las culturas minoritarias de una manera satisfactoria para ellos, a través de la selección de términos de clasificación e indización que permitan la recuperación de la información que necesiten, en el contexto y en el lenguaje de esos usuarios (ZAMBONI; FRANCELIN,

2016). Pero esta no es tarea sencilla, sobre todo porque las técnicas y los métodos de análisis e implantación de la garantía cultural todavía están en debate y construcción (WANG, 2019). En ese sentido, afirma Hansson que

the perhaps most complex kind of classificatory warrant, and the one least researched, is cultural warrant. The reason why this is problematic is simple; it targets the classification system in relation to conditions that go beyond its relation to the collections it helps organize as well as beyond the relatively stable structure of the scientific system (HANSSON, 2022, p. 102).

A diferencia de muchas otras garantías, la cultural es una de las que está sujeta a cierto nivel de interpretación y es, por lo tanto, controvertible. No suele ofrecer formulaciones semánticas precisas o indiscutibles, debido a que debe tomar posición en ámbitos donde pesan las creencias, los rituales y los valores de una comunidad. La garantía cultural trabaja con conceptos que se apoyan en valores y en símbolos, atravesados por una dimensión ética.

Para Barité (2017), algunas de las garantías que enumera en su cuadro podrían ser variedades de la garantía cultural: menciona en especial a la garantía ética, la garantía de género y la garantía indígena. En el presente trabajo se establece la hipótesis de que la garantía indígena es una variedad de garantía literaria, porque el conjunto de principios y métodos destinados a la representación temática están fundamentados en valores y símbolos. Si la garantía indígena es una forma de la garantía cultural, le corresponden y valen para la primera todos los presupuestos teóricos y metodológicos que son aplicables a la última.

Por otra parte, el conocimiento sobre poblaciones aborígenes puede ser organizado a partir de la similitud de sus contenidos fundamentales y por tanto, puede establecerse un modelo común a cualquier cultura indígena latinoamericana.

4 Procedimientos metodológicos

Los procedimientos metodológicos que se utilizó para el cumplimiento de los objetivos establecidos es de índole cualitativa y contó con las siguientes fases:

i) Revisión de literatura sobre garantía indígena. Esta revisión se realizó tomando como corpus la colección de las revistas Knowledge Organization Journal, Cataloging & Classification Quarterly, y la serie de Proceedings de las Conferencias Internacionales de la International Society for Knowledge Organization (ISKO), llamada Advances in Knowledge Organization. Reunidos los documentos que arrojó la búsqueda, se analizaron y seleccionaron las referencias bibliográficas pertinentes al tema que se encontraron en dichos documentos. Los resultados

de la revisión se pueden visualizar en la enumeración de referencias bibliográficas al final de este trabajo.

- ii) Acceso a los documentos originales. Se accedió a los documentos seleccionados en la biblioteca de la Facultad de Información y Comunicación de la Universidad de la República del Uruguay, en la biblioteca del grupo de investigación en terminología y organización del conocimiento (GTERM) de dicha facultad, y en archivos disponibles libremente en Internet.
- iii) Lectura e interpretación crítica del corpus. Se utilizó el análisis de contenido como método conforme a las siguientes categorías de análisis:
 - Criterios para la organización del conocimiento de las culturas indígenas.
 - Caracterización de la garantía indígena como forma de la garantía cultural.
 - Procesos de institucionalización de las culturas y los conocimientos indígenas.
 - Relación de métodos mencionados para la aplicación de la garantía indígena.
- iv) presentación textual de los resultados de la interpretación crítica, en función de las categorías de análisis mencionadas.

5 Resultados

Los resultados de la revisión crítica de literatura se distribuyen conforme a las cuatro categorías de análisis establecidas en la metodología.

5.1 Criterios para la organización del conocimiento de las poblaciones indígenas.

Ya en 2006, Doyle había propugnado la necesidad de ‘indigenizar’ a la CI, a través del desarrollo de marcos teóricos y conceptuales que permitieran adaptar las herramientas profesionales a las necesidades y los propósitos de las culturas indígenas (DOYLE, 2006, p. 110). Si Doyle planteaba indigenizar la disciplina, era porque esta claramente no trataba a estas culturas desde sus propias perspectivas. Como señalará una década más tarde en un trabajo en coautoría, las formas de organización de los materiales y documentos indígenas, así como la terminología utilizada para la representación temática provenían de las visiones y los valores de “newcomers to First Nations territories including early anthropologists, missionaries, government agents, and travellers, and not Indigenous perspectives or values” (DOYLE; LAWSON; DUPONT, 2015, p. 111).

En este segundo documento, Doyle, Lawson y Dupont plantean la necesidad de ‘indigenizar’ también la OC, analizando “possible intersections between Indigenous frameworks and the information professions” (DOYLE,

LAWSON Y DUPONT, 2015, p. 115). Y remarcan la bidireccionalidad de estos procesos: “We seek these intersections in order to explore ways in which KO might serve Indigenous interests, and ways in which to Indigenize the discipline of KO itself; this is both a critical and constructive undertaking” (DOYLE, LAWSON; DUPONT, 2015, p. 115). Una manera de desprenderse de las distorsiones provocadas por el análisis realizado desde perspectivas hegemónicas consiste, según estas autoras, en identificar y neutralizar las razones por las cuales se tiende a mostrar al conocimiento indígena como un conocimiento inferior.

Indigenizar la OC supone además no considerar las evidencias materiales e inmateriales de una cultura indígena, -de las que pueden formar parte de una biblioteca, un archivo o museo-, como expresiones del apogeo de pueblos que fueron luego absorbidos por los avances de una mal entendida civilización (como una expresión del pasado en definitiva), sino como pruebas de la vitalidad de culturas que vienen desde el fondo de la historia de nuestros países, y han demostrado una extraordinaria resiliencia, así como un apego extremo por la defensa de su identidad y sus tradiciones. Esto implica, en palabras de Littletree, Belarde-Lewis y Duarte (2020) practicar una intervención epistemológica en donde los artefactos, las reliquias y los documentos indígenas no sean vistos como objetos individuales, sino como elementos integradores de una tradición y una construcción cultural. En este sentido, las autoras proponen a la relacionabilidad (relationality) como principio organizador de esa construcción cultural, para la identificación, el discernimiento, la creación y la continuación de sistemas indígenas de conocimiento (LITTLETREE; BELARDE-LEWIS; DUARTE, 2020, p. 413).

Las autoras van más allá al proponer las bases genéricas de lo que denominan Indigenous Knowledge Organization (IKO), en oposición a las interpretaciones colonialistas del conocimiento que suele imponer la OC tradicional. Argumentan que

to understand IKO—that is, the methodologies and means by which Native and Indigenous peoples create protocols to cohere, name, articulate, collate, and make accessible objects that indicate Indigenous knowledge—requires that practitioners of KO appreciate the colonial history of KO. Furthermore, it requires that KO practitioners recognize that the work of IKO is fundamentally a practice of liberation (LITTLETREE; BELARDE-LEWIS; DUARTE, 2020, p. 413).

Y agregan: “This means that IKO may be paradigmatically distinct from the canon of KO that is tied to the project of US national expansion” (LITTLETREE; BELARDE-LEWIS; DUARTE, 2020, p. 413). Quizás esto lleve a aceptar que cada cultura indígena es un ecosistema sociocultural, y que la OC orientada a la cuestión indígena debe considerar la enseñanza, el aprendizaje y la investigación, para la más

adecuada interpretación y organización de sus fondos documentales y su tradición (DOYLE; LAWSON; DUPONT, 2015, p. 108). La participación de miembros de la cultura atendida, como integrantes o consultores de equipos interdisciplinarios, parece condición necesaria en cualquier proyecto que se proponga indigenizar los SOC.

En este sentido, en un documento reciente Farnel (2021) señala que un criterio para establecer metadatos culturalmente sensibles debe incorporar necesariamente a las estructuras de conocimiento utilizadas por los miembros de la cultura de que se trate (FARNEL, 2021, p. 8). Esto implica cumplir con los pasos que lleven a utilizar las maneras de pensar, organizar y nombrar de la cultura minoritaria, evitando cualquier sesgo lingüístico-ideológico que puedan provenir de la cultura hegemónica o dominante.

La referencia anterior corresponde a una pesquisa de Farnel (2021) sobre colecciones digitales de cultura autóctona, con la participación de miembros de la comunidad indígena Inuvialuit, situada en territorio canadiense. En ella apeló a tres teorías para organizar su marco de pesquisa, que pueden servir de soporte para otras investigaciones en el área: i) la teoría anticolonialista que “emphasizes the multiplicity of local Indigenous knowledges, and asserts their ability to resist colonial power structures and to go beyond dismantling colonial structures by building new and better structures based on those knowledges” (FARNEL, 2021, p. 3); ii) la teoría de las ontologías fluidas promovida por Srinivasan (2002, 2007), para establecer estructuras de conocimiento flexibles, que consideren los intereses de comunidades con una cultura propia; y, iii) la teoría sociolingüística de los códigos del lenguaje (FARNEL, 2021).

5.2 Caracterización de la garantía indígena como forma de la garantía cultural

La que se conoce como indigenous warrant o garantía indígena, es una elaboración progresiva que fue desarrollando la canadiense Ann M. Doyle (sola o en coautorías) en distintos documentos comprometidos con la necesidad de generar espacios de representación y reconocimiento de las culturas indígenas (DOYLE, 2006, 2013; DOYLE; LAWSON; DUPONT, 2015; BURNS; DOYLE; JOSEPH; KREBS, 2009; WEBSTER; DOYLE, 2008).

Sandy y Bossaller resumen la idea doyleana de garantía indígena al establecer que “terms and potentially classification structures are derived from the worldview of the indigenous peoples themselves, not from the dominant cultures who write about them or who search for information about them” (SANDY; BOSSALLER, 2017, p. 133).

En la misma línea, “indigenous literary warrant serves as evidence for classificatory structure and as a source

of terminology and is based on Indigenous-authored or Indigenous-informed literature guided by the primary principle of Indigenous authority” (DOYLE; LAWSON; DUPONT, 2015, p. 115). En un intento de operacionalizar esta garantía, agregan que “indigenous cultural warrant is used, for example, in identifying Indigenous self-representation of names of nations, tribal councils and other forms of governance, as well as contemporary terminology for issues and movements” (DOYLE; LAWSON; DUPONT, 2015, 115).

5.3 Procesos de institucionalización de las culturas y los conocimientos indígenas.

Tradicionalmente, la CI y la OC necesitan una referencia institucional para su quehacer profesional. Las bibliotecas, los archivos y los museos no son entidades aisladas, sino que se integran a instituciones que les dan soporte económico, material y humano, a cambio del desarrollo de un conjunto de acciones organizadas en pro del cumplimiento de la visión, la misión y las metas de esas instituciones.

En muchos países se aplican políticas de institucionalización que intentan establecer mecanismos de recomposición cultural. Hay dos formas tradicionales de institucionalización de la cuestión indígena bajo la órbita de otras culturas dominantes.

Por un lado, se crean organizaciones públicas y privadas para el estudio, la protección, defensa o cuidado de estas culturas minoritarias. Desde el punto de vista institucional los países latinoamericanos han logrado fundar, no sin dificultades, este tipo de instituciones. En Brasil, por ejemplo, fue creado en 2002 el llamado Comitê de Demografia dos Povos Indígenas, reuniendo demógrafos y antropólogos interesados en analizar los movimientos demográficos y la organización social de las comunidades indígenas brasileñas.

La segunda forma de institucionalización es la creación de centros e instituciones a las que se integran bibliotecas, archivos y museos. El Archivo General de Asuntos Indígenas, en Temuco, por ejemplo, busca reunir el patrimonio documental histórico de las diferentes comunidades indígenas de Chile. Por su parte, el Inuvialuit Cultural Resource Centre (ICC), fundado en 1998, promueve y preserva el lenguaje y la cultura de los Inuvialuit, que conservan sus territorios en el norte de Canadá (FARNEL, 2021). El Museo Cemi situado en Ponce, Puerto Rico, cumple por su vez con la finalidad de preservar y exponer las evidencias de la cultura taina. Son tres ejemplos de una extensa lista de centros similares a lo largo de toda Latinoamérica.

Estas políticas tienen, sin embargo, sus bemoles, pues antes que ser la consecuencia de un proceso de indigenización de las posturas institucionales, suelen ser formas estatales de disciplinamiento social y de convalidación de la idea de culturas indígenas como algo marginal y superado por el paso del tiempo

(LITTLETREE; BELARDE-LEWIS; DUARTE 2020, p. 412).

En el caso particular de los archivos y museos, el disciplinamiento se verifica a través de la organización de materiales y documentos bajo responsabilidad de expertos que generalmente no forman parte de la cultura indígena atendida, sino que expresan la voz de la autoridad académica impregnada de tintes colonialistas.

5.4 Relación de métodos mencionados para la aplicación de la garantía indígena.

Un abordaje metodológico que ha sido aceptado en OC para incorporar conceptos propios de una determinada cultura en esquemas de clasificación, es el de la hospitalidad cultural, propuesto por Beghtol (2002b). Como señalan Choi et al. (2022), el concepto de hospitalidad cultural constituye “an approach to improve information systems by providing ethical resource descriptions and access” (CHOI et al., 2022, p. 554). De este modo, “cultural hospitality refers to the ability of a system to connect existing knowledge with perspectives, expectations, and assumptions from different cultures and users” (CHOI et al., 2022, p. 554).

Podrían considerarse como modalidades de hospitalidad cultural tanto la inserción puntual de términos o facetas con cierta orientación cultural en clasificaciones preexistentes (por ejemplo, en los sistemas denominados universales), como la sustitución de esquemas completos de un área.

Con relación a la inserción de términos puntuales, corresponde señalar que muchos conceptos especializados van evolucionando a medida que evolucionan las mentalidades, y que los esquemas de clasificación, en sucesivas ediciones tratan de acompañar “as evoluções e revoluções epistemológicas, as mudanças de paradigmas e as alterações parcelares ocorridas nas mentalidades” (SIMÕES, 2010, p. 532).

En su tesis de doctorado, Simões (2010) ejemplifica esta idea a través del análisis de la representación del término ‘etnia’ en cinco ediciones sucesivas de la Clasificación Decimal Universal. En su trabajo se visualiza que la utilización de la hospitalidad cultural no es una mera herramienta formal, sino que debe alimentarse de los estudios de evolución de los conceptos y las terminologías.

Una forma particular de hospitalidad cultural es la creación de espacios paradójicos, propuesta por Olson y Ward (1998) para insertar terminología de género en las tablas del Sistema de Clasificación Decimal Dewey, una solución extensible a otras situaciones en que se requiera introducir términos culturalmente orientados, desde perspectivas diferentes en los esquemas de un SOC. La implementación de espacios paradójicos implica subdividir un concepto general a efectos de incorporar conceptos particulares que tengan determinado sesgo u orientación.

Otro elemento a considerar en el marco de la hospitalidad cultural es que la elección de descriptores en las temáticas aborígenes, debe realizarse desde el respeto por los conceptos tal como son construidos y nominados en su cultura de origen. Farnel (2021) establece que hasta el nombre de los lugares debe seguir esta regla, y ejemplifica con una decisión que tomaron los profesionales del Centro Cultural que recoge la herencia cultural de los Inuit. En ese caso, los nombres de los lugares geográficos del territorio Inuit se presentaban en una lista de referencia en el lenguaje Inuvialuktun y en inglés, hasta que luego de debatirlo en profundidad, decidieron utilizar solo las expresiones indígenas (FARNEL, 2021, p. 10).

Barité y Rauch por su parte, han propuesto considerar métodos comunes a algunas ciencias sociales y humanidades, como el análisis de contenido, el análisis terminológico, el análisis del discurso y, ya dentro de la OC, los métodos de análisis de dominio, para mejorar la aplicación de la garantía cultural (BARITÉ; RAUCH, 2020).

Una solución extrema para favorecer la representación temática integral de una cultura indígena es la creación de nuevos SOC. De este modo se neutraliza la hibridación con las concepciones provenientes de la cultura hegemónica, y se pueden delimitar de forma nítida las fronteras del ecosistema conceptual y terminológico de una cultura aborígen. En esta solución se plasma el criterio de implantación de sesgo positivo, por el cual los clasificacionistas imponen premeditadamente una orientación específica al sistema nocional (COLOMBO; BARITÉ, 2015).

Hay por lo menos tres ejemplos significativos en esta línea: la Indigenous Classification de Brian Deer, que es estudiada con cierto detalle por Cherry (2015), el Xwi7xwa Classification System de la Universidad de Columbia Británica, y la lista FNHL Subject Headings, de la First Nations House of Learning de Canadá.

Doyle, Lawson y Dupont (2015) resaltan la calidad del esquema Xwi7xwa en sí, como un caso de sistema hecho a medida para una biblioteca de cultura indígena, y mencionan además sus posibilidades didácticas, pues a través del trabajo en aula con sus estudiantes universitarios, logra poner en juego una postura crítica de los procesos de OC en contextos indígenas (DOYLE; LAWSON; DUPONT, 2015, p. 123). Cualquiera de estos tres instrumentos mediadores podrían disponer de hecho, de ese potencial didáctico que declaran las autoras.

6 Conclusiones

El concepto de garantía cultural no es neutro, como tampoco lo es el concepto de cultura. A través de la garantía cultural se relativiza la idea de que tanto los SOC como los procesos de clasificación y la indización

deben basarse en la objetividad y la neutralidad, siguiendo las premisas de la presunta objetividad de la ciencia. En este contexto, la garantía cultural orienta los frutos del saber acumulado hacia formas de representación del conocimiento no académico ni científico, y contribuye a introducir y fortalecer abordajes socioculturales a la OC.

La garantía cultural conduce además a una ética del respeto: respeto por las particularidades culturales de una comunidad, respeto por las determinaciones de las culturas locales, respeto por la integridad cultural de las minorías, respeto por la diversidad de creencias, respeto por los principios de igualdad o equidad de género, respeto por los desfavorecidos de cualquier categoría, respeto por la visibilidad de los ecosistemas culturales desde perspectivas no hegemónicas. Respeto, en definitiva, por valores universales que se expresan en las culturas locales y que cada vez más se asocian a los procesos de construcción de ciudadanía, de inclusión social y de democracia civil, política y económica.

Parece confirmarse la hipótesis de que la garantía indígena es una forma de expresión de la garantía cultural, en la medida en que lo expuesto encaja en los presupuestos y en las formas de concebir la organización del conocimiento de los pueblos aborígenes.

De forma más específica, la revisión crítica de literatura realizada permite identificar, en primer término, una base epistemológica y conceptual significativa para encuadrar la OC de las culturas indígenas. Esta base se ha ido construyendo con el aporte primordial de investigadores de Estados Unidos, Canadá y Australia. Dada la universalidad y la similitud de esos procesos de encuadre, marginación y recomposición institucional de las culturas indígenas en esos países, con los de los países latinoamericanos, los mismos pueden servir de modelo para encontrar soluciones a la OC indígena en nuestro subcontinente, y a soluciones más generales que pueda ofrecer la CI.

La revisión confirma además que existe una batería consistente de métodos para orientar la aplicación de la garantía indígena, tanto si se opta por crear nuevos SOC a medida, como si se elige enmendar, corregir o ampliar la cobertura de descriptores proporcionados por los sistemas o las listas de clasificación tradicionales, en una suerte de integración o convivencia entre visiones hegemónicas y alternativas. En cualquier caso, las soluciones metodológicas deben tender a respetar la integridad del conocimiento indígena en tanto ecosistema cultural, y exhibir las claves de su identidad.

El énfasis puede estar dado en la recuperación de su historia; también en la identificación de sus tradiciones (en especial su tradición oral), rituales y creencias, sus lenguas y dialectos, y los aspectos de diversidad en la unidad de la misma cultura.

Por último, cabe remarcar la urgencia de promover la producción académica en la materia desde el área de CI, y más específicamente desde la OC en América Latina, considerando el conjunto - hasta ahora extremadamente reducido - de antecedentes, todos ellos brasileños, que han comenzado a marcar el camino. La dignidad de nuestras culturas originarias así lo exige.

Referencias

- AGRAWAL, A. Indigenous Knowledge and the Politics of Classification. **International Social Science Journal**, v. 54, p. 287-297, 2002.
- ALBUQUERQUE, A. C.; MORAES, M. A. A organização e representação do conhecimento indígena a partir de um estudo nas publicações da área de Ciência da Informação no Brasil. *In: ENCUENTRO DE DOCENTES DE BIBLIOTECOLOGÍA Y CIENCIA DE LA INFORMACIÓN DEL MERCOSUR, EDDBCIM*, 23. **Ponencias (...)**. (Inédito). Montevideo: Facultad de Información y Comunicación, Udelar, 2023.
- BARITÉ, M.; RAUCH, M. Cultural warrant: old and new sights from Knowledge Organization. *In: KNOWLEDGE ORGANIZATION AT THE INTERFACE: SIXTEENTH ISKO CONFERENCE*, Aalborg University. **Proceedings (...)**. Baden-Baden: Ergon, 2020, p. 31-40.
- BARITÉ, M. Towards a general concept of warrants: first notes. **Knowledge Organization**, v. 46, n. 8, p. 647-655, 2019.
- BARITÉ, M. Literary warrant. **Encyclopaedia of Knowledge Organization**. ISKO, 2017. Disponible en: https://www.isko.org/cyclo/literary_warrant. Acceso el 29 mar. 2023.
- BEALL, J. Approaches to expansions: case studies from the German and Vietnamese translations. *In: WORLD LIBRARY AND INFORMATION CONGRESS: 69 IFLA GENERAL CONFERENCE AND COUNCIL 1-9 August 2003*, Berlin. **Papers (...)**. Disponible en: <http://www.ifla.org/IV/ifla69/papers/123e-Beall.pdf>. Acceso el: 29 mar. 2023
- BEGHTOL, C. Universal concepts, cultural warrant and cultural hospitality. *In: INTERNATIONAL ISKO CONFERENCE*, 7., 2002, Granada, 10-13 July 2002. **Proceedings (...)**. Würzburg: Ergon Verlag, 2002a, p. 45-49.
- BEGHTOL, C. A proposed ethical warrant for global knowledge representation and organization systems. **Journal of Documentation**, v. 58, n. 5, p. 507-532, 2002b.
- BEGHTOL, C. Relationships in classificatory structure and meaning. *In: BEAN, C. A.; GREEN, B. (eds.) Relationships in the organization of knowledge*. Dordrecht: Kluwer, 2001. p. 99-113.
- BEGHTOL, C. Semantic validity: concepts of warrant in bibliographic classification systems. **Library Resources & Technical Services**, v. 30, n. 2, p. 109-123, 1986.
- BURNS, K.; DOYLE, A.; JOSEPH, G.; KREBS, A. Indigenous Librarianship. *In: ENCYCLOPEDIA of Library and Information Science*, 3. ed. Francis & Taylor, 2009. p. 2330-2346.
- CARRON, T. **Integrating Indigenous Knowledge in Non-Indigenous Knowledge Systems and Institutions**. John Hopkins University: Museum Studies Digital Curation Certificate Program, 2019.
- CHERRY, A. A case study in Indigenous classification: revisiting and reviving the Brian Deer scheme. **Cataloging & Classification Quarterly**, v. 53, n. 5/6, p. 548-567, 2015. doi: 10.1080/01639374.2015.1008717
- CHOI, I. *et al.* Cross-cultural Information Access. *In: Association for Information Science and Technology. Proceedings*, v. 59, n. 1, p. 551-554, 2022.
- COLOMBO, S.; BARITÉ, M. Tres enfoques de bias en organización del conocimiento: bias neutro, bias negativo y bias positivo. **Brazilian Journal of Information Science**, v. 9, n. 2, p. 9-13, 2015. Disponible en: <http://basessibi.c3sl.ufpr.br/brapci/index.php/article/view/0000018577/c2fea535036c906d277c9ea1513600f4> Acceso el: 29 mar. 2023.
- CUSTER, B. Editors Introduction. *In: Dewey, Melvil. Dewey Decimal Classification and Relative Index*. 16th ed. N.Y.: Forest Press, 1958.
- DEI, G. J. S. Rethinking the role of Indigenous knowledges in the academy. **International Journal of Inclusive Education**, v. 4, n. 2, p. 111-132, 2000.
- DEI, G. J. S.; ASGHARZADEH, A. The power of social theory: the anti-colonial discursive framework. **The Journal of Educational Thought (JET) / Revue de la Pensée Éducative**. v. 35, n. 3, p. 297-323, 2001.
- DOYLE, A. M. **Naming, claiming, and (re)creating: indigenous knowledge organization at the cultural interface**. 2013. 401 f. Tese (Doutorado em Filosofia) – Faculdade de Estudos de Graduação - University of British Columbia, Vancouver, 2013. Disponible en: <https://open.library.ubc.ca/soa/cIRcle/collections/ubctheses/24/items/1.0073667>. Acceso el: 29 mar. 2023.
- DOYLE, A. M. Naming and Reclaiming Indigenous knowledges in public institutions: intersections of landscapes and experience in Knowledge Organization for a global learning society. *In: International ISKO Conference*, 19., 2006, Vienna Austria. **Anais [...]**. Vienna, Austria, 2006. p. 436-442.
- DOYLE, A. M.; LAWSON, K.; DUPONT, S. Indigenization of Knowledge Organization at the Xwi7xwa Library. **Journal of Library and Information Studies**, v. 13, n. 2, p. 107-134, dez. 2015. doi:10.6182/jlis.2015.13(2).107
- EAST, R. Indigenous subject access, Ngā Ūpoko Tukutuku the Māori subject headings thesaurus. *In: AUSTRALIAN AND NEW ZEALAND THEORETICAL LIBRARY ASSOCIATION (ANZTLA) CONFERENCE. Proceedings [...]*, 2008.
- FARNEL, S. Knowledge Organization in Support of an Indigenous Worldview: A Community Driven Metadata Framework for the Inuvialuit Digital Library. *Proceedings from NORTH AMERICAN SYMPOSIUM ON KNOWLEDGE ORGANIZATION. Proceedings*. v. 8, p. 1-17, 2021.
- FARNEL, S. *et al.* Of places and names: working with northern Canadian communities to enhance subject access to

- digital resources. *In: CONNECTIONS. COLLABORATION. COMMUNITY. IFLA WORLD LIBRARY AND INFORMATION CONGRESS, 82., 2016. Papers [...]* Columbia, OH. Disponible en: <http://library.ifla.org/1326/1/151-farnel-en.pdf> Acceso el: 29 mar. 2023.
- FRANÇA, A.; SILVEIRA, N. C. A representação descritiva e a produção literária indígena brasileira. **Transinformação**, v. 26, n. 1, p. 67-76, 2004. Disponible em: 10.1590/S0103-37862014000100007. Acceso el: 29 mar. 2023.
- GODBOLD, N. User-centred design vs. “good” database design principles: a case study, creating knowledge repositories for Indigenous Australians. **Australian Academic & Research Libraries**, v. 40, n. 2, p. 116-131, 2009.
- GUIMARÃES, J. A. C. Aspectos éticos em organização e representação do conhecimento (ORC): uma reflexão preliminar. *In: GONZÁLEZ DE GÓMEZ, M. N.; ORRICO, E. G. D. (Org.). Políticas de memória e informação: reflexos na organização do conhecimento.* Natal: EDUFRN, 2006. p. 237-264.
- GUIMARÃES, J. A. C.; FERNÁNDEZ-MOLINA, J. C.; PINHO, F. A.; MILANI, S. O. Ethics in the Knowledge Organization Environment: an overview of values and problems in the LIS literature. *In: CULTURAL and Identity in Knowledge Organization.* Würzburg: Ergon, 2008 p. 340-346. (Advances in Knowledge Organization, 11).
- HANSSON, J. Classification Systems for Jewish Libraries and Bibliographies: Philosophy, Warrants and Relation to non-Jewish Bibliographic Practice. **Journal of Religious & Theological Information**, v. 21, n. 3-4, p. 99-121, 2022. DOI: 10.1080/10477845.2022.203804
- KOLESAS, M. **Una introducción al rol de la biblioteca en la educación del siglo XXI.** Buenos Aires: Fondo de Cultura Económica, 2008.
- LEE, D. Indigenous Knowledge Organization: a study of concepts, terminology, structure and (mostly) indigenous voices. **Partnership: the Canadian Journal of Library and Information Practice and Research**, v. 6, n. 1, p. 1-33, 2011.
- EE, D. Indigenous Knowledge Organization: a study of concepts, terminology, structure and (mostly) indigenous voices. **Partnership: the Canadian Journal of Library and Information Practice and Research**, v. 6, n. 1, p. 1-33, 2011.
- LITTLETREE, S.; BELARDE-LEWIS, M.; DUARTE, M. Centering relationality: a conceptual model to advance Indigenous knowledge organization practices. **Knowledge Organization**, v. 47, n. 5, p. 410-426, 2020.
- MAIMONE, G. D.; MATOS, A. P. DE. Culturas indígenas sob a perspectiva da Ciência da Informação. **Brazilian Journal of Information Science: Research Trends**, v. 13, n. 3, p. 46–55, 2019. Disponible en: <https://doi.org/10.36311/1981-1640.2019.v13n3.05.p46>. Acceso el: 29 mar. 2023.
- MARTÍNEZ-ÁVILA, D.; BUDD, J. Epistemic warrant for categorizational activities and the development of controlled vocabularies. **Journal of Documentation**, v. 73, n. 4, p. 700-715.
- MELLO, R. P. S. A importância dos arquivos para a garantia dos direitos indígenas. **Arquivo & Administração**, v. 13, n. ½, p. 69-86, 2014. Disponible em: <http://hdl.handle.net/20.500.11959/brapci/50047>. Acceso el: 29 mar. 2023.
- NAKATA, M.; LANGTON, M. **Australian Indigenous Knowledge and Libraries.** Canberra: Australian Academic e Research Libraries, 2005.
- NAKATA, M. *et al.* Indigenous digital collections: An early look at the Organisation and culture interface. **Australian Academic and Research Libraries**, v. 39, n.4, p. 223-236, 2008.
- OLSON, H. A. The power to name: representation in library catalogs. **Signs: Journal of Women in Culture and Society**, v. 26, n. 3, p. 639-668, 2001.
- OLSON, H. A. Cultural discourses of classification: indigenous alternatives to the tradition of Aristotle, Durkheim and Foucault. *In: ASIS SIG/CR CLASSIFICATION RESEARCH WORKSHOP, 10., 1999. Proceedings [...].* 1999, p. 1107-1124, 1999. DOI: 10.7152/acro.v10i1.12484.
- OLSON, H. A.; WARD, D. Charting a journey across knowledge domains: feminism in the Decimal Dewey Classification. *In: STRUCTURES AND RELATIONS IN KNOWLEDGE ORGANIZATION; INTERNATIONAL ISKO CONFERENCE, 5., 1998. Lille, 25-29 August, 1998. Proceedings [...].* Würzburg: Ergon, 1998, p. 238-244. (Advances in knowledge organization, 6).
- PACEY, P. The Classification of Literature in the Dewey Decimal Classification: the primacy of language and the taint of colonialism. **Cataloging & Classification Quarterly**, v. 9, n. 4, p. 101-107, 1989.
- PUENTE HERNÁNDEZ, L. E. **Biblioteca pública, democracia y buen vivir: aportes para la definición de políticas en Ecuador.** Quito: FLACSO, 2013.
- SANDY, H. M.; BOSSALLER, J. Providing cognitively justsubject access to indigenous knowledge through knowledge organization systems. **Cataloging & Classification Quarterly**, v. 55, n. 3, p. 129-152, 2017.
- SIMÕES, M. G. M. **A representação de Etnia e sua evolução na Classificação Decimal Universal.** 2010. Tese (Doutorado em Biblioteconomia e Documentação) - Universidade de Salamanca, 2010. 600 p. Disponible en: <https://gredos.usal.es/bitstream/handle/10366/83331>. Acceso el: 12 maio 2023.
- SRINIVASAN, R. **Village Voice: expressing narrative through community-designed ontologies.** Thesis (Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Arts and Sciences, Massachusetts Institute of Technology, 2002.
- SRINIVASAN, R. Ethnomethodological architectures: Information systems driven by cultural and community visions. **Journal of the American Society for Information Science & Technology**, v. 58, n. 5, p. 723-733, 2007
- THORPE, K.; GALASSI, M. Rediscovering Indigenous languages: the role and impact of libraries and archives in cultural revitalisation. **Australian Academic & Research Libraries**, v. 45, n. 2, p. 81-100, 2014. DOI:10.1080/00048623.2014.910858.

TOMREN, H. **Classification, bias and American indian materials** (Unpublished research paper). San Jose: San Jose State University, 2004.

WANG, S. The Intellectual Landscape of the Domain of Culture and Ethics in Knowledge Organization: an analysis of influential authors and works, **Cataloging & Classification Quarterly**, v. 57, n. 4, p. 227-243, 2019. DOI: 10.1080/01639374.2019.1614710

WEBSTER, K; DOYLE, A. Don't Class me in antiquities! giving voice to native American materials. *In*: ROBERTO, K. R.; JEFERSON, N. C. **Radical cataloging: essays at the front**. Carolina do Norte: Greenwood, 2008. p. 189-197. DOI: <https://dx.doi.org/10.14288/1.0077956>.

YEH, T. Y; FROSIO, E. T. The treatment of the American Indian in the library of congress E-F Schedule. **Library Resources & Technical Services**, v. 15, n. 2, p. 1122-131, 1971.

ZAMBONI, R. C. V. **Organização do conhecimento, classificação e diversidade cultural: uma análise a partir do conceito de "garantias"**. 2018. Tese (Doutorado em Cultura e Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2018. doi:10.11606/T.27.2018.tde-17072018-171543.

ZAMBONI, R. C. V.; FRANCELIN, M. M. The location of classification: between the local and the global. *In*: KNOWLEDGE ORGANIZATION FOR A SUSTAINABLE WORLD: CHALLENGES AND PERSPECTIVES FOR CULTURAL, SCIENTIFIC, AND TECHNOLOGICAL SHARING IN A CONNECTED SOCIETY. INTERNATIONAL ISKO CONFERENCE, 14., 2016, Rio de Janeiro. **Proceedings** [...]. Rio de Janeiro: International ISKO Conference, 2016, p. 572-578. (Advances in Knowledge Organization, 15).

Intersecção de três lentes da análise de domínio para a caracterização de comunidades discursivas com focos decoloniais/críticos-sociais

Intersection of three domain analysis approaches for the characterization of discursive communities with decolonial critical-social focuses

Maíra Fernandes Alencar (1), Brígida Maria Nogueira Cervantes (2), Mario Barité (3)

(1) (2) Universidade Estadual de Londrina, Londrina (UEL) - Rodovia Celso Garcia Cid PR 445 Km 380 – Campus Universitário, Londrina, PR. (3) *Facultad de Información y Comunicación, Universidad de la República, Uruguay*
(1) alencarmfa@gmail.com (2) brigidacervantes@gmail.com (3) mario.barite@fic.edu.uy

Resumo:

Percebemos certa necessidade de caracterizar comunidades discursivas com focos decoloniais/críticos-sociais que visam mudanças explícitas estruturais em um contexto de pluralidade para abarcar grupos historicamente marginalizados. Necessidade esta no âmbito de uma dimensão de cunho sintético, que represente em algum nível sua especificidade social. Nesse sentido, o **objetivo** deste trabalho é apresentar diante de tal contexto, uma proposta de intersecção de três lentes da Análise de Domínio: epistêmica, de estruturação científica e terminológica para guiar a caracterização de uma comunidade discursiva com tais dimensões. A **metodologia** utilizada é de foco exploratório e descritivo, a partir da literatura da Organização do Conhecimento, nos estudos de Hjørland e Swales. Como **resultados** apresenta-se uma sistematização de dados para serem coletados, e as formas de análise para viabilizar tal caracterização, levando em conta a dimensão biográfica dos membros da comunidade, a escolha de um termo representativo para realizar a análise terminológica pontual, e, por fim, a identificação e análise da lente da estruturação científica, formando uma tríade: membros/termos/comunicação. A **conclusão** é de que as comunidades discursivas com focos decoloniais e críticos sociais produzem discursos pautados em mudanças epistêmicas com horizontes de libertação/emancipação em algum nível, considerando a necessidade de inclusão e justiça social/direitos humanos para grupos historicamente marginalizados. Nesse sentido, a intersecção dessas três lentes ajudam na caracterização de forma a elencar aspectos implícitos e explícitos mediante a constatação da formação e interesses dos membros, da pertinência/fundamentação em usar determinados termos para representar seus contextos de ações, e da lógica que assume a forma de comunicar essas sistematizações.

Palavras-chave: Análise de Domínio; Comunidades Discursivas; Decolonialidade; Estudos Críticos.

Abstract

It starts from the understanding that when dealing with discursive communities with decolonial/critical-social focuses, that is, which aim at explicit structural changes in a context of plurality to encompass historically marginalized groups, in different areas of knowledge, there is a need to characterize such communities to the scope of a dimension of a synthetic nature, which represents, at some level, its social specificity. In this sense, the **objective** of this work is to present, in this context, a proposal for the intersection of three lenses of Domain Analysis, epistemic, scientific and terminological structuring to guide the characterization of a discursive community with such dimensions. The **methodology** used has an exploratory and descriptive focus, based on the Knowledge Organization literature, in the studies of Hjørland and Swales. **Results and discussion:** a systematization of which data can be collected is presented., and the forms of analysis to enable such characterization, taking into account the biographical dimension of the community members, the choice of a representative term to carry out the punctual terminological analysis, and, finally, the identification and analysis of the lens of scientific structuring, forming a triad: members/terms/communication. **Conclusion:** Discursive communities with decolonial and social critical focuses produce discourses based on epistemic changes with horizons of liberation/emancipation at some level, considering the need for inclusion and social justice/human rights for historically marginalized groups, in this sense, the intersection of these three lenses help in the characterization in order to list implicit and explicit aspects by verifying the training and interests of the members, the pertinence/reasons in using certain terms to represent the contexts of actions, and the logic that assumes the form of communicating these systematizations.

Keywords: Domain analysis; Discursive Communities; Decoloniality; Critical Studies.

Resumen

Al hablar de comunidades discursivas cuyos enfoques son decoloniales y crítico-sociales, es decir, que apuntan a cambios estructurales explícitos en un contexto de pluralidad para abarcar a grupos históricamente marginados en diferentes áreas del conocimiento, se identifican dificultades en cómo estudiar tales comunidades en el ámbito de una caracterización de carácter sintético, que represente en algún nivel, su especificidad social. En ese sentido, el **objetivo** de este trabajo es presentar, en este contexto, una propuesta de intersección de tres enfoques de análisis de dominio, estructuración epistémica, científica y terminológica para orientar la

caracterización de una comunidad discursiva de tales dimensiones. La **metodología** utilizada tiene un enfoque exploratorio y descriptivo, basada en la literatura de Organización del Conocimiento, particularmente en los estudios de Hjørland y Swales. Como **resultados** presentamos una sistematización sobre los datos a recolectar y las formas de análisis para viabilizar tal caracterización, considerando la dimensión biográfica de los integrantes de la comunidad discursiva, la elección de un término representativo para realizar el análisis terminológico puntual y, finalmente, la identificación y análisis del enfoque de estructuración científica, formando una tríada: miembros/términos/comunicación. **Conclusión:** Las comunidades discursivas con enfoques decoloniales y críticos sociales producen discursos basados en cambios epistémicos con horizontes de liberación/emancipación en algún nivel considerando la necesidad de inclusión y justicia social/derechos humanos de los grupos marginados. En ese sentido, la intersección de estos tres enfoques ayuda en la caracterización para enumerar aspectos implícitos de la formación e intereses de los miembros, la pertinencia/razones en el uso de ciertos términos y la lógica que asume la forma de comunicar estas sistematizaciones.

Palabras clave: Análisis de dominio; Comunidades discursivas; Decolonialidad; Estudios Críticos.

1 Introdução

Visando a realização de pesquisas e formação acadêmica, destacamos duas contribuições que a Organização do Conhecimento (OC) tem proporcionado aos estudos da informação no Brasil: o uso teórico-metodológico da Análise de Domínio (AD), e o conceito de comunidades discursivas (CD). Tais contribuições acontecem, mais especificamente, no âmbito da educação superior em Programas de Pós-graduação em Ciência da Informação (CI), cuja formação dos discentes são, em sua maioria, em Biblioteconomia e Arquivologia.

O uso da AD é frequente em pesquisas de diferentes naturezas na Ciência da Informação (CI), englobando assuntos diversos, para fins de caracterização mediante diferentes abordagens proposta por Hjørland (2002).

Em relação às comunidades discursivas, esse conceito ficou conhecido na área da CI, em particular da OC, após do trabalho de Hjørland e Albretsen (1995). A CD é citada frequentemente a partir da fundamentação de Swales (1990), que vincula principalmente a perspectiva de que: em síntese, uma CD diz respeito a um grupo que compartilha interesses/objetivos em comum. Sendo assim, podemos entender as CDs sob focos de naturezas diversas. Para esse trabalho, buscamos chamar atenção para a dificuldade em pautar ou trazer para debate as comunidades discursivas que atuam no campo de focos decoloniais e críticos-sociais, isto é, que em seus objetivos centrais visam tratar propostas de mudanças estruturais; por exemplo, incluindo grupos historicamente excluídos.

Essa dificuldade tem a ver com o processo de caracterização de tais comunidades, pois, geralmente, apenas a análise de sua produção científica não contempla a síntese de seu aspecto social abordado. Nesse sentido, apresentamos como proposta a interseção de três lentes da análise de domínio para que uma comunidade possa ser caracterizada contemplando a dimensão social de seu foco: epistêmica, de estruturação científica, e terminológica.

O uso de tais lentes interseccionais é apresentado com alguns focos específicos: na epistêmica privilegia-se a coleta de dados de uma micro biografia dos membros

pertencentes às comunidades; na lente de estruturação científica pauta-se um caminho centrado na área que almeja-se analisar a partir de determinada base de dados ou domínio editorial, e por fim, para a lente terminológica, atribui-se a dimensão pontual, para que possa haver especificidade da análise, no contexto de uso de apenas um termo representativo de tais debates críticos da comunidade em foco.

A metodologia empregada é exploratória e descritiva, a partir da literatura da Organização do Conhecimento, e é resultado da dimensão metodológica da tese: “Os Arquivos Comunitários na pesquisa e educação arquivística: Uma análise a partir da *Archival Education Research Initiative*” de Alencar (2022).

A estruturação do trabalho apresenta uma discussão sobre as lentes da Análise de Domínio proposta por Hjørland em 2002 (2), seguida por: uma explanação a respeito das Comunidades Discursivas (3), em construção um entendimento de “Comunidades Discursivas com focos decoloniais e críticos-sociais” (3.1), metodologia (4), resultados (5), e por fim, considerações finais (6).

2 As lentes da Análise de Domínio proposta por Hjørland em 2002

O conceito de domínio a partir de Hjørland (s/p, 2017) está inserido em uma condição socialmente construída para atender diversos âmbitos, isto é, sendo um corpo de conhecimento compartilhado por pessoas de diferentes compromissos ontológicos e epistemológicos, ele poderá abarcar igualmente diferentes teorias e interesses. Na visão do autor podem ser desde disciplinas acadêmicas até *hobbies*. No entanto, se tratando da realização da AD, eles devem ter algum nível de estabilidade e infraestrutura (HJØRLAND, 2017). Complementa a visão da estabilidade, o entendimento de que um domínio é melhor compreendido enquanto uma "unidade de análise para construir um sistema de organização do conhecimento", possuindo também a característica de ser “um grupo com uma base ontológica que revela uma teleologia subjacente, um conjunto de hipóteses comuns, consenso epistemológico sobre abordagens metodológicas e semântica social”. (SMIRAGLIA, 2012, p. 114 - tradução nossa).

A respeito da concepção de domínio, podemos nos perguntar: afinal, por quais razões analisar um domínio?

Destacar caminhos possíveis e/ou necessários que poderão levar a cenários almejados dentro de um lócus estabelecido pressupõe conhecer áreas dos saberes e delimitar temas para serem vistos e estudados. Se tratando da diversidade que há nessas áreas científicas, elencamos a seguir quatro razões para se analisar um domínio.

A primeira razão para se analisar um domínio é o ganho de um panorama sintético diante da quantidade de informações que existem. Sob este ponto de vista, de uma dimensão sintética, fica explícito que a perspectiva analítica opera a partir e com a categorização, como observado por Guimarães (2014, p. 16).

A segunda razão é a riqueza semântica que esse tipo de análise possibilita. Assim, com a combinação de algumas abordagens, a qualidade dessa análise torna-se ampla em detrimento de apresentar diferentes aspectos do mesmo objeto. Por exemplo, é caso deste artigo: unir a parte epistêmica e terminológica de estruturação científica para caracterizar um determinado tipo de comunidade discursiva.

A terceira razão resulta em dar visibilidade ao que está sendo estudado em determinado campo. Por fim, a quarta razão trata-se de construir sistemas de organização do conhecimento (SOC).

Vistas as razões, passamos para as lentes propostas por Hjørland em 2002 enquanto um dos métodos de análise.

No texto *“Domain analysis in information science: Eleven approaches- traditional as well as innovative”* de Hjørland publicado em 2002 no periódico *Journal of Documentation*, o autor enfatiza que os domínios não são semelhantes, são específicos e diferentes de cada aspecto no mundo social do trabalho, dessa forma, apresenta 11 abordagens para se estudar um domínio.

O problema colocado pelo autor diz respeito aos tipos de conhecimentos que são necessários aos profissionais da informação, especialistas em diversos campos, como saúde, educação, entre outros. Importante destacar que o autor menciona que as abordagens apresentadas por ele podem já estar implícitas na literatura, mas que até então não haviam sido sistematizadas. (HJØRLAND, 2002). Nesse sentido sua contribuição, ao descrever a visão analítica de formas de se produzir conhecimentos com foco em domínios variados, tem sido utilizada aos pares na área da Ciência da Informação, de modo que, além de uma perspectiva teórica, também fornece um foco metodológico, uma vez que auxilia na percepção da forma em que determinado conhecimento científico está se construindo e socializando, ou seja, evoca aspectos do que tem sido considerado como *significado* em determinado campo. (GUIMARAES, 2014).

Essa característica de se restringir a determinado contexto ou comunidade para identificar padrões prevê alguns tipos de passos a serem realizados, o que Hjørland (2002) nomeou como abordagens. São 11 abordagens que podem ser combinadas: 1) Produzir guias de literatura ou portais de assunto; 2) Construir classificações especiais e tesouros; 3) Indexar e recuperar especialidades; 4) Estudo empírico do usuário; 5) Estudos bibliométricos; 6) Estudos históricos; 7) Estudos de gêneros e documentos; 8) Estudos epistemológicos e críticos; 9) Estudos terminológicos, linguagem para fim especial (LSP), semântica de base de dados e estudos do discurso; 10) Estruturas e instituições na comunicação científica; 11) Cognição científica, conhecimento especialista e inteligência artificial (IA). Cabe destacar que em 2015 Guimarães e Tognoli fizeram uma proposta de mais uma abordagem - a da proveniência arquivística. Nesse mesmo ano Smiraglia (2015) também propôs inclusão e modificação.

Neste trabalho estamos nos referindo a essas abordagens enquanto *“lentes”*, uma vez que, entendemos que elas possibilitam um determinado tipo de visão para a caracterização do domínio adotado. Essa visão, certamente, orientada para fins analíticos contém em si um aspecto de delimitação, isto é: toda visão sempre opera a partir de um recorte circunscrito em determinados aspectos selecionados. Outros conceitos que se relacionam com o de domínio, são os de comunidades discursivas e colégios invisíveis. Trataremos sobre as comunidades discursivas entendidas a partir de Swales (1990, 2016), para em seguida apresentar em um tipo específico delas, de focos críticos-sociais.

3 As comunidades discursivas

Para Smiraglia (2012, p. 113) os parâmetros funcionais de um domínio, comunidades discursivas e colégios invisíveis são divergentes. O conceito de "domínio" sugere fronteiras intelectuais e o conceito de "comunidade discursiva" sugere uma troca ativa de informações; já o termo "colégio invisível" tem sido usado para sugerir tanto a uniformidade intelectual, quanto o discurso ativo que ocorre em uma unidade socialmente estruturada". Tênnis (2003, p. 191 - tradução nossa) também falou sobre este ponto quando mencionou que “há uma série de conceitos similares ao de domínios na CI; como comunidades de prática, e comunidades epistêmicas [...]”.

Outra contribuição importante para esse debate é o trabalho de Evangelista; Grácio e Guimarães (2022) que analisou afinidades e especificidades dos conceitos de domínio, comunidade discursiva e comunidade epistêmica. Para os autores, as características de uma comunidade atuam enquanto núcleo desses três

conceitos. Sobre a escolha desses três termos, a justificativa dos autores é de que “analisam a ciência e o conhecimento a partir de seus indivíduos, e sob sua perspectiva social [...]” (EVANGELISTA; GRÁCIO; GUIMARÃES, 2022, p. 3 - tradução nossa).

Para os autores, as semelhanças entre esses três conceitos se fazem no contexto em que foram propostas, já que fazem menção à análise científica e sociológica de produzir conhecimento; contexto histórico e contexto da necessidade, servindo para analisar a lógica de como o conhecimento é compartilhado. Eles estabelecem, ainda, uma visão hierárquica com base em Smiraglia (2012), domínio, CD e CE. Citam outros autores que comentam a respeito da visão de Hjørland sobre relacionar a AD com a CD. Unem esses três conceitos à vertente do movimento helicoidal de Guimarães (2006). Outro ponto em comum destacado por eles é a base teórica. Há, por fim, uma discussão da relação entre CD e CE.

Sobre as distinções, os autores citados destacam: níveis de interação entre os membros e também o lugar. Em domínios ou CD, o consenso entre as ideias dos membros não é requisito, possuir os temas de interesse sob mesmo domínio basta; aspecto que não equivale para as CE, uma vez que, nessa configuração de comunidade a união se dá pela forma de pensar. Outra diferença é a maneira que se formam: enquanto um domínio pode ser diverso, uma CE necessita de grupos no âmbito do universo científico. Por fim, há uma tabela com 21 características para mostrar as principais semelhanças e distinções desses três conceitos. Sobre as CDs, os autores inserem:

[...] um domínio tem formas próprias de organizar o conhecimento, tanto do ponto de vista intelectual quanto social, e possui fronteiras difusas, o que pode ser explicado pelo seu alto nível de interação com outros semelhantes. **Uma comunidade discursiva, por sua vez, é composta por membros que compartilham práticas e atividades profissionais, revelam um sentimento especial de pertencimento e compartilham o conhecimento produzido principalmente por meio de textos.** Em uma configuração mais específica, uma comunidade epistêmica se caracteriza por ser restrita à esfera científica e por possuir membros especialistas com autoridade, reputação nacional e internacional na tomada de decisões institucionais, conhecimento relevante na área e, conseqüentemente, um poder implícito de controlar o conhecimento e a informação (EVANGELISTA; GRÁCIO; GUIMARÃES, 2022, p. 39 - grifo e tradução nossa).

Após debater sobre esses conceitos similares, focando nas CDs, escolhemos Swales (1990), autor muito citado nos trabalhos de OC, e em específico, seu entendimento no artigo de 1990, em que uma das partes apresentou seis características para identificar um grupo enquanto uma CD. Após 26 anos, em 2016, o autor publicou outro trabalho no qual, dentre outros aspectos, reconsiderou os critérios anteriores proposto por ele:

[...]o conceito de CD originalmente concebido era excessivamente estático. Embora isso talvez não importasse tanto em 1990, no mundo mais instável e incerto de hoje, parece um problema, em particular, o conceito não abordou diretamente como as pessoas ingressam ou saem dos CDs (SWALES, 2016, p. 4 – tradução nossa).

Um dos argumentos para tal revisão surgiu com o transcorrer do tempo e das mudanças, principalmente tecnológicas, uma vez que à medida em que o mundo se modificou, o próprio conceito de CD e os critérios se tornaram limitados, estáticos e idealistas. O autor inseriu, ainda, uma discussão sobre tipos de comunidades discursivas, categorizando-as em três tipos^[1]: comunidade discursiva local (grupo de pessoas que trabalham no mesmo local); comunidade discursiva focal (são, tipicamente, associações que designam região, nação ou, em âmbito internacional, podem ser recreativas ou profissionais) e comunidade discursiva “folocal” (são híbridas e contemplam os dois tipos anteriores). Outra mudança proposta por Swales foi a inserção de mais dois critérios e também a reflexão sobre os anteriores. Portanto, para o autor uma CD possui oito critérios:

- 1) Um conjunto de objetivos que podem ser formulados e reconhecidos de forma geral ou parcial por seus membros;
- 2) Mecanismos de intercomunicação entre seus membros que podem contar com os novos canais das tecnologias de informação;
- 3) Uso de seus mecanismos participativos para fornecer informações e feedbacks, e também para gerenciar operações próprias e promover recrutamento, mudanças e crescimentos;
- 4) Uma seleção evolutiva de gêneros para promover seus objetivos e gerir seus mecanismos participativos, que se desenvolvem conforme são utilizados;
- 5) Terminologia específica;
- 6) Uma hierarquia e/ou estrutura explícita ou implícita que gerencia o desenvolvimento da comunidade, fazendo com que haja um limite de membros com um grau de conteúdo e especialização;
- 7) Senso de “relações silenciosas”, que faz com que alguns aspectos não precisem ser ditos ou explicados detalhadamente;
- 8) Um horizonte de expectativas, ritmos definidos de atividades, percepção de sua história e sistemas de valores para o que é bom ou não no âmbito do trabalho. (SWALES, 2016 – tradução nossa).

Após inserir esses novos critérios, o autor insere um debate sobre o *status* das discussões da CD, do qual outras possibilidades poderiam fazer parte, mas se

posiciona argumentando a respeito do que para ele é mais relevante nos debates, a dimensão da visibilidade nos “princípios retóricos de organização, nas expectativas discursivas, nos símbolos linguísticos significativos e nos extratos textuais relevantes” (SWALES, 2016, p. 10 – tradução nossa). Para reforçar esse ponto, ele cita Porter (1992) abordando a utilidade do termo CD: descrever um espaço no qual não havia reconhecimento, uma vez que o termo ainda não existia, configurando, assim, de uma nova forma as unidades de escritor, público e texto (SWALES, 2016).

Considerando essas discussões de Swales sobre novos critérios e, principalmente tipos de comunidades, propusemos na sessão seguinte, um debate a respeito da especificidade dessas comunidades, não no sentido de uma classificação de espaços físicos, mas sim, no aspecto de intenções explícitas que movem as CDs, isto é, da dimensão de suas razões de um *locus*, explicitamente, de posicionamento.

3.1 Em construção um entendimento de “comunidades discursivas com focos decoloniais e críticos-sociais”

As especialidades das CDs podem ser diversas, como notadamente se presume, isto é, podem ser de caráter educativos quando circunscritos em ambientes de ensino, tais como escolas, universidades, entre outros. Podem se referir a ambientes comunitários, empresariais, de saúde, alimentícios e tudo que engloba as diferentes divisões do trabalho em sociedade. No entanto, consideramos que podemos ainda, especificar um pouco mais, delimitando o tipo de proposta vigente da comunidade, se esta visa rupturas com o estado vigente ou não. Há casos de CDs de cunho educativo preocupados com a dimensão decolonial, um exemplo seria no contexto brasileiro os membros da Pedagogia Griô. Outro exemplo de ruptura, no campo das ciências agrárias há a especificidade de cunho agroecológico com o uso de sistemas agroflorestais (SAFs) para produção de alimentos de forma ecológica e restauração de florestas, em oposição à visão das monoculturas, que se denomina agricultura convencional.

Essas especialidades também podem dialogar com a proposta de tipologia de domínios em que Barité (2020, p. 245) estuda: 1) os domínios emergentes, definido enquanto, “[...] *aquellos de desarrollo reciente (20 a 40 años), que son consecuencia del vertiginoso proceso de especialización y de intersección recíproca entre disciplinas y/o campos temáticos*” (BARITÉ, 2020, p. 242) e; 2) os domínios de emergência súbita, enquanto um tipo daquele, mas que por ser súbito, ainda tem poucos materiais produzidos.

Ainda que haja consenso entre domínios e CDs, considera-se a distinção apresentada na sessão anterior, e utiliza-se das CDs para tratar dessas especificidades. Ao se estudar sobre domínios, um dos pontos que fica evidente na literatura da análise de domínio e ou/ os

conceitos relacionados a ele é o de que possuem uma perspectiva social. Esse aspecto é explicitamente citado nos trabalhos da área de OC. Mas, afinal, qual o sentido do termo “social” usado na OC? Em outras palavras, quando nos referimos à existência de um paradigma ou dimensão social a partir do uso dessas teorias/metodologias, esse sentido visa algum aspecto de transformação ou ruptura com algo já estabelecido – isto é – com algumas características da sociedade atual? Ou se utiliza para se referir a uma forma genérica de sociedade? Precisamos pontuar o sentido do termo social para esses contextos. Longe de responder a essas perguntas, elas nos interessam enquanto guia para nossa especificidade deste trabalho.

Adotar o termo social sem uma descrição do que se entende por ele, pode resultar na falta de direcionamento de um foco. Ainda que seja um campo de muitas possibilidades teóricas, estabelecer um recorte orienta um posicionamento. Por exemplo, o contexto no qual alguns autores/as europeus falam de “social” pode não contemplar os aspectos do “social” para a região geográfica/cultural e política latino-americana. Considerando esse panorama, utilizamos o “social” aqui, tendo em vista, a descrição do verbete “questão social”² da enciclopédia contemporânea da América Latina e do Caribe (2006).

Nesse verbete, o uso do termo “questão social” está orientado levando em consideração a realidade latino-americana que passou por um processo histórico de colonização, gerando diversos tipos de opressões. As categorias propostas nesse verbete dizem respeito aos encontros e às manifestações das quais fazem parte as desigualdades sociais.

Podemos ainda elencar com essa dimensão da questão social as ideias dos estudos decoloniais. Sabemos que o campo dos estudos decoloniais é muito amplo e envolve diferentes campos do conhecimento, como ocorre com as teorias críticas. Se tratando de decolonialidade, o foco centra-se no reconhecimento de que há a colonialidade, isto é, essa forma de dominação ainda existe, enquanto uma “lógica global de desumanização”. Nesse sentido, a decolonialidade “refere-se à luta contra a lógica da colonialidade e seus efeitos materiais, epistêmicos e simbólicos”, sendo, portanto, um conceito que está alinhado com o conceito de libertação (MALDONADO-TORRES, 2018, p. 36).

Levando em conta as concepções acerca do decolonial e questão social e, ainda considerando o conceito de comunidade discursiva, entende-se inicialmente por CD com focos decoloniais e críticas sociais - quando um grupo de pessoas (em diferentes âmbitos, universitário, de movimentos sociais, ONGs, ou outras formas) têm o seguinte interesse em comum: atuar/ escrever e pesquisar sobre como transformar/pautar formas de promover uma sociedade menos injusta, com inclusão de grupos excluídos por “cultura, raça, etnia, religião,

Alencar, Máira Fernandes; Cervantes, Brígida Maria Nogueira; Barité, Mario. Interseção de três lentes da análise de domínio para a caracterização de comunidades discursivas com focos decoloniais/críticos-sociais. *Organização e Representação do Conhecimento em diferentes contextos: desafios e perspectivas na era da datificação*, 2023.

posição socioeconômica, gênero, identidade de gênero, sexualidade, deficiência e status de cidadania, bem como reconhecer as interseções entre eles” (PACG, 2011, p. 72). No contexto brasileiro e da América Latina esses grupos são os povos e comunidades tradicionais (os povos originários brasileiros - indígenas, quilombolas, ciganos, entre outros), a população em situação de vulnerabilidade social, a população negra, os retirantes nordestinos, a população LGBTQI+, os camponeses de movimentos sociais, entre outros.

Nesse sentido, pode-se considerar que essas comunidades produzem discursos em diferentes linguagens, sejam eles escritos (artigos, livros, revistas), orais (contos, músicas, *podcasts*), performáticos (teatros, e outras manifestações artísticas/políticas), audiovisuais (filmes, documentários), entre outros possíveis, para comunicar e direcionar suas atividades/estudos preocupadas com situações de opressões em diferentes contextos no âmbito do saber, como educação, memória, saúde, alimentação, entre outros.

A partir de tais focos, geralmente há as preposições explícitas mediante termos e práticas com a existência de projetos, ensino e pesquisas, onde há inserção de termos específicos pautando tais discussões.

4 Metodologia

A metodologia utilizada para esse trabalho é de cunho exploratório e descritivo, a partir da literatura da área da Organização do Conhecimento. A proposta em pauta teve origem em uma tese da Arquivologia na qual caracterizou-se uma comunidade discursiva no âmbito da pesquisa e educação pluralista, na qual a comunidade discursiva escolhida foi a AERI (*Archival Education and Research Initiative*) - representativa de um foco crítico/social da área arquivística.

5 Resultados: Análises e discussões

Foi analisada a forma como Hjørland descreve e caracteriza as abordagens citadas, em seguida, como podemos compreendê-las, e que foco podemos dar a elas se tratando do objetivo de ações de cunho pluralista e inclusivo, como é o caso das CDs com focos decoloniais e críticos-sociais. Por fim, apresentamos a sistematização na Figura 1³ intitulada: elementos para caracterização de comunidades discursivas com focos decoloniais e críticos/sociais, que detalha todos os tipos de dados para serem coletados, e a forma de análise.

5.1 Sobre os estudos críticos e epistemológicos enquanto lentes epistêmicas de dimensão biográfica

Hjørland (2002, p. 440) argumenta que os estudos críticos e epistemológicos “são estudos que examinam as suposições explícitas ou implícitas por trás de tradições de pesquisa. Tais suposições estão

frequentemente ligadas a suposições ontológicas relacionadas ao objeto em estudo. Elas representam uma análise das abordagens ou paradigmas em áreas de pesquisas”. Ele menciona ainda 5 categorias de base epistêmica: 1. empirismo/positivismo; 2. racionalismo; 3. historicismo/ hermenêutica/ fenomenologia; 4. pragmatismo/ funcionalismo/ Marxismo/ feminismo; 5. ecletismo, pós-modernismo e ceticismo. Como, afinal, se identificaria a abordagem epistêmica em um artigo ou um *corpus* de trabalho? Um exemplo de aplicação dessa lente foi a tese de Araújo (2019), que buscou responder sobre a concepção de epistemologia no domínio da OC. Para tanto, analisou dez artigos que abordaram a respeito do conceito de epistemologia na OC. Analisando um *corpus* de trabalhos na Arquivística, se formos aplicar essa lente, podemos mencionar enquanto base epistêmica outras categorias específicas da área para identificar nos trabalhos, que se referem de forma implícita ou explícita ao ciclo vital, ao *records continuum*, ou ainda ao pluralismo arquivístico.

Como vimos, utilizar essa lente é buscar a base epistêmica, ou seja, as suposições que estão na base do trabalho. No caso de um *corpus* com muitos trabalhos e um tempo bem delimitado, pode tornar inviável de realizar a análise epistêmica como foi abordada por Hjørland. Porém, considerando outro foco para além da análise dos artigos e passando para os membros; uma interpretação dessa abordagem pode ser conhecer a base biográfica, ainda que micro, considerando dados disponíveis em páginas institucionais ou mediante entrevistas. Dessa forma, inserimos a dimensão biográfica para descrever temas de pesquisa, formação e projetos dos membros/as pertencentes a determinadas comunidades. Outra lente que poderia ser adaptada a este foco seria a dos estudos históricos. A diferença é que na epistêmica pode-se ver a coerência da biografia dos membros/as refletida nas publicações, o que, se tratando de um dado que contextualiza as origens e interesses, ajudaria, no sentido de revelar a coerência da continuidade de propostas estudadas pelos membros/as. Na síntese final constará a categoria de Campo Biográfico.

5.2 Sobre os estudos terminológicos e linguagens para propósitos específicos/estudos de discurso enquanto lente terminológica pontual

Hjørland agrupou estudos terminológicos com as linguagens para fins especiais de análise do discurso. Desse modo, o autor não caracteriza de forma explícita como usar os estudos terminológicos para analisar um domínio, mas cita o uso de construções de tesaurus e enfatiza a dimensão linguística. Afirma, ainda, que “a linguagem e terminologia são objetos muito importantes para a CI, porque elas afetam nosso pensamento e, assim, as questões que colocamos nas bases de dados, bem como os textos de busca.

Alencar, Máira Fernandes; Cervantes, Brígida Maria Nogueira; Barité, Mario. Interseção de três lentes da análise de domínio para a caracterização de comunidades discursivas com focos decoloniais/críticos-sociais. *Organização e Representação do Conhecimento em diferentes contextos: desafios e perspectivas na era da datificação*, 2023.

As reflexões a partir da compreensão dessas abordagens são acerca de uma lacuna observada por Barité e Rauch (2022, p. 11-12) em um artigo que analisa os estudos terminológicos como uma modalidade particular de análise de domínio. Destacamos uma das posições dos autores a respeito do que está situado enquanto estudos terminológicos, afirmando que “compreendem um conjunto de trabalhos de natureza diferentes (descritivos, analíticos, especulativos) [...]” e ainda que:

é possível acrescentar três ideias: a primeira é que, por sua natureza, os estudos terminológicos podem necessitar de contribuições de outras áreas do conhecimento (lexicografia, semântica e lógica, por exemplo). A segunda é que nem todos os estudos terminológicos são diretamente relevantes para a OC. A terceira é que mesmo quando identificada a presença ou influência de estudos terminológicos na OC, isso não implica uma referência automática a nenhuma das modalidades de AD (BARITÉ; RAUCH, 2022, p. 11-12).

A terminologia compreende pesquisas do tipo sistemática e pontual. Essa última foi escolhida enquanto lente a ser trabalhada. “A pesquisa terminológica pontual consiste em “uma técnica para investigar um termo ou um grupo restrito de termos pertencentes a qualquer campo da atividade humana, com vistas a satisfazer uma necessidade imediata expressa por um usuário” (CÉLESTINET al. 1984, p. 17). Esses autores também estabelecem uma tipologia de pesquisa pontual, que inclui a análise do significado de um termo, o termo mais adequado para representar um novo conceito, a certificação ou validação do uso de um termo, ou a busca por um termo exato ou aproximado e a equivalência em outra língua. Dubuc aponta que há quem subestime a pesquisa pontual ao considerar que “*un trabajo 'por pedazos' [um 'trabalho fragmentado'] no aporta realmente soluções de conjunto a las necesidades terminológicas*” (DUBUC 1999 p. 65). No entanto, ele encontra as seguintes vantagens em se envolver neste trabalho: 1. a solução para problemas urgentes de tradução e escrita, 2. a oportunidade de usar esse tipo de pesquisa como escola de formação, colocando o futuro terminólogo em contato direto com os usuários e suas necessidades e 3. a possibilidade de aprofundar a documentação existente em um determinado domínio (DUBUC, 1999). A essas vantagens, deve-se acrescentar a possibilidade de resolver rapidamente problemas de designação, conceituação, tradução e precedência histórica ou etimológica, continuamente apresentados por neologismos (BARITÉ; RAUCH, 2022, p. 7-8 - tradução nossa).

Por fim, a razão dessa lente para fins de proposta de ações ocorre em detrimento de uma vantagem do tipo de precedência contextual, para compreender o contexto de uso de um conceito no âmbito de uma comunidade em foco. Na síntese final constará na categoria de Campo da Prática.

5.3 Sobre estudos de estruturas e instituições em comunicação enquanto lente de estruturação científica

Hjørland (2022) menciona que quando se estuda as estruturas da divisão interna do trabalho de cada domínio, bem como as trocas que ali acontecem, tem-se informações úteis para compreender a razão da existência de determinados tipos de documentos e serviços, o que poderá ajudar inclusive na construção de guias de literatura. Para realizar essa estruturação abrem-se alguns caminhos. Para o autor:

[...] Essa área é rica em questões que estão abertas para, entre outros tipos de estudos, investigações bibliométricas. [...] Muitos dados quantitativos e qualitativos são necessários, para mapear tal estrutura de modo preciso. Quem são os produtores? Quanto eles produzem e comunicam e como esta comunicação é distribuída em diferentes canais? Como a comunicação é filtrada e influenciada por diferentes meios de comunicação? Qual é a cobertura tanto quantitativa quanto qualitativamente em diferentes bases de dados e bibliotecas? Que tipo de normas epistêmicas guiam o processo de seleção? Quão interdisciplinares são os diferentes agentes e instituições e que tipos de vieses podem ser envolvidos por influências disciplinares? Que tipo de tradições nacionais ou geográficas, normas culturais e influências econômicas estão em jogo? (p. ex. diferenças entre organizações com fins lucrativos e sem fins lucrativos). ((HJØRLAND, 2002, p.448 – tradução nossa)

Essas abordagens permitem identificar quem são os atores, as instituições presentes em dado domínio (GUIMARES, 2014) de tal forma que, para o âmbito da caracterização de comunidades podem mostrar se há algum padrão de publicação, ou predominância de determinada revista, membro ou, ainda, se prevalece determinado período da comunicação científica, entre outras possibilidades.

O foco se dará na quantidade de publicações em geral e por autor(a); na descrição dos periódicos, anos, listagem dos coautores e temas dos trabalhos. Possivelmente, esses dados poderiam ser por outra lente, a bibliográfica também.

A vantagem dessa lente para caracterizar comunidades discursivas está vinculada à ideia de dar visibilidade aos espaços em que se encontram as comunicações científicas de tais grupos. Nesse sentido, nosso trabalho, está amparado por um foco de comunicação. Na síntese final constará a categoria de Campo da Comunicação.

5.4 A intersecção das lentes epistêmica-biográfica, de estruturação científica e terminológica pontual

A partir dos estudos de Swales (2016) e Hjørland (2002) com as lentes e suas intersecções, podemos notar que a proposta de caracterização para comunidades de focos decoloniais e críticos/sociais contribui para: a visibilidade de alguma prática veiculada a tal grupo, ajudar a conhecer a comunidade mediante um contexto analítico, isto é, a partir de relações vinculadas entre si,

possibilitando uma via de acesso mais detalhada e que pode servir para gerar propostas de ações, já que possibilita a síntese de dados combinados entre si.

Com a sistematização da lente epistêmica-biográfica é possível entender o contexto político, social, cultural, econômico que tal membro/a vive/viveu, que influencia/ou na sua produção discursiva, relativo a tal grupo de interesse.

Ao levar em consideração o aspecto de que há condições históricas/políticas/culturais/econômicas nas biografias dos membros é possível identificar se nos focos de publicações há coerência na trajetória do membro/a, aspecto que possibilita visualizar as dimensões implícitas e explícitas de tais propostas.

Outro aspecto é a identificação dos temas de engajamento, a partir da análise de contexto do uso de determinado termo. Em outras palavras, identificando e analisando a dimensão da terminologia pontual, certamente, chega-se também à identificação de temáticas de engajamento expresso mediante conceitos específicos. E, ainda, um mapeamento das fundamentações utilizadas, isto é, as garantias literárias (ou outros tipos de garantias) nas quais tais membros se apoiam para fundamentar seu *locus* de discurso/prática.

Assim, nesse processo de investigação de quanto uma lente tende a se beneficiar quando combinada com outra ou outras, reside o aspecto da riqueza semântica e da dimensão sintética que esse método possibilita.

6 Considerações Finais

As abordagens sugeridas por Hjørland (2002) parecem seguir atuais, do ponto de vista a serem aplicadas em diferentes contextos com foco na análise de domínios variados, como exemplificam os trabalhos de dissertações e teses que temos visto no Brasil e em outros lugares do mundo que utilizam tal método. No entanto, não é comum discussões que tratem das limitações e/ou formas de combinações possíveis para comunidades discursivas com focos decoloniais e críticos/sociais. Foi diante desse contexto que apresentamos a proposta de intersecção das três lentes: epistêmica-biográfica, de estruturação científica e terminológica pontual. Certamente, outras combinações também são possíveis. Essa proposta de intersecção a partir desse tipo de coleta de dados, tem a ver com a perspectiva que Hjørland (2017, s/p – tradução nossa) coloca quando afirma que “[...] é muito mais importante aprender sobre o desenvolvimento (teórico) de um domínio do que simplesmente mapear quais tópicos foram mais estudados [...]”. A escolha dessas três lentes se deu mediante a constatação de que são representantes de um desenvolvimento, levando em consideração os membros/termos, e formas de comunicação.

Por se tratar de um caminho, uma via de acesso para conhecer e caracterizar um determinado domínio ou comunidade discursiva, pode resultar em significado, do ponto de vista analítico, quando se descreve as consequências, ou desdobramentos de diferentes tipos de intersecções.

Como apresentado na Figura 1 (Apêndice A) elementos para caracterização de comunidades discursivas com focos decoloniais e críticos/sociais, podemos notar que esse tipo de intersecção atua de modo problematizador, uma vez que, ao invés de um mapeamento genérico, o que se visa fazer é, portanto, um roteiro para perguntar a respeito das condições que possibilitaram a sistematização desse conhecimento por tal comunidade -, a partir de tais **membros**, dos **termos** escolhidos para analisar pontualmente, e da **comunicação** escolhida com toda estrutura inerente, para divulgar tais ideias. De forma que essa tríade membros/termos /comunicação, elenca aspectos implícitos da formação e interesses dos membros, da pertinência/fundamentação em usar determinados termos e da lógica que assume a forma de comunicar essas sistematizações.

Os discursos das comunidades discursivas com focos decoloniais e críticos sociais são pautados em mudanças epistêmicas com horizontes de libertação/emancipação em algum nível considerando a necessidade de inclusão e justiça social/direitos humanos para grupos historicamente marginalizados. Nesses discursos há termos específicos que retratam formas de atuação. Passa a ser muito comum que essas comunidades integrem também grupos de movimentos sociais, podendo ser membros/as de universidades, institutos, ONGs, entre outros.

Considerando esse panorama, quando caracterizamos tais comunidades mediante a tríade “membros/termos/comunicação” ajudamos na visibilidade de seus discursos e na compreensão da lógica de como tal comunidade se organiza. Ou seja, diante do excesso de informações, de certa forma, a intersecção apresentada no trabalho se propõe a tornar essas comunidades mais visíveis a partir de um panorama sintético de como elas operam. Ainda que não seja uma tarefa fácil determinar exatamente se de fato uma comunidade discursiva opera a partir de uma visão de rupturas, no sentido apenas teórico, da dimensão de englobar grupos socialmente marginalizados, ou pautar algo desse tema, alguns pontos podem ajudar nessa determinação. Por exemplo: as preposições explícitas mediante termos e práticas, a existência de projetos, a continuidade de pesquisas com tal debate, ou ainda a inserção de um conceito pautando discussões. Tais pontos podem ser ampliados, servindo, aqui, enquanto caminho inicial. Ressaltamos que esse debate sobre tais comunidades, ainda, pedem novas reflexões e estudos.

Por fim, como último aspecto nos caberia, ainda, questionar outras vias de possíveis análises para

caracterizar essas comunidades. Por exemplo: até que ponto o uso metodológico da AD realmente pode ajudar a caracterizar comunidades discursivas com tais focos? Seria necessário, também, outros métodos para esse objetivo de caracterização? Certamente sim. Ou ainda, questionar a razão de uso desses conceitos que foram sistematizados a partir de outras realidades. No entanto, para fins estratégicos, já que essa teoria/metodologia da AD se encontra estabelecida no cerne da literatura brasileira e latino-americana da OC, e apresenta flexibilidade de debates com autores/as receptivos/as à ampliação, usá-la e criticá-la sob um ponto de vista para fins de mudanças epistêmicas nos parece resultar em algo estrategicamente necessário.

Notas

[1] Com base no trabalho de Rampazzo e Aranha (2019) cabe ressaltar que em 1998, Swales já havia iniciado as discussões de categorização de comunidades discursivas, nesse trabalho há a inserção de *comunidade discursiva de lugar e comunidade discursiva focal*.

[2] Outra referência que trata dessa dimensão é o livro de Alejandra Postorini (2010) “A categoria questão social em debate”, em que há um capítulo específico a respeito das manifestações da questão social na dimensão da América Latina.

[3] Na tese em que esse estudo se fundamentou é possível ver os resultados, e os exemplos das intersecções das lentes, além dos modelos das fichas.

Referências

ALENCAR, M. F. **Os Arquivos Comunitários na pesquisa e educação arquivística**: Uma análise a partir da *Archival Education Research Initiative*. 2022. 242 p. Tese. (Doutorado em Ciência da Informação). Universidade Estadual de Londrina (UEL). Londrina, 2022.

ARAÚJO, P. C. “**Epistemology of knowledge organization: a metatheoretical investigation**”. Tese de Doutorado. Universidade Estadual Paulista Júlio de Mesquita Filho, Programa de Pós-graduação em Ciência da Informação. 2019.

BARITÉ, M. “**Los árboles de dominio**”. Teoría y praxis en terminología. Coordinated by Sara Álvarez Catalá and Mario Barité. CSIC, 2017, p. 91-102.

BARITÉ, M. “**Los árboles de dominio**”. Teoría y praxis en terminología. Coordinated by Sara Álvarez Catalá and Mario Barité. CSIC, 2017, p. 91-102.

BARITÉ, M. La definición terminológica. In: CATALÁ, S. A.; BARITÉ, M. (Coord.). **Teoría y praxis en terminología**. Montevideo: Universidad de la República, 2017.p.45-67.

BARITÉ, M.; RAUCH, M. Terminological studies as domain analysis: a critical exploration. **Brazilian Journal of Information Science**, v. 16, 2022. DOI: 10.36311/1981-1640.2022.v16.e02140.

CÉLESTIN, T.; GODBOUT, G.; VACHON-L'HEUREUX, P. **Méthodologie de la recherche terminologique ponctuelle**. Québec: Office de la Langue Française, 1984.

DUBUC, R. **Manual de terminología**. Unión Latina, 1999.

EVANGELISTA, I. V; GRÁCIO, M. C. C; GUIMARÃES, J. A. C. The Concepts of Domain, Discourse Community and Epistemic Community: affinities and specificities. **Brazilian Journal of Information Science: Research Trends**, v.16, 2022.

GUIMARÃES, J.A.C. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, v. 43, n. 1, pp.13-21, 2014.

GUIMARÃES, J.A.C. “Aspectos éticos em Organização e Representação do Conhecimento (ORC): uma reflexão preliminar”. **Políticas de memória e informação: reflexos na organização do conhecimento**. Organizado por: Maria Nélida González de Gómez e Evelyn Goyannes Dill Orrico. EDUFRRN, 2006.

HJØRLAND, B. “Domain analysis in information science: eleven approaches: traditional as well as innovative”. **Journal of Documentation**, 58 (4): p. 422–462. 2002.

HJØRLAND, B. Domain analysis. **Knowledge Organization**, v. 44, n. 06, 2017, pp. 436-464.

HJØRLAND, B; ALBRECHTSEN, H. “Toward a new horizon in information science: domain-analysis”. **Journal of The American Society for information Science**, 46, n.6, p. 400-425, 1995.

MALDONADO-TORRES, N. Analítica da colonialidade e da decolonialidade: algumas dimensões básicas. In: BERNARDINO-COSTA, Joaze; GROSGOUEL, Ramón. (org.). **Decolonialidade e pensamento afrodiáspórico**. 2. ed. Belo Horizonte: Autêntica, 2020, p. 9-26.

PASTORINI, A. **A categoria “questão social” em debate**. 3. ed. São Paulo: Cortez, 2010.

PLURALIZING THE ARCHIVAL CURRICULUM GROUP (PACG) (A. Gilliland, corresponding author), Educating for the Archival Multiverse. **American Archivist**, Spring/Summer, 2011, p.68-102.

RAMPAZZO, L.; ARANHA, S. Revisitar o conceito de comunidade para discutir sua aplicação a contextos telecolaborativos. **ALFA: Revista de Linguística**, São Paulo, v. 63, n. 2, 2019.

SMIRAGLIA, R. Epistemology of domain analysis. In Smiraglia, Richard P. and Hur-Li Lee, eds. **Cultural frames of knowledge**. Würzburg: Ergon-Verlag, 2012.

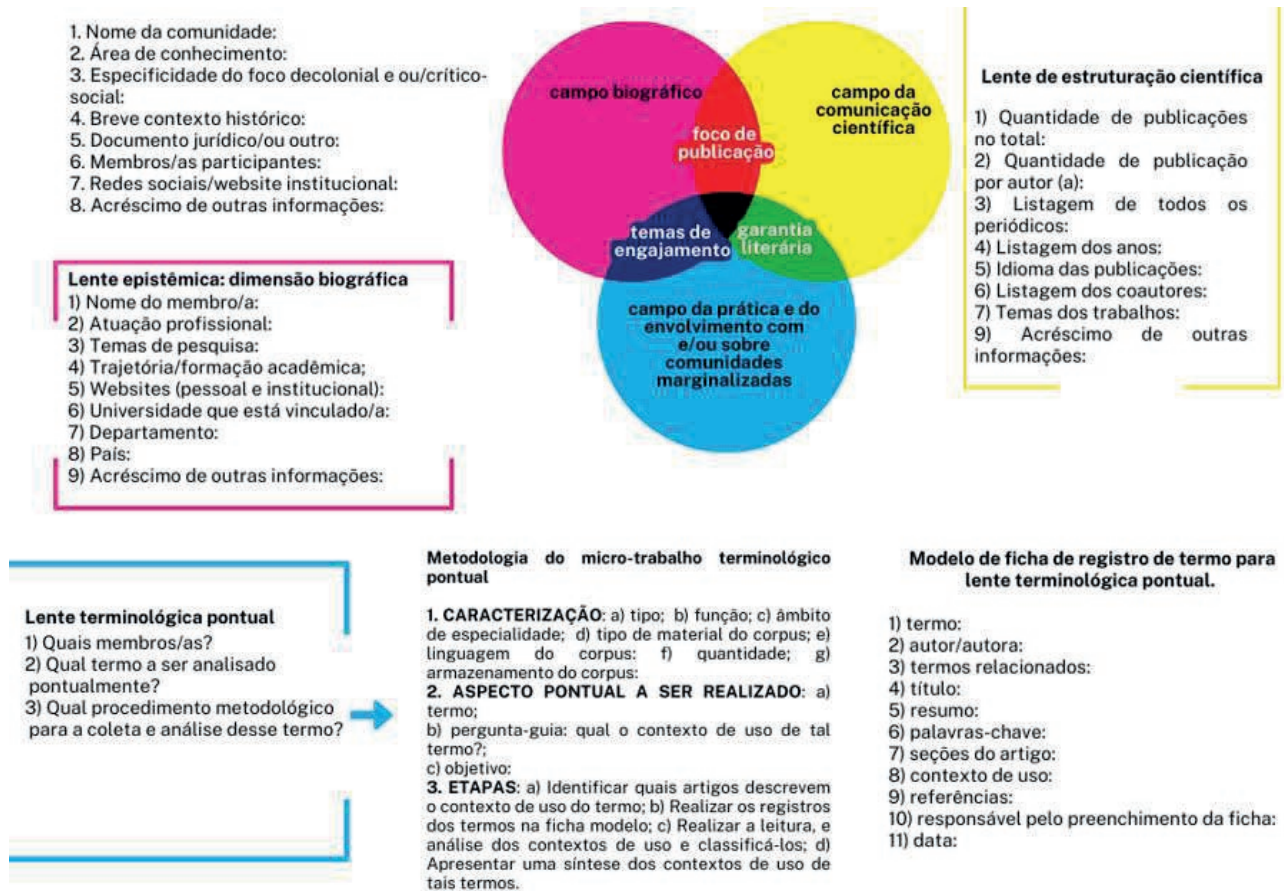
SMIRAGLIA, R. P. **Domain Analysis for Knowledge Organization**. New York: Springer, 2015.

SWALES, J. M. “The Concept of Discourse Community”. In Swales, Jhon M. **Genre analysis: english in academic and research settings**. Cambridge, 1990, pp. 21-32.

SWALES, J. M. Reflections on the concept of discourse community. **Asp**, Bordeaux v.69, p. 7-19, 2016.

TÊNIS, J. Two Axes of Domains for Domain Analysis. **Knowl. Org.** 30, 2, n.3/4. 2003.

Apêndice A – Figura 1. Elementos para caracterização de comunidades discursivas com focos decoloniais e críticos/sociais.



Fonte: Elaboração própria (2023).

Linguagem Documentária e Terminologia: fundamentos linguísticos para a organização do conhecimento

Documentary Language and Terminology: linguistic foundations for the knowledge organization

Alexandre Robson Martines (1), Wilson Roberto Veronez Júnior (2), Edmilson Alves dos Santos Júnior (3), Daniel Martínez-Ávila (4)

Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP-Marília), (1) alexandre.martines@unesp.br, (2) wilson.veronez@unesp.br. (3) Universidade Federal do Rio de Janeiro/IBICT, edmilson.junior@unesp.br, (4) Universidad de León, Espanha, dmarta@unileon.es.

Resumo

O processo de organização do conhecimento está intrinsecamente alinhado à condição lógica de precisão e de metalinguagem evocada pela linguagem documentária, aplicada para estabelecer a classificação e a categorização. Nessa linha, a linguagem especializada assume um papel fundamental de conduzir o processo classificatório entre os fatores linguísticos, cognitivos e técnico-científicos. Com isso, produtos de organização e representação podem orientar a elaboração mais complexa de objetos informacionais, tratando a informação para efetuar a comunicação entre sistemas informacionais e usuários. Assim, é objetivo desta pesquisa analisar os recursos da terminologia para que este dê condições ao tesouro para atuar como elementos conceitual e comunicativa, além de apontar os fundamentos linguísticos nesse processo de representação e mediação. Trata-se de uma pesquisa teórica bibliográfica com análise descritiva e natureza de método com abordagem qualitativa, a partir da leitura e interpretação de artigos científicos da área em bases de dados nacionais e internacionais. Portanto a terminologia atua em várias camadas da interação linguística, possuindo características complexas, as quais devem ser aprofundadas a fim de que se possa analisar as dimensões de aplicabilidade e como conduzir o estudo sobre o significado e a significação. Desse modo, torna-se importante estudos alinhados aos aspectos linguísticos a fim de aproximar o uso de terminologia ao público em geral, visto que é uma ferramenta essencialmente direcionada a especialistas, a cientistas e a estudantes, todavia a familiaridade com esse estrato linguístico formal técnico-científico significa a ampliação do conhecimento especializado de uma sociedade.

Palavras-chave: Terminologia. Linguagem Documentária. Linguística. Tesouro. Organização do Conhecimento.

Abstract

The knowledge organization process is intrinsically aligned with the logical condition of precision and metalanguage evoked by the documentary language, applied to establish classification and categorization. In this line, specialized language assumes a fundamental role in conducting the classification process between linguistic, cognitive and technical-scientific factors. With that, organization and representation products can guide the more complex elaboration of informational objects, treating the information to carry out the communication between informational systems and users. Thus, the objective of this research is to analyze the resources of terminology so that it gives conditions to the thesaurus to act as conceptual and communicative elements, in addition to pointing out the linguistic foundations in this process of representation and mediation. This is a theoretical bibliographical research with descriptive analysis and method nature with a qualitative approach, from the reading and interpretation of scientific articles in the area in national and international databases. Therefore, terminology operates in several layers of linguistic interaction, having complex characteristics, which must be deepened in order to analyze the applicability dimensions and how to conduct the study on meaning and significance. In this way, it becomes important to carry out studies aligned with linguistic aspects in order to bring the use of terminology closer to the general public, since it is a tool essentially directed at specialists, scientists and students, however, familiarity with this formal technical linguistic stratum -scientific means the expansion of a society's specialized knowledge.

Keywords: Terminology. Documentary Language. Linguistics. Thesaurus. Knowledge Organization.

Resumen

El proceso de organización del conocimiento está intrínsecamente alineado con la condición lógica de precisión y metalenguaje que evoca el lenguaje documental, aplicado para establecer clasificaciones y categorizaciones. En esta línea, el lenguaje especializado asume un papel fundamental en la conducción del proceso de clasificación entre factores lingüísticos, cognitivos y técnico-científicos. Con ello, los productos de organización y representación pueden orientar la elaboración más compleja de objetos informacionales, tratando la información para realizar la comunicación entre los sistemas informacionales y los usuarios. Así, el objetivo de esta investigación es analizar los recursos de la terminología para que ésta dé condiciones a los tesauros para actuar como elementos

conceptuales y comunicativos, además de señalar los fundamentos lingüísticos en este proceso de representación y mediación. Se trata de una investigación bibliográfica teórica con análisis descriptivo y carácter metodológico con enfoque cualitativo, a partir de la lectura e interpretación de artículos científicos del área en bases de datos nacionales e internacionales. Por lo tanto, la terminología opera en varias capas de interacción lingüística, teniendo características complejas, que deben ser profundizadas para analizar las dimensiones de aplicabilidad y cómo realizar el estudio sobre el significado y la significación. De esta manera, se vuelve importante realizar estudios alineados con los aspectos lingüísticos para acercar el uso de la terminología al público en general, ya que es una herramienta dirigida esencialmente a especialistas, científicos y estudiantes, sin embargo, la familiaridad con esta técnica formal estrato lingüístico-científico significa la expansión del conocimiento especializado de una sociedad.

Palabras clave: Terminología. Lenguaje Documental. Lingüística. Tesouro. Organización del conocimiento.

1 Introdução

O processo de organização do conhecimento está intrinsecamente alinhado à condição lógica de precisão e de metalinguagem evocada pela linguagem documentária, aplicada para estabelecer a classificação e a categorização. Nessa linha, a linguagem especializada assume um papel fundamental de conduzir o processo classificatório entre os fatores lingüísticos, cognitivos e técnico-científicos.

Organizar e representar o conhecimento é uma atividade laboral desenvolvida na Ciência da Informação (CI) e na Biblioteconomia, e até mesmo na Arquivologia, conforme defende Esteban Navarro (1993). No cenário da CI, é potencial que seus produtos de organização e representação possam orientar a elaboração mais complexa de objetos informacionais, tratando a informação para efetuar a comunicação entre sistemas informacionais e usuários, bem como ampliar essa interação para que o material presente nos objetos informacionais sejam assistentes na condução de mediação entre informação e usuários.

Ademais, o desenvolvimento do campo da Organização do Conhecimento permitiu que se desenvolvessem serviços, processos e produtos direcionados, com recursos especializados, para explorar os fatores metafísicos e ontológicos do conhecimento e, assim, classificá-los, com condição de enquadrá-los consoante o domínio, por conseguinte ajustar-se à linguagem que permitir o diálogo e a formatação do discurso entre os especialistas, a ciência, a produção científica, estudiosos e interessados no assunto. Nessa linha, a linguagem documentária está alinhada a um estrato lingüístico de formalidade técnico-científica: a terminologia.

Assim sendo, a Terminologia, como área de estudo, oferece subsídios ao processo de organização, representação e mediação da informação. Por um lado, atua como norteadora da elaboração de linguagens documentárias, em destaque o tesouro, a fim de que atuem como recurso da representação do conhecimento, através dos Sistemas de Organização do Conhecimento (SOCs); por outro, devido ao seu fundamento lingüístico, bem como sua formatação semântica, pragmática, semiótica, funcionalista e discursiva, pode servir de recurso lingüístico-comunicacional para que se efetue a organização e a representação da informação, ou seja, é possível que esse tesouro, cuja natureza

primária é ser o tesouro cognitivo, ou melhor, a diretriz da correlação conceitual que atua na constituição do conhecimento, também tenha condições de se tramitar em aparato comunicativo e, desse modo, ser instrumento aplicado na indexação.

1.1 Objetivos

Frente a isso, é objetivo desta pesquisa analisar os recursos da terminologia para que esta dê condições ao tesouro para atuar como elemento conceitual e comunicativo, além de apontar os fundamentos lingüísticos nesse processo de representação e mediação.

1.2 Metodologia

Trata-se de uma pesquisa teórica bibliográfica com análise descritiva e com método de abordagem qualitativa, a partir da leitura e interpretação de artigos científicos da área em bases de dados nacionais e internacionais.

A pesquisa teórica “é dedicada a reconstruir teorias, conceitos, ideias, ideologias, polêmicas, tendo em vista, em termos imediatos, aprimorar fundamentos teóricos e, em termos mediatos, aprimorar práticas...” (DEMO, 2000, p. 20), assim, a pesquisa teórica contribui para o conjunto do conhecimento bibliográfico registrado.

Segundo Gil (2002), a pesquisa bibliográfica baseada em materiais publicados, tem a vantagem de oferecer uma abrangência mais ampla do fenômeno em estudo do que das possibilidades que o próprio pesquisador deve investigar. Nesse sentido, outra vantagem da pesquisa bibliográfica é permitir que os pesquisadores conduzam extensas descobertas sobre os fatos ocorridos. No entanto, é importante que se aprofunde na análise e verifique a veracidade do conteúdo, em questão, os qualitativos.

A natureza da pesquisa é de método com abordagem qualitativa, desse modo entende-se que esta tem uma “[...] relação dinâmica entre o mundo real e o sujeito, isto é, um vínculo indissociável entre o mundo objetivo e a subjetividade do sujeito que não pode ser traduzido em números” (KAUARK *et al.*, 2010, p. 26). Nessa perspectiva, sobre o entendimento daquilo que não pode ser explicado de modo quantitativa, ou seja,

O pesquisador procura reduzir a distância entre a teoria e os dados, entre o contexto e a ação, usando a lógica da análise

fenomenológica, isto é, da compreensão dos fenômenos pela sua descrição e interpretação (TEIXEIRA, 2005, p. 137).

O percurso metodológico se deu pelo levantamento bibliográfico dos artigos científicos extraídos nas bases de dados *Library Information Science Abstracts* (LISA), Scopus, Base de dados de Periódicos em Ciência da Informação (BRAPCI), além de livros especializados, entre outras. Sobre uma abordagem de análise descritiva, buscou-se realizar a seleção dos artigos nessas bases para interpretação dos dados a serem analisados.

De acordo com Silva e Menezes (2000, p. 21) “[...] a pesquisa descritiva visa a descrever as características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis”. Desse modo, “envolve o uso de técnicas padronizadas de coleta de dados: questionário e observação sistemática. Assume, em geral, a forma de levantamento” (SILVA; MENEZES, 2000, p. 21). Em complemento, “as pesquisas descritivas têm como objetivo primordial a descrição das características de determinada população ou fenômeno, ou então, o estabelecimento de relações entre variáveis” (GIL, 1999, p. 46).

A princípio, a revisão bibliográfica buscou-se amparar nos campos da Biblioteconomia, Ciência da Informação, da própria Organização do Conhecimento, além dos aspectos linguísticos que permeiam a terminologia para compreensão mais abrangente acerca da discussão que este estudo se propõe, a fim de encontrar dados nas fontes seminais e também mais recentes, sobre diferentes discussões que possam colaborar para o construto deste estudo e uma discussão mais completa da pesquisa.

Desse modo, a importância da pesquisa é vista e entendida como, uma “atividade básica das ciências na sua indagação e descoberta da realidade. É uma atitude e uma prática teórica de constante busca que define um processo intrinsecamente inacabado e permanente” (MINAYO, 1994, p. 23).

2 Organização do Conhecimento

Com base em Hjørland (2003), é possível identificar que a Organização do Conhecimento (letra maiúscula) se trata de um campo científico, histórico, epistemológico e interdisciplinar. Já a organização do conhecimento (letra minúscula) tem relação com processos, funções e atividades desempenhadas no interior das unidades de informação por bibliotecários, arquivistas, documentalistas, museólogos, especialistas em informação e cientistas da informação por intermédio de instrumentos e sistemas de informação e de organização do conhecimento.

A evolução histórica da OC (campo e processo) está ligada à trajetória cultural, política, econômica, tecnológica, entre outras muitas, sendo um conceito em constante mudança (SAN-SEGUNDO, 2012, tradução nossa).

Dahlberg (2006, p. 12, tradução nossa), enquanto campo define a OC como

A ciência que estrutura e organiza sistematicamente unidades do conhecimento (conceitos) segundo seus elementos de conhecimento (características) inerentes e a aplicação desses conceitos e classes de conceitos ordenados a objetos/assuntos.

Na Biblioteconomia e na Ciência da Informação, a OC utiliza-se de instrumentos representados pelos sistemas de classificação, tesouros, vocabulário e outras linguagens de indexação. Como abordagens predominantes destacam-se as questões relativas a aspectos cognitivos, tecnológicos, lógico-conceituais, socioculturais e de gestão, o que sinaliza para os diálogos mais presentes na área (HJORLAND, 2003, GUIMARÃES *et al*, 2015).

Em termos de processo, para Amorim e Brascher (2017, p. 364), o objetivo da organização do conhecimento é entender o conhecimento em ação, isto é, o estudo de sua transferência, que se dá por meio de artifícios variados de informação, integrados num sistema informacional. Isso implica no usuário ter acesso garantido aos conhecimentos registrados nos documentos armazenados de forma física ou digital por meio dos SOCs.

Enquanto processo, a organização do conhecimento refere-se às atividades, tais como descrição de documentos, indexação e classificação, realizadas em bibliotecas, bases de dados bibliográficas, arquivos e instituições de memória, por bibliotecários, arquivistas, especialistas em informação, especialistas em assunto, assim como por algoritmos de computador e leigos (MELO; BRASCHER, 2014, p. 75).

Sales e Murguia (2015), lembram que a organização do conhecimento não se trata de uma atividade exclusiva aos profissionais da informação e não se resume às unidades de informação, como as bibliotecas, ela também pode ser desempenhada por outros profissionais, assim como por outras instituições, sobretudo aqueles que lidam com a organização e comunicação do conhecimento, como as escolas, instituições religiosas, políticas etc.

Sobre os tipos de organização, Hjørland (2003, p. 93, tradução nossa), recorda que

O campo da Organização do Conhecimento trabalha com dois: a) a organização intelectual do conhecimento também pode ser denominada como a organização cognitiva do conhecimento. Isso é basicamente a organização do conhecimento em sistemas conceituais e teorias. Se usarmos os sistemas periódicos da química ou a taxonomia

zoológica da biologia como bases dos sistemas de indexação, estamos usando sistemas intelectuais ou cognitivos de organização do conhecimento; tabelas hierárquicas são exemplos disso, como a tabela periódica de elementos químicos; b) A organização social do conhecimento, por outro lado, é basicamente as organizações das profissões, funções práticas, ofícios e disciplinas científicas. Se nos referirmos a disciplinas em nossas representações de conhecimento, estamos usando sistemas sociais de organização do conhecimento.

A organização do conhecimento vale-se de instrumentos como sistemas de classificação, tesouros, vocabulário e outras linguagens de indexação. Como abordagens predominantes destacam-se as questões relativas a aspectos cognitivos, tecnológicos, lógico-conceituais, socioculturais e de gestão, o que sinaliza para os diálogos mais presentes na área (GUIMARÃES *et al.*, 2015).

Segundo Ohly (2015), a Organização do Conhecimento é denominada com outros termos no campo da ontologia, taxonomia, terminologia, mapas de tópicos e arquitetura de informação. Aqui a Organização do Conhecimento não é identificada como um campo em si, como componentes básicos do conhecimento (classes, hierarquias, termos etc.) comumente tidos como certos. O que falta é uma terminologia comum e consistente nesse campo. O autor prefere a visão de dimensões de Organização do Conhecimento no sentido ontológico, epistemológico, pragmático etc.

2.1 Sistemas de Organização do Conhecimento

Os Sistemas de Organização do Conhecimento (SOCs), podem ser definidos como sistemas abstratos que dão sentido aos termos, definições, relações e atributos aos conceitos (CARLAN; MEDEIROS, 2011). Entende-se que os SOCs são sistemas estruturados que visam à sistematização do conhecimento, e que também procuram viabilizar o acesso e a recuperação da informação e do conhecimento.

Para Zeng e Mai Chan (2004), os SOCs é um termo que se refere aos instrumentos que apresentam a interpretação organizada das estruturas e sistematização do conhecimento, visando a facilitar o acesso e a recuperação da informação.

Carlan e Medeiros (2011) explicam que os SOCs cumprem a função de padronização dos termos, cujo objetivo é a orientação que facilita a indexação da maneira mais correta para o usuário acessar e recuperar a informação futura.

Nessa perspectiva, podemos entender que os SOCs,

São instrumentos que fazem a tradução dos conteúdos dos documentos originais e completos, para um esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, consequentemente,

facilitar a recuperação das informações contidas nos documentos (CARLAN; MEDEIROS, 2011, p. 55).

Os SOCs referem-se a uma ampla variedade de itens, que consiste em assunto, títulos, esquemas de classificação, ontologias, tesouros, classificação, dicionários, entre outros aspectos voltados para a representação e a organização da informação e do conhecimento (MAZZOCCHI, 2018, tradução nossa).

Essa instrumentalização e sistematização contribui no acesso e na recuperação da informação pelos usuários nos mais variados sistemas. Com relação ao conhecimento que deveria ser organizado ou seguir uma ordem previamente estabelecida, Mazzocchi (2018, tradução nossa) entende que qualquer tipo de conhecimento, cultura ou suas representações deveriam seguir algum tipo de ordem e que tenha algum tipo de relação com aquela sociedade.

No sentido metodológico e aplicado, Moreira (2018) menciona que os SOCs são representados pelos sistemas de classificação, as listas de cabeçalhos de assuntos, os tesouros, as taxonomias e as ontologias. Para o autor, em comum, tais instrumentos, identificam a intenção e a extensão dos conceitos, organizam as relações conceituais e primam pelo controle terminológico.

Um grande exemplo de SOCs são os tesouros, que, por sua vez, derivam do latim e do grego thesaurus que significava tesouro. Com o tempo passou a usar-se para designar o vocabulário de filologia, designadamente os dicionários especializados de uma determinada área de conhecimento. O primeiro emprego desta palavra data de 1531. Um tesouro constitui um sistema de indexação para representação conceitual, um vocabulário controlado, instrumento de organização, representação e recuperação do conhecimento (SANTOS, 2010, p. 18).

Os tesouros são instrumentos frequentemente utilizados na representação e na organização do conhecimento, além disso são constituídos à base de uma terminologia especializada. Para Hjørland (2010), o elemento básico de um tesouro pode ser considerado os conceitos, que são entendidos como termos que representam o sinônimo que tem sido visto de forma sem ambiguidade de palavras com diferentes significados. São organizados alfabeticamente, às vezes, também sistematicamente, com a subestrutura representando as relações entre conceitos (relações semânticas).

Segundo Cunha e Cavalcanti (2008, p. 362), o tesouro é um “vocabulário controlado e dinâmico de descritores relacionados semântica e genericamente, que cobrem de forma extensiva um ramo específico de conhecimento”. Cunha e Cavalcanti (2008, p. 362), o tesouro é considerado uma ferramenta de terminologia que tem por função a tradução de uma linguagem artificial, isto é, a linguagem documentária e a linguagem de indexação.

No âmbito da Documentação, podemos entender os tesouros como um instrumento de controle terminológico que apresenta eficácia na organização do conhecimento, e conseqüentemente uma importante ferramenta no que tange ao tratamento e a recuperação da informação (CARLAN; MEDEIROS, 2011, p. 56). São tratados como dispositivos para a organização e sistematização do conhecimento, tendo como objetivo final tratar e recuperar a informação registrada em seu suporte informacional.

2.2 Representação do Conhecimento

De acordo com Lara (1999), o termo representação está diretamente relacionado às narrativas da literatura na filosofia, terminologia, linguística e sociologia. O processo de organização e representação se faz presente na humanidade desde os primórdios, por meio de uma multiplicidade de conceitos evocados.

Para Caixeta e Souza (2008), a representação é um processo antigo quanto o advento da civilização, assume sua forma mais vívida na escrita e é ainda definida por regras de organização de elementos para a produção de significados, assim, “talvez um dos trabalhos mais angulares de significados da nossa civilização tenha sido a representação dos seres, das coisas, ideias e fenômenos pelo alfabeto” (CAIXETA; SOUZA, 2008, p. 35).

Na filosofia kantiana, a percepção humana há a captação das coisas por meio de sentidos, os quais, se originaram por sensações desordenadas e caóticas, contudo, o pensamento era responsável por ordenar os sentidos, quanto ao tempo e ao espaço, havendo uma relação e o estabelecimento da organização (MORAES; ARCELLO, 2000). As organizações se apresentam como representações organizadas da realidade.

Para Kant (1980), o conhecimento se constitui em formas de sistemas compreensíveis que podem ser determinados, e somente sobre a totalidade e a organização se proverá da validade e do valor de todos os elementos que o compõem.

Já para Barité (2015, p. 136) destaca que a Representação do Conhecimento:

É uma parte da organização do conhecimento na qual estão inseridos todos os processos de simbolização notacionais e conceituais do conhecimento humano dentro do campo específico de cada área. Classificação, indexação e todos os aspectos linguísticos e computacionais relacionados com a tradução simbólica do conhecimento estão incluídas na representação do conhecimento.

A exposição inteligente inclui adquirir, armazenar e inferir o conhecimento, para que o mesmo se acumule, é importante que ele seja expresso, as representações do conhecimento são a melhor forma de implementar soluções para os problemas mais complexos que possam existir, pois requer excesso de conhecimento e

mecanismos específicos para manipulá-los (RUSSELL; NORVIG, 2004).

A representação do conhecimento pode ser definida como um conjunto de convenções sintáticas e semânticas que permitem que algo seja descrito. Consiste em usar uma linguagem, frase ou número específico que corresponde a uma descrição ou estado do mundo (JOHN, 2000).

Após o fim da Segunda Guerra Mundial, à medida que o conhecimento científico e tecnológico aumentava, as dificuldades de tecnologia, armazenamento e recuperação da informação tornaram-se ainda maiores. Enfatizando a necessidade de mudar o foco no campo da expressão e do conhecimento e, assim, a criação de uma linguagem documental e outras ferramentas de representação (RAMALHO, 2009). Para o autor, as ontologias nos últimos anos têm se apresentado como uma nova categoria de ferramentas de representação, levando a inúmeros estudos que declaram suas vantagens sobre os modelos existentes.

Segundo Ramalho (2006, p. 97), a ontologia é conceituada como

Um artefato tecnológico que descreve um modelo conceitual de um determinado domínio em uma linguagem lógica e formal, a partir da descrição dos aspectos semânticos de conteúdos informacionais, possibilitando a realização de inferências automáticas por programas computacionais.

Desse modo, essa definição sugere que quando uma ontologia é usada para representar o conhecimento. Ela não tem "propriedades transcendentais" e pode ser usada para descrever formalmente o conteúdo da informação usando uma linguagem que pode ser processada computacionalmente, realizando inferência automática em um domínio pré-definido e limitado.

3 Os aspectos linguísticos da terminologia: um diálogo com a lexicologia e a lexicografia

A linguagem técnico-científica exige uma concepção diferenciada, por isso sua aplicabilidade requer estratégias específicas, as quais não devem e nem podem ser confundidas com o estrato social da linguagem natural, ou mais específico, com o estrato social de uma língua. Os aspectos que envolvem o trabalho linguístico técnico-científico são apurados e visam à representação de realidades sob a análise especializada, portanto a linguagem restrita a esse tipo de situação corrobora a dinamização do conhecimento específico.

Ao contrário do que se possa imaginar, a linguagem especializada não é um recurso gerador de exclusão social, pois sim, é esse tipo de linguagem que atua como orientação para que a população de modo geral comece a diferenciar o conhecimento à base do senso comum do

conhecimento especializado, ou seja, à proporção que pessoas passem a articular esse tipo de linguagem é um índice de que o nível de conhecimento social está se elevando, assim reconhece-se que o nível de estrato linguístico evidencia o nível cultural de uma sociedade.

Evidentemente, a cultura de uma nação não pode nem deve ser mensurada apenas pelo seu conhecimento técnico-científico, todavia essa análise contribuiu para auxiliar mecanismos sociais, culturais, históricos, políticos, científicos e pedagógicos no combate à *fake news* e à pós-verdade.

Ademais, a linguagem especializada atende às necessidades do profissional especializado, seja um cientista, um pesquisador, seja o profissional de atuação na área, seja o estudante nos níveis de graduação e pós-graduação. Através de um estrato linguístico formal técnico-científico, as pessoas envolvidas com um determinado domínio do conhecimento possuem maior autonomia e autocontrole acerca da realidade que estudam e investigam, pois, esse nível de linguagem busca especificar, em um cenário único, aquilo que a cognição humana é capaz de compreender sobre os objetos, fenômenos ou eventos.

De fato, a linguagem especializada aplicada às ciências decorrentes do núcleo da lógica e da natureza recebe maior especificidade e sua compreensão direciona a uma exatidão consoante aquilo que a ciência visou a estabelecer como verdade. Por outro lado, a maior dificuldade acerca da linguagem especializada recai sobre as ciências pertencentes ao universo das ciências humanas, visto que, em muitas circunstâncias, essa linguagem especializada acaba sendo confundida com o estrato social da linguagem natural, devido à aproximação entre a ciência e aos eventos sociais, culturais, políticos e históricos.

Por esses fatores, a linguagem especializada é denominada de terminologia, pois deve ser compreendida como sendo um nível técnico-científico frente ao estrato linguístico da linguagem natural, ou seja, a linguagem natural, em sua fundamentação, apresenta diversos níveis de estratos linguísticos e seu estudo é de interesse da lexicologia.

Frente a isso, o objetivo desta seção é analisar os fundamentos da lexicologia e da lexicografia referentes aos estratos linguísticos da linguagem natural e, assim compreender como se efetiva seus níveis, com o fito de refletir sobre a acepção acerca da terminologia e como esse fenômeno linguístico se desdobra em fatores filosóficos e de especialidade, tendo em vista que, embora a terminologia seja um nível do estrato linguístico, a sua atuação não se limita a esse aspecto.

Assim sendo, entende-se por lexicologia o estudo científico do léxico, assim, sob a perspectiva estruturalista, é conduzida pela comutação dos eixos paradigmático e sintagmático, ou melhor, “[...] a

palavra está integrada num sistema de relações e sua única realidade significante provém das delimitações que lhe impõe a existência desse sistema” (DUBOIS *et al.*, 1998, p. 373). Além disso, na visão moderna, “[...] o estudo lexicológico tomará por critério o rendimento funcional: a palavra será considerada uma unidade de significação realizada por fonemas [...]” e será identificada “[...] em função de suas possibilidades de comutação numa frase para formar novas frases” (DUBOIS *et al.*, 1998, p. 374).

Nessa linha, a lexicologia reconhece que a palavra apresenta vários níveis, por isso analisa a necessidade de “[...] distinguir unidades de significação superiores à palavra” (DUBOIS *et al.*, 1998, p. 374). Com isso, a lexicologia moderna se afasta da concepção saussuriana de linguagem e língua, visto que entende que a estaticidade da *langue* atrapalha o desenvolvimento da área, já que a evolução das palavras é dinâmica e também reflete o discurso, por isso, por um lado, sua atuação junto à lexicografia visa à unicidade para efetuar a correlação entre campo lexical e campo conceitual; por outro, aproxima-se dos fundamentos da semântica e visa a compreender a aplicabilidade da experiência de mundo.

Com isso, há a ampliação de materiais que exploram o formato onomasiológico, ou seja, não se limitando ao modelo estrutural da semasiologia. Outrossim, o estudo da lexicologia deve se aproximar dos fundamentos semânticos, pragmáticos, semióticos e discursivos, visto que “o estudo da unidade léxica isolada não deve ser privilegiado, na única medida em que a análise de vocabulário não constitui senão uma parte da análise do discurso e não pode ser isolado desta” (DUBOIS *et al.*, 1998, p. 377).

Diante disso, a lexicologia aborda os níveis de estrato linguístico, porém no senso comum esses aspectos não estão bem delineados, ou seja, “palavra, léxico, vocabulário e, em alguns contextos, termo, são utilizados como sinônimos, não havendo preocupação com suas aplicações terminológicas na análise científica” (MARTINES, MOREIRA, ALMEIDA, 2022, p. 48).

Entende-se, desse modo, que palavra é uma espécie de suporte da coisa-dita, ou ainda é pedra bruta a ser talhada conforme a aplicação, estrato ou discurso; o léxico “[...] é a representatividade de uma língua”, é o estrato que fomenta o idioma, ou seja, é o que define “[...] sua forma e sua formatação, assim como os processos morfológicos de revitalização das palavras, ampliando seu significado, definindo novos contextos de aplicação” (MARTINES, MOREIRA, ALMEIDA, 2022, p. 48).

Já o vocabulário se constitui “[...] inerente ao discurso e à sistematização de seu percurso argumentativo”, ou seja, está alinhado a fatores linguísticos responsáveis por tecer “[...] a rede semântica e a progressão textual,

bem como os elementos da concatenação textual” (MARTINES, MOREIRA, ALMEIDA, 2022, p. 48). Por outro lado, o vocabulário é uma aplicação mais individualizada, ou melhor, “o vocabulário marca a representação do indivíduo no texto, pois compreende que sua ação é performática”, tal fato indica que, na visão lexicológica, “o vocabulário pressupõe um léxico, assim como age na forma de palavra, já que evidencia a coisa-dita” (MARTINES, MOREIRA, ALMEIDA, 2022, p. 48).

Por sua vez, o termo é o nível formal do estrato linguístico de uma linguagem natural. Além de formal, apresenta uma funcionalidade específica, a qual está inserida no cenário da especialidade, ou seja, sua aplicação não ocorre em qualquer manifestação textual, mas sim em circunstâncias especiais, ou melhor, decorrente a cenários técnicos ou científicos, assim entende-se que o termo está inserido em um processo linguístico tão específico que, “[...] para que haja precisão na evocação do conceito, dele pressupõe um domínio” (MARTINES, MOREIRA, ALMEIDA, 2022, p. 48).

Frente a isso, o termo, elemento linguístico técnico-científico, objeto da terminologia, pressupõe algumas fundamentações:

a) qualquer campo científico ou de especialidades se utilizará de um conjunto de termos para expressar seu pensamento: essa é sua esfera de mediação da realidade. Sua função é ser específica e se diferenciar da linguagem cotidiana ou das descrições e explicações baseadas no senso comum. Nesse aspecto, a terminologia é recurso para o pesquisador nomear o objeto de estudo e descrever as suas particularidades, seus atributos de modo ontológico;

b) o termo é parte de uma interface de pensamento, enquanto o termo é a forma, o conteúdo é o conceito, o construto, que por sua vez é uma interface da realidade: essa é a sua esfera semiótica, pois há a constituição de uma interface cognitiva, composta por forma, conteúdo e objeto, ou melhor, termo, conceito e realidade. Importante destacar que o termo não é uma forma simples e usual, mas sim um recurso informacional de potencialidades complexas e discursivas;

c) o termo tem como função especial representar os objetos científicos: essa é sua esfera de representação da realidade. É através da terminologia que a compreensão sobre o objeto, fenômeno ou evento é elaborado. Nesse aspecto, o termo trabalha em conjunto com outros termos e, assim, fundamenta uma rede semântica, semiótica e cognitiva a fim de expressar, através da experiência, fundamentos que descrevam e expliquem particularidades da realidade em construtos lógicos pelo viés da generalidade;

d) o termo é o recurso linguístico utilizado por um especialista para elaborar seus relatórios sobre

resultados: essa é sua esfera discursiva. Importante observar que o termo a ser utilizado pelo produtor do texto deve ser pertencente a um repertório técnico-científico já consolidado e de domínio entre os pares de especialistas. Nesse aspecto, o termo está inserido no discurso, sendo assim, passível de significação e ampliação de sentido, por isso também pertence a uma esfera pragmática. Além disso, sob essa configuração, o termo executa a sua função de comunicação;

e) o termo é aplicado como recurso informacional para efetuar a indexação: essa é sua esfera documental, ou melhor, assume a função de linguagem documentária. Nesse cenário específico de organização e representação da informação, o termo é fundamento linguístico que serve de suporte para elaboração de linguagem documentária, como por exemplo um tesouro, cuja função é normalizar o significado aplicado ao registro da informação, por isso seu valor documental;

Frente a esse último item, é válido refletir que a terminologia possui uma função mais abrangente do que o tesouro, ou qualquer outro tipo de linguagem documentária, pois a terminologia é instrumento linguístico presente em qualquer campo do conhecimento, já que é a terminologia que vai viabilizar a representação, mediação, cognição e comunicação do objeto científico. Já o tesouro é ancorado na terminologia e é recurso específico de uma determinada especialidade: a organização e a representação do documento, ou seja, o tesouro é forjado à base da terminologia, mas possui uma função mais restrita.

Acerca das características do termo, este “[...] é a identidade representativa de um ser, de um pensamento, de uma ideia, de uma teoria ou de uma ciência [...]” (MARTINES, MOREIRA, ALMEIDA, 2022, p. 49), assim reconhece-se que a terminologia atua em várias camadas da interação linguística, possuindo características complexas, as quais devem ser aprofundadas a fim de que se possa analisar as dimensões de aplicabilidade e como conduzir o estudo sobre o significado e a significação.

Ademais, o termo exerce a função de “[...] âncora informacional que resgata e revitaliza o conceito, capaz de representar todo o conhecimento ou a área científica através de uma palavra, de um sintagma, ou de uma frase” (MARTINES, MOREIRA, ALMEIDA, 2022, p. 49). Consoante Cabré (2010, p. 398), o termo atua como a “[...] descrição e registro das diferentes terminologias utilizadas nos múltiplos domínios do conhecimento” e seu nível de formalidade permite que se trate da “[...] padronização, até então vista como a meta primordial do registro dos termos e a garantia da univocidade científica”.

Concomitantemente, a terminologia pode ser vinculada às preocupações da lexicografia, já que se enquadra no fomento direcionado ao estudo do léxico aplicados às áreas técnicas e aos campos científicos e, por isso,

desempenha um papel importante na dicionarização dos verbetes científicos.

Entende-se lexicografia como “[...] a técnica de confecção de dicionários e a análise linguística dessa técnica”, com isso há a análise do léxico em geral pertencente a uma língua natural, como também pode ser aplicado a uso especializado de léxicos científicos, no caso, a terminologia, assim estabelece-se que “o dicionário técnico distingue-se dos precedentes enquanto não reflete sobre as palavras do vocabulário geral, mas sobre os termos da ciência ou da técnica considerada” (DUBOIS *et al.*, 1998, p. 368).

É válido destacar que é no âmbito da lexicografia que se manifestam a polissemia, a homonímia e os tipos de ambiguidade, pois é necessário evidenciar que a palavra isolada não manifesta sua aplicabilidade discursiva, ou seja, enquanto na lexicografia a palavra está isolada, no máximo se recupera as possibilidades semânticas e em pouquíssimos aspectos o aparato semiótico, por isso a descrição dicionarizada requer definições de contextos, domínios e aplicações; na lexicologia, a palavra está inserida em um discurso e é este que conduz sua função linguística, semântica, pragmática e semiótica.

Em se tratando da lexicografia, sua aplicação à terminologia fundamenta a elaboração de dicionários especializados, os quais retratam os vocábulos inseridos em um contexto específico, evidentemente no cenário a qual se destina o dicionário. Nessa linha, aponta-se a grande revolução na lexicografia que é a “[...] introdução de técnicas baseadas no uso de corpora, e os dicionários modernos costumam, hoje em dia, basear-se em enormes corpora [...]” (TRASK, 2008, p. 156-157).

Com essa técnica, é possível extrair “[...] palavras, formas, gráficas, significados e comportamentos gramaticais, o que permite aos lexicógrafos apelar diretamente para os fatos de uso observados” (TRASK, 2008, p. 157), desse modo valorizando a construção de materiais e manuais a partir da abordagem onomasiológica.

Assim sendo, reconhece-se que a terminologia, tanto vista pelos aspectos lexicológicos, como aspectos lexicográficos, apresenta as condições de unidade linguística, unidade de comunicação e unidade cognitiva como aponta a teoria comunicativa da terminologia, sistematizada por Maria Teresa Cabré, em 1999.

Não obstante, é possível explorar com mais intensidade a relação entre esses três elementos à base da semiótica, da pragmática, da enunciação e do discurso e aproximá-los epistemologicamente. Nesse aspecto, seria necessário atualizar a concepção linguística aplicada, pois está ancorada nas fundamentações estruturalistas de Ferdinand de Saussure. Por sua vez, a proposta de Rita Temmerman, em 2000, observa a dinâmica semântica, pragmática, semiótica, discursiva, já que compreende

que a “delimitação do conteúdo toma como base o texto no qual o termo está inserido” (BARROS, 2006, p. 23).

Estando esse ponto epistemológico ajustado, a lexicografia e a lexicologia pode contribuir, como já tem contribuído significativamente para o desenvolvimento de dicionários especiais, visto que “com o acelerado desenvolvimento das ciências e técnicas, as terminologias das áreas de especialidade ocupam cada vez mais terreno na comunicação” e essa situação gera impacto na sociedade como um todo, pois a terminologia tem integrado “[...] o universo léxico das línguas e, por conseguinte, reivindicam maior espaço nos dicionários de língua geral” (BARROS, 2006, p. 23).

Já na documentação e organização do conhecimento, a terminologia está alinhada à lexicografia na elaboração de linguagem documentária, em especial ao tesauro; por outro lado, a terminologia se ajusta a lexicologia no tratamento documental em busca da extração de descritores manifestados no conteúdo do documento classificado, assim a lexicografia sustenta a concepção que será comunicada pela lexicologia e, assim, facilitar o acesso à informação.

Por fim, com a crescente busca por informações em sistemas informacionais e até mesmo realizado pelos meios digitais e pela rede mundial de informação e comunicação – internet, torna-se importante estudos alinhados aos aspectos linguísticos a fim de aproximar o uso de terminologia ao público em geral, visto que é uma ferramenta essencialmente direcionada a especialistas, a cientistas e a estudantes, todavia a familiaridade com esse estrato linguístico formal técnico-científico significa a ampliação do conhecimento especializado de uma sociedade.

4 A terminologia e linguagem documentária: uma construção da representação

A terminologia é um estrato linguístico técnico-científico e por isso é objeto de estudo da linguística. Visto que seu desempenho está atrelado à produção discursiva, comunicacional e significativa, a terminologia possui todos os fundamentos de uma linguagem natural, ou seja, a terminologia também possui o atributo de representar e mediar a realidade. No caso, em contextos mais específicos, já que é um processo inserido na cientificidade, ou seja, esse propósito de representação não é apenas ordenar os pensamentos, mas sim estabelecer critérios de padronizar o conhecimento sobre o mundo.

Nessa linha, é importante evidenciar que ao relacionar a terminologia ao universo da documentação, seus atributos como estrato linguístico são repassados à linguagem documentária aplicada. Com isso, essa linguagem documentária elaborada à base de discurso e estruturada por um repertório lexicográfico possui a

capacidade de condensar o conhecimento em palavras-chave na confecção de objetos informacionais e também no processo de indexação, sendo assim é possível representar o conhecimento, como também representar a informação, já que a terminologia aplicada como estrato e registro possui a condição discursiva de evidenciar um conhecimento já existente (BENVENISTE, 2014).

Ademais, é válido ressaltar que a terminologia possui um aparato epistemológico e funcional mais abrangente do que se tem apresentado nas discussões no cenário da Ciência da Informação, pois a terminologia está alinhada ao campo da Organização do Conhecimento, como também é utilizada na ciência da Documentação e na Biblioteconomia, já que utilizam linguagem documentária para efetivar o serviço de organização e representação do conhecimento e da informação.

Nesse propósito, inicialmente os campos mencionados se aproximaram das considerações de Wüster, na linha positivista do Círculo de Praga, e atualmente seguem as recomendações da teoria comunicativa da terminologia, estabelecida pelas pesquisas de Cabré (2003). Se por um lado a teoria geral da terminologia restringiu o uso da terminologia como um recurso formal, ou seja, uma fundamentação lógica, inerente à própria língua, sem a preocupação dos usos sociais, culturais, políticos e históricos; a teoria comunicativa da terminologia ampliou essa discussão explorando os fatores abandonados por Wüster, além de evidenciar o processo filosófico e especializado acordando com as áreas científicas a que se destina o tratamento documental.

No entanto, Cabré (1999, 2003) limita a função da linguística, indicando que a parte linguística da terminologia se resume ao aspecto formal, já que a análise de conteúdo ficaria, portanto, sob a égide da filosofia. Por seu turno, a linguística tem apresentado significativas contribuições para o avanço da terminologia (BARROS, 2006; ALMEIDA, 2009), explorando fatores conceituais, tratando do conteúdo, evidenciando os aspectos discursivos, semânticos, semióticos, pragmáticos e cognitivos. Contudo, o universo da Ciência da Informação, ao seguir os estudos de Cabré (1999, 2003), pouco tem acompanhado a evolução dos estudos sobre a terminologia realizados pela linguística.

Outro ponto no cenário da Ciência da Informação que se apresenta de modo problemático e carece de correções para que os estudos sobre esses objetos possam ser ampliados é que muitos estudos e pesquisas tratam terminologia, tesouro e vocabulário controlado como sinônimos.

Não devem ser vistos como sinônimos, mas sim como um proporcionando a elaboração do outro, já que cada um possui um aparato de funções diferentes e restringindo sua utilização como sinônimo compromete o acervo terminológico da ciência, visto que é através

deste que a elaboração conceitual e taxonômica de uma área deve atuar.

Frente a isso, entende-se que a terminologia, entre as suas funções como estrato linguístico técnico-científico, atua como insumo para a elaboração de linguagem documentária, por sua vez o tesouro ao ser aplicado em contextos rígidos de padronização da linguagem de representação assume o papel de vocabulário controlado.

Isso leva à seguinte reflexão, o tesouro é resultante da terminologia, portanto também é ancorado na lexicologia e na lexicografia, sendo assim, como produto da lexicologia o tesouro pode servir de referência ao uso científico da Organização do Conhecimento, Ciência da Informação, Documentação e Biblioteconomia, como também ao uso profissional da área; já como produto da lexicografia, o tesouro é recurso da prática profissional tanto para consulta, como para elaboração de objetos informacionais e para indexação, assim, nesse formato, a depender das exigências do sistema informacional, ou melhor, a depender da política de indexação desse sistema informacional, o tesouro atua como vocabulário controlado.

O problema de tratar essas situações como sinônimos é restringir os recursos informacionais aplicados no processo de representação da informação como algo estático e imutável, sendo que eles estão em constante evolução conforme ocorre a evolução nos ambientes técnicos, culturais e científicos. Frente a isso, é preciso compreender que o vocabulário controlado é um estágio extremamente específico, enquanto a terminologia está em um estágio um pouco mais textual, porém um registro técnico-científico, o qual diferentemente do vocabulário controlado pode estar presente em alguns discursos a depender do gênero textual e da intencionalidade da produção do documento.

Considerações Finais

Diante do exposto, este artigo buscou discutir os recursos da terminologia, a partir da Organização do Conhecimento (enquanto campo e processo), SOCs e da Representação do Conhecimento. Sendo assim, demonstrou a relação desses elementos com a linguagem, já que a organização e representação do conhecimento e da informação ocorre à base de uma linguagem documentária.

Ademais, esta pesquisa evidenciou que o estudo da terminologia está alinhado a estudos linguísticos e que a terminologia é um estrato técnico-científico e exerce diversas funções no cenário científico-cultural.

No cenário da Ciência da Informação, a terminologia atua como insumo para a elaboração de linguagem documentária, por isso não deve ser tratada como

sinônimo de tesouro ou vocabulário controlado, pois a terminologia não se restringe ao processo de representação da informação.

Além disso, a terminologia como estrato linguístico também está presente nas camadas textuais do documento, o que leva a entender que a terminologia também está presente no processo discursivo, portanto sendo elemento linguístico inserido no processo de significação.

Referências

- ALMEIDA, G. M. de B. A Teoria Comunicativa da Terminologia e a sua prática. ALFA: **Revista de Linguística**, São Paulo, v. 50, n. 2, 2009.
- AMORIM, I. S.; BRÄSCHER, M. O fora da organização do conhecimento: considerações para a arquivologia, biblioteconomia e museologia. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 22, n. 2, p. 359-376, 2017.
- BARITÉ, M. COLOMBO, S.; DUARTE BLANCO, A.; SIMÓN, L.; CABRERA CASTROMÁN, G.; ODELLA, M. L.; VERGARA, M. **Diccionario de organización del conocimiento**: clasificación, indización, terminología. 6. ed. aum. Montevideu: CSIC, 2015. Disponível em: <https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/9028>. Acesso 23 maio 2023.
- BARROS, L. A. Aspectos epistemológicos e perspectivas científicas da terminologia. **Cienc. Cult.**, São Paulo, v. 58, n. 2, p. 22-26, Jun. 2006.
- BENVENISTE, E. **Últimas aulas no Collège de France** (1968 e 1969). São Paulo: Editora Unesp, 2014.
- CABRÉ, M. T. **La terminología**: representación y comunicación. Barcelona, IULA, 1999.
- CABRÉ, M. T. Theories of terminology: their description, prescription and explanation. *Terminology*, v.9, n.2, p.163-200, 2003.
- CABRÉ, M. T. Norma y normas en terminología: concepto, tipología y justificación. In: ISQUERDO, A. N.; FINATTO, M. J. B. (org.). **As ciências do léxico**: lexicologia, lexicografia, terminologia. v. 4. Porto Alegre: Editora da UFRGS, 2010.
- CAIXETA, M.; SOUZA, R. R. Representação do conhecimento: história, sentimento e percepção. **Informação & Informação**, Londrina, v. 13, n. 2, p. 34-55, jul./dez. 2008. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/1815/0>. Acesso em: 27 mar. 2023.
- CARLAN, E.; MEDEIROS, M. B. B. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, v. 4, n. 2, p. 53-73, 2011.
- CUNHA, M. B da; CAVALCANTI, C. R de O. **Dicionário de biblioteconomia e arquivologia**. Briquet de Lemos Livros, 2008.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, Frankfurt, v. 33, n. 1, p. 11-19, 2006.
- DEMO, P. **Metodologia do conhecimento científico**. São Paulo: Atlas, 2000.
- DUBOIS, J. *et al.* **Dicionário de linguística**. São Paulo: Cultrix, 1998.
- ESTEBAN NAVARRO, M. Á. La representación y la organización del conocimiento en los archivos: los lenguajes documentales ante los procesos de clasificación, ordenación y descripción. In: Organización del conocimiento en sistemas de información y documentación, Madrid, 1993. **Actas [...]**, Madrid, 1993.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4ª ed. São Paulo: Atlas; 2002.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 5. ed. São Paulo: Atlas, 1999.
- GUIMARÃES, J. A. C.; MATOS, D. F de O; SANTOS, A. Y dos; SALES, R. dos. A dimensão conceitual da organização do conhecimento no universo científico da ISKO: uma análise de domínio a partir dos congressos de ISKO-Brasil, ISKO-Espanha, ISKO-América do Norte e ISKO-França. **Scire: representación y organización del conocimiento**, v.21, n.2. p. 13-26, Jul.Dec. 2015.
- HJORLAND, B. Fundamentals of knowledge organization. **Knowledge Organization**, v. 30, n. 2, p. 87-111, 2003.
- HJORLAND, B. Concepts, paradigm and knowledge organization. In: GNOLI, C.; MAZZOCCHI, F. (Ed.). **Paradigms and conceptual systems in knowledge organization**: proceedings of the eleventh International ISKO Conference. Rome: Ergon Verlag, 2010. p. 38-42.
- KANT, I. **Crítica da razão pura**. São Paulo: Abril Cultural, 1980.
- KAUARK, F. da S; MANHÃES, F. C; MEDEIROS, C. H. **Metodologia da pesquisa: um guia prático**. 1. Ed. Itabuna: Via Litterarum, 2010.
- LARA, M. L. G. de. **Representação e linguagens documentárias: bases teórico metodológicas**. São Paulo, 1999. Tese (Doutorado em Comunicação) --Universidade de São Paulo, São Paulo, 1999.
- MARTINES, A. R.; MOREIRA, W.; ALMEIDA, C. C. Do signo ao tesouro: contribuições de três correntes da linguagem. **Ciência da Informação**, [S. L.], v. 51, n. 1, 2022. DOI: 10.18225/ci.inf.v51i1.5543. Disponível em: <https://revista.ibict.br/ciinf/article/view/5543>. Acesso em: 03 mar. 2023.
- MAZZOCCHI, F. Knowledge organization system (KOS): an introductory critical account. **Knowledge Organization**, v. 45, n. 1, p. 54-78, 2018.
- MELO, M. A. F; BRÄSCHER, M. Termo, conceito e relações conceituais: um estudo das propostas de Dahlberg e Hjørland. **Ciência da Informação**, v. 43, n. 1, 2014.
- MINAYO, M. C. de S. (Org.). **Pesquisa social: teoria, método e criatividade**. 17ª Ed. Petrópolis, RJ: Vozes, 1994. 80 p.

MORAES, A. F. de.; ARCELLO, E. N. O conhecimento e sua representação. **Informação & Sociedade: estudos**, João Pessoa, v. 10, n. 2, 2000.

MOREIRA, W. **Sistemas de Organização do Conhecimento: aspectos teóricos, conceituais e metodológicos**. 164 fls. Tese (Livre-Docência em Sistemas de Organização do Conhecimento)- Programa de Pós-Graduação em Ciência da Informação, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Marília, São Paulo, 2018.

OHLY, H. P. Knowledge organization and ISKO: state, demands, ideals. **Scire: representación y organización del conocimiento**, p. 13-19, 2015.

RAMALHO, R. A. S. Representação do conhecimento e ontologias: reflexões interdisciplinares. *In*: ENANCIB - ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 10., 2009, João Pessoa. **Anais** [...] João Pessoa, Universidade Federal Da Paraíba, 2009. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/171942>. Acesso em: 29 mar. 2023.

RAMALHO, R. A. S. **Web Semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação**. 2006. 120 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006.

RUSSELL, S; NORVIG, P. **Inteligência Artificial**. Editora Campus, tradução da segunda edição, 2004.

SALES, R de; MURGUIA, E. I. Instaurações discursivas da organização do conhecimento: HE Bliss e a International Society for Knowledge Organization (ISKO). **Ciência da informação**, v. 44, n. 3, 2015.

SAN SEGUNDO, R. **Visibilidad científica y diálogos internacionales en Organización del conocimiento**. *In*: Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade. [recurso eletrônico] José Augusto Chaves Guimarães, Vera Dodebei (Org.). – Marília: ISKO-Brasil: FUNDEPE, 2012.

SANTOS, C. da S. A. **Terminologia e Ontologias: metodologias para representação do conhecimento**. 2010. 245 f. Dissertação (Doutor em Linguística) – Departamento de Línguas e Culturas da Universidade de Aveiro, 2010.

SILVA, E. L. da.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. Florianópolis: UFSC/PPGEP/LED, 2000, 118 p.

SOWA, J. F. **Knowledge Representation: logical, philosophical, and computational foundations**. Pacific Grove, CA: Brooks Cole Publishing Co., 2000.

TEIXEIRA, E. **As três metodologias: acadêmica, da ciência e da pesquisa**. Petrópolis, RJ: Vozes, 2005.

TEMMERMAN, R. **Towards new ways of terminology description**. The sociocognitive approach. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2000.

TRASK, R. L. **Dicionário de linguagem e linguística**. Trad. Rodolfo Ilari. 2ª ed. São Paulo: Contexto, 2008.

LEI ZENG, M; MAI CHAN, L. Trends and issues in establishing interoperability among knowledge organization systems. **Journal of the American Society for information science and technology**, v. 55, n. 5, p. 377-395, 2004.

Mediação cultural da informação e organização do conhecimento: uma abordagem crítica de enfrentamento ao epistemicídio

Cultural information mediation and knowledge organization: a critical approach to face epistemicide

Carlos Robson Souza da Silva (1), Luciane de Fátima Beckman Cavalcante (2)

(1) Programa de Pós-graduação em Ciência da Informação, Universidade Estadual de Londrina, Rodovia Celso Garcia Cid (PR 445), KM 380, Londrina, Paraná, CEP 86057-970; crobsonss@gmail.com. (2) Programa de Pós-Graduação em Ciência da Informação, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rua Lauro Muller, 5 – 4º andar, Botafogo (RJ), CEP 22290-160; luciane.cavalcante@facc.ufrj.br

Resumo

Trata do estudo do desenvolvimento teórico-prático da Organização do Conhecimento, buscando entender como se dá a sua relação com a manutenção de estruturas hegemônicas de poder e da proposição de uma abordagem crítica para a Organização do Conhecimento a partir da Mediação Cultural da Informação. Desenvolve-se a partir de uma pesquisa exploratória, de abordagem qualitativa, que utiliza como instrumento de coleta de dados a pesquisa bibliográfica e como análise de dados a revisão de literatura. Realiza-se por meio de pesquisa bibliográfica na Base de Dados em Ciência da Informação (BRAPCI) e da seleção de forma intencional de materiais resultantes dessa pesquisa para compor a argumentação do trabalho, julgada de acordo com os objetivos da presente pesquisa. Conclui que a proposta teórica da Mediação Cultural da Informação contribui para a defesa de uma abordagem crítica da Organização do Conhecimento, uma vez que colabora para a superação de estruturas hegemônicas de poder e propõe maneiras plurais de enfrentamento ao epistemicídio.

Palavras-chave: Organização do Conhecimento. Mediação Cultural da Informação. Epistemicídio.

Abstract

It deals with the study of the theoretical-practical development of the Organization of Knowledge, seeking to understand how its relationship with the maintenance of hegemonic structures of power occurs and the proposition of a critical approach to the Organization of Knowledge from the Cultural Information Mediation. It develops from an exploratory research, with a qualitative approach, which uses bibliographical research as data collection instrument and literature review as data analysis. It is carried out through bibliographic research in the Base de Dados em Ciência da Informação (BRAPCI) and the intentional selection of materials resulting from this research to compose the work's argumentation, judged according to the objectives of the present research. It concludes that the theoretical proposal of the Cultural Information Mediation contributes to the defense of a critical approach to the Organization of Knowledge, since it collaborates to overcome hegemonic structures of power and proposes plural ways of facing epistemicide.

Keywords: Knowledge Organization. Cultural Information Mediation. Epistemicide.

Resumen

Se trata del estudio del desarrollo teórico-práctico de la Organización del Conocimiento, buscando comprender cómo se da su relación con el mantenimiento de las estructuras hegemónicas de poder y la proposición de un abordaje crítico de la Organización del Conocimiento desde la Mediación Cultural de la Información. Se desarrolla a partir de una investigación exploratoria, con enfoque cualitativo, que utiliza la investigación bibliográfica como instrumento de recolección de datos y la revisión bibliográfica como análisis de datos. Se lleva a cabo a través de la investigación bibliográfica en el Base de Datos em Ciência da Informação (BRAPCI) y la selección intencional de materiales resultantes de esta investigación para componer la argumentación del trabajo, juzgado de acuerdo con los objetivos de la presente investigación. Concluye que la propuesta teórica de la Mediación Cultural de la Información contribuye a la defensa de un abordaje crítico de la Organización del Conocimiento, ya que colabora a superar las estructuras de poder hegemónicas y propone modos plurales de enfrentar el epistemicidio.

Palabras clave: Organización del Conocimiento. Mediación Cultural de la Información. Epistemicidio.

1 Introdução

O ato de organizar está associado à humanidade desde o início de sua trajetória no mundo, enquanto espécie de seres inteligentes e capazes de moldarem a realidade

para atender às suas necessidades. A organização reflete então a busca dos seres humanos por classificar a vida, mas também por encontrar uma forma de distinguir os modos de viver e de conceber o mundo.

Quando associada ao conhecimento registrado (aquele inscrito fora da memória humana), pode-se traçar a trajetória da organização desde as pinturas rupestres. Entretanto é importante ressaltar que é com a criação de bibliotecas, arquivos e museus na Antiguidade, pelo menos a que irradia o que hoje considera-se como Ocidente, que a preocupação com a organização do conhecimento se torna um saber sobre o saber e um instrumento político que passa a determinar os modos de pensar e criar, quem detém e de quem será destituído o conhecimento.

Nhacuongue e Dutra (2016, p. 138) apontam que o uso político do conhecimento está presente desde os primórdios da história da sua organização. Podem ser tomados como exemplo, o Império Sumeriano, quando Assurbanipal abastece sua biblioteca com “[...] livros apreendidos dos adversários derrotados”, o Egito, quando Ptolomeu I confisca “[...] todos os livros a bordo dos navios que entravam no Porto de Alexandria [...]” e a China, onde Shi Huangdi ordena a destruição de todos os livros até então guardados “[...] só para construir uma biblioteca que refletia os seus interesses”.

O que se percebe nas histórias destas primeiras bibliotecas e também na história de unidades de informação contemporâneas é que sua organização não é neutra, mas obedece e muitas vezes se alinha aos interesses políticos e econômicos das classes dominantes, ou seja, das classes que possuem poder sobre as dinâmicas do Estado e do mercado.

Além disso, é possível identificar um duplo movimento, que, por um lado preserva conhecimentos, classificando-os como bons e necessários para o progresso da sociedade, e por outro lado, oculta ou mesmo destrói conhecimentos outros, considerados sem valor e inúteis, mas que pertencem a povos, nações, grupos e comunidades sujeitadas ao jugo dos autoconsiderados vencedores.

Esse movimento pode ser identificado como epistemicídio. Moura (2018, p. 9) afirma que o epistemicídio funciona como “[...] um aparato social dinâmico que coloca em ação processos e instrumentos dedicados à destituição ‘da racionalidade, da cultura e da civilização do Outro’”. Na mordenidade, esse processo de assassinato sistemático de outras culturas e formas de pensar se revelam associados ao patriarcado, à cisheteronormatividade, ao racismo, ao colonialismo e, principalmente, aos interesses do capitalismo. Para manter a hegemonia do pensamento dominante (que se materializa através da racionalidade cientificista de base cultural europeia) é necessário mostrar que outras formas de pensar (como os pontos de vista de mulheres, pessoas negras, indígenas, pessoas com deficiência, e da própria classe trabalhadora) sejam tidos como sinônimo de morte.

Tal realidade não se restringe apenas à guarda e à preservação de documentos e de informação, mas inclui também a sua organização dentro das unidades de informação (bibliotecas, arquivos, museus, centros de documentação ou sistemas de informação). Aqui a organização do conhecimento deixa de ser prática do senso comum e se identifica como um campo científico voltado a “[...] sistematização e organização do conhecimento com a finalidade de recuperação da informação, considerada também como um campo que exerce um papel de comunicação, mediação e transferência de conhecimento [...]” (NASCIMENTO; PINHO, 2019, p. 107).

Através dos estudos desenvolvidos neste campo e a ele atrelados, pode-se discutir o que é e como o conhecimento é criado, assim como pode-se definir instrumentos, métodos e técnicas que colaborem na sua preservação e recuperação. Tais instrumentos se materializam através de Sistemas de Organização do Conhecimento (SOCs), como “[...] os esquemas de classificação, as listas de cabeçalhos de assunto, os tesouros, as taxonomias, os *gazeteer*, as ontologias e as redes semânticas” (COSTA; MIRANDA, 2019, p. 157).

Suas primeiras manifestações não surgem com o interesse do acesso ao público em geral aos documentos inseridos no contexto de uma unidade de informação específica, mas com o desenvolvimento do chamado paradigma pós-custodial passaram a ter como objetivo “[...] garantir que o item de informação fosse consultado, chegasse ao seu usuário [...]” (BRENTAN JUNIOR; MARTINS; SANTOS NETO, 2018, p. 9).

Neste sentido, os profissionais que produzem tais instrumentos e a própria Organização do Conhecimento possuem um papel mediacional. Mediacional no sentido de que articula o encontro e a apropriação dos sujeitos aos conhecimentos organizados ou em processo de organização. Entretanto mesmo assumindo um papel de mediação em um contexto pós-custodial, não significa, porém, que tal processo tenha apenas propostas benéficas para a sociedade. A organização do conhecimento já é herdeira e reflete o pensamento que determina a guarda e a preservação de documentos imersas em epistemicídio. É necessário superar tal realidade e propor enxergar a Organização do Conhecimento e própria produção de SOCs “[...] em uma perspectiva plural e inclusiva” (MOURA, 2018, p. 9).

Uma dessas possibilidades é atrelar-se ao que Bezerra e Cavalcante (2020) chamam de Mediação Cultural da Informação. Nessa proposta, a pluralidade de saberes e modos de ser e viver no mundo não são descartados, mas têm “[...] o condão de operar uma espécie de ‘reencantamento do mundo’, permitindo que a ciência se beneficie dessa pluralidade de saberes sem precisar abrir mão dos métodos e rigores científicos que lhe são caros” (BEZERRA; CAVALCANTE, 2020, p. 4).

Ou seja, na Mediação Cultural da Informação as propostas dominantes e os conhecimentos “outros” são postos lado a lado, evidenciando suas contradições e as provocando ao embate. Para a Organização do Conhecimento isso pode significar que, se assumir esta perspectiva, deve percorrer um caminho que se inicia no questionamento de quais informações e conhecimentos trata quando afirma que tem como objetivo permitir o acesso de qualidade à informação e se, na verdade, a sua intervenção na realidade colabora para apagamentos culturais e a perpetuação de epistemicídios.

Diante deste contexto, o presente artigo questiona: *de que maneiras a Mediação Cultural da Informação pode contribuir para uma abordagem crítica da Organização do Conhecimento que esteja alinhada a um processo de enfrentamento ao epistemicídio?* Definindo como objetivos específicos: identificar no desenvolvimento teórico da Organização do Conhecimento como se dá a sua relação com a manutenção de estruturas hegemônicas de poder; propor uma abordagem crítica para Organização do Conhecimento tendo como fundamento a proposta teórica da Mediação Cultural da Informação.

2 Procedimentos metodológicos

O presente artigo é resultado de uma pesquisa de natureza teórica realizada através de revisão da literatura de que intersecciona os temas Organização do Conhecimento, Mediação Cultural da Informação e Epistemicídio. O material utilizado como referência foi recuperado na Base de Dados em Ciência da Informação (BRAPCI). A escolha desta base se deu por seu acervo apresentar uma abrangência representativa da pesquisa brasileira sobre Ciência da Informação e Organização do Conhecimento.

Uma pesquisa de natureza teórica é aquela que é voltada “[...] para discussões conceituais e/ou filosóficas ou [...] para discussões de dados da pesquisa básica e aplicada, mas que não descrevam dados obtidos por observação direta ou indireta do comportamento [...]” (STRAPASSON; MAGALHÃES; CUSTÓDIO, 2013, p. 119). Tal afirmação corrobora com os objetivos da presente pesquisa, uma vez que estão relacionados à proposição da inter-relação entre Mediação Cultural da Informação e Organização do Conhecimento contra o Epistemicídio.

Outro ponto a se destacar é que, tendo como hipótese que a inter-relação acima citada ainda não possui grande capilaridade no campo da Ciência da Informação, a presente pesquisa teórica se deu através de um interesse exploratório, mais do que descritivo ou explicativo. Classificando-a como exploratória,

destacamos que “[...] o conhecimento acumulado sobre um problema ou fenômeno particular é limitado [...]”, ou seja, “[...] quando a literatura sobre o tema ainda acumula um conjunto limitado de resultados que permita descrever fatos e relações entre variáveis implícitas em situações” (GOMES; GOMES, 2020, p. 16). Aqui pode-se considerar que as variáveis são o Epistemicídio e a Mediação Cultural da Informação e a situação é a Organização do Conhecimento.

Nesse sentido, é que para a pesquisa teórica do tipo exploratória realizada decidiu-se adotar como instrumento de coleta de dados a pesquisa bibliográfica realizada na já mencionada BRAPCI. Dessa forma, é possível também considerar o repositório como uma fonte de informação confiável e atualizada também sobre Organização do Conhecimento no Brasil.

A pesquisa aqui baseou-se no resultado quatro buscas. A quatro buscas tinham como filtros: obras publicadas entre 2018-2022, em cujos títulos estivessem expressos de maneira clara e direta os termos aplicados. Cada busca disse respeito a um termo: “Organização do Conhecimento”, que recuperou 110 materiais; “Epistemicídio”, que recuperou 3 materiais; “Mediação Cultural”, que recuperou 35 materiais; e “Mediação da Informação”, que recuperou 98 materiais.

Após o levantamento bruto, decidiu-se realizar mais uma filtragem. Por se entender esta pesquisa como qualitativa, ou seja, uma pesquisa que se empenha mais na busca de conceitos e significados na realidade do que de dados empíricos e que dá maior liberdade à pessoa que pesquisa na seleção de temas de interesse, essa segunda filtragem se deu por meio de um processo de amostragem intencional, de modo eu foi possível selecionar unidades de estudo específicas, dispondo “[...] daquelas que geram os dados mais relevantes e fartos, considerando o [...] tema de estudo” (YIN, 2016, p. 79).

Dessa forma, realizou-se uma leitura dinâmica dos títulos, dos resumos, das palavras-chave, dos títulos das seções e subseções e da introdução de cada artigo visando delimitar ainda mais os resultados para que condissessem com os interesses argumentativos aqui propostos. Nesse sentido, o resultado contou com 18 artigos para Organização do Conhecimento ficou, para Epistemicídio mantiveram-se 2, para Mediação da Informação e Mediação Cultural apenas 3. Apesar dos números reduzidos, ressalta-se que tais resultados condizem com a pesquisa, uma vez que os temas abordados já estavam presentes nas obras sobre Organização do Conhecimento. Além disso, considerou-se a possibilidade de agregar à argumentação do presente artigo referências outras para complementar principalmente no que diz respeito à introdução e à metodologia.

3 A dimensão sociocultural da Organização do Conhecimento

A Organização do Conhecimento em sua confluência com o desenvolvimento da Ciência da Informação desempenha um papel importante na criação de estratégias e instrumentos que potencializam a armazenagem, a recuperação e o uso da informação, se ocupando

[...] em identificar que termos são representativos para denominar as entidades em determinado domínio, organizá-los sistematicamente como taxonomias de conceitos mais genéricos, mais específicos, além de outros tipos de relacionamentos, com vistas a padronizar um vocabulário através do qual documentos (suas representações) neste domínio pudessem ser recuperados (MARCONDES, 2021, p. 268).

Dentre os principais produtos gerados pela Organização do Conhecimento (OC) estão os chamados Sistemas de Organização do Conhecimento ou, como são amplamente conhecidos na literatura, os SOCs. Tais sistemas têm a função de gerir, normalizar, padronizar e controlar os termos que definem os conceitos atrelados a um ou mais domínios científicos dentro de uma estrutura de relações e hierarquias, “[...] trabalhando como uma interface entre os conteúdos dos documentos e usuários, [permitindo] diminuir as falhas comunicacionais entre pesquisadores, bancos de dados, documentos e usuários” (SONO; FRANCELIN, 2022, p. 5).

Segundo Nascimento e Pinho (2019, p. 108), os principais tipos de SOCs encontrados na literatura e na prática social da Organização do Conhecimento podem ser distinguidos em: “[...] taxonomias, tesouros e ontologias, sistema de classificação, mapa conceitual, *topic maps* e folksonomia”.

Cada um desses SOCs é estruturado de acordo com regras próprias, estabelecendo redes de relações e hierarquias que permitam a unidades de informação e a sistemas de recuperação de informação a estabelecer uma lógica que facilita aos usuários encontrarem as informações de que necessitam, evitando erros ou ambiguidades na representação dos documentos. A ideia é que os SOCs identifiquem assuntos, pessoas ou outras entidades que representem os documentos, traduzam-nos em termos e insiram em um sistema de recuperação da informação, funcionando tais termos como mediadores entre o documento e os potenciais usuários.

Historicamente, a construção dos SOCs esteve atrelada a princípios filosóficos específicos (MOURA, 2018). O Sistema de Classificação Decimal de Dewey (CDD) é herdeiro, por exemplo, de uma longa tradição que tem suas origens em Aristóteles, na divisão medieval entre Trivium e Quadrivium e na “[...] estrutura departamental das universidades européias e norte-

americanas do século XXI [...]” (MARCONDES, 2021, p. 267). Ao atrelar-se a filosofia e ao desenvolvimento científico, a CDD poderia tomar para a si uma ideia de universalidade e de neutralidade, assumindo-se como capaz de representar o conhecimento historicamente produzido pela humanidade, sem atender a interesses externos que não condigam com os ditames da filosofia e da ciência.

Essa corrente de pensamento universalista e que se autodenomina neutra percorreu toda a história da Organização do Conhecimento, entretanto vem sendo discutida com o desenvolvimento de uma abordagem sociocultural para a área. Essa abordagem possui entre suas raízes uma necessidade de mudança no foco das práticas de OC, tirando-o do acervo ou da custódia dos documentos e deslocando-o “[...] para perspectivas mais abertas e inclusivas, focadas no acesso pleno ao conhecimento registrado e na promoção das condições para a sua livre circulação e produção” (MOREIRA; SABBAG, 2022, p. 36).

Tal deslocamento põe em xeque a universalidade e a neutralidade dos SOCs, conclamando à necessidade de revisitá-los ou de propor SOCs outros que façam jus às realidades vividas pelos mais diversos sujeitos em produção ou em processo de interação com os conhecimentos produzidos.

A Análise de Domínio, por exemplo, é uma das primeiras correntes da Organização do Conhecimento que assume uma perspectiva social. Através dela, é possível construir uma crítica à naturalidade da divisão social do conhecimento já hegemônica e compreender as condições que as deram origem, “[...] as condições pelas quais o conhecimento se constrói e se socializa [...]” (GUIMARÃES et al, 2019, p. 27).

Partindo-se então do pressuposto que o conhecimento é resultante de uma grande quantidade de variáveis e condições sociais, às vezes limitantes, às vezes progressistas, a perspectiva social da Organização do Conhecimento descarta a suposta universalidade que até então a caracterizava e encaminha a crítica à discussão sobre a neutralidade de seus processos e instrumentos.

Tanto Moura (2018) quanto Guimarães *et al* (2019) concordam que a neutralidade na Organização do Conhecimento é impossível. Ao se proporem universais os processos e instrumentos de Organização do Conhecimento acabam por validar apenas os discursos hegemônicos (principalmente, europeias e norte-americanas) sobre o que é conhecimento e sobre como esse conhecimento se estrutura na história.

A crítica à neutralidade trazida pela perspectiva social da Organização do Conhecimento busca não somente debater seus processos atuais, mas como esses processos se estabelecem dentro de uma estrutura social mais ampla associada à colonização, ao

imperialismo, à dominação, à discriminação e à violência contra grupos não-hegemônicos e aos seus saberes. Essa crítica social pode ser associada a uma perspectiva cultural, que destitui o pensamento hegemônico de seu trono universalista e neutro, e sugere o desenvolvimento de uma dimensão sociocultural para a Organização do Conhecimento, que

[...] consiste em demonstrar a relevância dos aspectos identitários na criação de esquemas e instrumentos de classificação. Esses aspectos perpassam pelo domínio ético, na medida em que sistemas de classificação, os intermediários individuais e coletivos que se apartam da identidade cultural têm o potencial de estimular a marginalização e os preconceitos contra os grupos que se encontram fora da esfera epistemológica hegemônica (MANHIQUE; CASARIN, 2019, p. 3)

Portanto saberes e formas de conhecer não hegemônicas, como aqueles produzidos por trabalhadores e trabalhadoras, pessoas negras, indígenas, mulheres, pessoas LGBTQIA+, pessoas oriundas de países periféricos herdeiros de processos de colonização ou afetados por políticas imperialistas externas devem ser trazidos à tona nas discussões sobre a Organização do Conhecimento, de maneira que ela se torne mais inclusiva e mais condizente com a realidade material concreta das pessoas que à ela recorrem para aceder a informações e a conhecimentos.

4 Epistemicídio: a OC diante do apagamento e da negação de saberes não-hegemônicos

Por estabelecer uma crítica à universalidade e a neutralidade na Organização do Conhecimento e propor uma nova perspectiva que inclua os saberes e as formas de conhecer de grupos não-hegemônicos, a abordagem sociocultural da OC evidencia e põe em discussão a dimensão epistemológica dos processos sociais de dominação e exclusão.

A dimensão epistemológica é destacada por Guimarães (2015 *apud* MOREIRA; SABBAG, 2022, p. 37) como o eixo da Organização do Conhecimento que mais diz respeito à sua construção, definindo seus paradigmas, bases teóricas e bases metodológicas. Isso acontece porque a Epistemologia, campo da filosofia que se consolidou entre os séculos XVII e XVIII, é a disciplina cuja principal preocupação está em teorizar como se dá o processo de conhecimento (GARCEZ; SALES, 2021). Essa preocupação se intensifica quando é compreendida também como a filosofia que tem como objetivo a análise crítica do desenvolvimento das ciências e das formas de relação entre elas (MARTÍNEZ-ÁVILA; MELLO, 2023). Ou seja, ambas Epistemologia e Organização do Conhecimento debruçam-se sobre o conhecimento científico e sobre os processos de significação dentro e através dele.

Todavia, segundo Martínez-Ávila e Mello (2023), essa conexão entre a Epistemologia e a Organização do Conhecimento, mais do que uma neutra teorização sobre o conhecimento e suas inter-relações terminológicas internas e externas, revela, na verdade, relações de poder que ditam o que deve ser ou não considerado conhecimento e, conseqüentemente quem dele poderia se apoderar ou mesmo ter acesso. Em cima dessas relações, existe um véu adornado pelo discurso sobre o acesso à informação, que, na verdade, esconde o fato de que a dimensão epistemológica da Organização do Conhecimento tradicional não busca a mudança social, mas a manutenção dos interesses econômicos e ideológicos das classes dominantes, utilizando como apoio os desenvolvimentos científicos e tecnológicos

Mas como isso acontece efetivamente? Miranda e Costa (2019) ao fazer uma análise da representação bibliográfica da Umbanda (religião brasileira de matriz africana) no catálogo *web* da Fundação Biblioteca Nacional identificou que tal religião é classificada sob a notação da CDD para outras religiões, enquanto o Cristianismo ocupa a maior parte da classe 200 (Religião). Identificou-se também que a notação específica na CDD (299.672) para a Umbanda não a associa às demais religiões brasileiras de matriz africana e que os documentos a ela associadas na Fundação Biblioteca Nacional não são indexadas com o termo “Umbanda” por não existir cabeçalhos de assunto composto por ele. Tal situação ocorre por dois motivos: a adoção de um sistema de classificação que não corresponde à realidade brasileira; e a um processo de organização do conhecimento que se adapta aos interesses e às formas de pensamento das classes dominantes brasileiras, que se identificam com as estruturas patriarcais de poder associadas a distinção racial branca e à religiosidade cristã. Tal situação revela um processo histórico de dominação que não se restringiu apenas a corpos, mas aos saberes e aos modos de pensar, que não se restringe apenas ao genocídio, mas também ao epistemicídio.

De acordo com Moura (2018, p. 129), o epistemicídio lança mão de uma série de “[...] processos e instrumentos dedicados a destituição da ‘racionalidade, da cultura e da civilização do Outro’, promovendo uma continuada indigência cultural e intelectual”. Através dele é possível destronar saberes locais e dominar os sujeitos que deles se utilizam para conhecer o mundo, obrigando-os a rejeitá-los e desprezá-los. Isso não significa porém dar oportunidade de se apropriar dos conhecimentos hegemônicos, mas tutelar para sempre os sujeitos dominados, distribuindo a eles o mínimo necessário para manter de pé a lógica de dominação.

Cria-se, portanto, com o epistemicídio um processo persistente de indigência cultural que se evidencia pelo fato de que “[...] não é possível desqualificar as formas

de conhecimento dos povos dominados sem desqualificá-los também, individual e coletivamente, como sujeitos cognoscentes” (CARNEIRO, 2005, p. 97). Se os sujeitos são desprezados, suas formas de conhecer também o são, em um processo contínuo que “[...] fere de morte a racionalidade do subjugado [...]” (CARNEIRO, 2005, p. 97).

Utilizando como referência a Biblioteconomia e a Ciência da Informação, Silva e Silva (2022, p. 53) apontam que o epistemicídio deixam marcas na história da ciência, que se apresentam

a) no silenciamento forçado sobre a discussão racial e sobre colonialismo na ciência, pesquisa e esferas acadêmicas em BCI; b) apagamento e/ou exclusão de autorias negras, mulheres, indígenas e LGBTQIA+ na construção de acervos e currículos de cursos; c) violência intelectual e epistêmica contra pessoas negras e indígenas, via inferiorização do conhecimento por elas produzido; d) dependência epistêmica da produção intelectual criada pela ciência e grupo racial dominante do norte global para validação de conhecimentos produzidos no sul global; e, por fim, e) a utilização da (pseudo)neutralidade acadêmica e profissional para retirada de responsabilidade social de docentes e profissionais de Biblioteconomia do compromisso éticopolítico para com a justiça social e informacional de povos em vulnerabilidades informacionais, sociais, educacionais, históricas e econômicas (SILVA; SILVA, 2022, p. 53)

Entretanto, destaca-se aqui, que a situação do epistemicídio tão marcadamente presente na sociedade não conflui para uma negação da Organização do Conhecimento e de sua história, atrelando a ela uma negatividade e uma maldade como se ela mesma fosse uma entidade dotada de mentalidade humana. O reconhecimento do epistemicídio na Organização do Conhecimento contribui, pelo contrário, para se dar início a uma postura crítica, que se inicia por compreender a própria Organização do Conhecimento como produto do trabalho humano, que reflete os interesses da humanidade e que media o acesso da humanidade ao conhecimento por ela mesma produzido.

O que se deve buscar é propor uma Organização do Conhecimento que não se alinhe mais à ótica da dominação, mas em um movimento de usurpação do poder colonial e de distribuição deste poder para todos, propor uma alternativa subalterna que “[...] constrói o conhecimento contra-hegemônico (Grosfoguel, 2009), estabelecendo um compromisso ético-político que elabora o conhecimento na perspectiva dos povos, gêneros, raças e culturas que o sujeito colonizado foi submetido [...]” (GONÇALVES; MUCHERONI, 2021, p. 8).

Dentre as propostas contra-hegemônicas encontradas na presente pesquisa, destaca-se aqui: a desclassificação (GARCIA GUTIÉRREZ, 2018), a epistemografia (GARCIA GUTIÉRREZ, 2018), a mobilização do conhecimento (PÉREZ MORA; INGUANZO ARIAS,

2018) e a reparação taxonômica (MOURA, 2018, p. 128). Neste trabalho, tais propostas serão analisadas como opções decoloniais de Organização do Conhecimento, mas é importante salientar que as opções decoloniais não se esgotam nelas, uma vez que elas refletem apenas os resultados da pesquisa bibliográfica que deu origem ao presente trabalho.

A Desclassificação é uma alternativa proposta por García Gutiérrez (2018), que parte do princípio da negatividade incorruptível na Organização do Conhecimento. Tal princípio reflete o fato de que a Desclassificação ao mesmo tempo em que nega, subverte as chamadas tecnologias unificadoras, de maneira que não as descarta, mas as subordina “[...] às opiniões de reabilitação democrática e social, como implica, por exemplo, a descolonização (GARCÍA GUTIÉRREZ, 2018, p. 110). Dessa forma, a organização do conhecimento pode ainda recorrer a hierarquização do conhecimento, mas essa hierarquização nunca pode reproduzir “[...] hierarquias epistemológicas ou sociais determinadas ou que fique suficientemente explícita tal origem” (GARCIA GUTIÉRREZ, 2018, p. 114).

A Epistemografia também foi desenvolvida por García Gutiérrez, mas, por outro lado, se volta mais especificamente à epistemologia, tendo como objetivo “[...] apontar as contradições da epistemologia, identificando os conhecimentos propositalmente excluídos dos processos contemporâneos de inscrição e fluxo, e que são produzidos por comunidades consideradas não epistemológicas” (GARCIA GUTIERREZ, 2006, p. 104 *apud* MORAES, 2018, p. 5). Para efetivar o seu projeto, García Gutiérrez lança mão de mais dois conceitos: exomemória (GARCIA GUTIÉRREZ, 2018) e etnoconhecimento (MORAES, 2018). O primeiro evidencia o fato de que os documentos são resultantes de produção humana na história, possuindo sentidos éticos e políticos na sua formação; e o segundo denota a necessidade de que, ao lado dos conhecimentos hegemônicos, sejam também considerados aqueles conhecimentos chamados tradicionais, desenvolvidos por povos indígenas, por exemplo, e quase sempre transmitido por via oral.

Entretanto, quando transportada para a Organização do Conhecimento, a Epistemografia assume um novo sentido, pois torna-se

a representação ou mapa conceitual de um único documento, a partir de categorias autonarrativas ou heteronarrativas (sempre que a heteronarração seja descolonizante), como a construção de todo um sistema no qual as estruturas e conteúdos daqueles documentos seriam ligados e articulados para outro sistema de representação que lhes daria visibilidade mundial através de uma língua franca digital (outra ferramenta epistemográfica). (GARCÍA GUTIÉRREZ, 2018, p. 110).

Dessa forma, evidencia-se o que Moraes (2018) destaca como a “[...] redistribuição da presença e força dos conhecimentos e culturas excluídas na rede digital em igualdade de condições”. Os conhecimentos chamados dominantes e os conhecimentos chamados subalternos passam a ser submetidos a uma rede de inter-relações que possibilitam um pluralismo narrativo e ético no processo de representação do conhecimento.

A proposta da Mobilização do Conhecimento desenvolvida por Pérez Mora e Inguanzo Arias (2018) é realizada a partir da percepção de que a Organização do Conhecimento pode ser tanto vista como estrutura quanto como sistema. A visão estrutural identifica as relações de poder que a Organização do Conhecimento revela e que são condicionantes para sua realização. Já a visão sistêmica possibilita compreender que os conhecimentos em embate podem ser reconhecidos como igualmente válidos e com importante função social. Diante dessa dialética, a Mobilização do Conhecimento permite “[...] geração de ações que envolvam os diversos atores de maneira participativa e democrática para a geração, apropriação e aplicação de tais conhecimentos na solução de problemas sociais.” (PÉREZ MORA; INGUANZO ARIAS, 2018, p. 209, *tradução nossa*). A dimensão estrutural e a dimensão sistêmica são tomadas, portanto, como referências para se compreender em que ponto os atores podem interferir para mudar a realidade.

Por fim, a Reparação Taxonômica é uma proposta idealizada por Melissa Adler, que tem como objetivo propor um “conjunto de iniciativas que visa identificar ‘certas estruturas imutáveis que continuam a apoiar o discurso eugenista’ [...]” (MOURA, 2018, p. 128) presentes nos Sistemas de Organização do Conhecimento. Tal proposta influenciou Silva e Silva (2022) a constituir o conceito de Reparação Epistêmica, que busca estabelecer uma justiça informacional através da disponibilização de informação científica e literária produzidas por povos originários e afrodescendentes, de maneira que se possibilite não apenas o acesso igualitário a informação, mas ao desenvolvimento pleno dos sujeitos em sociedades herdeiras de processos de colonização.

O que se identifica nas propostas acima elencadas, é uma luta por evidenciar a necessidade de se negar a universalidade e a neutralidade na Organização do Conhecimento, na sua dimensão epistemológica e nos Sistemas de Organização do Conhecimento desenvolvidos a partir de suas reflexões teórico-metodológicas, denunciando o seu papel de manutenção das violências epistêmicas perpetradas pelas classes dominantes, ao mesmo tempo em que se dispõe ambiente propício para se desenvolver teorias, metodologias e ferramentas que facilitem a valorização dos conhecimentos subalternizados, assim como sua

adequada representação, disponibilização e recuperação.

Ou seja, o que se identifica nas propostas acima é a necessidade de convocar os profissionais da informação a se entenderem como mediadores da informação e, portanto, conscientes de seu papel de interferência na realidade e no processo de apropriação da informação, devendo optar em manter as estruturas dominantes ou subvertê-las. Por isso que, a partir deste ponto, se iniciará uma proposta de abordagem crítica de enfrentamento ao epistemicídio através da identificação da Organização do Conhecimento com a Mediação Cultural da Informação.

5 Mediação Cultural da Informação: caminhos para uma abordagem crítica de enfrentamento ao epistemicídio na Organização do Conhecimento

A discussão sobre a Mediação no contexto da Biblioteconomia e da Ciência da Informação vem se estruturando ao longo dos anos de maneira a se criar uma base teórica capaz de compreender as dinâmicas relacionadas ao processo de facilitação do acesso, uso e apropriação da informação, principalmente no contexto dos chamados equipamentos informacionais, como bibliotecas, arquivos, museus, centros de documentos e informação. De acordo com Almeida Júnior (2015, p. 25 *apud* SILVA; CAVALCANTE, 2022, p. 7), a Mediação da Informação pode ser entendida como:

Toda ação de interferência – realizada em um processo, por um profissional da informação e na ambiência de equipamentos informacionais –, direta ou indireta; consciente ou inconsciente; singular ou plural; individual ou coletiva; visando a apropriação de informação que satisfaça, parcialmente e de maneira momentânea, uma necessidade informacional, gerando conflitos e novas necessidades informacionais.

Tendo como objetivo satisfazer as necessidades de apropriação da informação dos sujeitos, a Mediação da Informação acaba por ter um caráter plural, abrangendo não apenas os serviços tais como a referência ou a animação cultural, mas incluindo todos os processos realizados pelos profissionais da informação. A Organização do Conhecimento, por exemplo, é classificada pelos teóricos da Mediação como uma atividade de mediação implícita da informação, uma vez que não ocorre na relação direta com os usuários, estando ao lado de outras como a seleção de materiais, a preparação física do acervo e as ações de gestão (GOMES; REIS; JESUS, 2022).

A partir da ótica da Mediação da Informação, os profissionais da informação, mesmo e principalmente aqueles que se dedicam à Organização do Conhecimento e aos seus processos, podem assumir uma atitude crítica e consciente da potencialidade de suas práticas, compreendendo que o fazer de tais

profissionais pode colaborar para transformação da realidade ou para a sua manutenção e continuidade, caso continue alinhado aos interesses das classes hegemônicas.

Segundo Gomes (2020), porém, se tomar como fundamento a Mediação da Informação em todas as suas dimensões (ética, estética, dialógica, política e formativa), o profissional da informação tenderá a assumir uma postura consciente do seu papel na sociedade. Castro e Almeida Júnior (2022) corroboram com Gomes (2020) afirmando que a mediação da informação atrelará ao fazer do profissional um caráter emancipatório, que o oriente, por exemplo, a

[...] cooperar para a desmistificação dos sistemas de organização do conhecimento na esfera pública, discutindo e justificando porque os sistemas de organização do conhecimento e sua funcionalidade são importantes para o público. Ou seja, os profissionais que medeiam a informação precisam argumentar que esses sistemas fazem uma diferença na sociedade e mostram como eles afetam nossas atividades profissionais e diárias (CASTRO; ALMEIDA JÚNIOR, 2022, p. 16).

Desta maneira, o profissional da informação age diretamente junto à sociedade, assumindo uma postura protagonista na defesa dos Sistemas de Organização do Conhecimento e das Unidades de Informação, evidenciando a sua importância para a emancipação coletiva dos sujeitos a quem elas atendem ou podem atender. Aqui se evidencia que a relação de protagonismo junto a sociedade destaca um componente cultural, na prática do mediador da informação. Os profissionais na Organização do Conhecimento não estão lidando apenas com informação e conhecimento registrados, mas artefatos culturalmente produzidos e que refletem disputas epistêmicas de poder. A mediação da informação é também mediação cultural.

Todavia a mediação cultural deve ser entendida para além da viabilização aos sujeitos o acesso aos objetos da cultura, devendo provocar a possibilidade de construção da realidade pelos sujeitos tomando como ponto de partida a apropriação da “[...] informação circunscrita em um aspecto cultural presente em espaços e ações que promovem a cultura [...]” (CAVALCANTE, 2022, p. 8).

Tal postura corrobora com as reflexões desenvolvidas no âmbito da dimensão sociocultural da Organização do Conhecimento acima discutidas, mas agregam a elas uma abordagem crítica, entendida aqui como práxis transformadora que promove o pluralismo cultural e informacional, alinhando-se a um processo de emancipação e transformação da realidade em favor dos grupos não-hegemônicos e enfrentando diretamente as práticas epistemicidas vigentes.

Uma possibilidade rumo a essa abordagem crítica é a de que o profissional da Organização do Conhecimento

passe a se reconhecer como mediador a partir do que é entendido como Mediação Cultural da Informação. Para Bezerra e Cavalcante (2020, p. 6), a Mediação Cultural da Informação

[...] ainda que estabeleça interlocução tanto à mediação da informação quanto à mediação cultural, se diferencia ao englobar e propiciar compreensões interculturais dos fenômenos inseridos na lógica dos processos das relações informacionais/culturais que circundam os sujeitos, para além somente de um contato stricto com a informação ou objeto cultural.

Ela se alinha ao pensamento decolonial e propõe que o processo mediacional não se enfoque apenas no conhecimento científico, mas possibilitem o contato com as diversas formas de conceber o mundo. As vivências de povos indígenas, pessoas negras, pessoas LGBTQIA+, mulheres, pessoas periféricas, trabalhadoras e/ou oriundas de nações que foram ou ainda são sujeitas a dominação colonial e imperialista sejam elevadas a pé de igualdade ao pensamento de tradição racionalista, branco, europeu e norteamericano, cisheteronormativo, dominante e patriarcal.

Para efetivar sua prática intercultural e inclusiva, a Mediação Cultural da Informação toma como uma de suas principais ferramentas a “Tradução” (BEZERRA; CAVALCANTE, 2020). Com a “Tradução” os diversos saberes (hegemônicos, não hegemônicos ou contra-hegemônicos) são postos em diálogo, de maneira que se possa permitir criar uma inteligibilidade recíproca entre as diversas formas de conceber o mundo.

Aplicada à Organização do Conhecimento, a Mediação Cultural da Informação pode se agregar a outras propostas já desenvolvidas especificamente para a área (como a Reparação Taxonômica, a Mobilização do Conhecimento, a Epistemografia e a Desclassificação) e servir como abordagem e postura crítica mais ampla e plural que permite uma visão sistêmica e estrutural em relação à produção, à representação, à distribuição e à apropriação da informação e do conhecimento.

A partir dela também, os Sistemas de Organização do Conhecimento, sejam eles listas de cabeçalhos de assunto, sistemas de classificação ou mesmo ontologias, poderão ser revisitados e recriados com o objetivo de serem mantidas vivas, válidas e celebradas as diversas formas de conhecer e representar o mundo. Utilizando-os como instrumentos de diálogo intercultural, os profissionais da Organização do Conhecimento agora também Mediadores Culturais da Informação terão a possibilidade de promover a cultura de paz e a realização de um mundo melhor, mais inclusivo, “re-encantado” e que preza pela historicidade da concepção coletiva de mundo a partir da vivência de todos.

6 Considerações Finais

A Organização do Conhecimento historicamente tem se preocupado na produção de instrumentos e ferramentas que permitam a gestão da informação registrada, assim como a viabilização para o seu acesso e uso. Sua relação com o desenvolvimento da Epistemologia no Ocidente foi essencial para o estabelecimento de hierarquização entre os saberes e inter-relação entre eles, resultando-se assim nos Sistemas de Organização do Conhecimento (SOCs).

O surgimento nas últimas décadas de um movimento entre os pesquisadores da área para o estabelecimento de uma perspectiva sociocultural para Organização do Conhecimento tem colocado em pauta a suposta naturalidade dessa hierarquização de saberes, relutando-se a aceitar que os processos de representação da informação e do conhecimento como são amplamente realizados hoje em dia possam continuar sendo aceitos acriticamente como neutros e universais e denunciando a possibilidade de os SOCs estarem contribuindo para o epistemicídio dos saberes e modo de conceber o mundo de grupos marginalizados, como mulheres, indígenas, negros e pessoas LGBTQIA+, principalmente em países que sofreram ou sofrem dominação colonial e/ou imperialista.

Visando superar tal realidade, pesquisadores da área vêm construindo propostas contra-hegemônicas que possibilitem uma revisão e/ou o desenvolvimento de teorias e metodologias que contribuam para a valorização dos conhecimentos subalternizados, assim como sua adequada representação, disponibilização e recuperação. Através da pesquisa bibliográfica aqui realizada, pôde-se identificar ao menos quatro: a desclassificação, a epistemografia, a mobilização do conhecimento e a reparação taxonômica.

Tais propostas revelam uma preocupação dos profissionais da Organização do Conhecimento não somente em representar e alimentar bases de dados, mas em atuarem como mediadores de informação e de cultura. De maneira consciente e crítica, a mediação constitui-se como ação de emancipação coletiva que inclui não somente a facilitação do acesso e da apropriação da informação e dos artefatos da cultura, mas também a busca por estratégias de representação que condigam com as diferentes formas de ser e conceber o mundo e que apontem para um mundo mais inclusivo e plural.

Por isso que a presente pesquisa se enfocou atrelar a Organização do Conhecimento à Mediação Cultural da Informação. Como proposta decolonial e contra-hegemônica, a Mediação Cultural da Informação confere aos profissionais da informação a possibilidade de assumirem uma postura consciente em relação ao

seu fazer e adotarem uma abordagem crítica que coopere para o enfrentamento ao epistemicídio.

Encerra-se aqui apontando a necessidade de se dar continuidade aos estudos teóricos aqui desenvolvidos, buscando ampliar a discussão sobre a Mediação Cultural da Informação como abordagem crítica de enfrentamento ao Epistemicídio na Organização do Conhecimento, assim como gerar ferramentas e instrumentos para efetivar tal enfrentamento.

Referências

- BEZERRA, Arthur Coelho; CAVALCANTE, Luciane de Fátima Beckman. Mediação cultural da informação para o reencantamento do mundo. **Encontros Bibli**, Florianópolis, v. 25, p. 1-19, 2020. Disponível em: <https://brapci.inf.br/index.php/res/download/145984>. Acesso em 18 abr. 2023.
- BRENTAN JUNIOR, Elisio Custodio; MARTINS, Bianca Rodrigues; SANTOS NETO, João Arlindo. A mediação cultural e a análise de assunto: mais que discursos, unindo comunidades. **PontodeAcesso**, Salvador, v. 12, n. 3, p. 3-27, dez. 2018. Disponível em: <https://brapci.inf.br/index.php/res/download/120018>. Acesso em 18 abr. 2023.
- CARNEIRO, Aparecida Sueli. **A construção do outro como não-ser como fundamento do ser**. 2005. 339f. Tese (Doutorado em Educação) – Programa de Pós-graduação em Educação, Universidade de São Paulo, 2005. Disponível em: <https://negrasoulblog.files.wordpress.com/2016/04/a-construc3a7c3a3o-do-outro-como-nc3a3o-ser-como-fundamento-do-ser-sueli-carneiro-tese1.pdf>. Acesso em 15 maio 2023.
- CASTRO, Jetur; ALMEIDA JÚNIOR, Oswaldo Francisco de Almeida. Mediação da informação e reconhecimento intersubjetivo: aproximações. **Encontros Bibli**, Florianópolis, v. 27, p. 1-22, 2022. Disponível em: <https://brapci.inf.br/index.php/res/download/212603>. Acesso em 18 abr. 2023.
- CAVALCANTE, Luciane de Fátima Beckman. A violência contra a mulher sob o olhar Mediação Cultural da Informação: análise da exposição “Retratos Relatos”. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 18, p. 1-19, 2022. Disponível em: <https://brapci.inf.br/index.php/res/download/204786>. Acesso em 18 abr. 2023.
- COSTA, Deniz; MIRANDA, Marcos. A organização do conhecimento sobre Umbanda e sua representação bibliográfica: uma análise exploratória a partir de registros bibliográficos. **Informação & Informação**, Londrina, v. 24, n. 3, p. 154-182, set./dez. 2019. Disponível em: <https://brapci.inf.br/index.php/res/download/134243>. Acesso em 18 abr. 2023.
- GARCEZ, Dirnéle Carneiro; SALES, Rodrigo de. Decolonizando a Organização do Conhecimento: um olhar do periódico Knowledge Organization (2000-2020). **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 14, 2021. Disponível em:

<https://brapci.inf.br/index.php/res/download/195226>. Acesso em 18 abr. 2023.

GARCÍA GUTIÉRREZ, Antonio. Por uma economia política da organização do conhecimento. **Liinc em Revista**, Rio de Janeiro, v. 14, n. 2, p. 109-117, nov. 2021. Disponível em: <https://brapci.inf.br/index.php/res/download/109818>. Acesso em 18 abr. 2023.

GOMES, Alex Sandro; GOMES, Claudia Roberta Araújo. Classificação dos tipos de pesquisa em Informática na Educação. *In*: JACQUES, Patrícia Augustin; PIMENTEL, Mariano; SIQUEIRA, Sean; BITTENCOURT, Ig. (Orgs). **Metodologia de Pesquisa em Informática na Educação: concepção de pesquisa**. Porto Alegre: SBC, 2020. Disponível em: https://metodologia.ceie-br.org/wp-content/uploads/2019/06/livro1_cap4.pdf. Acesso em 01 abr. 2023.

GOMES, Henriette Ferreira. Mediação da informação e suas dimensões dialógica, estética, formativa, ética e política: um fundamento da Ciência da Informação em favor do protagonismo social. **Informação & Sociedade: Estudos**, v. 30, n. 4, p. 1-23, out./dez. 2020. Disponível: <https://brapci.inf.br/index.php/res/download/155808>. Acesso em 18 abr.2023.

GOMES, Henriette Ferreira; RESI, Dayane Patrícia Silva dos; JESUS, Jéssica Natália de. Mediação explícita e implícita: atividades, atributos e zonas de interseção para o alcance das dimensões da mediação da informação. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 18, n. esp., p. 1-23, 2022. Disponível em: <https://brapci.inf.br/index.php/res/download/204746>. Acesso em 18 abr. 2023.

GONÇALVES, Robson de Andrade; MUCHERONI, Marcos L. O que é epistemicídio? Uma introdução ao conceito para a área da Ciência da Informação. **Liinc em Revista**, Rio de Janeiro, v. 17, n. 2, p. 1-12, nov. 2021. Disponível em: <https://brapci.inf.br/index.php/res/download/169571>. Acesso em 18 abr. 2023.

GUIMARÃES, José Augusto Chaves *et al.* A dimensão cultural da organização do conhecimento: uma análise no contexto internacional da Ciência da Informação. **Scire**, v. 25, n. 1, p. 35-43, en./jun 2018. Disponível em: <https://brapci.inf.br/index.php/res/download/196362>. Acesso em 18 abr. 2023.

MANHIQUE, Ilídio Lobato Ernesto; CASARIN, Helen de Castro Silva. Abordagem cultural da organização do conhecimento na Ciência da Informação brasileira. **Encontros Bibli**, Florianópolis, v. 24, n. 56, p. 1-20, set./dez. 2019. Disponível em: <https://brapci.inf.br/index.php/res/download/124081>. Acesso em 18 abr. 2023.

MARCONDES, Carlos H. Fundamentos da Organização do Conhecimento. **PontodeAcesso**, Salvador, v. 15, n. 3, p. 249-282, dez. 2021. Disponível em: <https://brapci.inf.br/index.php/res/download/169386>. Acesso em 18 abr. 2023.

MARTÍNEZ-ÁVILA, Daniel; MELLO, Mariana Rodrigues Gomes de. Epistemologias, gênero e dogmatismo científico: desdobramentos na Organização do Conhecimento. **LOGEION**, Rio de Janeiro, v. 9, n. 1, p. 182-194, set.

2022/fev. 2023. Disponível em: <https://brapci.inf.br/index.php/res/download/204862>. Acesso em 18 abr. 2023.

MORAES, Míriam Gontijo de. Tensão identitária e organização do conhecimento: olhar epistemográfico. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 11, n. 2, 2018. Disponível em: <https://brapci.inf.br/index.php/res/download/151993>. Acesso em 18 abr. 2023.

MOREIRA, Walter; SABBAG, Deise. Diálogos entre as questões socioculturais e os sistemas de organização do conhecimento. **Scire**, v. 28, n. 2, p. 35-43, jul./dic 2022. Disponível em: <https://brapci.inf.br/index.php/res/download/212942>. Acesso em 18 abr. 2023.

MOURA, Maria Aparecida. Organização social do conhecimento e performatividade de gênero: dispositivos, regimes de saber e relações de poder. **Liinc em Revista**, Rio de Janeiro, v. 14, n. 2, p. 118-135, nov. 2018. Disponível em: <https://brapci.inf.br/index.php/res/download/109799>. Acesso em 18 abr. 2023.

NASCIMENTO, Felipe Mozart de Santana; PINHO, Fábio Assis. Sistemas de organização do conhecimento: semelhanças e diferenças. **ConCI: Convergências em Ciência da Informação**, v. 2, n. 3, p. 104-122, set./dez. 2019. Disponível em: <https://brapci.inf.br/index.php/res/download/141027>. Acesso em 18 abr. 2022.

NHACUONGUE, Januario Albino; DUTRA, Moisés Lima. De Paul Otlet à Web Semântica: aportes teóricos sobre a organização do conhecimento. **Informação & Tecnologia (ITEC)**, Marília/João Pessoa, v. 3, n. 1, p. 138-158, jan./jun. 2016. Disponível em: <https://brapci.inf.br/index.php/res/download/54218>. Acesso em 18 abr. 2023.

PÉREZ MORA, Ricardo; INGUANZO ARIAS, Blanca Lizbeth. La organización del conocimiento como proceso: la movilización del conocimiento. Rio de Janeiro, v. 14, n. 2, p. 199-212, nov. 2021. Disponível em: <https://brapci.inf.br/index.php/res/download/109829>. Acesso em 18 abr. 2023.

SILVA, Carlos Robson Souza da; CAVALCANTE, Luciane de Fátima Beckman. Da mediação à mediação cultural da informação: percursos e questionamentos. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais [...]**. Porto Alegre: UFRGS, 2022. Disponível em: <https://brapci.inf.br/index.php/res/v/200600>. Acesso em 18 abr. 2023.

SILVA, Franciéle Carneiro Garcês da; SILVA, Rubens Alves da. Da ausência à evidência: notas teórico-críticas sobre o Princípio da Ausência, Epistemicídio e Reparação Epistêmica em bibliotecas e Biblioteconomia. **InCID**, Ribeirão Preto, v. 13, n. 1, p. 47-72, mar./ago. 2022. Disponível em: <https://brapci.inf.br/index.php/res/download/205600>. Acesso em 18 abr. 2023.

SONO, Rodrigo Akio Siqueira; FRANCELIN, Marivalde Moacir. Avaliação de sistemas de organização do conhecimento: uma análise na literatura da área. **Revista**

Brasileira de Biblioteconomia e Documentação, São Paulo, v. 1, p. 1-27, 2022. Disponível em: <https://brapci.inf.br/index.php/res/download/213348>. Acesso em 18 abr. 2023.

STRAPASSON, Bruno Angelo; MAGALHÃES, Fernanda Gutierrez; CUSTÓDIO, Joice Kalfeld. Comunicação entre a pesquisa básica, aplicada e teórica na Análise do

Comportamento no Brasil: uma análise bibliométrica.

Interação em Psicologia, v. 17, n. 1, p. 117-128, jan./abr. 2013. Disponível em:

<https://revistas.ufpr.br/psicologia/article/view/26496/21326>. Acesso em 01 abr. 2023.

YIN, Robert K. **Pesquisa qualitativa do início ao fim**. Porto Alegre: Penso: 2016.

Meninas e mulheres na ciência: o caso das pesquisadoras brasileiras e em Ciência da Informação no Norte e Nordeste do Brasil

Girls and women in science: the *case* of brazilian researchers in Information Science in the North and Northeast of Brazil

Virgínia Bentes Pinto (1), Eduarda Kelly Teixeira Ricardo (2), Francilania Lima de Sousa (3)

Universidade Federal do Ceará, Av. da Universidade, 2762 - Benfica, Fortaleza - CE, 60020-181, vbentes@ufc.br (1), eduardakellyteixeira@alu.ufc.br (2), francilania@alu.ufc.br (3).

Resumo

No decorrer da história da humanidade, as mulheres são confrontadas com personificações arquetípicas de toda espécie, sejam familiares, sociais e, naturalmente, na ciência, onde sua representação não é percebida por sua competência, demonstrando pré-conceitos enraizados na sociedade. Observando essa realidade, desenvolvemos uma pesquisa que tem como **objetivo geral**: identificar a presença e as temáticas de pesquisadores brasileiros que atuam nas áreas de Biblioteconomia e Ciência da Informação em universidades do norte e nordeste do Brasil. **Metodologia**: estudo exploratório multicase nas universidades do norte e nordeste do Brasil, identificando-se as instituições, os nomes dos pesquisadores e suas produções no período de 2020-2022, consultando-se o Lattes. **Resultados**: os achados da pesquisa refletem a presença de 169 pesquisadores, sendo que 103 são mulheres e 66 homens. Concernente à produção foram identificados 1210 artigos de periódicos. Desse total, 60,16% são mulheres pesquisadoras das 14 instituições que possuem os cursos de graduação em Biblioteconomia e Arquivologia, pós-graduação em Biblioteconomia, Ciência da Informação e gestão da informação e do conhecimento. Concernentes as temáticas estudadas no contexto da organização e representação destacam-se: Análise da informação, Análise documental, Catalogação, Indexação, Classificação, Organização da informação, Organização do conhecimento, Organização documental, Representação da informação, Representação da informação temática, Representação de mulheres, Representação descritiva, Representação do conhecimento e Representação temática da informação. **Conclusão**: A produção das meninas e mulheres pesquisadora das áreas de Biblioteconomia, Ciência da Informação e Arquivologia nas regiões norte e nordeste brasileira já se destaca na maioria das universidades, com excessão do Curso de Arquivologia na Universidade Estadual da Paraíba.

Palavras-chave: pesquisadoras na ciência da informação-Brasil-Norte e Nordeste. visibilidade da mulher na Ciência da Informação-Brasil- Norte e Nordeste. meninas e mulheres na pesquisa.

Abstract

Throughout human history, women are confronted with archetypal personifications of every kind, whether familiar, social and, of course, in science, where their representation is not perceived by their competence, evidencing preconceptions rooted in society. Analyzing this reality, we developed a research that has the **general objective**: to identify the presence and themes of Brazilian researchers who work in the areas of Librarianship and Information Science in universities in the North and Northeast of Brazil. **Methodology**: exploratory multicase study in the universities of the north and northeast of Brazil, identifying the institutions, the names of the researchers and their productions in the period between 2020-2022, consulting the Lattes System (CNPq). **Results**: the research findings show the presence of 169 researchers, of which 103 are women and 66 men. When it comes to production, 1210 journal articles were identified. Of this total, 60.16% are women researchers from the 14 institutions that have graduation courses in Librarianship and Archival Science, Library post-graduation in Librarianship and Information Science and work in several themes that include the intention of demining Information Science and Librarianship and they work on different themes, standing out in the context of organization and representation; Information analysis, Documentary analysis, Cataloguing, Indexing, Classification, Information organization, Knowledge organization, Document organization, Information representation, Thematic information representation, Representation of women, Descriptive representation, Knowledge representation, Thematic representation of information. **Conclusion**: The production of girls and women researchers in the areas of Library Science, Information Science and Archival Science in the North and Northeast regions already stands out in most universities, with the exception of the Archival Science Course at the State University of Paraíba.

Keywords: researchers in information science-Brazil-North and Northeast. visibility of women in Information Science-Brazil-North and Northeast. girls and women in research.

Resumen

En el transcurso de la historia humana, la mujer se enfrenta a personificaciones arquetípicas de todo tipo, ya sea familiar, social y, por supuesto, en la ciencia, donde su representación no es percibida por su competencia, demostrando prejuicios arraigados en la sociedad. Observando esta realidad, desarrollamos una investigación que tiene como objetivo general: identificar la presencia y las temáticas de investigadores brasileños que actúan en las áreas de Biblioteconomía y Ciencias de la Información en universidades del norte y nordeste de Brasil. Metodología: estudio exploratorio multicaso en universidades del norte y noreste de Brasil, identificando las instituciones, los nombres de los investigadores y sus producciones en el período 2020-2022, consultando a Lattes. Resultados: los hallazgos de la investigación reflejan la presencia de 169 investigadores, de los cuales 103 son mujeres y 66 son hombres. En cuanto a la producción, se identificaron 1210 artículos de revistas. De este total, el 60,16% son investigadoras de las 14 instituciones que ofrecen carreras de grado en Biblioteconomía y Archivología, posgrados en Biblioteconomía, Ciencias de la Información y Gestión de la Información y el Conocimiento. En cuanto a los temas estudiados en el contexto de la organización y representación, se destacan: Análisis de la información, Análisis documental, Catalogación, Indización, Clasificación, Organización de la información, Organización del conocimiento, Organización de los documentos, Representación de la información, Representación de la información temática, Representación de mujeres, Representación descriptiva, Representación del conocimiento y Representación temática de la información. Conclusión: La producción de niñas y mujeres investigadoras en las áreas de Bibliotecología, Ciencias de la Información y Archivística en las regiones Norte y Nordeste de Brasil ya se destaca en la mayoría de las universidades, con excepción del Curso de Archivología de la Universidad Estadual de Paraíba.

Palabras clave: mujeres investigadoras en Ciencias de la Información-Brasil-Norte y Nordeste. visibilidad de las mujeres en las Ciencias de la Información-Brasil-Norte y Nordeste. niñas y mujeres en la investigación.

1 Introdução

Desde o início da evolução humana, a mulher carrega na sua história uma representação imprecisa, estruturada numa terminologia estereotipada, de modo a ser desacreditada na sociedade, independentemente se nos campos artísticos, religiosos, relações sociais e, posteriormente, também na ciência. Todo esse arquétipo coloca a figura feminina como uma sub-representação que é fundamentada em preconceitos “tatuados” por uma sociedade machista e patriarcal.

Schiebinger (1999, p.26) afirma que “ao romantizar a feminilidade tradicional, o feminismo da diferença pouco faz para superar os estereótipos convencionais de homens e mulheres”. A autora continua refletindo que essa visão de diferenças não é nada mais que uma hipótese, uma vez que “não é óbvio que o gênero tenha uma influência mais forte sobre a ciência do que outras divisões políticas e culturais [...]” (SCHIEBINGER, 1999, p.37).

Particularmente, no âmbito da ciência brasileira, essa realidade, embora ainda permaneça, está sendo desconstruída pouco a pouco em decorrência da conscientização e a autonomia conquistadas pela mulher, conforme argumenta Bandeira e Siqueira (2002, p.270) devido ao Movimento Feminista, “[...] que abriu novas perspectivas e que trouxe novas questões aos campos disciplinares, à produção do conhecimento e da ciência [...]”.

Nessa lógica, a década de noventa pode ser considerada um grande marco na ciência brasileira, em consequência do prêmio da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) dedicado às mulheres, conquistado em 2013 pela cientista Márcia Barbosa, física brasileira, que pesquisa sobre as estruturas complexas das moléculas da água.

Na mesma esteira, a Organização das Nações Unidas (ONU) chama a atenção para a necessidade de igualdade de gênero entre a ciência como condição *sine qua non* para que os objetivos do desenvolvimento sustentável definidos na Agenda 2030 sejam atingidos.

[...] nos últimos 15 anos, a comunidade global fez um grande esforço para inspirar e envolver mulheres e meninas na ciência. No entanto, elas continuam sendo excluídas de participar plenamente da ciência. (ONU NEWS, 2020, online).

Por sua vez, Oliveira e Bufrem (2019) argumentam em oposição ao discurso de muitos pesquisadores que culpabilizam as mulheres por não produzirem o suficiente e que isso seria responsável pela sua parca visibilidade. Para essas pesquisadoras, isso não é verdade, posto que

[...] a mulher, há muitos anos, vem exercendo seu papel de pesquisadora, mas sua visibilidade e seu reconhecimento na condição de cientista têm sido ratificados a partir de estudos específicos [...] (ibidem, p.3).

Observando essa realidade, empreendemos esta pesquisa nos pautando no seguinte **problema**: como se evidencia o destaque das mulheres pesquisadoras dos campos da Ciência da Informação, da Biblioteconomia e Arquivologia das Universidades do norte e nordeste do Brasil? Visando encontrar respostas a esse questionamento definimos como **objetivo geral**: investigar a relação entre a produção científica de pesquisadoras e pesquisadores brasileiras/os no campo da Ciência da Informação e da Biblioteconomia no norte e nordeste brasileiro. Os **objetivos específicos** são:

- a) Mapear as Universidades do norte e nordeste que possuem graduação em Biblioteconomia e Arquivologia, pós-graduação em Biblioteconomia e Ciência da Informação.

- b) Levantar, na plataforma Lattes, pesquisadoras/res das referidas instituições.
- c) Verificar o número de produções por pesquisador(a) no período de 2020 a 2022, bem como as temáticas investigadas.

Estes são os aspectos fundamentais tratados neste artigo, e acreditamos que possam comprovar a representação das pesquisadoras nas instituições analisadas, bem como despertar o interesse para que mais pesquisas nessa temática venham a ser desenvolvidas.

2 Algumas considerações sobre gênero na ciência

Os estudos com ênfase nas questões de gênero no contexto da pesquisa são essenciais para o entendimento sobre a discriminação e o preconceito vivenciados pelas mulheres pesquisadoras em todos os campos de conhecimentos.

Espírito Santo (2009) traz em sua pesquisa o movimento de emancipação feminina como exemplificação do ponto de partida para estudos voltados às mulheres. A perspectiva mais conhecida, conforme a autora, é a de gênero, na qual define a construção do que é ser “homem” ou ser “mulher” de um ponto de vista cultural, social, político e econômico, não havendo relação, assim, com questões biológicas.

É sob essa ótica que o papel da mulher na sociedade foi construído. Vistas como frágeis, obrigatoriamente delicadas, e “donas de casa”, por muito tempo a mulher foi excluída das políticas que regem a população, do espaço de trabalho, da ciência e da sociedade.

Para se entender o problema que existe entre a ciência e as mulheres é preciso, inicialmente, se entender que se trata de um problema de relações sociais de gênero, uma vez que a ciência tem se caracterizado como masculina, ora excluindo as mulheres, ora negando os seus feitos científicos, através de discursos e métodos nada neutros. (RODRIGUES DA SILVA, 2008, p.2).

De acordo com Schiebinger (2008), as discussões sobre a participação da mulher na ciência começaram no Brasil na década de 70, e, no entanto, ainda não se verificam mudanças significativas nos dias atuais.

[...]apesar da ampla disponibilidade de dados, as políticas científicas do país ainda não assimilaram a importância de estabelecer indicadores sobre relações de gênero, que perpassam também o sistema de ciências e tecnologia. (ibidem, p.269).

Leta (2003, p. 283) afirma que “historicamente, a ciência sempre foi vista como uma atividade realizada por homens”. Ainda nesse ínterim, a autora diz que somente a partir dos anos 1980 e 1990 é que houve o aumento da participação feminina na pesquisa brasileira.

A ciência, teoricamente, possui caráter neutro, uma vez que não devem ser envolvidos valores sociais em trabalhos científicos. Sabemos, no entanto, que esse princípio é complexo e não é cumprido, não ocorre como deveria.

É necessário compreender a Ciência como um empreendimento humano e as/os cientistas como produtos dessa cultura, reconhecendo que a Ciência afeta e é afetada por vários elementos e esferas intelectuais da cultura em que está inserida. (HEERDT; BATISTA, 2016, p. 43).

Em sua pesquisa acerca das questões de gênero e da natureza na formação docente, Heerdt e Batista criaram uma Unidade Didática, intitulada “construção do conhecimento científico e a visibilidade de gênero na Ciência”, cujo público-alvo eram docentes da rede estadual. As falas dos participantes foram gravadas em vídeo e, em um desses registros, foi identificada a fala de um docente (masculino) que acredita ser o machismo uma espécie de amuleto a pouca visibilidade das mulheres na ciência:

Já parou para pensar que pode ser um amuleto, esse negócio de machismo... porque, veja bem, chega lá vê que é pauleira, acha esse amuleto de desculpa... segundo ponto e:: veja bem, é muito difícil para a mulher desistir de ir num salão de beleza, de ficar uma hora na frente do espelho, de desistir de ir na academia. PORQUE um cientista tem que ser dedicado. Se ela for ((pesquisadora)) ela NÃO PODE TER ((marido, filhos, vida familiar)) ela tem que abdicar disso, ela tem que abdicar. (ibidem, p.42).

Além de uma visão estereotipada de quais são realmente as prioridades da mulher, há uma imagem equivocada da ciência e do ser cientista, sendo absurda a ideia de abdicar de toda uma vida social em prol de pesquisas científicas: se trata de um equilíbrio entre ambas. É importante ressaltar, também, que as desistências das mulheres nesse âmbito, não se trata da dificuldade de ser cientista, mas sim da falta de apoio e discriminação que elas vivenciam durante toda a sua jornada na vida acadêmica e como pesquisadoras.

Dessa forma, sob a visão de uma ciência pautada na supervalorização do homem, as produções de mulheres cientistas são invisibilizadas, seja operante o discurso de que elas não produzem o suficiente ou pela ideia que não conseguem lidar com as dificuldades científicas, não percebendo, assim, a existência de muitas mulheres na ciência que acabam não encontrando solo fértil o suficiente para crescer em meios que as reprimem.

3 Metodologia

Adotamos para a realização desse estudo o caráter exploratório de cunho fundamental, tendo em vista que objetivamos maior familiaridade com o nosso tema, através de levantamentos bibliográficos e compreensão da visibilidade das meninas e mulheres na ciência.

De acordo com Severino (1980), a pesquisa exploratória está voltada ao mapeamento das condições do objeto a ser trabalhado através do levantamento de informações.

Concernente, a pesquisa fundamental está focada na melhoria das teorias científicas, ou seja, busca a contextualização do objeto a ser estudado com o intuito de produzir novos conhecimentos, no entanto, sem aplicação prática. Para Godelier (1996), essa categoria de investigação tem como característica a incitação de um conjunto de áreas específicas em prol de um objetivo.

Referente a pesquisa empírica, em um primeiro momento, mapeamos, por meio da plataforma *Google*, as universidades públicas do norte e nordeste brasileiro a fim de identificar em quais delas haviam cursos de graduação em Biblioteconomia e Arquivologia, assim como Programas de Pós-graduação em Biblioteconomia, Arquivologia e Ciência da Informação. Identificadas as instituições, listamos os docentes que lecionam nos referidos cursos.

Em seguida, realizamos um mapeamento, na plataforma Lattes, com ênfase na representatividade da mulher pesquisadora brasileira nos campos da Biblioteconomia, Ciência da Informação e Arquivologia, no norte e nordeste do Brasil, identificando aqueles que possuem mestrado e doutorado nas áreas da pesquisa.

Nessa plataforma, pesquisadores, estudantes e profissionais alimentam seu currículo completo, com atividades em várias categorias: pesquisa, ensino, extensão e outros. Trata-se de um ambiente de dados abertos.

Observados os docentes, ainda nesta plataforma, separando-os por gênero, verificamos aqueles considerados como pesquisadores, em razão do seu envolvimento com a pós-graduação nas referidas áreas, bem como a intensidade da produção acadêmica no período 2020–2022.

Ainda nessa empiria, cotejamos a produção de todos os professores e professoras, bem como aquelas temáticas relacionadas à representação e organização do conhecimento, as linguagens documentárias de maneira geral e terminologias.

4 Resultados e Discussões

Para a estruturação da análise dos dados e discussão dos resultados, nos pautamos nas seguintes categorias:

- a) Instituições que possuem curso de graduação em Biblioteconomia e Arquivologia e pós-graduação em Biblioteconomia ou Ciência da Informação.

Os achados da pesquisa retratam que nas regiões norte e nordeste do Brasil, das 608 universidades e faculdades existentes, 14 delas possuem cursos de graduação e pós-graduação nas áreas de Biblioteconomia, Arquivologia ou Ciência da Informação, conforme pode ser observado no quadro-1 no apêndice.

Com relação à oferta de cursos de pós-graduação, constatou-se que a Universidade Federal do Maranhão (UFMA), Universidade Federal de Rondônia (UNIR) e Universidade Federal do Amazonas (UFAM) não possuem cursos de pós-graduação *stricto sensu*.

A Universidade Federal do Cariri (UFCA) possui Mestrado Profissional em Biblioteconomia, por sua vez, a Universidade Federal de Sergipe (UFS) tem mestrado Profissional em Gestão da Informação e do Conhecimento e a Universidade Federal do Rio Grande do Norte (UFRN) conta com o mestrado profissional em Gestão da Informação e do Conhecimento

A Universidade Federal do Ceará (UFC) possui mestrado em Ciência da Informação, as Universidades Federais da Paraíba (UFPB), de Pernambuco (UFPE) e da Bahia (UFBA) contam com mestrado e doutorado em Ciência da Informação.

- b) Quantidade de pesquisadoras/res na plataforma Lattes, e que atuam nas referidas instituições.

Atinentes a pesquisadores por universidade, a empiria na plataforma Lattes, evidenciou a existência de 169 que atuam nos referidos cursos, desse total, 103 são pesquisadoras. Tal resultado significa que 60,16% do corpo científico das instituições aqui retratadas são mulheres. Vejam-se a tabela 1

Tabela 1. Número de pesquisadores por Universidade

Universidades	Mulheres	Homens
UFAM	3	1
UNIR	4	4
UFPA	9	7
UFAL	6	5
UFBA	22	10
UFCA	5	4
UFC	8	5
UFMA	7	1
UEPB	1	3
UFPB	19	8
UFPE	7	11
UESPI	?	?
UFRN	6	6
UFS	6	1
<i>Total</i>	<i>103</i>	<i>66</i>

Fonte: Dados da pesquisa.

Tais dados evidenciam que as mulheres aparecem com mais evidência na pesquisa nas Universidades que têm pós-graduação *Stricto Sensu*, naturalmente, com exceção da UFPE.

Outro fator que chama a atenção é o fato de que as Universidades: UNIR, UFAM, UFMA e UEPB, ainda que não possuam cursos de pós-graduação, se sobressaem com pesquisadores. Ressalta-se que na UNIR e a UFRN, os homens e as mulheres estão empatados na pesquisa. As demais, com exceção da UEPB, as mulheres se destacam. No que de respeito a UEPB, tais achados podem estar atrelados ao fato de que ela tem curso de arquivologia em que muitas disciplinas são ministradas por pesquisadores da área de direito.

Ressaltamos que, embora esteja corroborado no quadro-1 que a Universidade Estadual do Piauí possui curso de graduação em Biblioteconomia, no *site* dessa Universidade não foi possível identificarmos o referido curso e, naturalmente, os pesquisadores.

- c) Relacionada ao destaque no número de produções por pesquisador no período de 2020 a 2022 e temáticas que se enquadram.

Os achados da pesquisa, demonstram um total de 1210 publicações, sendo que as mulheres se notabilizam com um total de 728 produções. Esse resultado positivo, comprova que de certa forma, as conquistas femininas pouco a pouco estão permitindo que as mulheres adentrem progressivamente, na atividade de pesquisa.

Concernente às temáticas que as pesquisadoras mais se dedicam no contexto da organização e representação: Análise da informação, Análise documentária, Catalogação, Indexação, Classificação, Organização da informação, Organização do conhecimento, Organização documental, Representação da informação, Representação da informação temática, Representação de mulheres, Representação descritiva, Representação do conhecimento, Representação temática da informação.

Chama atenção o fato de que somente duas pesquisadoras se dedicam a investigação sobre a temática das meninas e mulheres na ciência no âmbito da Ciência da Informação, Biblioteconomia e Arquivologia.

5 Conclusão

Tínhamos um pressuposto de que os homens seriam destaque da pesquisa nos domínios de conhecimento investigados, pois a institucionalização da Ciência da Informação deu-se na década de 60, demonstrando a presença de 59 pessoas, sendo 52 homens e 7 mulheres. Entretanto, tal pressuposto, foi negado. Consideramos que essa positividade dos resultados esteja associada a

evolução da presença feminina nas pesquisas, a partir das décadas de oitenta e noventa.

Também, não podemos deixar de expor que as áreas de Biblioteconomia e Arquivologia, tem grande representatividade na atuação de mulheres, no contexto brasileiro o que pode ter contribuído para esses resultados positivos.

Ressaltamos que, embora nesta pesquisa as mulheres tenham se destacado, não significa que em todos os estados do país tal fato seja real. Contudo, é um dado significativo para que as meninas e mulheres sejam cada vez mais reconhecidas na ciência, não somente pelo fato de que no Brasil pareça ser uma cultura de atuação nesses campos, porém que, efetivamente elas possam ser cada vez mais reconhecidas.

Não podemos deixar de lembrar que em se tratando de gênero feminino, embora ainda esteja associado ao trabalho doméstico, foi perceptível na pesquisa que a participação das meninas e mulheres na ciência da Informação, Biblioteconomia e Arquivologia, nas universidades pesquisadas das regiões norte e nordeste se distinguem na pesquisa.

Inferimos ainda que mesmo se enfatizando na pesquisa, a literatura corrobora que as mulheres continuam a desempenhar inúmeras atividades dentre as quais a pesquisa.

No tocante as temáticas em que as pesquisadoras se sobressaem, encontram-se aquelas relacionadas com o tronco “duro” das áreas de Biblioteconomia, Ciência da Informação e Arquivologia. Tal constatação espelha que as áreas de organização e representação da informação cada vez mais se fortalecem, ratificando a necessidade de maiores investimentos em pesquisa nesse assunto.

Também não podemos deixar de dizer que, mesmo sendo destaque na pesquisa aqui exibida, o preconceito ainda se faz presente em relação à questão de gênero na ciência, em particular, de considerar tipicamente que o seu trabalho está relacionado às tarefas tidas como tipicamente femininas.

Finalmente, consideramos que nossa questão foi respondida e que os objetivos do estudo foram atingidos e abrem possibilidades para outros estudos pautados nos dados inseridos na plataforma Lattes.

Notas

As autoras agradecem ao CNPQ, o financiamento desta pesquisa, por meio do PIBIC.

Referências

- BANDEIRA, L. SIQUEIRA, D. A perspectiva feminista no pensamento moderno e contemporâneo. **Sociedade e Estado**, v.12, n. 02, p. 263–284, 2002. Disponível em: <https://periodicos.unb.br/index.php/sociedade/article/view/44156>. Acesso em: 11 mai. 2023.
- ESPÍRITO SANTO, P. Os estudos de gênero na Ciência da Informação. **Em Questão**, Porto Alegre, v. 14, n. 2, p. 317–332, 2009. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/6389>. Acesso em: 06 mar. 2023.
- GODELIER, M. **O enigma da dádiva**. Lisboa: Edições 70, 1996.
- HEERDT, B., BATISTA, I. de L. Questões de gênero e da natureza da ciência na formação docente. **Investigações Em Ensino De Ciências**, v.21, n.2, p.30–51, 2016. Disponível em: <https://doi.org/10.22600/1518-8795.ienci2016v21n2p30>. Acesso em: 07 mar. 2023.
- LETA, J. As mulheres na ciência brasileira: crescimento, contrastes e um perfil de sucesso. **Estudos Avançados**, São Paulo, v. 17, n. 49, set./dez. 2003. Disponível em: <https://www.scielo.br/j/ea/a/F8MbrypqGsJxTzs6msYFp9m/?lang=pt>. Acesso em: 20 de abril de 2023.
- OLIVEIRA, A. L. T.; BUFREM, L. S. Visibilidade da mulher como fonte de informação: mapeamento das produções científicas apresentadas no encontro nacional de pesquisa em ciência da informação (2009 – 2018). *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 20, 2019. Florianópolis. **Anais [...]** Florianópolis: UFSC, 2019. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/123781>. Acesso em: 06 mar. 2023.
- ONU NEWS. Online. 2020. Disponível em: <https://news.un.org/pt/story/2020/02/1703721>. Acesso em: 12 mar. 2023.
- RODRIGUES DA SILVA, E. A (in)visibilidade das mulheres no campo científico. **Travessias**, Cascavel, v. 2, n. 2, 2008. Disponível em: <https://e-revista.unioeste.br/index.php/travessias/article/view/3026>. Acesso em: 12 mar. 2023.
- SCHIEBINGER, L. **Has feminism changed science?** Edição brasileira: O feminismo mudou a ciência? Trad. Raul Fiker. Bauru: EDUSC, 2001.
- SCHIEBINGER, L. Mais mulheres na ciência: questões de conhecimento. **História, Ciências, Saúde – Manguinhos**, Rio de Janeiro, v.15, supl., jun. 2008, p.269-281. Disponível em: <https://www.scielo.br/j/hcsm/a/LZcRqYbsQR4cxYkgfCGyjr/?lang=pt>. Acesso em 13 mar. 2023
- SEVERINO, A. J. **Metodologia do trabalho científico**: diretrizes para o trabalho científico-didático na universidade. 5.ed. São Paulo: Cortez & Moraes, 1980. Disponível em: <https://www.ufrb.edu.br>.

Apêndice A - Universidades e referidos cursos de graduação e pós-graduação

Quadro 1. *Universidades, Cursos de graduação e Pós-graduação*

Região	Universidade	Cursos	
		Graduação	Pós-Graduação
Norte	Universidade Federal do Amazonas - (UFAM)	Biblioteconomia	-
	Universidade Federal de Rondônia - (UNIR)	Biblioteconomia	-
	Universidade Federal do Pará - (UFPA)	Arquivologia, Biblioteconomia	Mestrado em Ciência da Informação
Nordeste	Universidade Federal de Alagoas - (UFAL)	Biblioteconomia	Mestrado em Ciência da Informação
	Universidade Federal da Bahia - (UFBA)	Arquivologia, Biblioteconomia e documentação	Mestrado e doutorado em Ciência da Informação
	Universidade Federal do Cariri - (UFCA)	Biblioteconomia	Mestrado Profissional em Biblioteconomia
	Universidade Federal do Ceará - (UFC)	Biblioteconomia	Mestrado em Ciência da Informação
	Universidade Federal do Maranhão - (UFMA)	Biblioteconomia	-
	Universidade Estadual da Paraíba - (UEPB)	Arquivologia	-
	Universidade Federal da Paraíba - (UFPB)	Arquivologia, Biblioteconomia	Mestrado e doutorado em Ciência da Informação
	Universidade Federal de Pernambuco - (UFPE)	Biblioteconomia	Mestrado e Doutorado em Ciência da Informação
	Universidade Estadual do Piauí - (UESPI)	Biblioteconomia	-
	Universidade Federal do Rio Grande do Norte - (UFRN)	Biblioteconomia	Gestão da Informação e do conhecimento - PPGIC
	Universidade Federal de Sergipe - (UFS)	Biblioteconomia e documentação	Mestrado Profissional em Gestão da Informação e do conhecimento

Fonte: Autores.

Organizar para quem? Uma reflexão sobre a mediação da informação

Organizing for whom? A reflexion on information mediation

Giulia Crippa (1), Lucia Sardo (2)

(1) Università di Bologna, Via degli Ariani 1 – 48121 Ravenna (RA) Itália, giulia.crippa2@unibo.it,

(2) AIB – grupo CILW, Viale di Castro Pretorio 105 – 00185 Roma, sardo.lucia@gmail.com

Resumo

Objetivo: Propor uma reflexão sobre a situação da mediação da informação nos instrumentos de organização do conhecimento, com enfoque nos usuários e na marginalidade dos instrumentos de organização do conhecimento utilizados pelas bibliotecas em comparação com os motores de busca. **Metodologia:** análise bibliográfica e utilização de estudos de caso, tanto no campo da biblioteconomia como no da sociologia. **Resultados:** os instrumentos disponíveis parecem ser incapazes de fornecer respostas eficazes às necessidades de informação dos usuários. De fato, são ferramentas altamente organizadas, mas o crescimento exponencial dos dados disponíveis e o monopólio da pesquisa de informação online por parte de empresas privadas cria um curto-circuito informacional frustrante para os usuários. **Conclusões:** É proposta uma reflexão relacionada com uma melhor utilização das tecnologias disponíveis, que considera uma maior e flexibilidade e interoperabilidade dos instrumentos, um aumento dos estudos sobre *users interaction* e sobre suas necessidades reais de informação, um maior enfoque na participação ativa das comunidades alvo, bem como uma abertura interdisciplinar.

Palavras-chave: Organização do conhecimento. Usuários. Mediação da informação.

Abstract

Aim: To propose a reflection on the situation of information mediation in knowledge organisation tools, with a focus on users and the marginality of knowledge organisation tools adopted by libraries, compared to search engines. **Methodology:** literature analysis and use of case studies, both in library science and sociology fields. **Results:** the available tools seem to be unable to provide effective answers to users' information needs. In fact, they are highly organised tools, but the exponential growth of available data and the monopoly of online information research by private companies creates a frustrating information short-circuit for users. **Conclusions:** A reflection is proposed related to a better use of the available technologies, a greater flexibility and inter-operability of the tools, an increase in surveys on users' interaction and their actual information needs, a greater focus on the active participation of the target communities as well as an interdisciplinary openness.

Keywords: Knowledge organization. Users. Information Mediation.

Resumen

Objetivo: Proponer una reflexión sobre la situación de la mediación de la información en las herramientas de organización del conocimiento, centrándose en los usuarios y en la marginalidad de las herramientas de organización del conocimiento utilizadas por las bibliotecas en comparación con los motores de búsqueda. **Metodología:** revisión bibliográfica y utilización de estudios de casos, tanto en el ámbito de la biblioteconomía como en el de la sociología. **Resultados:** las herramientas disponibles parecen incapaces de dar respuestas eficaces a las necesidades de información de los usuarios. De hecho, son herramientas muy organizadas, pero el crecimiento exponencial de los datos disponibles y el monopolio de la búsqueda de información en línea por parte de empresas privadas crean un frustrante cortocircuito informativo para los usuarios. **Conclusiones:** Se propone una reflexión relacionada con un mejor uso de las tecnologías disponibles, que contemple una mayor flexibilidad e interoperabilidad de las herramientas, un aumento de los estudios sobre la interacción de los usuarios y sobre sus necesidades reales de información, un mayor enfoque en la participación activa de las comunidades destinatarias, así como una apertura interdisciplinar.

Palabras clave: : Organización del conocimiento. Usuarios. Mediación de la información.

1 Introdução

O papel das bibliotecas na organização do conhecimento sofreu mudanças radicais devido ao surgimento do mundo digital e de concorrentes impensáveis em um mundo analógico. No início dos anos 2000, as expectativas sobre o potencial do mundo digital, seu papel e o da WWW eram, em geral, bastante otimistas, mesmo no que diz respeito

às bibliotecas. Muitos estudiosos, entre eles Tim Berners Lee, considerado o criador da Internet, previam que as ferramentas digitais e a Rede permitiriam uma verdadeira revolução "democrática" no acesso à informação e ao conhecimento. É assim que Berners-Lee (2020) se lembra dos ideais no nascimento da Rede, destacando, porém disso, a assimetria com a realidade atual:

The web was always meant to be a platform for creativity, collaboration, and free invention — but that's not what we are seeing today. Today, business transformation is hampered by different parts of one's life being managed by different silos, each of which looks after one vertical slice of life, but where the users and teams can't get the insight from connecting that data. Meanwhile, that data is exploited by the silo in question, leading to increasing, very reasonable, public skepticism about how personal data is being misused. That in turn has led to increasingly complex data regulations.

Pouco mais de vinte anos depois, na verdade, o otimismo parece ter permanecido apenas no mercado das idéias neoliberais, enquanto a maioria dos estudiosos se confronta com uma situação em que a perspectiva é qualquer coisa menos favorável a tal revolução. [1]

Sempre Berners Lee (2019)

I broadly see three sources of dysfunction affecting today's web:

- Deliberate, malicious intent, such as state-sponsored hacking and attacks, criminal behaviour, and online harassment.
- System design that creates perverse incentives where user value is sacrificed, such as ad-based revenue models that commercially reward clickbait and the viral spread of misinformation.
- Unintended negative consequences of benevolent design, such as the outraged and polarised tone and quality of online discourse

Entre os setores envolvidos nas previsões cor-de-rosa estava, como mencionado acima, também o das bibliotecas, que, afetadas pelas mudanças provocadas pelas infraestruturas digitais, se viu progressivamente mais e mais à margem do universo da informação, tendo que se adaptar a essas infraestruturas e perdendo cada vez mais sua capacidade de assumir a liderança ao propor mudanças nas estruturas de organização da informação. Como um exemplo entre muitos, o surgimento de gigantes de busca como o Google mudou radicalmente não apenas a forma de busca, mas também as expectativas dos usuários (amplamente entendidos como aqueles com necessidades de informação de natureza e complexidade diferentes e para diferentes propósitos).

Isto, entretanto, não levou a uma profunda reflexão sobre o papel, funções, potencial e reformulação das bibliotecas e seus sistemas de organização do conhecimento, exceto com declarações gerais sobre sua importância, a necessidade de ter instituições para a preservação (e, espera-se, consequente uso) dos recursos documentais e a importância de ter espaços genericamente 'qualitativos'.

2 Transformações sociotecnológicas entre Modernidade e Contemporaneidade

Consideremos, por um momento, a ideia de 'princípio' e 'razão' através dos quais as estruturas de informação da Modernidade se desenvolvem idealmente, com base em algumas proposições sociológicas. Marcuse (1999) vê a tecnologia como um conjunto de dispositivos de controle. Na medida em que é um modo de produção, o dispositivo de dominação é guiado pelas diretrizes de eficiência, que o autor observa que foram delineadas no século XVIII. Antes, a tecnologia era o resultado do que ele chama de racionalidade individual, cuja existência é sustentada por valores éticos que, mesmo quando em conflito com autoridades externas, devem ser preservados. O desenvolvimento da vida social e pessoal entre os séculos XVI e XVIII propõe uma nova racionalidade, baseada na liberdade de pensamento como agente de transformação da realidade. É um desenho social que se opõe à liberdade de pensamento às restrições milenares impostas pelo pensamento teológico teocêntrico.

Os princípios do individualismo moderno são baseados no reconhecimento do pensamento autônomo que, por sua vez, responde às necessidades sociais - a não ser confundido, portanto, com o interesse individual. O indivíduo é livre em sua crítica da realidade e, através desta crítica, encontra as ferramentas para sua transformação racional. A conduta social é medida pelo trabalho, onde a satisfação individual é calculada quando seus produtos se tornam necessidades sociais.

Por outro lado, a tecnologia permitiu primeiro o desenvolvimento da mecanização e, mais recentemente, da digitalização. São processos que, pouco a pouco, tendem a suplantam a ideia do "sujeito econômico livre" (MARCUSE, 1999, p. 76) com o que são, hoje, os mega conglomerados multinacionais, cujo princípio de eficiência competitiva está centrado no lucro e que, inseridos nos processos de tecnologização, concentram o poder econômico em verdadeiros impérios financeiros e industriais. A tecnologia favorece o poder das empresas que criam redes transnacionais cada vez mais padronizadas e sem desperdício. Em uma palavra: eficientes.

Em uma visão reconhecidamente idealista demais, Marcuse afirma que a racionalidade individual era o compromisso entre as necessidades individuais e o bem-estar comum. A transição para uma nova racionalidade, que chama de tecnológica, é caracterizada por: "[...] rentabilidade e competitividade [...] verdadeiros determinantes da inovação tecnológica e do crescimento da produtividade". (CASTELLS 2002, p. 136).

Para autores como Castells (2009), o poder tecnológico estabelece padrões e atitudes que predispoem os atores envolvidos a aceitar as regras do sistema. As percepções do conhecimento enquanto qualidade individual devem tornar-se padronizadas, enquanto o desempenho é medido por disposições que estabelecem tarefas e funções para cada um. Mais uma vez, podemos dizer que

[...] a empresa tende a reduzir custos, aumentando a exploração de seus funcionários. Assim, em geral, as 'máquinas de informação' desempenham as mesmas funções que as máquinas da grande fábrica, mesmo sendo uma ferramenta utilizada para a produção de valor. Na realidade, os trabalhadores da economia do conhecimento correspondem apenas em pequena escala à imagem que lhes é frequentemente atribuída de "super técnicos" criativos capazes de produzir conhecimentos sempre novos. Na maioria das vezes, eles são meros apêndices das novas máquinas de TICs, realizando um trabalho repetitivo vinculado a regras específicas. (CODELUPPI 2012, p. 27).

O indivíduo eficiente é, portanto, aquele que satisfaz as exigências do aparelho, enquanto sua liberdade é concebida dentro das seleções dos meios disponíveis para atingir os objetivos do sistema, que não coincidem necessariamente com os do indivíduo. O progresso do indivíduo corresponde a sua eficiência em termos de desempenho para o sistema. A eficiência aparece assim como um elemento em contradição com a racionalidade individual que, ao racionalizar o mundo, fez dela um poder ao qual os indivíduos devem se adaptar. De fato, para que o sistema funcione, seus atores devem se submeter aos ritmos e linguagens do sistema. O ideal da máquina automática é assim estabelecido, para cujo funcionamento são necessários treinamento, prática e uma atitude factual: categorias de conhecimento definidas como "técnicas", que tendem a excluir interpretações.

É uma condição que não permite uma saída individual de um sistema que tende à padronização: obedecer a instruções produz os resultados desejados, o sucesso é o resultado da adaptação. Esta condição vincula o progresso científico ao desenvolvimento tecnológico, já que a competitividade tecnológica é um elemento chave da concorrência. O comportamento individual se adapta à racionalidade do sistema, que podemos agora afirmar ser algo com um conteúdo social definido.

Este sistema de dominação se estende da ordem tecnológica à ordem social na aplicação do treinamento como método privilegiado de aprendizagem, bem como na estrutura dos tempos de vida, estabelecendo parâmetros de desempenho no trabalho, no repouso e no lazer.

Nesta apresentação, abordamos principalmente o papel das bibliotecas acadêmicas e de pesquisa,

instituições que deveriam lidar com a organização do conhecimento como uma atividade destinada a fomentar a comunicação nesta comunidade supostamente global (também se poderia usar o termo comunicação científica, mas num sentido amplo e não limitado às 'ciências duras'), a difusão do acesso aberto e as práticas de disseminação do que é produzido na esfera acadêmica para partes mais amplas da sociedade.

Muitas propostas informativas são estruturadas para dar conta de uma possível solução para um sistema supostamente eficiente desenvolvido por indivíduos treinados, ou seja, aqueles que aprenderam práticas que os colocam em razoável submissão, pois garantem a coexistência pacífica com a ordem e a instituição da tecnologia, cujo objetivo é desenvolver maior eficiência, como descrito acima.

3 Mudanças estruturais das buscas dos usuários

Para que servem hoje as bibliotecas? Uma ideia compartilhada por muitos profissionais e uma grande parte da comunidade acadêmica é a da biblioteca como um lugar para a produção da cidadania, especialmente se considerarmos as bibliotecas públicas. Esta é a direção tomada pelos estudos e pesquisas da nova biblioteconomia proposta por Lankes (2011) e da biblioteconomia social e crítica, que não se concentram mais no gerenciamento de documentos ou processos internos das bibliotecas, mas nos usuários e comunidades de referência.

Ao lado disso, e muitas vezes interligada, está a linha de pensamento que levou ao desenvolvimento da informação científica e tecnológica dentro dos processos de produção e do mercado capitalista. Ao colocar o controle das informações mais estratégicas e valiosas nas mãos de grupos relativamente pequenos e altamente especializados, a mobilidade social é impedida, ou pelo menos impedida, na medida em que o acesso ao conhecimento é reduzido. É evidente que as duas perspectivas foram, aqui, esquematizadas ao extremo e que não existe uma fronteira rígida entre elas.

Entre os séculos XVII e XIX, a biblioteconomia não é nem científica, nem axiomática, nem experimental, na medida em que sua tarefa não é a investigação e o conhecimento de uma realidade externa. A biblioteconomia, entendida como uma disciplina que trata do desenvolvimento de coleções e sua catalogação e classificação, apresenta as ambiguidades de um conhecimento heurístico, apresentando tanto formas intuitivas e elementares como manifestações de alta complexidade e articulação intelectual. Uma vez eliminado o equívoco de uma ideia única e integrada de biblioteconomia - dentro da qual os mesmos

protocolos regulam todas as operações, procedimentos e atividades da biblioteca - pode ser interessante estudar os critérios para avaliar as relações de competência, funcionalidade e eficiência das estruturas que medeiam entre as coleções e os usuários.

A biblioteca moderna, de fato, está se reformulando devido a mudanças nas estruturas do conhecimento, que são determinadas simultaneamente pelo desenvolvimento da ciência e tecnologia do livro impresso e, portanto, na lógica da organização da informação, principalmente através da produção bibliográfica e de catálogos. A organização bibliográfica, que está destinada a se tornar, através de catálogos, uma estrutura física em bibliotecas, é, acreditamos, baseada em uma nova disposição das partes envolvidas na discussão do conhecimento.

Um aspecto crítico a não ser subestimado na organização do conhecimento está relacionado ao conceito de cânon no sentido bibliográfico e de grupo social. Se mesmo um destes fatores estiver faltando, ou se for vago e difícil de identificar e circunscrever, torna-se impossível de organizar, pois há uma falta de parâmetros que permitam um compartilhamento sistêmico.

Quando pensamos na biblioteca, nos referimos a um local onde o conhecimento é ordenado, classificado e armazenado em vários registros, desde livros a periódicos até meios audiovisuais, em forma analógica ou digital. É um ambiente no qual existe uma hierarquia tanto em termos de classificação quanto em termos da relação bibliotecário/usuário. Na biblioteca física, por exemplo, o serviço de referência, a interface entre a biblioteca e o usuário, tem uma estrutura "rígida", na qual o bibliotecário tem mais conhecimento do que o usuário, a fim de atender às suas necessidades. Todas as ferramentas necessárias para obter as informações procuradas enfatizam o caráter de "busca da verdade" do qual a própria biblioteca se torna, em teoria, garante. Em uma biblioteca, encontramos teoricamente os documentos publicados que representam a qualificação das diversas perspectivas "canônicas" do conhecimento. Não apenas reconhecemos a autoridade do conhecimento em indivíduos, mas também em livros, instituições, ferramentas organizacionais e de pesquisa. Para usuários 'ingênuos', todas as publicações podem ter autoridade, mas na verdade foi estabelecida uma hierarquia que permitiu à comunidade científica e à instituição social da ciência adquirir um papel, tornando sua produção confiável, em termos de informação, na medida em que foram 'certificadas' pelos sistemas de produção e controle da própria comunidade científica. Trata-se de uma interação intelectual complexa, que inclui fatores relacionados à vontade, necessidades e expectativas e que, ao lado

de oportunidades de aprendizagem e habilidades, produz efeitos sobre a visão do mundo.

Embora a maioria da produção especializada não chegue ao público em geral, tornando-se o domínio da informação científica, existe uma indústria cultural e de conhecimento dedicada à produção de livros para popularização e entretenimento, destinados a uma ampla gama de leitores e que devem ser encontrados em bibliotecas públicas. Também neste campo, nos tempos modernos, foram produzidos e procurados sinais para garantir a qualidade da informação, identificáveis nas garantias oferecidas por certas editoras, nas publicações apresentadas por organizações profissionais ou institucionais, agências oficiais ou impressores estatais, e nos prêmios recebidos. Tudo isso, teoricamente, garantia a qualidade das informações oferecidas.

Este modelo entrou em crise quando o acesso, agora potencialmente ilimitado à informação, não é mais mediado, mas é teoricamente possível a qualquer hora e em qualquer lugar, desde que o acesso à web esteja disponível. O grande equívoco foi pensar que isto democratizaria e impulsionaria o surgimento de diferentes culturas e pontos de vista, enquanto o que aconteceu foi a progressiva concentração da mediatização de informação nas mãos de algumas empresas de Internet e a criação de câmaras de eco que criam curto circuitos cognitivos. Isto acontece porque só se tem confirmações das próprias crenças, pois os algoritmos dos mecanismos de busca e redes sociais (cujo objetivo é tornar o usuário fiel e gastar o máximo de tempo possível on-line) tendem a apresentar coisas semelhantes àquelas já vistas e isto leva a uma visão distorcida da realidade. Estes mecanismos também limitam as possíveis fontes de informação, pois se tornam particularmente difíceis de descobrir e alcançar se não forem apresentados entre os primeiros resultados das páginas de busca.

A confiabilidade da biblioteca se baseia na ideia de que os documentos passaram por processos de elaboração dentro dos campos em que foram produzidos, dando-lhes uma designação bibliográfica. Isto não significa que a biblioteca contenha verdades absolutas, mas sim que ela é o lugar onde, com o tempo, as informações foram consolidadas a partir de sistemas que procuram fornecer bases socialmente compartilhadas para a formação de opiniões estruturadas e, com elas, operacionalizar escolhas que geram conhecimentos. Estes documentos também eram submetidos a práticas de indexação que os incorporavam aos sistemas de organização do conhecimento que teoricamente permitiam a contextualização e facilitavam o desenvolvimento de novas pesquisas.

As garantias são ideais e nem todas as bibliotecas as respeitam. Entretanto, de acordo com Wilson (1983),

não há dúvida de que na base de sua existência está a possibilidade de acesso a materiais selecionados e organizados para estimular a transformação de opiniões instáveis em escolhas canônicas de informação geradoras de conhecimento, na medida em que tenham sido institucionalizadas.

O papel das bibliotecas parece ter sido o de oferecer diferentes alternativas para diferentes assuntos, por meio do desenvolvimento de novos conhecimentos e da busca de alternativas para as questões de informação colocadas pelos usuários, ou seja, pode ser necessário que um usuário pesquise vários textos para encontrar o que está procurando, e as bibliotecas se desenvolveram nessa direção.

Pode-se dizer que embora a biblioteca não seja um lugar de verdade absoluta, nela trabalham profissionais que, dentro de um sistema mais amplo de produção de conhecimento, estabeleceram metodologias que lhes permitem selecionar e fornecer informações que podem ser consideradas confiáveis, pois é assim estabelecida pelos processos estabelecidos nos campos que a produziram. (GORMAN, 2004).

Podem parecer irrelevantes reiterar estes aspectos que caracterizam a constituição das bibliotecas na modernidade, mas estas são questões atuais, pois estão passando por uma fase de descrédito na medida em que são colocadas em uma falsa competição com ofertas múltiplas e aparentemente muito mais amplas, através da rede global. Nesta rede "rica", parece-nos, existe uma armadilha muito perigosa: a questão da seleção e confiabilidade das fontes está em aberto e ainda sem soluções satisfatórias. Pensemos, de fato, em como funciona a lógica da informação estruturada quando se trata dos problemas de comunicação do ponto de vista da eficiência, mas de forma indiferente à qualidade das mensagens, cuidando da perspectiva mais estritamente informativa e não-informativa. A grande preocupação dos sistemas informatizados estava centrada na velocidade e confiabilidade dos sinais, mas não necessariamente na seleção e qualidade do que esses sinais transportam. O importante era conceber sistemas em que o que era transmitido correspondesse ao que era recebido (a base da informação é a teoria matemática de Shannon (1948)), mas qualquer informação transmitida, nesta lógica, tem o mesmo valor, independentemente de sua origem.

A preocupação dominante desta perspectiva é que a informação recebida é a informação necessária, e não que ela corresponda ao resultado produzido na esfera do conhecimento. Além disso, pode-se argumentar que existem situações em que nem mesmo o conteúdo transmitido é importante, mas a quantidade de tráfego e de compartilhamento que o que é apresentado on-line produz. (Cfr. EICHORN 2022).

Da mesma forma, pelo menos até meados do século 19, período em que o conceito de referência nasceu e se desenvolveu, o aspecto da competência informativa do usuário, que pode ou não ter a capacidade de buscar informações adequadas, não foi levado em conta. Para resolver este tipo de problemas é necessário recorrer a especialistas, e as habilidades necessárias são, aliás, precisamente as dos bibliotecários como bibliógrafos e prestadores de serviços de referência.

Cada informação adquire este valor dentro de um contexto de busca individual no qual os filtros disponíveis são desconhecidos dos usuários. A distinção entre verdadeiro e falso, representada na história pela própria biblioteca e suas divisões internas, perde seus limites. Com o advento da World Wide Web, entretanto, esta relação hierárquica tende a se dissolver, pois o que é validado e classificado pela instituição da biblioteca, à qual Latour (2004) atribui um papel fundamental na rede de produção de conhecimento, é equiparado a informações efêmeras que, através dos mecanismos de busca, são encontradas com o mesmo peso que as informações 'certificadas'. Toda informação adquire assim valor dentro de um contexto de busca individual, no qual os filtros disponíveis são muitas vezes desconhecidos. A distinção entre verdadeiro e falso perde seus limites, que eram representados, na história, pelas próprias bibliotecas e suas divisões internas (RADFORD, 1998): nos espaços da rede mundial, o princípio de autoridade e divisão dos campos do conhecimento transcende os limites dos campos disciplinares e o conhecimento torna-se, para usar a metáfora de Bauman (2002), "líquido".

A ideia de conhecimento, como vimos, é marcada pelo processo de construção da autoridade de grupos e instituições, tornando-se um dos elementos capazes de articular a própria ideia de progresso. O modelo de comunicação observado neste horizonte de culturalização positiva não é diferente do delineado nas bibliotecas tradicionais, e pode ser esquematizado em uma sequência linear e unidirecional que vai desde o mito-condutor, passando pelo meio, até o receptor. Em um dispositivo deste tipo, o receptor só é considerado no momento em que ocorre o "transbordo" de informações. A seleção e o controle dos significados são a realização exclusiva do remetente, enquanto o meio é o foco da comunicação. Neste sentido, o termo "disseminação" representa bem o ato comunicativo no qual o usuário é um receptor passivo. É um modelo que oferece um 'imperativo moral': a função da informação deve ser transferida, enquanto os processos desenvolvidos para este fim visam 'administrar' a informação a ser assimilada. O princípio de autoridade pertence ao remetente, que parece estar dotado do legítimo conhecimento a ser distribuído ao público, o qual, por

sua vez, ainda não é visto da perspectiva da cultura como um "sistema concreto de significação" (WILLIAMS, 2000), ou seja, com atenção aos aspectos da vida social que contribuem para a construção de significados.

As concepções modernas do conhecimento e dos atores que o produzem e lhe dão acesso eram, portanto, positivas e visavam representar ideais enciclopédicos e universais, visíveis na própria separação dos espaços das bibliotecas, onde os estudiosos eram colocados em oposição ao público em geral. Neste ponto, podemos considerar a mudança trazida pelas tecnologias digitais, que possibilitam a coleta e organização de dados de informação fora dos quadros disciplinares tradicionais, e que permitem o desenvolvimento do conhecimento sem fronteiras claras entre o real e o fantástico. Isto destaca a perda de orientação na apropriação do co-conhecimento que a rede global produz, especialmente quando a informação é procurada sem filtros adicionais, como acontecia na biblioteca. Entretanto, não é possível defender hierarquias pré-existentes, pois estas não são mais capazes de refletir as necessidades atuais de construção do conhecimento. Se as narrativas da modernidade são agora pouco convincentes, notamos com preocupação que o que as substitui está essencialmente ligado às necessidades do mercado, com as contradições que isso implica. As empresas do sistema capitalista, por exemplo, têm acesso extraordinário a informações que, com a ajuda de tecnologias de armazenamento, podem ser recombinações e aplicadas para todos os fins e em todos os contextos como *commodities*.

Por outro lado, isto implica uma grande pressão sobre o trabalho. A *e-economy* não pode funcionar sem trabalhadores capazes de lidar, tanto tecnologicamente como em termos de conteúdo, com este enorme fluxo de informações, organizando-o e transformando-o em conhecimento específico, adaptado ao objetivo e à finalidade do processo de produção. A "força de trabalho intelectual" que lida com a infoesfera deve ser altamente instruída, treinada para tomar iniciativas (CASTELLS 2009). As empresas, grandes ou pequenas, dependem da qualidade e da autonomia da força de trabalho. Os trabalhadores ligados a esta circulação de informações, uma "força de trabalho" altamente especializada, devem ser capazes de reprogramar suas próprias competências e conhecimentos e pensar na base de objetivos em rápida evolução, frequentemente no âmbito de um ambiente empresarial em constante mutação. Esta capacidade de programação, ou autoprogramação, requer um certo tipo de treinamento e o patrimônio de conhecimento e informação acumulado deve ser constantemente expandido e modificado (CASTELLS 2009). Aprender a aprender torna-se uma necessidade, assim como

adquirir a capacidade de transformar a informação obtida no processo de aprendizagem em conhecimento específico: são duas necessidades que afetam todas as áreas do conhecimento. Em um sistema no qual o conhecimento se tornou hiper-estruturado por uma especialização excessiva, é necessário estabelecer até que ponto o aprendizado é necessário para a manutenção de um sistema vinculado à lógica da informação como mercadoria e capital, entregando-se a uma racionalidade puramente tecnológica. É um quadro que atrai trabalhadores como "executores" de ordens, eficientes e atualizados, mas ainda meros executores de políticas neoliberais.

Por outro lado, a pesquisa científica crítica reacende o debate sobre as instituições relacionadas à informação; os responsáveis por uma rede de informação para a pesquisa, anteriormente bibliotecários, não conseguiram reivindicar para si mesmos um papel influente não apenas na discussão sobre como organizar o conhecimento, mas também sobre o porquê, para quem e para quê, tentando mudar o foco de saber como fazer para saber como dizer (CERTEAU 1996).

Nesta discussão, há alguns elementos que podemos destacar:

- 1) A ideia de ampliar o acesso à informação, já presente, por exemplo, em Naudé (NAUDÉ 1627), choca-se com a alta seletividade dos usuários, pois de fato apenas uma porcentagem muito pequena deles tem acesso à informação estratégica gerenciada por sistemas ligados à pesquisa científico-tecnológica. E esta minoria não tem necessariamente as habilidades necessárias para poder realizar pesquisas eficazes.
- 2) Ao recorrer à informação científica e tecnológica, um nível muito restrito e especializado de acessibilidade é sistematizado em primeiro lugar.
- 3) Os processos de informação hoje podem ser claramente inscritos nas transformações da racionalidade individual para a racionalidade tecnológica. Um exemplo disso é a eliminação ou, pelo menos, a limitação severa dos procedimentos heurísticos em favor da aplicação decisiva da métrica.

3.1 Case Study: o Discovery Tool da Universidade de Bolonha.

As ferramentas disponibilizadas pelas bibliotecas (e, portanto, como elas organizam o conhecimento) hoje são certamente complexas e refinadas, mas é provável que elas não sejam funcionais na contemporaneidade ou sejam apenas funcionais para uma parte cada vez mais marginal mesmo de usuários fortes em potencial. Um estudo realizado sobre o uso

da ferramenta de descoberta na Universidade de Bolonha produziu resultados inconclusivos. Constatou-se que

Mais de 92% das buscas começam com uma busca base e menos de 11% utilizam operadores booleanos, caracteres especiais ou busca por campos específicos no registro, tais como título ou assunto. Isto significa que o modo usado para 9 de cada 10 buscas é o mesmo que nos motores de busca, um conjunto de palavras digitadas em um único campo. Esta é provavelmente a única maneira possível para os usuários, pois eles não possuem as habilidades de informação necessárias para estruturar uma consulta mais detalhada enquanto, para usuários mais experientes, pode-se deduzir que eles não sentem a necessidade de pesquisar usando campos específicos ou operadores booleanos, provavelmente porque eles já obtêm respostas adequadas para suas necessidades. (QUAQUARELLI 2021, p. 111)

Embora fora do escopo da pesquisa deste trabalho, resultados semelhantes podem ser deduzidos das palavras de busca dos usuários de sistemas de bibliotecas públicas. Os padrões de busca no catálogo veem uma prevalência quase absoluta de buscas simples e com poucas palavras destinadas a identificar a presença de recursos conhecidos, a fim de obter acesso a eles [2]. Nas mesmas áreas, já que as respostas dadas ao usuário não são contextualizadas dentro da ferramenta como um todo (um usuário de um catálogo físico, seja ele nominal ou classificado, sabe onde está com relação ao catálogo como um todo), é fácil criar situações de alienação e frustração com relação aos resultados obtidos, que não parecem relevantes ou pelo menos marginais com relação à busca realizada.

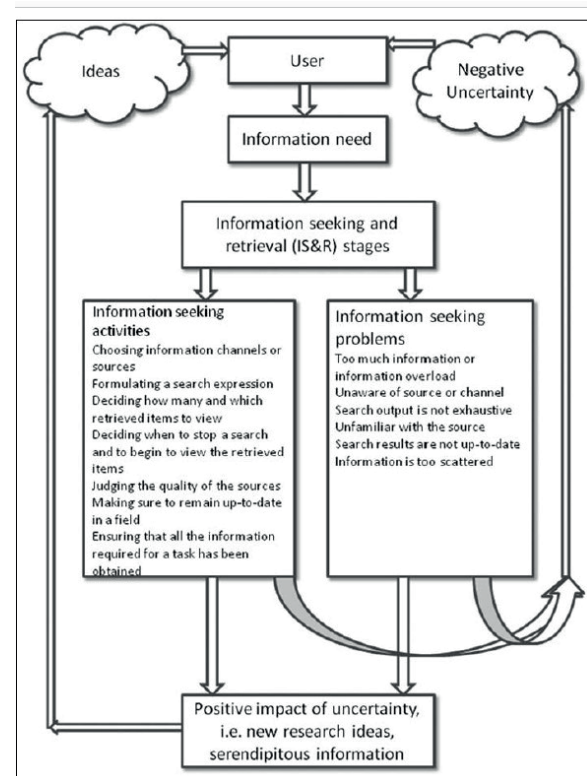
Das chamadas funções de usuário propostas nos Princípios de Catalogação da IFLA (IFLA 2016) e nos modelos conceituais da IFLA (IFLA 2017), o que realmente parece ser relevante é "obter" (SARDO 2022). Como corolário do acima, deve-se lembrar também da crescente "impaciência" dos usuários, pela qual se não se obtém o que se está procurando em um tempo muito curto, a busca é interrompida.

Neste contexto, a fragmentação das interfaces de busca, tanto do ponto de vista de sua diversidade como do ponto de vista da necessidade dos usuários de consultar diferentes catálogos, bancos de dados, bibliotecas digitais (em resumo, qualquer coisa que forneça acesso a metadados estruturados que se liguem a recursos possuídos ou acessíveis - gratuitamente ou mediante pagamento), e das formas de codificação dos dados considerados de interesse não facilita a busca e recuperação de informações, dados, recursos. Impõe, por outro lado, buscas frustrantes em diferentes sistemas, sempre deixando a dúvida de que algo fundamental foi perdido, assim como o sentimento de *overload* de informação, geralmente considerado um dos aspectos mais problemáticos nas buscas on-line, juntamente com a

confiabilidade da informação. Isto pode levar a uma situação de frustração e, poderíamos dizer, FOMO (*fear of missing out*) informativo, por um lado, e, por outro lado, de abandono da busca após algumas tentativas, limitando-se ao que é facilmente recuperado, mas sem a certeza de que é realmente o que interessa.

Neste sentido, o modelo de busca e recuperação de informações proposto por Chowdhury (2011, p. 69) confirma o acima exposto. Os problemas de busca de informações destacados acima podem facilmente levar ao que é chamado de incerteza negativa.

Figura 1. O modelo da incerteza na busca e na recuperação de informações



Fonte: Chowdhury (2011, p. 69).

O advento e predominância de corporações com interesses econômicos no mundo digital criou uma lacuna entre as bibliotecas e aqueles que realmente "possuem" os recursos e que inevitavelmente impõem modelos de organização do conhecimento, codificação de dados e práticas de pesquisa não necessariamente baseadas nas formas tradicionalmente implementadas pelas bibliotecas: os hábitos de busca não são baseados em instrumentos do tipo bibliográfico-catalográfico, mas em outros tipos de sistemas aparentemente amigáveis, mas tão traiçoeiros, na medida em que não são administrados por instituições públicas, mas por instituições cujos objetivos não são facilitar a pesquisa, mas sim econômicos e de certa forma manipuladores.

4 Conclusões

O acima exposto se propõe a ser uma primeira reflexão sobre as crescentes críticas colocadas pelo desenvolvimento tecnológico, cuja direção parece ser oposta àquela imaginada nos primeiros anos de desenvolvimento da web. Os desafios que os sistemas de organização do conhecimento têm que enfrentar partem da necessidade de sair de uma situação de marginalidade, principalmente devido a fatores econômicos e de visibilidade. Como a complexidade da busca e da recuperação de informações aumentou significativamente no mundo digital, um dos principais desafios é de “understand users’ information seeking behaviour and develop a new information service or improve existing services” (CHOWDHURY 2011, p. 76). Na sociedade da informação (ou desinformação) do novo milênio, o problema não é apenas responder às necessidades de informação, mas também fornecer respostas qualitativas e contextuais (um recurso de qualidade que não é corretamente contextualizado é completamente idêntico a um recurso sem qualidade). Finalmente, qualquer tentativa de propor formas alternativas de organizar o conhecimento tem que contar com o fato de que ele utiliza uma tecnologia e ferramentas com fortes conotações das quais, no momento, parece muito difícil prescindir.

A complexidade das ferramentas implementadas não deve ser questionada, mas deve ser avaliada sua eficácia em responder às exigências dos usuários incapazes de interagir com essas ferramentas usando-as em todo o seu potencial.

O regime de verdade da Modernidade, refletido no foco exclusivo do conhecimento científico, continha um modelo ingênuo daquela parte da sociedade que expressava a cultura hegemônica. Além disso, foi implantado como uma prescrição social, sem qualquer interesse ou negociação sobre sua validade ou aceitabilidade. A imposição completamente irrefletida destas premissas apenas polarizou o debate em torno da distração do valor verdadeiro das proposições científicas e polêmicas sobre a suposta irracionalidade das classes populares e a corrupção dos cientistas e das instituições reguladoras. Um processo de aprendizagem reflexiva teria reconhecido as condições nas quais as conclusões científicas se baseiam, destacado as questões sociais envolvidas e as teria examinado com o benefício, entre outras coisas, das diferentes formas de conhecimento detidas por outras pessoas que não os cientistas. Este processo de aprendizagem reflexiva teria significado a negociação entre diferentes epistemologias e formas culturais, entre diferentes discursos, e como tal teria envolvido o desenvolvimento das identidades sociais ou morais dos atores envolvidos (BECK 1992). O tema da modernização reflexiva corresponde a um processo encontrado nos conflitos entre os regimes

contemporâneos da verdade e aprofunda a crise de legitimação das instituições modernas. O novo regime de verdade no qual nos encontramos implica o abandono em massa dos modos de pensamento científico-instrumental, ao qual a modernidade confere grande poder. A modernização reflexiva enfrenta e tenta acomodar as dificuldades humanas, como se vê na tentativa incessante, mas sempre aberta, de renegociar narrativas coerentes. As pessoas lutam para conciliar pontos de vista conflitantes, promovidos em redes diferentes e sobrepostas (GEHL R.W.; LAWSON, S.T. 2022). Sua ambivalência em relação às afirmações científicas como fonte reflete esta situação social com múltiplas narrativas incoerentes.

Quando lidamos com sistemas de organização do conhecimento sem nos preocuparmos em como a Rede está mudando substancialmente os hábitos e padrões de busca dos usuários, mudando o foco das bibliotecas para as empresas que fornecem acesso aos recursos, estamos nos desviando de uma questão-chave, que é: como garantir que o que é feito pelas bibliotecas e outras instituições similares não seja marginalizado para as comunidades que elas servem.

Linhas de pesquisa e tópicos a serem desenvolvidos podem ser hipotetizados, tais como:

- * melhor utilização das tecnologias (Web semântica, ontologias, classificações facetadas, em perspectivas de interoperabilidade e compartilhamento)
- * melhoria da inclusão e flexibilidade das ferramentas de catalogação e mediação bibliográfica (por exemplo, do ponto de vista da interculturalidade, do multilinguismo, etc.)
- * aumento das pesquisas sobre os usuários e suas necessidades
- * criação de bibliotecas como conversas, ou seja, incluindo, tanto quanto possível, comunidades de referência nos processos de tomada de decisão
- * abertura à interdisciplinaridade com campos como sociologia, antropologia, estudos de *users interaction* e estudos de usabilidade
- * fazer o melhor uso das competências tradicionais relacionadas à seleção e acesso a recursos (resolver ou, pelo menos, tentar frear o *information overload*).

Notas

- [1] Como exemplos, podem ser propostos os trabalhos do coletivo italiano Ippolita, de Morozov, Lovink e o Onlife Manifesto.

[2] Isto é baseado na análise das cordas de pesquisa dos usuários, disponibilizada aos autores pelos gerentes do software do catálogo

Referências

BAUMAN, Zygmunt. **Modernità líquida**. Roma-Bari: Laterza, 2002.

BECK, Ulrich. **Risk Society: towards a new Modernity**. London: Sage, 1992.

BERNERS-LEE, Tim. 30 years on, what's next #ForTheWeb? Web Foundation, March 12, 2019. Disponível em: <https://webfoundation.org/2019/03/web-birthday-30/>. Acesso em: 24 mar. 2023.

BERNERS-LEE, Tim. A New Era of Innovation and Trust in Data Nov. 9, 2020, Disponível em: https://medium.com/@timberners_lee/a-new-era-of-innovation-and-trust-in-data-781f3b52a1e4. Acesso em: 24 mar. 2023.

CASTELLS, Manuel. **A sociedade em rede**. São Paulo: Paz e Terra, 2002.

CASTELLS, Manuel. **Comunicazione e potere**. Milano: Università Bocconi, 2009.

CERTEAU, Michel de. **A invenção do cotidiano**. Vol. 1. Artes de fazer. Rio de Janeiro: Vozes, 2000.

CHOWDHURY G.G.; CHOWDHURY S., **Information Users and Usability in the Digital Age**. London: Facet Publishing, 2011.

CODELUPPI, Vanni. **Ipermondo: dieci chiavi per capire il presente**. Bari: Laterza, 2012

EICHORN, Kate. **Content**. Boston, MIT Press, 2022.

GEHL R.W.; LAWSON, S.T. **Social Engineering**. How Crowdmasters, Phreaks, Hackers, and Trolls Created a New Form of Manipulative Communication. Cambridge, Mass.; London: The MIT Press, 2022

GORMAN, Michael. **La biblioteca come valore: tecnologia, tradizione e innovazione nell'evoluzione di un servizio**. Udine: Forum, 2004.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS - IFLA. **Statement of International Cataloguing Principles**

(ICP). Netherlands: IFLA 2016. Disponível em: https://repository.ifla.org/bitstream/123456789/80/1/icp_2016-en.pdf. Acesso em: 24 mar. 2023.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS - IFLA. **IFLA library reference model: a conceptual model for bibliographic information**. Netherlands: IFLA, 2017. Disponível em: https://repository.ifla.org/bitstream/123456789/40/1/ifla-lrm-august-2017_rev201712.pdf. Acesso em: 24 mar. 2023.

LANKES, David, **The Atlas of new librarianship**. Boston, MIT press, 2011. Disponível em: <https://davidlankes.org/new-librarianship/the-atlas-of-new-librarianship-online/>

LATOURET, Bruno. Redes que a razão desconhece: laboratórios, bibliotecas, coleções. In: PARENTE, André (org.). **Tramas da rede: novas dimensões filosóficas, estéticas e políticas da cognição**. Porto Alegre: Sulina, 2004, p. 39-63.

MARCUSE, Herbert. **Tecnologia, guerra e fascismo**. São Paulo: UNESP, 1999.

NAUDÉ Gabriel. **Advis pour dresser une bibliotheque**. Paris: Targa, 1627.

QUAQUARELLI, Maura. **Il discovery tool AlmaStart: un'analisi quantitativa e qualitativa dei dati di ricerca**. Tesi (Laura magistrale in Scienze del libro e del documento). AMS Università di Bologna. Campus di Ravenna, 2021.

RADFORD, Gary P. Flaubert, Foucault, and the Bibliotheque Fantastique. Toward a Postmodern Epistemology for Library Science, **Library Trend**, v. 46, n. 4, p. 616-634, Spring 1998.

SARDO, Lucia. I modelli concettuali e gli utenti. **JLIS.it**, v.13, n. 3, p. 90-102, 2022.

SHANNON, Claude E. A mathematical theory of communication. **The Bell System Technical Journal**, v. 27, n. 3, p. 379-423, July 1948.

WILLIAMS, Raymond. **Cultura**. Rio de Janeiro, Paz e Terra, 2000

WILSON, Patrick. **Second-hand knowledge: an inquiry into cognitive authority**. Westport, Conn.; London: Greenwood press, 1983.

Representação arquivística: percepção dos arquivistas das Instituições Federais de Ensino Superior sobre a neutralidade

Archival representation: perception of archivists at Federal Institutions of Higher Education about neutrality

Jefferson Higino da Silva (1), Luciane Paula Vital (2)

(1) Universidade Federal de Santa Catarina, Florianópolis, jeffersonarquivista@gmail.com, (2) luciane.vital@ufsc.br

Resumo

Este artigo parte do problema que a adoção da posição de neutralidade pode impedir que os profissionais da arquivologia considerem sua própria realidade e reconheçam que suas ações são influenciadas por diferentes fatores. Por meio de uma pesquisa quali-quantitativa e de natureza exploratória e descritiva, com arquivistas de Instituições Federais de Ensino Superior, foi analisada a percepção sobre a neutralidade na classificação e na descrição de documentos, utilizando um questionário misto como coleta de dados. Os resultados mostram que a neutralidade para a maior parte dos arquivistas depende do contexto e que o profissional pode tomar decisões para cumprir objetivos pessoais, políticos ou institucionais. A ética, os princípios arquivísticos, os usuários, os instrumentos e normas, a instituição e a sociedade são pontos que mediam essas atividades, com impactos positivos e negativos, de acordo com a avaliação dos profissionais. Conclui-se que existem diferentes opiniões sobre a neutralidade e que os arquivistas, seja diretamente ou indiretamente, compreendem que o seu fazer é mediado por escolhas e por intervenções políticas, culturais, sociais e econômicas.

Palavras-chave: Neutralidade. Arquivista. Representação arquivística. Organização do conhecimento.

Abstract

This article addresses the issue that adopting a neutral position may prevent professionals in archival science from considering their own reality and recognizing that their actions are influenced by different factors. Through a qualitative-quantitative, exploratory, and descriptive research with archivists from Federal Institutions of Higher Education, their perception of neutrality in document classification and description was analyzed using a mixed questionnaire as data collection method. The results show that neutrality for most archivists depends on context and that professionals may make decisions to fulfill personal, political, or institutional objectives. Ethics, archival principles, users, instruments and norms, institution, and society are points that mediate these activities, with positive and negative impacts according to the evaluation of professionals. It is concluded that there are different opinions about neutrality and archivists, whether directly or indirectly, understand that their work is mediated by choices and political, cultural, social, and economic interventions.

Keywords: Neutrality. Archivist. Archival representation. Knowledge organization.

Resumen

Este artículo aborda el problema de que adoptar una posición de neutralidad puede impedir que los profesionales de la archivística consideren su propia realidad y reconozcan que sus acciones están influenciadas por diferentes factores. A través de una investigación cualicuantitativa y de naturaleza exploratoria y descriptiva con archivistas de Instituciones Federales de Educación Superior, se analizó la percepción sobre la neutralidad en la clasificación y descripción de documentos, utilizando un cuestionario mixto como método de recopilación de datos. Los resultados muestran que la neutralidad para la mayoría de los archivistas depende del contexto y que el profesional puede tomar decisiones para cumplir objetivos personales, políticos o institucionales. La ética, los principios archivísticos, los usuarios, los instrumentos y normas, la institución y la sociedad son puntos que median estas actividades, con impactos positivos y negativos, según la evaluación de los profesionales. Se concluye que existen diferentes opiniones sobre la neutralidad y que los archivistas, ya sea de manera directa o indirecta, comprenden que su labor está mediada por elecciones e intervenciones políticas, culturales, sociales y económicas.

Palabras clave: Neutralidad. Archivista. Representación archivística. Organización del conocimiento.

1 Introdução

O termo "neutro" tem origem na palavra latina "neuter" e significa "nem uma coisa e nem outra". Segundo o Houaiss (2001), "neutro" é definido como algo que não toma partido nem se envolve em questões polêmicas, mantendo-se imparcial e isento. Essa ideia de neutralidade não se restringe à profissão de arquivista e à Arquivologia e está presente em outras áreas do conhecimento como herança do positivismo.

O positivismo, como corrente filosófica, surgiu em um momento histórico em que a Europa estava passando por grandes transformações industriais entre os séculos XIX e XX, o que gerou diversos problemas sociais. Dessa forma, o positivismo apresentou uma forma de solução racional e científica para esses problemas, por meio do uso do método científico para investigar e para compreender a sociedade, buscando uma compreensão objetiva da realidade (REALE; ANTISERI, 2007).

A concepção do arquivista como um sujeito neutro diante dos arquivos tem sido questionada por autores pós-modernos, tendo em vista a complexidade da sociedade contemporânea. A perspectiva pós-moderna argumenta que os documentos e os arquivos estão sujeitos a influências que condicionam sua produção. Isso significa que não são capazes de transmitir uma verdade absoluta, mas apresentar uma visão particular de quem os produziu ou registrou. Assim, esses fatores que condicionam a produção dos documentos e dos arquivos também condicionam o que podemos conhecer por meio deles (NESMITH, 2018).

Cook (2018) questiona a relação entre os arquivos e os indivíduos, levando em consideração a construção social, cultural e política. O autor canadense nos convida a compreender melhor as políticas de memória e as bases que influenciaram as teorias arquivísticas, com o objetivo de ter arquivos que reflitam a complexidade das sociedades que pretendem representar.

Nesse aspecto, destacam-se os estudos da Organização do Conhecimento também estão ligados às dimensões socioculturais, baseados em representações que contextualizam e que respeitam os domínios do conhecimento (GARCÍA GUTIÉRREZ, 2002). Na mediação desses domínios, as questões abordadas ultrapassam o aspecto profissional e concentram-se no tratamento da informação dentro das comunidades discursivas, levando em consideração a garantia cultural presente na representação da informação e do conhecimento.

A Arquivologia tem estabelecido diálogos com a Organização do Conhecimento, buscando meios de ampliar seus instrumentos sem perder de vista seus princípios e seus aspectos orgânicos. Essa aproximação tem sido fundamental para o desenvolvimento de metodologias que auxiliam na organização e no acesso a diferentes contextos (BARROS; SOUSA, 2019).

Observa-se que os arquivos refletem, cada vez mais, a inclusão de indivíduos, de histórias e de narrativas documentais. Os agentes que integram esses espaços são protagonistas frente à realização das representações dos documentos. A esses profissionais,

[...] confere uma responsabilidade ética, que prioriza o atendimento ao usuário, e tal qual as relações sociais, ao representar precisa considerar o contexto e respeitar a diversidade e as diferenças (MARTINS; CORTÊS, 2019, p. 159-160).

Nesse cenário, este artigo tem como objetivo analisar a percepção dos arquivistas das Instituições Federais de Ensino Superior (IFES) acerca da neutralidade na classificação e na descrição de documentos, atividades essenciais no processo de organização do conhecimento arquivístico. Por meio dessa análise, pretende-se fomentar uma reflexão crítica sobre o tema e seus possíveis impactos nas decisões tomadas pelos arquivistas, destacando o papel dos arquivos como fontes de informação para a democratização da sociedade.

2 Neutralidade e o Arquivista

Na filosofia antiga pode-se dizer que se fazia presente uma noção de neutralidade. Aristóteles (1991) não abordou especificamente o tema em seus escritos, de forma explícita. No entanto, argumentou que a virtude moral está no meio entre dois extremos, ou seja, é um equilíbrio entre dois opostos.

Assim como Aristóteles (1991), no século XVII, Descartes (2009) não usa o termo "neutralidade", mas uma ideia de imparcialidade e de objetividade, a partir de suas regras metodológicas para alcançar o conhecimento. Isso é explicitado na medida em que a busca pela verdade se afasta de opiniões preconcebidas que podem interferir no processo de construção do conhecimento.

A noção de neutralidade, na história da ciência, surgiu em oposição ao conhecimento religioso, que era visto como não-neutro devido a sua capacidade de intervir na sociedade e converter pessoas através de suas crenças. O iluminismo questionou a neutralidade ao interrogar o pensamento religioso, enquanto o positivismo reforçou-a ao pressupor que a subjetividade deveria estar contida dentro dos limites da objetividade (DAGNINO, 2002).

Augusto Comte, considerado o idealizador do positivismo defendeu a neutralidade na ciência. Para Comte (1983), a busca pelo conhecimento objetivo e a aplicação de um rigor científico permitiriam à ciência alcançar a verdade sobre as coisas.

Portanto, a partir do exposto, percebe-se que a ideia de neutralidade na filosofia e na ciência tem motivação diretas na forma como profissionais de diversas áreas do conhecimento se relacionam com suas práticas, incluindo os arquivistas. O distanciamento da realidade

e o discurso sobre um agente neutro afetam as formas como os arquivistas se relacionam com suas atividades de diferentes maneiras.

Por isso, muitos teóricos clássicos na área da Arquivologia apresentam a visão do arquivista como um sujeito imparcial e distante de certas funções, como é o caso do Manual de Administração e Arquivos de 1922 escrito pelo britânico Hilary Jenkinson, que teve influência do positivismo, o que era coerente com sua época, e, por isso, via o papel do arquivista como um aplicador de técnicas, de forma distanciada, na avaliação de documentos.

Contudo, autores pós-modernos questionam a neutralidade, influenciados por filósofos como Derrida e como Foucault [1], os quais protestam a criação tradicional de arquivos como locais neutros de armazenamento de informações imparciais. Derrida (2001), por exemplo, destaca os processos seletivos e interpretativos dos arquivos, enfatizando que eles são construídos por estruturas de poder e controle. Já Foucault (1979) ressalta a natureza filtrada dos arquivos, que são influenciados pelas relações de poder que existem na sociedade e controlam os arquivos.

Nesse ponto, Schwartz e Cook (2002) argumentam que os arquivos são instituições que têm o poder de silenciar informações e de controlar o acesso aos documentos, evidenciando a capacidade de registrar, de nomear e de preservar documentos, entre outros meios. Essa perspectiva relaciona os arquivos a locais de construção de memória e de narrativas, formados pelas relações de poder que existem na sociedade. Desse modo, as escolhas feitas pelos arquivistas em relação aos documentos têm um impacto direto na configuração como a sociedade lembra do passado e como constrói sua história.

Jimerson (2009) defende que os arquivistas devem se comprometer com a responsabilidade social e com a ética, reconhecendo que suas escolhas e suas decisões interferem em seu trabalho. Em outro escrito o autor destaca:

Embora os arquivistas considerem seu papel na sociedade, é importante reconhecer que objetividade não é o mesmo que neutralidade. O compromisso com a honestidade, justiça e a verdade lhes dá a base necessária para a "objetividade". Se nos mantivermos fiéis a estes valores éticos e profissionais, podemos nos envolver de forma responsável em debates sobre política pública e procurar assegurar que arquivos documentem as perspectivas de todos os segmentos da sociedade. Isso não é "neutro" porque arquivistas tomarão partidos e fiarão sua expertise profissional para os problemas e questões sociais. No entanto, isso pode ser realizado objetivamente, seguindo padrões aceitáveis da prática arquivística. Isto pode requerer uma definição mais ampla da ética arquivística, a qual deve incluir participação ativa para assegurar os direitos de todos os cidadãos da sociedade. (JIMERSON, 2008, p. 15).

Com essa afirmação, defende que os arquivistas devem ser comprometidos com valores éticos e participar ativamente em debates de políticas públicas para garantir que os arquivos documentem as perspectivas de todos os segmentos da sociedade. Essa abordagem não é neutra, mas ainda assim pode ser objetiva seguindo os padrões aceitáveis das práticas arquivísticas. Dessa forma, ressalta a realização um trabalho mais consciente, considerando perspectivas e experiências diversas registradas nos documentos, bem como os contextos históricos, culturais e políticos que moldaram essas experiências, em prol da justiça social.

Ao buscar referências na literatura arquivística, nesse contexto, acerca da justiça social, Tognoli e Rocha (2021) observaram que esta pode ser entendida como uma forma de reconhecer e de resistir às desigualdades sistêmicas perdidas que permeiam as atividades mantidas nos arquivos, com o objetivo de combater as estruturas de poder. Segundo os autores, isso é uma forma de estimular reflexões teóricas sobre a ética na prática arquivística.

Silva, Tognoli e Guimarães (2017), destacaram a importância de refletir sobre os valores éticos relacionados às funções de classificação e de descrição de documentos. Para os autores, essas funções são exercidas por um profissional que assume a responsabilidade de decidir sobre questões como, por que, para que e para quem esses documentos serão reconhecidos, determinando os valores que devem ser seguidos.

Nesse sentido, é crucial considerar as influências sobre as condutas dos profissionais e da estrutura dos códigos de ética da profissão. O código de ética do arquivista deve resumir os princípios, os valores e as condutas esperadas durante sua atuação, levando em consideração as diversas relações que podem ser protegidas (SILVA, 2021).

Logo, o tratamento dos arquivos pode influenciar a forma como as informações são mediadas nas instituições para a sociedade. Por isso, a imparcialidade e a neutralidade são conceitos enganosos, pois a informação é composta por concepções e por significados que estão imbricados em ideologias, interesses psicológicos e culturais. O profissional encarregado de lidar com a informação não pode ser considerado neutro (ALMEIDA JÚNIOR, 2009).

Reflexões que nos levam a analisar a organização do conhecimento arquivístico, a partir da classificação e da descrição de documentos, já que essas funções geram substitutos documentais. De acordo com Olson (2002), esses substitutos podem apresentar *biases* [2], tanto negativas, ao excluir características importantes de indivíduos, grupos ou sociedade, quanto positivas, ao atender às necessidades e às perspectivas da comunidade de usuários.

Grout (2019) destaca que os arquivos são comumente visualizados como espaços neutros que guardam a verdade dos fatos, contudo temos de reconhecer a possibilidade da ocorrência de preconceitos e de visões culturais por parte dos arquivistas.

Em resumo, enfatiza-se que as atividades de classificação e de descrição podem ser usadas como controle de poder, sendo importante que os profissionais estejam conscientes de seu papel de intervenção e adotem uma postura crítica, informada aos seus valores éticos, bem como na busca por decisões justas e democráticas.

3 Instituições Federais de Ensino Superior

As Instituições Federais de Ensino Superior (IFES), vinculadas ao Ministério da Educação (MEC), compreendem as universidades, institutos e centros federais de educação tecnológica, que promovem a formação de profissionais capacitados em diferentes áreas, através da produção de conhecimento científico, tecnológico e cultural para o desenvolvimento da sociedade, por meio do ensino, da pesquisa e da extensão.

As IFES são compostas por um contingente de aproximadamente 800 profissionais, entre arquivistas e técnicos em arquivos, distribuídos nas cinco regiões do país. Além disso, tais instituições possuem uma Rede denominada ARQUIFES [3], a qual está em funcionamento há mais de uma década. Por meio de grupos de trabalho específicos, essa Rede produz propostas e as discute para serem integradas nas instituições (SILVA, 2019).

A presença de profissionais especializados em Arquivologia nesses espaços no gerenciamento e na preservação dos documentos, visa assegurar o acesso à informação e o cumprimento das obrigações legais. Isso garante a consolidação de políticas e de práticas que promovam a transparência e a reponsabilidade frente à Administração Pública Federal.

As IFES estão subordinadas ao Poder Executivo Federal e integram o Sistema Nacional de Arquivos (SINAR), o qual segue as políticas protegidas pelo Conselho Nacional de Arquivos (CONARQ). O SINAR atua em colaboração com o Conselho, que é o órgão central responsável pela definição e pela regulamentação da política nacional de arquivos públicos e privados no país (BRASIL, 2002).

A articulação entre as IFES e esses órgãos estimula a aplicação de boas práticas de gestão documental, por meio da realização de atividades como a classificação, a avaliação e a descrição de documentos.

No que se refere às atividades de classificação e de avaliação nas IFES, os profissionais são orientados pelo Código de Classificação de Documentos e Tabela de Temporalidade e Destinação de Documentos, elaborados pelo CONARQ, tanto para as atividades-fim

quanto para as atividades-meio. Uma medida central na realização dessas atividades é a criação de Comissões Permanentes de Avaliação de Documentos (CPAD), responsáveis pela análise de documentos classificados e avaliados por arquivistas nas IFES. No entanto, a eliminação desses documentos, sejam eles digitais ou não digitais, só pode ocorrer após a análise da CPAD e autorização do órgão competente.

As práticas adotadas pelas IFES para a descrição de documentos seguem padrões e normas nacionais e internacionais, dentre elas a Norma Brasileira de Descrição Arquivística (NOBRADE), que apresenta diretrizes específicas para a descrição de documentos arquivísticos. Além disso, visando ampliar a disseminação de documentos no meio digital, muitas instituições vêm adotando o software ATOM [4], o qual é compatível com as normas de descrição.

As atividades arquivísticas nas IFES envolvem governos, empresas, comissões, instituições e indivíduos, o que impactam nas decisões dos arquivistas. Esses aspectos nos levam a refletir sobre o discurso da neutralidade por parte dos arquivistas diante dos documentos na realização das atividades de classificação e de descrição.

4 Metodologia

Quanto à abordagem, adotou-se uma direção mista, quantitativa e qualitativa, em relação aos objetivos exploratórios e descritivos. Dentro do nosso campo de recorte, utilizou-se a exploração para compreender as discussões e obter conhecimento sobre o entendimento de neutralidade nos arquivos, juntamente com aspectos descritivos, para apresentar o fenômeno observado.

Realizou-se uma pesquisa de campo nas Instituições Federais de Ensino Superior (IFES), considerando como sujeitos da pesquisa os arquivistas que trabalham nesses órgãos. Para coleta de dados, foi enviado questionário eletrônico para um grupo de WhatsApp composto por 281 arquivistas da IFES.

O formulário aplicado continha uma pergunta fechada que questionava se o arquivista pode ser neutro na elaboração e na aplicação da classificação e da descrição de documentos, contendo quatro alternativas para resposta, sendo três delas definidas e uma com a opção "outro", considerando possíveis respostas que não se enquadrassem nas definições das alternativas fornecidas. A segunda pergunta aberta permitia que os participantes respondessem livremente, questionando se esses profissionais entendem que a mediação do arquivista pode afetar a classificação e a descrição de documento.

Para a análise dos dados obtidos a partir das respostas dos participantes, foram adotadas técnicas específicas para cada tipo de questão. No caso das questões fechadas, optou-se pela apresentação dos resultados em per-

centuais, com o objetivo de mensurar o nível de compreensão dos participantes sobre a neutralidade. Para a questão aberta, utilizou-se a técnica de análise de conteúdo proposta por Bardin (2013), que envolveu três etapas: pré-análise, exploração do material e tratamento dos resultados. Nessa etapa, incorporou-se as respostas em uma planilha, identificando palavras-chave e ideias recorrentes. Posteriormente, agrupadas em categorias mais abrangentes, observando-se as ideias que se repetiam. Em um terceiro momento, identificou-se pontos mais específicos relacionados às atividades de classificação e de descrição dos documentos.

5 Resultados

A pesquisa contou com a participação de arquivistas das IFES de diferentes regiões do Brasil, totalizando 43 respondentes de 29 instituições (ver Apêndice A), o que representa uma amostra de 15% dos 281 profissionais pertencentes ao grupo para o qual o formulário foi enviado. Notou-se ainda a presença de múltiplas respostas de profissionais de uma mesma instituição.

Considerando a distribuição das instituições por região, constatou-se que a região nordeste concentrou a maior quantidade de respondentes, totalizando 34,9%, seguida pela região sudeste com 25,6%, sul com 16,3%, norte com 14% e centro-oeste com 9,3%. É válido ressaltar que a heterogeneidade geográfica e institucional dos participantes contribuiu para fornecer uma visão representativa da realidade dos arquivistas nas IFES.

Com relação à primeira pergunta, os resultados indicam uma variação de opiniões sobre a neutralidade na elaboração e na aplicação da classificação e da descrição de documentos.

Quadro 1. Primeira questão – Neutralidade

<i>Você acredita que o arquivista pode ser neutro no desenvolvimento e na aplicação da classificação e descrição de documentos?</i>	<i>Percentual de Participantes</i>
Sim, acredito que o arquivista pode ser neutro em suas atividades profissionais, sem permitir que suas influências culturais, políticas ou sociais afetem as atividades.	30,2%
Não, porque as ações do Arquivista sempre influenciarão na sua realidade.	27,9%
Depende do contexto em que o arquivista está inserido, pois em algumas situações ele pode ter que tomar decisões para cumprir objetivos pessoais, políticos ou institucionais.	34,9%
Outros	6,9%

Fonte: Os autores, 2023.

Como apresentado no Quadro 1, a maioria dos participantes (34,9%) afirmou que a neutralidade é contextual, sugerindo que as atividades do arquivista podem ser influenciadas por interesses diversos que surgem ao longo do processo. Outra parcela considerável dos entrevistados (30,2%) pensa ser possível alcançar a neutralidade eliminando influências culturais, políticas ou sociais no desenvolvimento das atividades. Já uma minoria desacredita na possibilidade de neutralidade, argumentando que as ações do arquivista sempre interferem em sua realidade.

Os participantes que não marcaram nenhuma alternativa correspondem a 6,9%, indicando outras possibilidades: uma delas sugerindo que a neutralidade está envolvida em atividades que abrangem certo grau de subjetividade, e a outra afirmando que o arquivista deve buscar a neutralidade, mas nem sempre consegue atingi-la.

A segunda questão abordada diz respeito ao impacto da mediação do arquivista nas funções de classificação e de descrição de documentos. A análise das respostas dos participantes permite identificar que a mediação do arquivista em tais atividades está relacionada a seis pontos fundamentais que estão interligados, a citar: ética, princípios arquivísticos, usuários, instrumentos e normas, instituição arquivística e sociedade. As respostas indicam que ao elaborar ou desenvolver a classificação e a descrição, o arquivista, em sua mediação, deve observar esses pontos citados, os quais apresentam impactos tanto positivos, quando seguidos, quanto negativos, se não seguidos.

Quadro 2. Segunda questão - Categorias de impactos do arquivista na mediação da classificação e da descrição

<i>Categorias</i>	<i>Impactos Positivos, se seguidos.</i>	<i>Impactos negativos, se não seguidos</i>
<i>Ética</i>	Atuar de acordo com os valores que regem a profissão e que norteiam as ações tomadas em relação ao acesso aos documentos.	Agir de maneira desonesta, prejudicial ou ilegal pode comprometer o acesso aos documentos.
<i>Princípios arquivísticos</i>	A aplicação dos princípios arquivísticos da proveniência e da organicidade permite que os documentos sejam recuperados de acordo com a sua origem e o seu contexto de produção.	Afetar a recuperação dos documentos.

<i>Usuários</i>	Seguir as demandas e dos usuários permite que os documentos de interesse sejam utilizados por diferentes contextos.	Não considerar os usuários pode gerar menor eficiência na gestão dos documentos.
<i>Instrumentos e normas</i>	Apresentam uniformidade nas atividades, auxiliam nos procedimentos legais, apoiam a gestão e as decisões administrativas.	Sem normas e instrumentos, pode haver prejuízo na organização e na recuperação dos documentos.
<i>Instituição</i>	Considerar a instituição ajuda a identificar o contexto de criação dos documentos e ajudar nos ritos administrativos.	Não considerar a instituição pode ocasionar uma gestão ineficiente.
<i>Sociedade</i>	Garantir a preservação da memória, promover à justiça, bem como o acesso à informação e à pesquisa.	Dificuldades no acesso à informação por parte dos cidadãos e violação de direitos.

Fonte: Os autores, 2023.

Esses pontos elencados no Quadro 2 foram os mais gerais citados como pertinentes às duas atividades. No que diz respeito aos aspectos mais específicos voltados às atividades, os respondentes apontam que ambas são *funções subjetivas* e que exigem uma *análise criteriosa* e um *conhecimento amplo da instituição*. Em relação à classificação, pode *auxiliar na identificação e na seleção de documentos*, além de ser importante para a avaliação de documentos. Porém, é necessário tomar cuidado para *evitar a utilização de códigos errados* que possam prejudicar a formação dos conjuntos documentais e *comprometer os vínculos arquivísticos*. Já em relação à descrição, é importante *controlar os níveis de descrição* e não deixar que *fatores internos e externos* possam influenciar na descrição dos documentos.

Por fim, cruzamos os dados dos respondentes que afirmaram ser possível a neutralidade na questão 1, que foi de 30,2%, e os impactos mencionados na questão 2. Embora acreditem que os valores pessoais, políticos e sociais não afetem as atividades de classificar e de descrever mencionadas na questão 1, as respostas dadas na questão 2 demonstram que esses profissionais pensam que essas duas atividades podem sofrer interferências. Os arquivistas afirmam que é possível realizá-las com má-fé, sem seguir preceitos éticos ou sem considerar os princípios da área.

6 Conclusão

A maioria dos profissionais reconhece que questões pessoais, políticas e institucionais exercem influência sobre o curso das atividades de classificação e descrição, dependendo do contexto em que ocorrem. No entanto, parte dos entrevistados parece indicar uma falta de compreensão em relação à subjetividade inerente a essas atividades, confundindo a neutralidade com a busca pela objetividade.

Entendemos que, mesmo buscando uma idealização “neutra”, os arquivistas ainda serão influenciados por suas próprias crenças, valores e experiências, bem como pelas demandas e pelos interesses da instituição e da sociedade em que trabalham. Os efeitos apresentados mostram como a mediação dos arquivistas é atravessada por diversos pontos, como a ética, os princípios arquivísticos, os usuários, os instrumentos e normas, a instituição e a sociedade, que apresentam consequências.

Compreende-se que o discurso de neutralidade nas atividades arquivísticas pode ser perigoso, uma vez que estimula uma falsa posição sem intervenção com a realidade. Para evitar esse problema, é preciso buscar uma atuação consciente e crítica, o que implica em uma reflexão responsável e comprometida com os entornos que envolvem a atuação profissional.

Na representação arquivística, é possível que ocorram *biases*, resultando em influências culturais e políticas que, por sua vez, podem levar a exclusões e omissões de informações. Por esse motivo, é crucial que o arquivista adote uma posição, como Lopes (2009) denomina de arquivista hermeneuta, sendo capaz de interpretar sua realidade e os diversos desafios na sociedade em que atua.

No contexto das IFES, esse trabalho crítico dos arquivistas se torna essencial, uma vez que esses profissionais têm um papel representativo nessas instituições e estão submetidos aos princípios da administração pública, como a transparência e o acesso. Assim, espera-se que outras pesquisas adicionais possam aprofundar as questões relacionadas à neutralidade nas práticas do arquivista, incluindo as diferentes formas de interferência possíveis.

Notas

- [1] Para Foucault (1979), assim como para Derrida (2001), não há sentido na história. As grandes narrativas se entrelaçam em crise. Não existe o real para os pós-modernos, apenas narrativas sobre a realidade que não são finalísticas e nem totalizantes, ou seja, a história não é conduzida por uma linearidade.
- [2] Milani (2014) considera em sua pesquisa o termo bias(es) na língua inglesa, pois em português abrange diferentes conotações: negativas (tendência, inclinação ou desvio) e positivas (direcionamentos ou viés).

[3] Histórico da REDE-ARQUIFES descrito por meio do site: <https://arquifes.com.br/>.

[4] Lista das instituições que implantaram o Atom, disponível através do site: <https://wiki.accesstomemory.org/>

Referências

ALMEIDA JÚNIOR, O. F. de. Mediação da informação e múltiplas linguagens. **Pesquisa Brasileira em Ciência da Informação**, Brasília, v.2, n.1, p.89-103, jan./dez. 2009.

ARISTÓTELES. **Ética a Nicômaco**. São Paulo: Nova Cultural, 1991. (Coleção Os pensadores).

BARROS, T. H. B.; SOUSA, R. T. B. de. Organização do Conhecimento e Arquivologia: abordagens metodológicas. **Informação & Informação**, [S. l.], v. 24, n. 2, p. 76–92, 2019. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/38290>. Acesso em: 26 maio. 2023.

BARDIN, L. **Análise de conteúdo**. São Paulo: Edições, 2013.

BRASIL. Lei nº 8.159, de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. **Diário Oficial da União**, Brasília, DF, 9 jan. 1991.

COMTE, A. **Curso de Filosofia Positiva**. Tradução de José Arthur Giannotti. 2. ed. São Paulo: Abril Cultural, 1983. (Coleção Os pensadores).

COOK, T. O passado é prólogo: uma história das ideias arquivísticas desde 1898 e a futura mudança de paradigma. In: HEYMANN, L.; NEDEL, L. (orgs.). **Pensar os arquivos: uma antologia**. Rio de Janeiro: FGV, 2018. p. 17-81.

DESCARTES, R. **Discurso do Método**. Tradução de Ciro Mioranza. São Paulo: Editora Escala, 2009.

DAGNINO, R. P. Enfoques sobre a relação ciência, tecnologia e sociedade: neutralidade e determinismo. **DataGram-Zero**: Revista de Ciência da Informação, Rio de Janeiro, v. 3, n. 6, 2002.

DERRIDA, J. **Mal de Arquivo: uma impressão freudiana**. Rio de Janeiro: Relume Dumará, 2001.

FOUCAULT, M. **Microfísica do poder**. Rio de Janeiro: Graal, 1979.

HOUAISS, A. **Dicionário Houaiss da Língua Portuguesa**. Rio de Janeiro: Ed. Objetiva, 2001.

GARCÍA GUTIÉRREZ, A. Knowledge Organization from a “culture of the border”: towards a transcultural ethics of mediation. In: LÓPEZ-HUERTAS, M. J. (ed.). **Challenges in Knowledge representation and Organization for the 21st century: integration of knowledge across boundaries**. Würzburg: ERGON, 2002. p. 516-522.

GROUT, H. Archiving critically: exploring the communication of cultural biases. **SPARK: UAL Creative Teaching and Learning Journal**, Londres, v. 4, n. 1, p. 71-74, 2019.

JIMERSON, R. C. **Archives Power: Memory, Accountability, and Social Justice**. Chicago: Society of American Archivists, 2009. 466 p.

JIMERSON, R. C. Arquivos para todos: a importância dos arquivos na sociedade. **Arquivo & Administração**, Rio de Janeiro, v. 7, n. 2, 2008.

JENKINSON, H. **A manual of archive administration including the problems of war archives and archive making**. Londres: Oxford, 1922.

MARTINS, G. K.; CÔRTEZ, G. R. A representação da informação e do conhecimento e as representações sociais: interseções e limites. In: ALBUQUERQUE, M. E. B. C.; MARTINS, G. K.; MOTA, D. A. R. (orgs.) **Organização e representação da informação e do conhecimento: Interseções teórico-sociais**. João Pessoa: Editora UFPB, 2019.

MILANI, S. O. **Bias na Representação de Assunto: Uma Discussão de Oposições Binárias nos Functional Requirements for Subject Authority Data (FRSAD)**. 2014. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

NESMITH, T. Relendo os arquivos: novas contextualidade para a teoria e a prática arquivística. In: HEYMANN, L.; NEDEL, L. (orgs.). **Pensar os arquivos: uma antologia**. Rio de Janeiro: FGV, 2018. p. 17-81.

OLSON, H. A. **The power to name: locating the limits or subject representation in libraries**. Dordrecht: Kluwer Academic Publisher, 2002.

REALE, G.; ANTISERI, D. **História da Filosofia**. 8. ed. São Paulo: Paulus, 2007. v. 2.

SCHWARTZ, J.; COOK, T. Archives, records, and power: the making of modern memory. **Archival Science**, Dordrecht, v. 2, n. 1, p. 1-19, 2002.

SILVA, A. P. **Códigos de ética profissional do arquivista: contribuições da crítica diplomática para caracterização de uma espécie documental**. 2021. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2021.

SILVA, R. M. R. D. **Rede Nacional ARQUIFES: Uma análise de sua constituição**. 2019. Dissertação (Mestrado em Gestão de Documentos e Arquivos) – Centro de Ciências Sociais e Humanas, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2019.

SILVA, A. N. D.; TOGNOLI, N. B. T.; GUIMARÃES, J. A. C. G. Os valores éticos na organização e representação do conhecimento arquivístico. **Brazilian Journal of Information Science: Research Trends**, Marília, v. 11, n. 1, p. 44-53, 2017.

TOGNOLI, N. B.; ROCHA, G. M. Q. A justiça social na Arquivologia: uma revisão de literatura. **ÁGORA: Arquivologia em Debate**, Florianópolis, v. 31, n. 63, p. 1-23, 2022.

Apêndice A – Lista das Instituições participantes

Quadro 1. *Lista das Instituições participantes*

	<i>INSTITUIÇÃO</i>	<i>SIGLA</i>
1	Centro Federal de Educação Tecnológica de Minas Gerais	CEFET-MG
2	Instituto Federal de Mato Grosso	IFMT
3	Instituto Federal de Sergipe	IFS
4	Instituto Federal Farroupilha	IFF
5	Universidade Federal da Bahia	UFBA
6	Universidade Federal da Paraíba	UFPB
7	Universidade Federal de Alfenas	UNIFAL-MG
8	Universidade Federal de Goiás	UFG
9	Universidade Federal de Juiz de Fora	UFJF
10	Universidade Federal de Mato Grosso do Sul	UFMS
11	Universidade Federal de Minas Gerais	UFMG
12	Universidade Federal de Ouro Preto	UFOP
13	Universidade Federal de Pelotas	UFPEL
14	Universidade Federal de Pernambuco	UFPE
15	Universidade Federal de Santa Catarina	UFSC
16	Universidade Federal de São Carlos	UFSCAR
17	Universidade Federal de Sergipe	UFSE
18	Universidade Federal do Amazonas	UFAM
19	Universidade Federal do Cariri	UFCA
20	Universidade Federal do Estado do Rio de Janeiro	UNIRIO
21	Universidade Federal do Maranhão	UFMA
22	Universidade Federal do Oeste do Pará	UFOPA
23	Universidade Federal do Pará	UFPA
24	Universidade Federal do Paraná	UFPR
25	Universidade Federal do Rio de Janeiro	UFRJ
26	Universidade Federal do Rio Grande	FURG
27	Universidade Federal Fluminense	UFF
28	Universidade Federal Rural de Pernambuco	UFRPE
29	Universidade Tecnológica Federal do Paraná	UTFPR

Fonte: Os autores, 2023.

A representação colaborativa do conhecimento no WikiFavelas - Dicionário de Favelas Marielle Franco

The collaborative knowledge representation in WikiFavelas - Dictionary of Favelas Marielle Franco

Ana Carla Florêncio (1), Juliana de Assis (2)

(1) Universidade Federal do Rio de Janeiro, Avenida Pasteur, 250, Urca, Rio de Janeiro, ana.carla459@gmail.com,
(2) juliana.assis@facc.ufrj.br

Resumo

Contextualização: aborda o “Dicionário de Favelas Marielle Franco”, também chamado “Wikifavelas”, uma plataforma colaborativa, que utiliza a hibridização entre três tipologias distintas de sistemas de organização do conhecimento: dicionário, folksonomia e taxonomia, com foco nas contribuições da folksonomia para a representação do conhecimento, **objetivo:** investigar o papel da folksonomia na construção do conhecimento coletivo e compreender a potencialidade da coletividade na produção de conhecimentos, **método:** trata-se de uma pesquisa exploratória e descritiva com caráter qualitativo, que utilizou como técnica de coleta de dados a observação participante. Caracteriza-se ainda como uma pesquisa bibliográfica e documental. O objeto empírico é a plataforma “Dicionário de Favelas Marielle Franco – Wikifavelas”. Buscou-se dialogar sobre os principais conceitos da Organização do Conhecimento pertinentes ao estudo. Abordou-se o conceito de Web 2.0, suas origens e os seus benefícios para a representação e construção coletiva do conhecimento. Observou-se a aplicabilidade das *tags* nos verbetes, atribuídas livremente pelos sujeitos informacionais, **conclusão:** concluiu-se que, a plataforma colaborativa analisada revela a hibridização dos sistemas de organização do conhecimento no contexto digital e a sua importância, não somente para a representação da informação e do conhecimento, mas também para a construção da memória coletiva incluindo a possibilidade de agregar diversas vivências e visões de mundo. Este estudo se torna relevante à medida que explora as características das práticas colaborativas de criação do conhecimento coletivo em uma comunidade pouco explorada no campo de estudos da Informação e da Organização do Conhecimento. Espera-se que este trabalho possa contribuir para despertar o interesse de outros pesquisadores da área em estudar plataformas similares.

Palavras-chave: Dicionário de Favelas Marielle Franco. Organização do Conhecimento. Representação do Conhecimento. Web 2.0. Folksonomia.

Abstract

Contextualization: it addresses the “Dicionário de Favelas Marielle Franco”, also called “Wikifavelas”, a collaborative platform, which uses the hybridization between three distinct typologies of knowledge organization systems: dictionary, folksonomy and taxonomy, focusing on the contributions of folksonomy to the representation of knowledge, **objective:** to investigate the role of folksonomy in the construction of collective knowledge and to understand the potential of the collective in the production of knowledge, **method:** This is an exploratory and descriptive research with a qualitative character, which used participant observation as a data collection technique. It is also characterized as a bibliographical and documentary research. The empirical object is the platform “Dicionário de Favelas Marielle Franco – Wikifavelas”. We sought to dialogue about the main concepts of the Organization of Knowledge relevant to the study. The concept of Web 2.0, its origins and its benefits for the representation and collective construction of knowledge were addressed. The applicability of the tags in the entries, freely assigned by the informational subjects, was observed, **conclusion:** it was concluded that the analyzed collaborative platform reveals the hybridization of knowledge organization systems in the digital context and its importance, not only for the representation of information and knowledge, but also for the construction of collective memory, including the possibility of aggregating diverse experiences and worldviews. This study becomes relevant as it explores the characteristics of collaborative practices for creating collective knowledge in a community little explored in the field of Information and Knowledge Organization studies. It is hoped that this work can contribute to arouse the interest of other researchers in the area in studying similar platforms.

Keywords: Marielle Franco’s Dictionary. Knowledge Organization. Knowledge Representation. Web 2.0. Folksonomy.

Resumen

Contexto: se trata del “Dicionário de Favelas Marielle Franco”, también conocido como “Wikifavelas”, una plataforma colaborativa que utiliza la hibridación de tres tipologías distintas de sistemas de organización del conocimiento: diccionario, folksonomía y taxonomía, con enfoque en las contribuciones de la folksonomía para la representación del conocimiento, **objetivo:** investigar el papel de la folksonomía en la construcción del conocimiento colectivo y comprender el potencial de la colectividad en la producción de conoci-

miento, **método:** se trata de una investigación exploratoria y descriptiva con carácter cualitativo, que utilizó como técnica de recolección de datos la observación participante. Se caracteriza también como una investigación bibliográfica y documental. El objeto empírico es la plataforma "Dicionário de Favelas Marielle Franco - Wikifavelas". Se buscó dialogar sobre los principales conceptos de la Organización del Conocimiento relevantes para el estudio. Se abordó el concepto de Web 2.0, sus orígenes y sus beneficios para la representación y construcción colectiva del conocimiento. Se observó la aplicabilidad de las etiquetas en los artículos, asignadas libremente por los sujetos de información, **conclusión:** se concluyó que la plataforma colaborativa analizada revela la hibridación de los sistemas de organización del conocimiento en el contexto digital y su importancia, no solo para la representación de la información y el conocimiento, sino también para la construcción de la memoria colectiva, incluyendo la posibilidad de agregar diversas experiencias y visiones del mundo. Este estudio se vuelve relevante a medida que explora las características de las prácticas colaborativas de creación de conocimiento colectivo en una comunidad poco explorada en el campo de los estudios de Información y Organización del Conocimiento. Se espera que este trabajo pueda contribuir a despertar el interés de otros investigadores en el área para estudiar plataformas similares.

Palabras clave: Dicionario de Favelas Marielle Franco. Organización del Conocimiento. Representación del Conocimiento. Web 2.0. Folksonomía.

1 Introdução

A Web 2.0 possibilitou diversos processos de interação social e de compartilhamento de informações e conhecimentos ao maximizar significativamente as atividades colaborativas em ambientes digitais. O'Reilly (2005), afirma que a Web 2.0 não tem fronteiras, sendo um centro gravitacional no qual circulam várias práticas e princípios, uma plataforma que explora a experiência do usuário como seu principal modelo de negócios.

Atualmente na web, o sujeito possui um papel ativo como produtor e disseminador da informação, sendo criador informacional em diversos dispositivos e ambientes colaborativos, faz parte dos ciclos de construção da informação e conhecimento. Diante dessa perspectiva, os dados, fluxos de informação e serviços, se tornam cada vez mais personalizáveis e baseados no aproveitamento da inteligência coletiva.

Com o advento da Web 2.0, a partir da década de 2000, os sites, portais e plataformas adquiriram uma forma inovadora de representação e organização da informação e do conhecimento, baseada na filosofia da colaboração: a folksonomia.

Para Sturtz, (2004) a folksonomia é o conjunto completo de *tags* (também conhecidas como etiquetas ou palavras-chave), que os sujeitos em uma plataforma de gerenciamento de conteúdo compartilhado, aplicam para agrupar, classificar e difundir conteúdos, gerada a partir de uma modalidade de indexação, denominada indexação social.

Guedes e Dias (2010, p. 41) definem a indexação social como “[...] a ação de etiquetagem desempenhada por usuários de ferramentas sociais em ambientes Web”. Sendo assim, tem-se uma abordagem de indexação que parte dos utilizadores do sistema e não dos profissionais indexadores.

A indexação social é um modelo orientado pelos usuários da *web*, de modo que envolve pessoas de diferentes esferas mediante o uso da linguagem natural em suas práticas de descrição e representação de conteúdos informacionais de diversas naturezas.

Diante disso, tais práticas colaborativas incentivam os sujeitos informacionais a desempenharem um papel ativo e contribuem para uma memória coletiva por meio da livre expressão contida em suas representações de conteúdo. Esse contexto implica diversos fatores como: a pluralidade dos pontos de vistas; a questão ética; as formas de representação da informação e do conhecimento; etc.

Tendo em vista as evoluções e mudanças da organização da informação e do conhecimento em contextos digitais, buscou-se analisar o processo da representação do conhecimento da Wikifavelas - Dicionário de Favelas Marielle Franco - projeto que possui apoio da Fundação Oswaldo Cruz e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ).

A Wikifavelas é uma plataforma virtual, de acesso aberto, que produz e reúne informações e memórias sobre favelas e periferias. O dicionário se caracteriza, como uma plataforma colaborativa, editada pelos sujeitos informacionais e gestores do domínio.

O objetivo geral deste trabalho é investigar o papel da folksonomia na construção do conhecimento coletivo, na Wikifavelas. Em diálogo com o objetivo geral, foram definidos os seguintes objetivos específicos: a) compreender a potencialidade da coletividade na produção de conhecimentos na Wikifavelas; b) descrever a contribuição da representação colaborativa do conhecimento na Wikifavelas; c) apontar vantagens e desvantagens da adoção dos modelos colaborativos na representação do conhecimento.

A motivação para a escolha do campo empírico, se dá pela importância de se investigar as estruturas das plataformas colaborativas que utilizam a folksonomia e compreender como sua atuação pode ir além da representação da informação e do conhecimento em contextos digitais.

Assim, na perspectiva das construções culturais ao longo da globalização e do papel da sociedade na produção do conhecimento, torna-se importante

conhecer e visibilizar verbetes e termos condizentes com a sua pluralidade e diversidade.

A relevância do objeto empírico fundamenta-se na necessidade de preservar a memória dos moradores das periferias e seus acontecimentos, bem como na importância das práticas colaborativas para os próprios sujeitos que desempenham um papel cooperativo em que participam ativamente da produção de informações e conhecimentos, além de proporcionar a visibilidade dos diferentes territórios.

A pesquisa realizada é de natureza exploratória. Para Gil (2008, p. 27) “as pesquisas exploratórias têm como principal finalidade desenvolver, esclarecer e modificar conceitos e ideias, tendo em vista a formulação de problemas mais precisos ou hipóteses pesquisáveis para estudos posteriores”. Essa abordagem possui como uma de suas características, proporcionar uma visão geral acerca de determinado fato, de modo a auxiliar na exploração do campo de pesquisa.

Possui ainda um caráter descritivo, em que o “[...] objetivo primordial é a descrição de características de determinada população ou fenômeno ou o estabelecimento de relações entre variáveis.” (GIL, 2008, p. 28). Neste trabalho buscou-se também realizar a descrição dos elementos da plataforma Wikifavelas.

A abordagem da pesquisa é qualitativa, pois permite uma compreensão melhor acerca do objeto empírico a ser analisado. Essa abordagem busca compreender um universo de crenças, significados, valores, aspirações etc., que para Minayo (2003), é uma perspectiva metodológica que se aprofunda nos mundos dos significados das relações humanas.

O estudo se qualifica como pesquisa bibliográfica, pois foi desenvolvida a partir de livros e artigos científicos anteriores de autores da área da Biblioteconomia e Ciência da Informação e ainda como pesquisa documental, pois utiliza o próprio dicionário como uma fonte de dados.

2. O dicionário enquanto Sistema de Organização do Conhecimento e sua hibridização

A Organização do Conhecimento (OC) é um campo de pesquisa e atuação profissional que instaura meios para se representar e assim compartilhar o conhecimento humano mediante suas unidades constituintes mais básicas, que são os conceitos (DAHLBERG, 2006). Tais unidades são formalizadas por símbolos e estruturadas de modo a compor esquemas compreensivos que retratam aspectos da realidade.

Em consonância com Lima e Alvares (2012), podemos definir a representação do conhecimento como toda a tentativa de simulação do pensamento humano a partir de seus processos e estruturas cognitivas com o

propósito de transformar o conhecimento individual em conhecimento socializado.

A OC lida com representações de unidades de conhecimento em diferentes níveis de complexidade mediante sistemas simbólicos denominados Sistemas de Organização do Conhecimento (SOC). Estes, são definidos por Bratková e Kučerová (2014) como esquemas que modelam estruturas de um conjunto sistematizado de conhecimento. Neles os conceitos são os elementos estruturais mais elementares e suas funções mais básicas são conferir suporte aos processos de organização e acesso ao conhecimento.

Conforme a taxonomia das tipologias de SOC definida por Zeng (2008), os dicionários são os tipos menos complexos de SOC no que tange à sua estrutura conceitual. Considerados como listas de termos, apresentam uma estruturação do conhecimento de forma alfabética e por serem dotados de definições de termos contribuem para a eliminação de ambiguidades.

Observa-se no contexto digital ocorrências de hibridização entre dois ou mais SOC visando melhor atender às especificidades que envolvem os processos de representação da informação e do conhecimento, bem como a complexa tarefa de prover a recuperação da informação.

É o que ocorre, por exemplo, com o objeto empírico desta pesquisa, uma plataforma colaborativa em que um dicionário eletrônico agrega em sua estrutura, outros dois tipos de SOC, a folksonomia e uma taxonomia, composta por categorias temáticas que possibilitam recuperar a informação também por meio da navegação (*browsing*), conforme exemplificado no Anexo A.

Desse modo, embora o foco do estudo seja a contribuição da indexação social e conseqüentemente, da folksonomia, para a representação do conhecimento em um dicionário, considera-se relevante ressaltar a hibridização dos SOC que atuam na referida plataforma. Não se trata de um fato isolado, mas de algo recorrente e que tende a aumentar na *web* atual visto que, no contexto digital, as imperfeições de um SOC podem ser contornadas pelas potencialidades de outro, conforme apontam os estudos contemporâneos da OC, tanto em âmbito nacional, quanto internacional.

3 O Contexto da Web 2.0, Indexação Social e Folksonomia

Os avanços tecnológicos que envolvem a *web* contribuíram significativamente para uma nova forma de interação dos sujeitos com as informações, de modo que a sociedade começou a participar e colaborar ativamente no desenvolvimento da informação e do conhecimento; essa mudança foi possibilitada pela fase denominada Web 2.0.

Para Davis (2005, p. 196 *apud* BLATTMANN; SILVA, 2007), “[...] a Web 1.0 direcionava as pessoas para as informações enquanto a Web 2.0 intencionava a levar as informações para as pessoas”.

O conceito de Web 2.0 iniciou-se em 2004 a partir de um evento promovido pelas empresas *O'Reilly Media* e *MediaLive International*. Na conferência discutiu-se a ideia de uma *web* mais dinâmica e interativa para seus usuários, na qual as pessoas também poderiam contribuir com os conteúdos (BLATTMANN; SILVA, 2007).

Com as inovações que o modelo de negócios da Web 2.0 proporcionou, surgiram as plataformas colaborativas, com serviços e funcionalidades diferentes, que viabilizaram práticas de produção de conteúdo, bem como o compartilhamento e a organização da informação e do conhecimento no contexto digital (GONÇALVES; ASSIS, 2016).

Um exemplo é a plataforma interativa de fotografias: *Flickr*, sistema colaborativo em que os próprios usuários criam conteúdos imagéticos e atribuem e estipulam arranjos de *tags* nas fotos. Nessa plataforma a representação da informação é pautada pela representação imagética, abordando aspectos intrínsecos e extrínsecos das imagens.

Antes da Web 2.0 propiciar o surgimento da folksonomia, existiam os sistemas de representação e organização do conhecimento, considerados tradicionais, caracterizados pelo controle terminológico, tais como tesouros e classificações bibliográficas. Atualmente, os SOC baseados em linguagem natural e em vocabulários controlados coexistem e são passíveis de hibridização no contexto digital.

Diante da produção das informações e da evolução dos ambientes e plataformas na *web*, surge a necessidade do modelo da indexação social que permite às atividades colaborativas de indexação, em que o usuário é também colaborador do conteúdo produzido e compartilhado. O caráter social da indexação possibilita a formação de grupos virtuais em torno de assuntos dos mesmos interesses (CATARINO; BAPTISTA, 2009). Esses grupos são denominados comunidades virtuais de prática.

Em consonância com Lira *et al.* (2021, p.104) podemos compreender esse tipo de comunidade como “[...] formadas por grupos de pessoas que possuem uma relação por meio da internet, com a finalidade de aprofundar discussões, debates, de maneira que possa enriquecer e contribuir, de forma privada ou pública”. Atuam, portanto no compartilhamento de conhecimentos conforme interesses comuns por meio de tecnologias de suporte às atividades de colaboração (LIRA *et al.*, 2021)

A abordagem da Web 2.0, instaura uma metodologia de representação de conteúdos, que inclui o usuário na produção e compartilhamento da informação e do conhecimento. Nesse cenário os atores sociais que se utilizam da Internet deixam de ser apenas consumidores, para serem criadores e produtores dentro de um ambiente colaborativo digital.

Para O'Reilly (2005, p. 7) a Web 2.0 alavancou o autosserviço do consumidor e algoritmos que possibilitaram atingir a rede em toda sua extensão e não apenas o centro. Essa lógica distribuída está na base do modelo de classificação por trás das folksonomias.

A palavra folksonomia surgiu em 2004, declarada pelo arquiteto de informação Thomas Vander Wal, segundo o autor, “A folksonomia pode ser definida como uma classificação popular que se origina das ações de representação da informação desempenhadas por usuários de diversos serviços na web atual” (VANDER WAL, 2007, *apud* ASSIS; MOURA 2013, p. 87).

O termo folksonomia é a tradução de *folksonomy*, neologismo criado a partir da junção de *folk* (povo, pessoas) com *taxonomy*, que significa classificação (CATARINO; BAPTISTA, 2007).

A partir da indexação social, a folksonomia, pode ser entendida como a sua classificação colaborativa resultante. Para Assis e Moura (2013), o sujeito informacional, estabelece sua subjetividade por meio de percursos informacionais na *web*, assim como explicita Tim O'Reilly, que confirma que nesse contexto o sujeito adiciona valor, fazendo parte de uma inteligência coletiva.

O tipo de descrição semântica propiciada pela folksonomia pode ser aplicada a conteúdos como documentos textuais, imagéticos, sonoros, audiovisuais, no todo ou ainda, em partes. O autor Sturtz (2004), disserta:

Em termos práticos, uma folksonomia é o conjunto completo de tags - uma ou duas palavras-chave — que os usuários de um sistema de gerenciamento de conteúdo compartilhado aplicam a partes de conteúdo para agrupar ou classificar essas partes para recuperação (STURTZ, 2004, p 1).

Desse modo, contempla uma característica da *web* como camada hipertextual e hiperídia da Internet: a recuperação de conteúdos completos ou de partes/trechos de documentos.

A folksonomia permite aos sujeitos uma participação ativa na criação de um tipo de sistema de classificação em que não se preconiza um sistema terminológico controlado, como nos instrumentos tradicionais da Biblioteconomia, esse caráter acessível, com linguagem natural, viabiliza a criação de comunidades virtuais de prática que podem atuar com diferentes propósitos. “Em termos estruturais, as folksonomias constituem espaços sociais semânticos em que ocorre a agregação de

representações subjetivas, objetivas e práticas dos objetos informacionais” (QIN, 2008 *apud* ASSIS; MOURA, 2013, p. 89). Sendo assim, a riqueza semântica e cultural das descrições dos conteúdos se dá pela variedade de sujeitos e abordagens possíveis mediante a indexação social.

Há plataformas mundialmente conhecidas que utilizam a folksonomia, como é o caso dos sites: *Flickr*, *LastFm* e *Skoob*. Observa-se a sua ocorrência em ferramentas de redes sociais digitais a fim de possibilitam a organização e a recuperação da informação produzida exponencialmente nesses contextos.

Além do caráter de representar e recuperar a informação, a folksonomia apresenta uma perspectiva de representatividade e inclusão social. Seu comportamento de cooperação entre diferentes pessoas impulsiona a criação de plataformas sobre temáticas sociais pouco discutidas e urgentes. Nesse entendimento, um exemplo é o Dicionário de Gêneros – “Só quem sente pode definir”, que possibilita a construção colaborativa do conhecimento e apresenta terminologias que permitem representar a diversidade de gênero e oferece um potencial emancipatório (SANTOS *et al.*, 2017).

A folksonomia reflete relações flexíveis entre a linguagem e a comunidade, resultando em práticas e vocabulários diferentes daqueles presentes em discursos autorizados e normalizados (como os tesouros, as taxonomias e as ontologias).

4 Dicionário de Favelas Marielle Franco (Wikifavelas)

A plataforma Wikifavelas - Dicionário de Favelas Marielle Franco - foi lançada no ano de 2019, sendo um projeto criado pela Fundação Oswaldo Cruz (Fiocruz) e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), com o objetivo de ser um espaço para reunir conhecimentos e construir saberes acerca da experiência de quem vivencia o cotidiano das favelas e resgatar as memórias das favelas e da cidade do Rio de Janeiro.

A plataforma tem como personalidade homenageada, a vereadora carioca Marielle Franco, que foi eleita em 2016 com mais de 46 mil votos, pela coligação “Mudar é possível”, formada pelo Partido Socialismo e Liberdade (PSOL) e pelo PCB. Marielle Franco foi covardemente executada em 14 de março de 2018 em um bárbaro crime político.

Dentro da plataforma a seguinte definição é estabelecida sobre a mesma:

O projeto do Dicionário de Favelas tem por objetivo favorecer a preservação da memória e identidades coletivas dos moradores das favelas, como parte do nosso compromisso com a expansão da cidadania e do direito à cidade e criar

um espaço virtual que reúna o conhecimento sobre estes territórios de forma interdisciplinar e interinstitucional (WIKIFAVELAS, 2022, *on-line*).

Essa definição expõe uma iniciativa social que se caracteriza por ações que não visam o lucro ou interesse próprio, visam potencializar a cidadania por meio do registro e compartilhamento da informação e do conhecimento. Assim, a plataforma é um espaço virtual de produção colaborativa que reúne atores e conteúdos transdisciplinares.

Figura 1. Página principal do dicionário.



Fonte: Wikifavelas, 2022.

O dicionário é organizado em quatro eixos de análise: 1) Sociabilidade e Cultura; 2) Estado e Mercado; 3) Coronavírus e 4) Associativismo e Memória. Cada eixo tem o objetivo de desdobrar múltiplos assuntos de forma que o sujeito acesse as temáticas de interesse. Os desdobramentos se dão por categorias temáticas e posteriormente, os verbetes com manifestações autorais.

Os verbetes são pequenos textos ou comentários que registram o significado das palavras e termos, delimitando seu conceito. Etimologicamente a palavra vem do latim *verbu-*, «palavra» +-etc. Na plataforma Wikifavelas, os termos são explicitados como "Verbetes", que vem de verbo, o que se quer falar, enunciar, depor, demonstrar, denunciar, mostrar, afirmar, explicitar etc. O ato de se expressar tem a intenção de comunicar, de compartilhar com outros seu ponto de vista, suas ideias e suas histórias.

Para cada eixo de análises existem os verbetes, que atribuíram todo o conjunto de informações, que emanam demonstrações referentes aos territórios das favelas, bem como a cultura, realidades e vivências.

Os verbetes estão distribuídos pelas seguintes tipologias: verbetes textuais, verbetes de filmes e verbetes de livros. Cada um desses conjuntos possui distintos critérios de formatação, para preparar a publicação com suas diferentes narrativas.

O verbete textual necessita possuir de 400 a 4.000 palavras, autoria, bibliografia e palavras-chave. O verbete de filmes, solicita a sinopse do filme, ficha técnica, autoria, link dos *trailers* e palavras-chave. O verbete de livros deve conter a autoria, resumo, imagem

da capa, seguir a recomendação de citação e autoria, de acordo com o *Creative Commons* e palavras-chave.

Os diferentes tipos de verbetes evidenciam a riqueza do conhecimento produzido e compartilhado pela comunidade responsável pela sua criação e atualização.

Quadro 1. *Tipos e critérios de verbetes*

<i>Verbetes textuais</i>	<i>Verbetes de filmes</i>	<i>Verbetes de livros</i>
Devem conter entre 400 a 4.000 palavras.	Autoria no topo da página.	O(a) autor(a) do livro no topo.
Autoria do verbete na abertura do texto	Sinopse e ficha técnica	Resumo
Linguagem simples e objetiva	Link para <i>trailer</i> ou a referência original	Imagem da capa
Bibliografia no final do texto		Seguir a recomendação de citação e autoria (<i>Creative Commons</i>)
Podem conter fotos e/ou tabelas		
Palavras-chave	Palavras-chave	Palavras-chave

Fonte: Wikifavelas, 2022.

Os verbetes possuem *tags* que visam relacionar conceitos ao conteúdo e representá-lo, a fim de recuperar a informação. Representar um verbete a partir do seu conteúdo, significa retirar informações importantes, que visam organizar e recuperar o conteúdo futuramente.

As *tags* presentes são atribuídas pelos sujeitos, mediante uma etiquetagem livre, baseada na linguagem natural, dispensando o uso de vocabulários controlados conforme preconiza a folksonomia enquanto tipo de SOC.

Para que as informações sejam recuperadas de modo categórico, é necessário realizar um conjunto de regras e procedimentos que visam facilitar o acesso. Nessa perspectiva, a folksonomia consiste na descrição do conteúdo de forma concisa, por meio de palavras-chave que exercem o papel de pontos de acesso aos verbetes.

Como estratégia de busca, a plataforma apresenta alguns recursos que visam estruturar os termos para a pesquisa, para que o sujeito obtenha os melhores resultados, de maneira a suprir sua necessidade informacional.

A plataforma possui diferentes recursos estratégicos de busca sendo descritos como: a busca por categorias de verbetes; consulta por ordem alfabética; busca simples; busca avançada; navegação por categorias temáticas e

nuvem de *tags*. As principais estratégias de recuperação da informação no âmbito da Wikifavelas são descritas a seguir:

A. Consulta por ordem alfabética: nota-se que neste modelo de busca, é possível visualizar os verbetes existentes, em que o sujeito acessa cada assunto por ordem alfabética, que estão divididos por seu eixo de análise.

B. Busca Simples: a busca simples visa recuperar a propriedade e o conceito da palavra ou termo expresso no campo de busca. Em que se observa um campo simples de pesquisa, para escrever qualquer palavra ou termo.

C. Busca Avançada: a busca avançada visa personalizar a busca do usuário, para obter resultados mais específicos, temos como exemplos os: operadores de proximidade e truncamento, os campos da base dados e os operadores booleanos.

D. Nuvem de tags: a nuvem de *tags* ou *tag cloud* possui hiperlinks para a página do verbete selecionado, possibilita uma rápida e simples identificação dos temas para assim serem recuperados por meio da navegação.

Ressalta-se que Moreira e Romão (2009), informam que mediante a nuvem de *tags*, é possível condensar quais os assuntos e ou informações estão em foco dentro de um site.

Com o propósito de exemplificar e destacar as palavras-chave mais acessadas da plataforma por meio da nuvem de *tags*, será apresentado uma tabela com 15 *tags* que apresentaram um número elevado de acessos até o dia 10 de dezembro de 2022.

Tabela 1. *Palavras-Chave*

<i>Palavras-Chave</i>	<i>Número de Acesso</i>	<i>Número de Páginas</i>
Direito à Cidade	12.814	41
“Verbetes em Inglês”	2.545	37
Educação	1.880	33
Urbanização	1.396	22
Coletivo	970	37
Marielle Franco	953	20
Cidade de Deus	937	26
Filmes	893	68
Complexo do Alemão	898	40
Mangueira	897	40
Chacinas	844	67
Branquitude	845	67
Borel	807	20
Comunicação popular	768	31
Coronavírus	766	11

Fonte: Dados da pesquisa, 2023.

5 Diagnóstico da plataforma

Buscou-se desenvolver um quadro avaliativo da plataforma. Para isso foi necessária a observação

participante na plataforma, na ocasião foram recolhidos dados e informações, para análise.

Os dados apresentados a seguir possuem o período de observação do dia 01 de setembro de 2022 até o dia 03 de novembro de 2022, no qual constam a quantidade de 1.423 verbetes e 4.732 páginas, 2.704 arquivos acessados.

Em relação a quantidade de pessoas, a plataforma conta com 596 sujeitos colaboradores, o perfil dos sujeitos que colaboram compreende desde o cidadão comum até pesquisadores, institutos, núcleos e grupos parceiros.

A observação participante possibilitou, foi observar as vantagens e desvantagens de adotar um modelo colaborativo para produção de conhecimentos.

Vantagens:

- A. Dispensa profissionais especializados;
- B. Uso da linguagem natural;
- C. Fácil manuseio, dispensando técnicas;
- D. Construção coletiva;
- E. Captura dos vocábulos dos sujeitos.

Desvantagens:

- A. Ambiguidade;
- B. Erros ortográficos;
- C. Variações de números;
- D. Tags sinônimas;
- E. Tags duplicadas.

Percebe-se que as desvantagens identificadas acima ocorrem em função da natureza da linguagem natural. Contudo, uma outra desvantagem a ser pontuada é dado pelo caráter digital da plataforma que exclui como potenciais colaboradores pessoas que não possuem acesso à Internet, bem como competências para utilizá-la.

A partir das informações e dados coletados elaborou-se ainda um quadro com perguntas e apontamentos, para identificar o perfil dos usuários, os tipos de interação entre os usuários na plataforma, linguagem e regras e critérios dos verbetes e *tags*.

Quadro 2. Análise da plataforma

Perguntas	Sim	Não
Os sujeitos podem Criar um verbete?	X	
Os sujeitos podem editar um verbete?	X	
Os sujeitos podem discutir um verbete?	X	
A pessoa que desenvolveu o site influencia a <i>tag</i> ?		X

Perguntas	Sim	Não
Possuem editores?	X	
Possui revisão de indexação?		X
Possui regras editoriais?	X	
Existe critério para os verbetes?	X	
Linguagem controlada?		X
Linguagem natural?	X	
Possui sugestão de termos?		X
Existe uma regra de etiquetagem?	X	
O verbete pode ser etiquetado pelas mesmas palavras-chave?		X
Existem limites de palavras-chave para o verbete?		X
Folksonomia ampla?	X	
Folksonomia restrita?		X
Possui nuvem de <i>tags</i> ?	X	
O verbete precisa ser original?		X
Dentro da plataforma é possível criar um grupo social?		X

Fonte: Dados da pesquisa, 2023.

De posse dos resultados encontrados, foi destacado como relevante o resgate dos conteúdos dos verbetes por meios da aplicação de *tags* atribuídas livremente pelos sujeitos participantes da comunidade.

A ação de criar conteúdo e empregar *tags* para serem pontos de acesso e recuperação de verbetes, fortalece a construção de conhecimento coletivo na plataforma Wikifavelas por possibilitam abordar diferentes aspectos do verbete.

Por se tratar de uma plataforma colaborativa, qualquer pessoa, com acesso à Internet, pode contribuir com o seu conteúdo, basta apenas criar uma conta e criar conteúdo sobre diferentes assuntos. Foi perceptível também a possibilidade de editar um verbete de modo a corrigir ou atualizar o conhecimento explicitado e compartilhado no dicionário.

Dessa forma, por meio da análise e interpretação dos resultados, pôde-se conhecer como a comunidade busca, constrói e contribui para a representação do conhecimento por meio da classificação popular: a folksonomia.

Ao contrário de ambientes como o *Twitter* em que a indexação social pode ser feita com a participação de perfis falsos e robôs, na Wikifavelas a contribuição é de

fato realizada livremente por meio de uma linguagem natural desempenhada pelos sujeitos, como contextualizam Assis e Moura (2013) ao caracterizarem a folksonomia:

Pode ser descrita como uma inovação que explora o potencial das redes sociais na organização e no compartilhamento dos recursos informacionais. Desse modo, ela agrega as manifestações da linguagem contextualizada e, por vezes, caótica de sujeitos em colaboração (ASSIS; MOURA 2013, p. 87)

O potencial enquanto tecnologia social é notório, uma vez que o Dicionário de Favelas Marielle Franco - Wikifavelas desempenha um papel social, de preservação de memória, de produção de conhecimentos acerca da vivência das favelas, esse ato é uma manifestação plural, de transformação e reflexão.

A memória é essencial na sociedade, é uma parte importante da experiência de vida, e as plataformas colaborativas podem ser uma ferramenta valiosa para ajudar a preservá-la e compartilhá-la.

Os autores Felipe, Medeiros e Pinho (2020, p. 133) explicam que “[...] é através da memória que se produzem os hábitos e costumes, por via das experiências vividas, sendo isso fundamental para a formação dos grupos sociais”. Neste contexto para representar a memória e armazená-la, existem diferentes recursos, como a música, imagens estáticas ou em movimento e também os registros de verbetes que podem ser explorados.

Existem muitas outras maneiras diferentes de se usar plataformas colaborativas para compartilhar e preservar memórias como as ferramentas de redes sociais digitais que possibilitam aos usuários o compartilhamento de fotos e vídeos de suas experiências.

A Wikifavelas remete ainda ao pioneirismo do modelo colaborativo da Wikipedia, que é uma plataforma colaborativa que permite que as pessoas criem e editem páginas sobre uma ampla variedade de tópicos, incluindo eventos históricos e figuras culturais.

A plataforma Wikifavela, mapeia, registra e documenta informações e conhecimentos construídos pelos sujeitos atrelados a um contexto social, sendo possível escrever sobre um determinado assunto e aplicar de palavras-chave para classificar e organizar informações.

Nela as pessoas podem criar uma rede de conexões entre diferentes conteúdos e ideias, permitindo que outros usuários descubram e explorem essas conexões, contribuindo assim para a produção de uma memória coletiva.

Adler (2009) *apud* Santos e Albuquerque (2021, p. 390), apontam que são:

[...] duas características importantes da representação colaborativa do conhecimento: a) a existência de um grau de consenso por parte dos usuários quanto à etiquetagem dos

recursos; b) alcance visível e negociável das expressões de vozes minoritárias nos ambientes colaborativos, uma vez que as tags atribuídas são estabelecidas pelas convenções sociais da linguagem e da ideologia dos usuários. (ADLER, 2009) *apud* (SANTOS; ALBUQUERQUE, 2021, p. 390)

Na representação colaborativa do conhecimento, diferentes comunidades podem discutir, mapear e registrar conteúdos para o registro da memória, é visto assim, que, as plataformas colaborativas podem ser uma ferramenta valiosa para preservar e compartilhar memórias, seja em nível pessoal, coletivo/comunitário ou histórico.

6 Conclusão

Os resultados da pesquisa evidenciam que, a comunidade analisada é heterogênea, possibilitando aos diversos sujeitos informacionais colaborarem na plataforma o que implica na amplificação da diversidade de vivências e visões de mundo.

Verificou-se que a partir da versatilidade da indexação social e conseqüentemente da folksonomia, os sujeitos podem colaborar e atribuir *tags*, dispensando um vocábulo controlado, o que possibilita contemplar melhor a garantia cultural enquanto um dos princípios fundamentais a qualquer tipo de SOC.

Este estudo traz à luz as práticas colaborativas de informações e conhecimentos de uma comunidade pouco explorada no campo.

Espera-se que este trabalho possa contribuir para despertar o interesse de outros pesquisadores da área em estudar esta comunidade, a fim de que o Dicionário de Favelas Marielle Franco - Wikifavelas, um sistema híbrido de organização do conhecimento, se torne cada vez mais pauta em outras discussões e estudos científicos.

Percebeu-se que a hibridização dos SOC possibilita maior expressividade e versatilidade semântica na representação da informação e do conhecimento e amplia as opções de recuperação das informações, seja via *searching* ou via *browsing*.

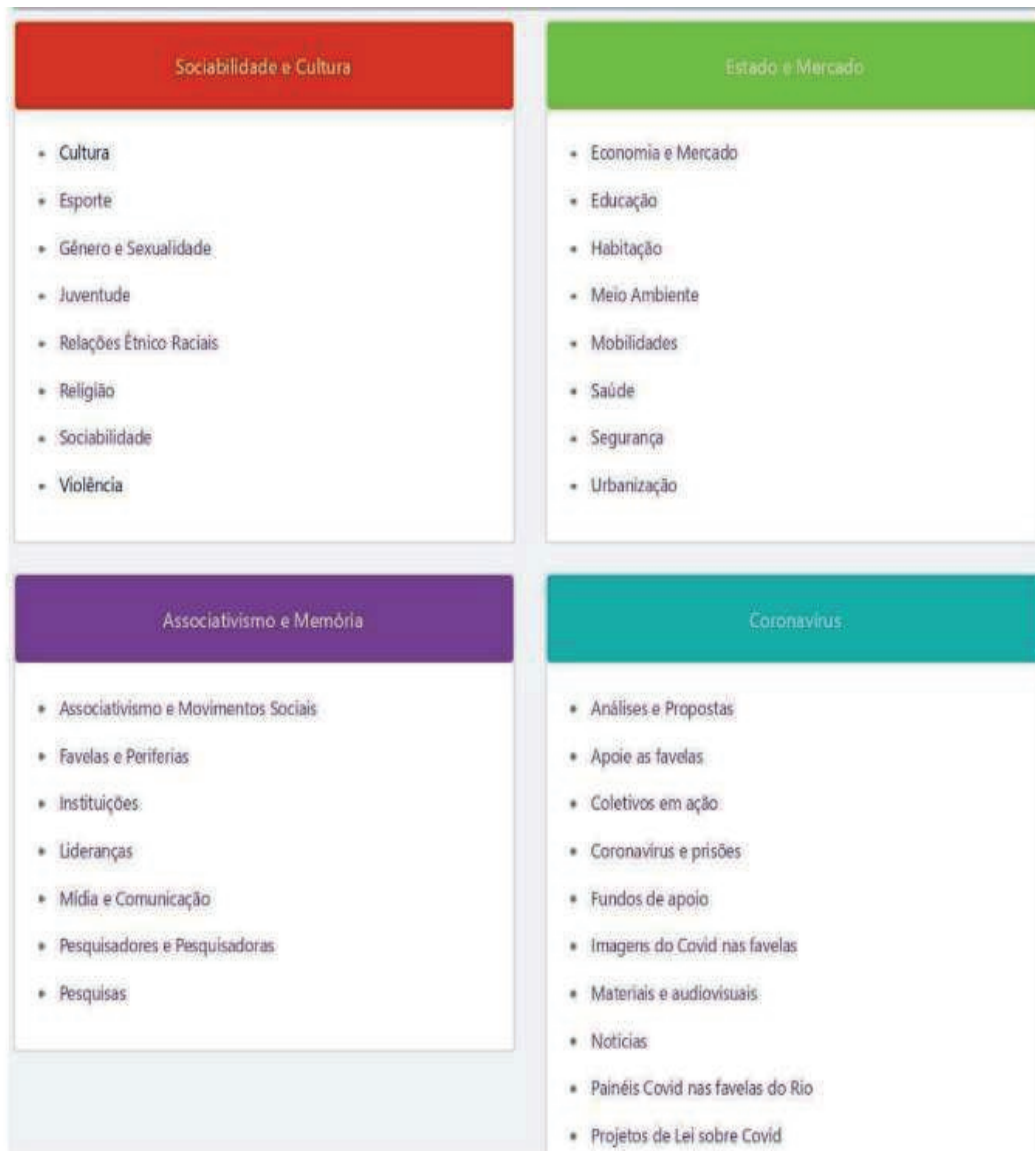
Como sugestões para trabalhos futuros indica-se o aprofundamento dos aspectos não abordados neste trabalho como a construção da memória coletiva da comunidade sob uma perspectiva que possa identificar os fatores sociais e culturais que influenciam essa prática, ou seja, não apenas numa perspectiva informacional, mas também sociológica.

Referências

ASSIS, J. de.; MOURA, M. A. Folksonomia: a linguagem das tags. Encontros Bibli: **Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 18, n. 36, p. 85-106, 2013.

- BLATTMANN, U.; SILVA, F. C. C. Colaboração e interação na Web 2.0. **Revista ACB: Biblioteconomia em Santa Catarina**, v.12, n.2, p. 191-215, jul./dez., 2007.
- BRATKOVÁ, E.; KUCEROVÁ, H. Knowledge Organization Systems and Their Typology. **Revue of Librarianship**, v. 25, n. 2, 2014. Disponível em: https://full.nkp.cz/nkkr/knihovna142_suppl/1402sup01.htm. Acesso em: 28 mar, 2023.
- CATARINO, M. E.; BAPTISTA, A. A. Folksonomias: características das etiquetas na descrição de recursos da web. **Inf. Inf.**, Londrina, v. 14, n. esp, p. 46 – 67. 2009.
- DAHLBERG, Ingetraut. Knowledge Organization: A New Science? **Knowledge Organization**, v. 33, n. 1, 2006.
- FELIPE, C. B. M.; MEDEIROS, W. O. de.; PINHO, F. A. Representação colaborativa de registros imagéticos da memória social: uma reflexão sobre a página the commons na plataforma Flickr. **Páginas a&b**, n.13, p.131-141. 2020.
- FUJITA, M. S. L., BOCCATO, V. R. C; RUBI, M. P.; Gonçalves, M. C. **A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias**. Um estudo de observação do contexto sociocognitivo com protocolos verbais [online]. São Paulo: Editora UNESP; São Paulo: Cultura Acadêmica, 2009. 149 p.
- GUEDES, R. M.; DIAS, E. J. W. Indexação social: abordagem conceitual social **indexing**: conceptual approach. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 15, n. 1, p. 39-53, 2010. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/72859>. Acesso em: 30 mar. 2023.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 2008.
- GONÇALVES, J. L. C. S.; ASSIS, J. H. A indexação social enquanto prática de representação colaborativa da informação imagética: a construção da memória na plataforma Flickr. **Revista Conhecimento em Ação**, v. 1, n. 2, 2016.
- HALBWACHS, M. **A memória coletiva**. São Paulo: Vértice, 1990.
- KRAP, Juliana. O Dicionário de favelas Marielle Franco é lançado na Fiocruz. **Portal Fiocruz**, 2019. Disponível em: <https://portal.fiocruz.br/noticia/dicionario-de-favelas-marielle-franco-e-lancado-na-fiocruz>. Acesso em: 20 set. 2022.
- LIMA, J. L. O.; ALVARES, L. Organização e representação da informação e do conhecimento. In: ALVARES, L. (Org.) **Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações**. São Paulo: B4, 2012.
- LIRA, S.; LLARENA, M. A.; SILVA, E. B. F.; MORENO, D. H. da S. Comunidades Virtuais de Prática em **evidência na gestão social do conhecimento**. **Revista Folha de Rosto**, n. 2, v. 7, p. 98-122, 2021.
- MINAYO, M. C. S. (Org.). **Pesquisa Social: teoria, método e criatividade**. Petrópolis, RJ: Vozes, 2003.
- MOREIRA, V. L.; ROMÃO, L. M. S. O funcionamento discursivo das nuvens de tags na rede eletrônica: sentidos sobre a capitolina. **DataGramZero**, v. 10, n. 6, 2009. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/7028>. Acesso em: 03 nov. 2022.
- SANTOS, R. F. dos. **Modelos colaborativos de indexação social e a sua aplicabilidade na base de dados referencial de artigos de periódicos em Ciência da Informação (BRAPCI)**. Dissertação (Mestrado em Ciência da Informação) - Universidade Federal de Pernambuco, Recife, 2016.
- SANTOS, R. F. dos, *et al.* A Representação colaborativa da informação e a construção de linguagens documentárias sobre diversidade de gêneros: análise das contribuições do dicionário de gêneros - “só quem sente pode definir”. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18, Marília, 2017. Anais [...], Marília: Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2017.
- SANTOS, R. F. dos.; ALBUQUERQUE, M. E. B. C. de. Aplicabilidade da Folksonomia na Construção e Atualização de Sistemas Híbridos de Organização do Conhecimento. **Ponto de Acesso**, v. 15, n.3. 2021.
- STURTZ, D. N. Communal categorization: the folksonomy. 2004. Disponível em: <http://www.davidsturtz.com/drexel/622/communal-categorization-the>. Acesso em 28 mar. 2023.
- O'REILLY, T. **What Is Web 2.0: design patterns and business models for the next generation of software**. 2005. Disponível em: <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>. Acesso em 28 mar. 2023.
- WIKIFAVELAS **Dicionário de favelas Marielle Franco**, 2022. Disponível em: https://wikifavelas.com.br/index.php/Dicion%C3%A1rio_de_Favelas_Marielle_Franco. Acesso em 28 mar. 2023.
- ZENG, M. L. Knowledge Organization Systems (KOS). **Knowledge Organization**, v. 35, n. 2/n.3, 2008.

Anexo A – Estrutura taxonômica da Wikifavelas



Fonte: Wikifavelas, 2022.

Representação do conhecimento na prisão: palavras e significados em jogo

Knowledge representation in jail: words and meanings at stake

Amabile Costa (1), Rodrigo de Sales (2)

Universidade Federal de Santa Catarina (UFSC), R. Eng. Agrônomo Andrei Cristian Ferreira, s/n - Trindade,
Florianópolis - SC, 88040-900 (1) amabilecosta.m@gmail.com, (2) rodrigo.sales.s@gmail.com

Resumo

O **objetivo** deste estudo é investigar, com base no vocabulário utilizado dentro das prisões, uma representação do conhecimento no contexto carcerário. A **abordagem metodológica** está caracterizada como descritiva e exploratória, com o intuito de registrar e descrever informações sobre uma determinada população carcerária no que se refere à invenção e à ressignificação de palavras. Os dados foram coletados por meio de técnicas etnográficas, especificamente por meio da observação *in loco* e entrevistas (foram entrevistadas vinte e seis pessoas que estão presas na Penitenciária Masculina de Florianópolis, SC). **Resultado:** quatrocentos e sessenta termos foram identificados como fruto do processo de ressignificação ou invenção de palavras, das quais algumas serão apresentadas neste trabalho. Constatou-se a criação de um vocabulário/linguagem próprio da prisão que, para além da comunicabilidade, viabiliza resistência, proteção e sobrevivência das pessoas presas. Verifica-se, assim, uma forma de representação do conhecimento baseada em trocas de significados (ou ressignificações) ou mesmo invenção de palavras compreensíveis apenas às pessoas presas, e que serão aqui representadas por meio de doze categorias.

Palavras-chave: Organização do conhecimento. Representação do conhecimento. Linguagem. Prisão.

Abstract

The aim of this study is to investigate, based on the vocabulary used within prisons, a knowledge representation in the jail context. The **methodological** approach is characterized as descriptive and exploratory, with the aim of registering and describing information about a certain jail population with regard to the invention and redefinition of words. Data were collected through ethnographic techniques, specifically through on-site observation and interviews (twenty-six people who are imprisoned in the Male Penitentiary of Florianópolis, SC were interviewed). **Result:** four hundred and sixty terms were identified as a result of the process of resignification or invention of words, some of which will be presented in this paper. It was verified the creation of a jail vocabulary/language that, in addition to communicability, enables resistance, protection and survival of prisoners. There is thus a way knowledge representing based on exchanges of meanings (or resignifications) or even the invention of words understandable only to prisoners, and which will be represented here through twelve categories.

Keywords: Knowledge organization. Knowledge representation. Language. Jail.

Resumen

El **objetivo** de este estudio es investigar, a partir del vocabulario utilizado dentro de las prisiones, una representación del conocimiento en el contexto penitenciario. El enfoque **metodológico** se caracteriza por ser descriptivo y exploratorio, con el objetivo de registrar y describir información sobre una determinada población penitenciaria en lo que se refiere a la invención y redefinición de palabras. Los datos fueron recolectados a través de técnicas etnográficas, específicamente a través de observación *in situ* y entrevistas (fueron entrevistadas veintiséis personas que se encuentran encarceladas en la Penitenciaria de Varones de Florianópolis, SC). **Resultado:** cuatrocientos sesenta términos fueron identificados como resultado del proceso de resignificación o invención de palabras, algunos de los cuales serán presentados en este trabajo. Se verificó la creación de un vocabulario/lenguaje penitenciario que, además de comunicabilidad, posibilita la resistencia, protección y sobrevivencia de los presos. Hay así una forma de representar el conocimiento a partir de intercambios de significados (o resignificaciones) o incluso de la invención de palabras comprensibles sólo para los presos, y que serán representadas aquí a través de doce categorías.

Palabras clave: Organización del conocimiento. Representación del conocimiento. Idioma. Prisión.

1 Introdução

Este estudo investiga formas de representação do conhecimento no ambiente prisional a partir dos jogos e dos usos das palavras inventadas ou ressignificadas pelas pessoas presas, levando em conta os aspectos que dizem respeito às pessoas presas que fazem parte do processo comunicativo; as peculiaridades no que se refere ao vocabulário utilizado nas prisões; como este vocabulário pode estar organizado e; como as palavras podem ser empregadas e percebidas diante de cada acontecimento.

Contextualizando o universo desta pesquisa, evidencia-se que em diversos países, especialmente na Europa e na América, o sistema prisional e suas estruturas estavam em progressão, cada um com seu contexto histórico e suas peculiaridades. Os Estados Unidos da América (EUA), por exemplo, experimentaram e reinventaram maneiras de punição e novos sistemas judiciários, sendo os mais famosos os da Filadélfia (Pensilvânia) e de Auburn (Alabama), ambos focados no trabalho, na mão de obra das pessoas encarceradas, visando uma “função punitiva ou organizada segundo padrões produtivos e competitivos” (MELOSSI; PAVARINI, 2017, p. 17). Com isso, “o debate sobre os méritos e os defeitos dos dois sistemas têm lugar num terreno predominantemente ideológico e moralista”. Os sistemas da Filadélfia e de Auburn foram reproduzidos em diversos lugares do mundo como, por exemplo, no Brasil (Ibidem).

O cárcere é um arquétipo organizacional vinculado ao poder disciplinar e se estabelece como um “modelo que assume aspectos quase metafísicos e que perde, exatamente devido à sua generalização e abstração, uma dimensão histórica precisa” (MELOSSI; PAVARINI, 2017, p. 12).

A privação de liberdade “tornou-se a principal forma de punição no mundo ocidental no exato momento em que o fundamento econômico da casa de correção foi destruído pelas mudanças industriais” (RUSCHE; KIRCHHEIMER, 1939, p. 111; MOTTA, 2011).

No Brasil, como em outros lugares no mundo, as prisões são frutos “de toda sua conjuntura social, econômica, cultural e religiosa, conjunto que dita o modo de produção e as relações sociais” (OLIVEIRA, 2013, p. 39). Ao longo dos anos, o sistema prisional brasileiro perpassou por diversas reformas até chegar na estrutura atual.

A partir disso, a linguagem utilizada na prisão faz parte da rotina desses estabelecimentos, na comunicação dos sujeitos, nas definições de legislações informais, “códigos, repertórios linguísticos, comportamentos, práticas culturais e formações de grupos” (OLIVEIRA, 2013, p. 18), representando as formas de ser, estar e agir, e a sociabilidade presente nas prisões.

A linguagem está estruturada por meio de signos. Esses signos são palavras e gestos que representam objetos de um determinado contexto, sendo assim, cada palavra dispõe de um significado. Ou seja, “as palavras do nosso cotidiano já chegam até nós carregadas de sentidos que não sabemos como se constituíram e que, no entanto, significam em nós e para nós” (ORLANDI, 2020, p. 18).

Dado o exposto até aqui, o objetivo deste trabalho é investigar, com base no vocabulário utilizado dentro das prisões, a representação do conhecimento no contexto carcerário.

Não é descabido afirmar que, tradicionalmente, a organização acadêmica do conhecimento vem lançando olhar para formas de representação do conhecimento em espaços considerados não-marginalizados. O foco desta pesquisa é evidenciar que há também nos espaços marginalizados formas de representação do conhecimento que lhes são peculiares. Assim, acreditamos ser possível ampliar o debate da representação do conhecimento e, também, promover um olhar mais crítico ao sistema em que as pessoas privadas de liberdade estão submetidas.

Representar o conhecimento no campo da Organização do Conhecimento parece convergir com a ideia essencialista de que os conceitos são os verdadeiros representantes do conhecimento ou, como prefere Dahlberg (1978), as verdadeiras unidades de conhecimento. Por essa razão, é bastante difundida a ideia de que para se organizar o conhecimento é necessário partir de conceitos descritos e de suas relações formais (BRASCHER; CAFÉ, 2010). Não queremos, obviamente, contrapor esta perspectiva, mas gostaríamos de nos apoiar também no aspecto etimológico de que representar corresponde a ‘apresentar novamente’. No caso do ambiente prisional, este processo de apresentar novamente demanda manobras vocabulares distintas das que estamos habituados, pois não são amparadas por campos de conhecimentos formais. No entanto, a nosso ver, tais manobras, ou jogos de palavras, não deixam de ser formas de representações de conhecimentos.

Entendemos que respeitar e reconhecer que a população carcerária produz e representa seus próprios conhecimentos é uma das formas de se promover uma organização do conhecimento mais atenta às diversas realidades sociais.

2 Linguagem na prisão e representação do conhecimento

A comunicação nas prisões é realizada para expressar seu cotidiano e estruturada em um contexto baseado em gírias e ressignificação dos termos, criados pelos detentos para descrever os acontecimentos (GOFFMAN, 2015). A cada momento, em cada

estabelecimento ou em cada região, a linguagem é alterada para adaptação à sua realidade, para sua compreensão, interpretação e organização de mundos.

A linguagem proporciona a interação em sociedade, e cada grupo social conta com características específicas que podem ser distinguidas pelo modo que a pessoa se comunica. A língua está em constante mutação, visto que a “necessidade de buscar novas expressões para designar novos objetos, conceitos e formas de relação social” (MARTELOTTA, 2018, p. 19) são perceptíveis ao longo das décadas, e estão presentes em aspectos culturais de cada grupo.

Goffman (2015) afirma que, nas prisões, gírias institucionais são criadas pelos internados para descrever acontecimentos. Uma parte da equipe dirigente conhece essa linguagem e com isso, comunica-se com os internos.

As mensagens nas instituições totais, circulam por meio de sistemas ocultos de comunicação. Um desses sistemas são, técnicas de falar sem mover os lábios e sem olhar para a pessoa que a mensagem está sendo direcionada (GOFFMAN, 2015).

Assim, a linguagem nas prisões é adaptada a partir dos contextos e ambiências em que as pessoas estão inseridas, ou seja, os termos possuem determinados significados e são ressignificados diante das circunstâncias impostas. Geralmente, a rotatividade no sistema prisional é vasta, com isso, os vocabulários vão sendo ampliados e as palavras possuem mais de um significado, dependendo da situação utilizada ou região. Os códigos são modificados de acordo com a necessidade ou descoberta de seus significados.

Tomamos como exemplo a palavra ‘jumbo’ no contexto prisional. Jumbo tem como significado “bolsa transparente em que os visitantes levam os alimentos e os produtos de limpeza e higiene” (ARAUJO, 2019, p. 19). Para as pessoas inseridas de alguma forma, seja como visitante, como funcionário ou como detento, no universo do sistema prisional, a palavra jumbo, quando pronunciada, será compreendida. Essa palavra e seu significado estão integrados a um jogo de linguagem estabelecido pelas comunidades carcerárias e é utilizada em um determinado contexto, o sistema prisional.

A linguagem compreende diversas essências e, conseqüentemente, várias formas de uso. Para compreender essa usabilidade, é preciso reconhecer a multiplicidade das línguas e como esses fenômenos estão relacionados uns com os outros de inúmeras maneiras. Com isso, o uso da palavra está associado ao seu significado, ou seja, entende-se o significado de uma expressão quando se compreende as regras de seu uso (GRAYLING, 2001).

Isso ocorre, por exemplo, no aprendizado de uma língua, pois, normalmente, se aprende as palavras por

meio de seus significados e, principalmente, as regras de como utilizar tais palavras aprendidas. Portanto, “a palavra *conformidade* e a palavra *regra* são aparentadas, são primas. Se ensino a alguém o uso de uma delas, ele aprende também o uso da outra” (WITTGENSTEIN, 1975, p. 96, grifo nosso).

A Organização do Conhecimento (OC) possui enquanto objeto de estudo o conhecimento, e está estruturada em duas ações, organizar e representar. A representação do conhecimento está constituída “numa estrutura conceitual que representa modelos de mundo, os quais [...] permitem descrever e fornecer explicações sobre os fenômenos que observamos” (MOREIRA, 2018, p. 29).

Representar o conhecimento “reflete um modelo de abstração do mundo real, construído para determinada finalidade” (MELO, 2010, p.181). É justamente essa abstração do mundo real construída para determinada finalidade que nos inspira a compreender um pouco da representação do conhecimento construída nos ambientes prisionais.

3 Procedimentos metodológicos

Este estudo está caracterizado como descritivo e exploratório, visto que o intuito é registrar e descrever informações sobre uma determinada população ou fenômeno sem interferir nos dados. Geralmente, a pesquisa descritiva usa técnicas de questionário e observação (PRODANOV; FREITAS, 2013). Descrevemos aqui aspectos e ações presentes nas instituições penais relacionadas às palavras utilizadas especificamente pela população carcerária.

A pesquisa exploratória proporciona “maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a construir hipóteses. A grande maioria dessas pesquisas envolve [...] entrevistas com pessoas que tiveram experiências práticas com o problema do pesquisado” (GERHARDT; SILVEIRA, 2009, p. 35).

Baseia-se, também, em pesquisa bibliográfica, ou seja, “elaborada a partir de material já publicado, constituído principalmente por: livros, revistas, publicações em periódicos e artigos científicos, jornais, boletins, monografias, dissertações, teses”, entre outros (PRODANOV; FREITAS, 2013, p. 54). A abordagem do problema está imbricada na pesquisa qualitativa, que “considera que há uma relação dinâmica entre o mundo real e o sujeito [...] interpretação dos fenômenos e a atribuição de significação são processos básicos desta pesquisa” (Idem, p. 70).

O universo desta pesquisa está em torno da Penitenciária Masculina de Florianópolis, com vinte e seis sujeitos participantes e 460 termos coletados. Para a coleta de dados, utilizamos técnicas provenientes da

etnografia. A etnografia proporciona, por meio de observação, entrevista e anotações em diário de campo, o conhecimento do cotidiano das instituições, a linguagem e os sentimentos dessa comunidade. Adentrar à Penitenciária de Florianópolis, enquanto pesquisadores, foi fundamental para a coleta dos dados, pois ao estar em companhia dos sujeitos da pesquisa tivemos condições de compreender os aspectos citados anteriormente.

Para identificar as palavras modificadas e utilizadas pelas pessoas privadas de liberdade, foi elaborado um questionário pré-estabelecido com cinco perguntas abertas, sendo elas: 1- Como é a comunicação utilizada no dia-a-dia da prisão? 2- Você sente que os significados das palavras foram alterados por conta de seu convívio aqui? Ou seja, existem palavras que lá fora você utilizava com significados diferentes dos significados que agora você usa? 3- Você poderia dar algum exemplo de palavras e significados que você começou a usar depois que entrou na prisão? 4- Tem algo a mais que você queira falar sobre o vocabulário da prisão? 5- Uma música que você gostaria que eu escutasse depois. O questionário foi aplicado presencialmente por meio de entrevistas realizadas com vinte e seis pessoas privadas de liberdade.

Para o registro dessas informações, utilizamos um diário de campo, visto que a utilização de gravadores, muito provavelmente, provoca um sentimento de intimidação e um desconforto para o entrevistado (BIONDI, 2018), especialmente em um ambiente tão inóspito e opressor.

As anotações realizadas em diários de campo proporcionam a transformação “de acontecimentos passados, que existe apenas em seu próprio momento de ocorrência, em um relato, que existe em sua inscrição e que pode ser consultado novamente” (GEERTZ, 2019, p. 14).

Ressalta-se que a interpretação e compreensão dos dados são individuais, visto que, cada pesquisador tem um entendimento e observação distinto de momentos, contextos, leituras e trajetórias de vida. Portanto, estar em campo é interpretar qualquer coisa como, por exemplo, uma situação, um poema, estórias, pessoas, rituais e instituições (GEERTZ, 2019).

Como mencionado anteriormente, para esta pesquisa as técnicas utilizadas foram observação e entrevista. A observação é “o ato de perceber as atividades e os inter-relacionamentos das pessoas no cenário de campo através dos cinco sentidos do pesquisador” (ANGROSINO, 2009, p. 56).

A observação é composta por diversas particularidades, algumas delas são: realizar anotações sobre o campo rotineiramente, explicar durante a coleta de dados o contexto em que os dados foram coletados, descrição dos participantes, dos seus comportamentos e

interações ao decorrer da pesquisa, registrar conversas formais e informais de uma maneira cronológica e identificar os participantes por meio de códigos, assim, preservando sua identidade e auxiliando na recuperação das informações desejadas para posterior escrita (ANGROSINO, 2009).

Já a entrevista, “é um processo que consiste em dirigir a conversação de forma a colher informações relevantes” (ANGROSINO, 2009, p. 61). As perguntas podem ser trocadas ao longo do desenvolvimento da pesquisa e das necessidades encontradas. A entrevista se torna uma parceria entre o pesquisador e o membro da comunidade que pode, inclusive, auxiliar no remanejamento das questões.

O objetivo da entrevista é “sondar significados, explorar nuances, capturar as áreas obscuras que podem escapar às questões de múltipla escolha que meramente se aproximam da superfície de um problema” (ANGROSINO, 2009, p. 62).

Quadro 1. *Compilação dos procedimentos metodológicos*

<i>Característica</i>	<i>Procedimento metodológico</i>
Descritiva	Questionário e observação
Bibliográfica	Atendendo aos objetivos desta pesquisa
Qualitativa	Interpretação a partir da observação e entrevistas realizadas
Entrevista	Questionário com perguntas abertas, anotações em diário de campo
Comitê de ética	Aprovado no Comitê de Ética da Universidade Federal de Santa Catarina e no Comitê de Ética da Secretaria de Administração Prisional

Fonte: Elaborado pelos autores (2023).

Para melhor compreensão do andamento e das etapas da pesquisa em campo, relatamos aqui algumas experiências vivenciadas. Os trâmites foram iniciados em novembro de 2021 com uma reunião na Penitenciária de Florianópolis, para expor a intenção da pesquisa e para acertar os detalhes do seu desenvolvimento na unidade. Nesta reunião foram definidos o dia do início da pesquisa, o setor que a acolheria, os limites determinados pela unidade prisional, os procedimentos de segurança e quem seria o responsável pelo projeto dentro da unidade.

A observação em campo e as entrevistas iniciaram no dia 29 de novembro de 2021 e finalizaram no dia 02 de fevereiro de 2022, alocada ao setor de saúde, juntamente com a Unidade Básica de Saúde (UBS) da Penitenciária Masculina de Florianópolis, ambiente, diga-se por sinal, com pessoas acolhedoras e dispostas a auxiliar no desenvolvimento da pesquisa. A Unidade ficou à disposição da pesquisa, mas a circulação dos pesquisadores no ambiente prisional esteve vinculada ao acompanhamento realizado pelos profissionais de saúde. Dado o ensejo, a parte da pesquisa dedicada à observação foi levada a cabo durante os atendimentos

que os enfermeiros e técnicos em enfermagem realizavam ao longo das semanas em toda à Penitenciária.

Com a etapa das observações, percebeu-se o andamento da unidade e dos procedimentos administrativos, a estrutura da Penitenciária Masculina de Florianópolis, os protocolos de saúde e a linguagem utilizada dentro daquele ambiente prisional. As observações em diferentes prédios da Instituição foram realizadas no período vespertino, e no período matutino foram realizadas as entrevistas. Geralmente, chegamos no ambiente da entrevista e iniciamos uma conversa com o plantonista dos policiais penais que estão responsáveis pelo dia e com o chefe da mesa. Cada prédio e cada plantonista determina como será o andamento da entrevista.

Por exemplo, na Padaria da Penitenciária, a conversa com os sujeitos da pesquisa foi realizada a partir da cela que fica acoplada na unidade. Quando chegamos pela primeira vez nos apresentamos, explicamos a pesquisa, como funcionava, assinamos o Termo de Consentimento Livre e Esclarecido (TCLE) e iniciamos a entrevista. Já na parte Interna, outro prédio da instituição destinado às pessoas que cumprem o regime fechado, dependendo do plantão, as entrevistas aconteceram na “passagem”, um local onde as pessoas presas passam para ir às suas celas, ou na parte da frente do Complexo Penitenciário, onde ocorriam os atendimentos de saúde e jurídico. Entretanto, outros plantonistas optaram por encaminhar a pessoa à sala da UBS, que fica na parte frontal, para que a entrevista transcorresse tranquilamente.

Nenhum policial penal acompanhou as entrevistas, apenas tiraram e colocaram as pessoas em suas celas. No entanto, os pesquisadores e os entrevistados ficaram separados por grades, fossem elas por “canudos” (1), fossem por grades nas janelas ou por grades na sala da UBS. Dependendo da galeria da Penitenciária de Florianópolis, conversávamos com o Voz (2), explicávamos o intuito e o desenvolvimento da pesquisa e ele organizava quem gostaria e estaria disponível para participar.

Depois desses procedimentos, uma pessoa por vez foi entrevistada. Houve entrevistas que duraram 10 minutos e entrevistas que duraram três horas e meia, e essa diferença dependia muito da disponibilidade da pessoa entrevistada, dos procedimentos administrativos que estavam sendo desenvolvidos no prédio e da compreensão do plantonista. Em seguida, apresentamos as perguntas e a conversa fluíu com os entrevistados respondendo aos poucos e com calma para que pudéssemos anotar suas respostas. Entrevistamos pessoas que estão há nove meses no sistema prisional e pessoas que estão a quinze anos.

As entrevistas mais longas foram além das perguntas que estavam estipuladas no questionário. Os

entrevistados contaram suas experiências de quando estavam em liberdade, suas histórias de vida, sobre seus familiares e sobre as atividades que realizavam em suas celas para passar o tempo.

Ressalta-se que essa pesquisa passou pelo Comitê de Ética da Universidade Federal de Santa Catarina, bem como pelo Comitê de Ética da Secretaria de Administração Prisional do Estado de Santa Catarina.

4 Discussões e resultados

A partir das entrevistas realizadas, percebeu-se que a linguagem utilizada na prisão é fruto não somente do processo de comunicação, mas também fruto de uma salvaguarda de sobrevivência, resistência e proteção.

De acordo com as respostas coletadas, a criação da linguagem utilizada na prisão ocorre para que a administração prisional não consiga identificar o assunto tratado nas conversas ou nas “pipas” (3) enviadas. Outro aspecto perceptível nas respostas, envolve a utilização da linguagem para compreender o funcionamento e as regras das pessoas privadas de liberdade, bem como para não virar chacota entre os próprios detentos. Dependendo da palavra utilizada dentro da prisão, seu significado não corresponde ao que estamos acostumados a identificar. Esses são dois dos motivos que nos leva a considerar que a linguagem na prisão é uma forma de sobrevivência e resistência.

Com a descoberta das palavras, outras formas de comunicação começaram a surgir, tais como: libras da cadeia (o que os sujeitos desta pesquisa identificam como linguagem de sinais) e; troca de sílabas, por exemplo, “você” torna-se “cêvo”.

Diante disso, alguns das palavras que compõe a linguagem na prisão envolvem a representação imagética da palavra, como por exemplo: “a que brilha” é atribuído ao acendedor. Outras são referentes às iniciais das palavras, “PH” é papel higiênico. Além disso, algumas palavras em inglês são utilizadas, como “Boot” significando “tênis”.

Para além disso, algumas palavras são utilizadas por grupos específicos dentro da prisão. A comunidade LGBTQIAP+ usa o vocabulário Pajubá, e nesse vocabulário, “Aqué” significa “Dinheiro”. Os funcionários da Penitenciária utilizam termos específicos, por exemplo: “Alta” significa dar alta para que as pessoas voltem as suas celas de origem depois da saída de sete dias.

As 460 palavras identificadas como formas específicas de se representar o conhecimento produzido no cotidiano da Penitenciária serão integralmente apresentadas em publicação que está sendo produzida pelos autores desta pesquisa. Mas, para esta publicação, elaboramos um quadro ilustrativo contendo

algumas palavras que foram criadas ou ressignificadas pelas pessoas presas entrevistadas:

Quadro 2. *Palavras utilizados na linguagem prisional*

<i>Palavra</i>	<i>Significado</i>
Cabeça branca	Pessoa que tem dinheiro
Cabuloso	Pessoa pesada
Cacau	Chocolate
Cagueta	Informante da polícia
Cair	Morrer
Caiu	Ser pego
Caixa	Cela
Camelo	Espeto
Caminhada	Trajatória da vida
Camiseta do Corinthians	Palheiro
Campana	Ficar de olho
Canela	Cigarro
Canjica	Dentes
Carrapato	Parafuso
Carroça	Espeto
Castigo	Ficar de castigo quando faz alguma coisa que não está na
Chinelo	Empanado
Confere	Chamada para conferência das pessoas que estão presas
Ducha	Banho
Estrutural	Revista nas celas
Folha	Autorização para saída de sete dias
Linha podre	Falar mal de alguém

Fonte: Elaborado pelos autores (2023).

Com base nas palavras identificadas em toda a pesquisa, doze categorias foram definidas, com o respaldo das próprias entrevistas realizadas, a saber: 1) Instrumentos improvisados e criados dentro da Penitenciária: com objetos adquiridos dentro da instituição; 2) Procedimentos realizados por funcionários da unidade; 3) Alimentação: incluindo alimentos, procedimentos para que a alimentação aconteça e objetos relacionados a alimentação; 4) Vocabulário Pajubá: termos utilizados por mulheres transexuais e travestis; 5) Objetos; 6) Cela: estrutura das celas; 7) Características das pessoas presente naquele ambiente; 8) Necessidades fisiológicas; 9) Procedimentos de segurança da unidade realizado por funcionários; 10) Status relacionados aos sentimentos das pessoas; 11) Procedimentos e regras realizadas pelas pessoas que estão privadas de liberdade e; 12) Materiais ilícitos.

Quadro 3. *Categorias e palavras*

<i>Categorias</i>	<i>Palavras</i>
Instrumentos improvisados e criados dentro da Penitenciária	Carrapato Carroça
Procedimentos realizados dentro da instituição	Castigo Folha
Alimentação	Cacau
Vocabulário Pajubá	Chinelo Linha apodre
Objetos	Camelo
Cela	Caixa
Características das pessoas presentes naquele ambiente	Cabeça branca Cabuloso Canjica
Necessidades fisiológicas	Ducha
Procedimentos de segurança da unidade realizado por funcionários	Confere Estrutural
Status relacionados aos sentimentos das pessoas	Cagueta
Procedimentos e regras realizadas pelas pessoas que estão privadas de Liberdade	Cair Caiu Caminhada Campana
Materiais ilícitos	Camiseta do Corinthians Canela

Fonte: Elaborado pelos autores (2023).

As categorias formuladas acima podem, por exemplo, e para estudos futuros, servir de base ou ponto de partida para a construção de algum sistema de organização do conhecimento, como vocabulários controlados ou classificações, especializados na linguagem nos estabelecimentos prisionais, salvaguardadas as distintas realidades que existem nas diferentes instituições penais.

5 Considerações

Diante do exposto, afirma-se que sim, há formas de organização e representação do conhecimento na linguagem utilizada dentro dos espaços de privação de liberdade. Tal afirmação é perceptível a partir dos dados coletados em entrevistas e observação *in loco*. Com base nisso, as categorias mencionadas anteriormente reafirmam a possibilidade de criação de um sistema de organização do conhecimento, ou mesmo de uma melhor compreensão a respeito de como o conhecimento pode ser representado em

contextos marginalizados para outros fins, que não necessariamente a comunicação técnico-científica.

A organização e representação do conhecimento estão atreladas ao constructo de modelos de mundo, que constituem aspectos da realidade e para além dela. As formas de comunicação dentro das prisões, não somente a oral, mas também a escrita, realizada por meio de trocas de sílabas, de sinais gestuais, de alterações de significados, e mesmo de invenções vocabulares, são formas específicas de se representar o conhecimento.

O que procuramos argumentar com esta pesquisa é que ressignificar, modificar, inventar e ‘maquiar’ palavras e significados faz parte de um jogo de linguagem capaz de promover não somente a comunicação codificada, ou categorizada, mas principalmente capaz de propiciar uma forma de representação do conhecimento que protege e resiste, em um exercício de sobrevivência social e física.

Notas

1 Canudo refere-se aos corredores por onde as pessoas privadas de liberdade caminham dentro dos estabelecimentos prisionais.

2 Pessoa responsável por ser o porta voz de determinada ala da Penitenciária Masculina de Florianópolis.

3 Pipa são bilhetes encaminhados aos presos ou a administração.

Referências

- ANGROSINO, Michael. **Etnografia e observação participante**. São Paulo: Artmed, 2009. 130 p.
- ARAUJO, Ericka. **Um dia de visita**: o drama de familiares de detento. São Paulo: Independently Published, 2019. 120 p.
- BIONDI, Karina. **Junto e Misturado**: uma etnografia do pcc. São Paulo: Editora Terceiro Nome, 2018. 248 p.
- Dahlberg, I. Teoria do conceito. **Ciência da Informação**, v. 7, n. 2, 1978. Disponível em <https://doi.org/10.18225/ci.inf.v7i2.115>. Acesso em: 17 abr. 2023.
- GEERTZ, Clifford. **A interpretação das culturas**. Rio de Janeiro: LTC, 2019. 323p.
- GERHARDT, Tatiana Engel; SILVEIRA, Denise Tolfo (org.). **Métodos de pesquisa**. Rio Grande do Sul: UFRGS Editora, 2009. 120 p.
- GOFFMAN, Erving. **Manicômios, prisões e conventos**. São Paulo: Perspectiva, 2019. 320 p.
- GRAYLING, A C. **Wittgenstein**. São Paulo: Loyola, 2002.
- MARTELOTTA, Mário Eduardo. **Manual de linguística**. São Paulo: Contexto, 2018. 254 p.
- MELO, Fabio Dantas de. Categorização lingüística como esteio da organização do conhecimento. In: Jaime Robredo; Marisa Bräscher (Orgs.). **Passeios no Bosque da Informação**: Estudos sobre Representação e Organização da Informação e do Conhecimento – EROIC. Brasília DF: IBICT, 2010, 335 p. Capítulo 9, p. 177-182. Edição eletrônica. Disponível em: <http://www.ibict.br/publicacoes/eroic.pdf>. Acesso em: 17 abr. 2023.
- MELOSSI, Dario; PAVARINI, Massimo. **Cárcere e fábrica**: as origens do sistema penitenciário (séculos XVI-XIX). Rio de Janeiro: Revan, 2017. 272 p.
- MOREIRA, Walter. **Sistemas de organização do conhecimento**: aspectos teóricos, conceituais e metodológicos. 2018. 164 f. Tese (Doutorado) - Curso de Ciência da Informação, – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Marília, 2018. Disponível em: <https://repositorio.unesp.br/bitstream/handle/11449/190878/0/00908213.pdf?sequence=1>. Acesso em: 17 abr. 2023.
- MOTTA, Manoel Barros da. **Crítica da razão punitiva**: nascimento da prisão no Brasil. São Paulo: Forense Universitária, 2011. 380 p.
- OLIVEIRA, Hilderline Câmara de. **A linguagem no cotidiano prisional**: enigmas e significados. Jundiaí: Paco Editorial, 2013. 272 p.
- ORLANDI, Eni P. **Análise de discurso**: princípios e procedimentos. São Paulo: Pontes Editores, 2020. 98 p.
- PRODANOV, Cleber Cristiano; FREITAS, Emami Cesar de. **Metodologia do Trabalho Científico**: métodos e técnicas da pesquisa e do trabalho acadêmico. Novo Hamburgo: Editora Feevale, 2013. 276 p.
- RUSCHE, Georg; KIRCHHEIMER, Otto. **Punição e estrutura social**. Rio de Janeiro: Revan, 1939. 288 p.
- WITTGENSTEIN, Ludwig. **Investigações filosóficas**. São Paulo: Abril Cultural, 1975. 226 p.

Sistemas de Organização do Conhecimento inclusivos: a importância das notas de escopo, notas de uso e definições

Inclusive Knowledge Organization Systems: the importance of scope notes, use notes and definitions

Suellen Oliveira Milani

Universidade Federal Fluminense, suellenmilani@id.uff.br

Resumo

Objetivo: Essa pesquisa apresenta subsídios para a discussão de ações de promoção de recursos dos sistemas de organização do conhecimento para inclusão de diferentes áreas do conhecimento e comunidades de usuários, via linguagem, nas representações que compõem catálogos de bibliotecas e outros sistemas de informação. **Metodologia:** Pesquisa qualitativa que investiga de que maneira as notas de escopo, notas de uso e definições são utilizadas na Lista de Cabeçalho de Assuntos da Fundação Biblioteca Nacional, no Vocabulário Controlado Básico e no Vocabulário Controlado USP. Após pré-teste, um questionário foi respondido pelos gestores dos sistemas de organização do conhecimento analisados. Para a análise dos resultados foi utilizada a técnica de close reading. **Resultados:** Destacam-se: o uso das notas de escopo e as orientações para sua elaboração; a importância da política para sugestões de inclusões e alterações de termos justificadas a partir de fontes de informação defensáveis; os procedimentos para o recebimento de sugestões de inserção ou revisão de termos, notas de escopo e definições, as quais têm origem mais a nível interno; a atenção dos gestores dada à manutenção e à sua responsabilidade em relação aos usuários; a necessidade do mapeamento e caracterização das comunidades de usuários que não têm as suas necessidades de informação contempladas pelo sistema de informação. **Conclusão:** Ao realizar o uso e a manutenção de sistemas de organização do conhecimento, um enfoque pode ser dado às notas de escopo, notas de uso e definições com vistas a viabilização de catálogos equitativos e inclusivos em uma perspectiva terminológica e de garantia cultural.

Palavras-chaves: Sistemas de organização do conhecimento brasileiros. Nota de escopo. Nota de uso. Definição

Abstract

Goals: This paper presents subsidies for the discussion of actions to promote resources within knowledge organization systems for inclusion of different disciplines and user communities in representations into library catalogs and other information systems through language. **Methodology:** Qualitative research that investigates how scope notes, usage notes and definitions are used in the Brazilian National Library Foundation Subject Heading List, in the Basic Controlled Vocabulary and in the USP Controlled Vocabulary. After a pretest, a survey was answered by the managers of the knowledge organization systems analyzed. For the result analysis, the close reading technique was applied. **Results:** It is highlighted: the use of scope notes and guidelines for their construction; the importance of the policy for suggestions regarding inclusion and changes in terms from defensible information sources; the procedures for receiving suggestions for insertions and reviews of terms, scope notes and definitions, which generally come from internal requests; the attention given by managers to maintenance and their responsibility towards users; the need to map and characterize user communities that do not have their information needs covered by the information system. **Conclusion:** During the use and maintenance of knowledge organization systems, a focus should be given to scope notes, usage notes and definitions with the aim to enable more equitable and inclusive catalogs from a terminological and cultural warrant perspectives.

Keywords: Brazilian Knowledge Organization System. Scope note. Usage note. Definition

Resumen

Objetivos: Esta investigación presenta subsidios para la discusión de acciones para promover recursos de sistemas de organización del conocimiento dirigidos a la inclusión de diferentes áreas del conocimiento y comunidades de usuarios, a través del lenguaje, en las representaciones que componen los catálogos de bibliotecas y otros sistemas de información. **Metodología:** Investigación cualitativa que indaga cómo se utilizan las notas de alcance, notas de uso y definiciones en la Lista de Encabezamientos de Materia de la Fundación Biblioteca Nacional del Brasil, en el Vocabulario Controlado Básico y en el Vocabulario Controlado de la USP. Después de una prueba previa, los gerentes de los sistemas de organización del conocimiento analizados respondieron un cuestionario. Para el análisis de los resultados se utilizó la técnica de la lectura atenta. **Resultados:** Se destacan: el uso de notas de alcance y lineamientos para su elaboración; la importancia de la política para las inclusiones sugeridas y los cambios justificados en términos de fuentes de información defensibles; los procedimientos para recibir sugerencias de inserción o revisión de términos, notas de alcance y definiciones, que tienen un origen más interno; la atención de los gestores a la actualización y su responsabilidad frente a los usuarios; la necesidad de mapear y caracterizar comunidades de usuarios que no tienen sus necesidades de información cubiertas por el sistema de información.

Conclusión: Al usar y actualizar sistemas de organización del conocimiento, se puede dar énfasis a las notas de alcance, notas de uso y definiciones con el objetivo de posibilitar catálogos equitativos y inclusivos desde una perspectiva terminológica y de garantía cultural.

Palabras clave: Sistemas de organización del conocimiento brasileños. Nota de alcance. Nota de uso. Definición

1 Introdução

Uma vez que promover equidade e a inclusão em bibliotecas e sistemas de informação envolve esforços para potencializar as vozes das comunidades de usuários que utilizam os seus produtos e serviços, os postulados da Organização do Conhecimento devem ser continuamente estudados. Os sistemas de organização do conhecimento, que são construídos a partir destes postulados, são os instrumentos utilizados para o processo de representação do conhecimento e figuram como objetos centrais desse debate. Isso ocorre tanto pelo fato de um sistema de organização do conhecimento promover a organização e recuperação de informações e documentos, quanto por ser composto por um sistema de conceitos (também denominado sistema nocional) que representa determinada área do conhecimento ou determinada comunidade de usuários neste contexto.

Observa-se uma trajetória de pesquisa no bojo do que se denomina Organização do Conhecimento Crítica, isto é, pesquisas que defendem que ao representar conhecimentos dispostos como informações, os sistemas de organização do conhecimento expandem a sua função de oferecer “endereços” para estantes e para objetos digitais e passam a figurar como um “modelo de mundo” que será percorrido pelos usuários durante a busca e recuperação de informações e documentos.

De forma a amparar teoricamente essa pesquisa, destacam-se alguns estudos importantes acerca das tendências e preconceitos que podem ocorrer na Organização do Conhecimento, quais sejam: a. na Library of Congress Subject Headings: Berman (1971), Olson (2000, 2002), Fischer (2005), Knowlton (2005), Strotman (2007), Adler (2012), Adler e Tennis (2013); b. nos sistemas de classificação: Furner e Dunbar (2004), Olson (1999), Miranda (2009), Mai (2010, 2013), Martínez-Ávila e Guimarães (2013), Lee (2015), McTavish (2015), Fox (2016); e c. acerca do entendimento dos conflitos que existem no interior dos domínios: Hjørland (2008), López-Huertas e López-Pérez (2010), Martínez-Ávila, Semidão e Ferreira (2016), Fox, Martínez-Ávila e Milani (2017), Milani e Guimarães (2017), Pinho (2017). Esses estudos apontam para a responsabilidade dos profissionais ao viabilizarem a recuperação relevante de informações por meio da sua representação, preservação e disseminação.

Por serem artificiais, provisionais e deterministas, conforme já alertava Barité (2001), os sistemas de organização do conhecimento refletem, muitas vezes, o

modelo de mundo de quem os construiu e/ou mantém, impondo limites à representação das diferenças.

Nessa perspectiva, essa pesquisa propõe-se a oferecer subsídios para uma discussão sobre ações de promoção dos recursos presentes nos sistemas de organização do conhecimento que podem incluir diferentes áreas do conhecimento e comunidades de usuários, via linguagem, nas representações que compõem catálogos de bibliotecas e outros sistemas de informação. Para tanto, investigou-se de que maneira as notas de escopo, notas de uso e definições podem ser utilizadas em três importantes sistemas de organização do conhecimento brasileiros a partir da perspectiva dos seus gestores. O foco é pensar sobre esses recursos enquanto ferramentas que podem viabilizar catálogos mais equitativos e inclusivos, atendendo às diferentes comunidades de usuários, principalmente as que não pertencem aos contextos dominantes, em uma perspectiva terminológica e de garantia cultural.

2 Sistemas de organização do conhecimento: Estruturação e vieses culturais

Os sistemas de organização do conhecimento, também denominados linguagens documentárias, linguagens de indexação, instrumentos de representação da informação e do conhecimento, possuem as seguintes características, explicadas por Cintra *et al.* (2002):

- a. são vocabulários controlados, isto é, linguagens não naturais, artificiais, intermediárias, metalinguagem;
- b. possuem mecanismos interpretativos próprios;
- c. realizam o controle terminológico, permitindo a construção de uma linguagem de natureza monossêmica, quando um significante tem apenas um significado;
- d. podem incorporar elementos de léxicos, vocabulários, nomenclaturas e terminologias, porém não devem ser confundidos com eles.

As etapas para a estruturação de um sistema de organização do conhecimento incluem a escolha dos termos que comporão o léxico, isto é, a lista de elementos descritores, o estabelecimento das classes e, por fim, o estabelecimento das relações exteriores entre os termos. A estruturação é norteada por princípios claros, lógicos, consistentes e desenvolvidos na prática bibliotecária a partir de manuais, normas e diretrizes, conforme orientam Cintra *et al.* (2002).

Os processos de identificação, categorização e nomeação de conceitos (isto é, os termos) podem tornar-se tendenciosos mediante:

- falta de cuidado: tem-se menos conhecimento do que se imagina e faz-se suposições incorretas e más escolhas na representação;
- escolha ou formulações de termos sem pensar nas suas possíveis conotações;
- apropriação ilegítima de vieses culturais;
- apropriação ilegítima de vieses pessoais;
- a estrutura do mapa de construção do índice de assuntos (relações entre os descritores; estrutura das referências cruzadas) pode representar estruturas ou crenças que não sejam apropriadas aos conceitos que estão sendo representados e/ou às sensibilidades dos usuários;
- as hierarquias de um vocabulário controlado ou de um esquema de classificação podem conter vieses;
- dimensões do “politicamente correto” podem ser usadas sem pensar ou não serem usadas de forma alguma (JACOBS, 2007, p. 164).

No que diz respeito aos vieses culturais terminológicos, destacam-se as dificuldades que advêm da falta de termos disponíveis para descrever a vida e os relacionamentos humanos, gerando “limitações das ideias que podem ser expressas, particularmente porque nossa linguagem evolui mais lentamente do que as mudanças da sociedade” (BELL, 1991, p. 175).

Nesta oportunidade, o léxico dos sistemas de organização do conhecimento será objeto de análise, mais especificamente as notas de escopo, de uso e as definições, que são utilizadas para esclarecer o significado dos elementos descritores vinculados a um sistema de conceitos.

Esteban Navarro (1997) aponta três princípios de justificação que podem orientar a seleção e a representação léxica dos conceitos: a. baseado no conhecimento, quando o termo se adequa a um domínio da realidade que transmite a disciplina; b. baseado na literatura, também denominado garantia literária, isto é, o termo é utilizado nas fontes de informação de um domínio; e c. baseado no uso. No que diz respeito ao uso, destaca-se o princípio da garantia cultural.

Segundo Beghtol (2002, p. 511), a garantia cultural propõe que qualquer tipo de sistema de organização do conhecimento “pode ser maximamente apropriado e útil para os indivíduos em alguma cultura apenas se for baseado nas premissas, valores e predisposições dessa mesma cultura. Por outro lado, se um sistema não for baseado nessas premissas, será apropriado e útil em menor grau para os indivíduos da cultura.” Em geral, as premissas nas quais os sistemas são baseados são invisíveis àqueles que compartilham da mesma cultura que o sistema e, por isso, qualquer comunidade de usuários que seja marginalizada por uma cultura dominante não será adequadamente representada naquele sistema.

Uma pessoa possui diferentes camadas de subjetividade e pode pertencer a diferentes culturas, sendo assim para que os sistemas de informação cumpram os seus propósitos de representação, recuperação e disseminação da informação, as pessoas são agrupadas em comunidades de usuários a partir de uma característica em comum ou semelhante contextualizadas em um tempo e espaço. As comunidades de usuários, também denominadas comunidades discursivas, buscarão informações baseando-se em suas estruturas cognitivas sedimentadas em suas visões de mundo, que envolvem a divisão do conhecimento, do tempo, do sistema político, religioso etc.

Assim, um sistema que “não esteja estabelecido a partir de uma garantia cultural apropriada não será adotado para a busca e recuperação da informação porque os seus usuários entenderão que este sistema não corresponde às suas visões de como o mundo funciona” (BEGHTOL, 2005, p. 904).

Nenhum sistema de organização do conhecimento, até o momento, é capaz de prever todos os possíveis significados que a informação recuperada terá para cada usuário. Por essa razão, uma abordagem para inclusão não será universal, pois envolve uma reflexão teórica, ética e contextualizada dos gestores dos sistemas de organização do conhecimento.

O uso do qualificador na expressão “sistemas de organização do conhecimento inclusivos” está alinhado aos preceitos do conceito de hospitalidade cultural, proposto por Clare Beghtol, o qual se baseia em pesquisas sobre garantia literária de E. Wyndham Hulme e sobre garantia cultural de Joel M. Lee. A hospitalidade cultural funciona como uma estrutura teórica com uma mensagem ética, isto é, o conceito evoca as dimensões éticas envolvidas na representação, acesso, recuperação e apropriação da informação.

Para cunhar o conceito de hospitalidade cultural, Beghtol (2002, p. 518) apropriou-se do conceito de hospitalidade da Classificação que consiste na “habilidade de uma notação incorporar novos conceitos e estabelecer relações semânticas e sintáticas apropriadas entre os conceitos velhos e os novos.” O objetivo da autora, ao propor esse conceito, é sinalizar que um sistema de informação pode abrigar diferentes comunidades de usuários baseando-se no poder de escolha desses usuários que, nesse contexto, depara-se com algumas limitações de linguagem.

Para possibilitar a recuperação da informação, o sistema de conceitos construído para estruturar o sistema de organização do conhecimento deve controlar fenômenos linguísticos como a ambiguidade, a polissemia, a homonímia, a sinonímia, a hiponímia, o eufemismo, bem como as nuances entre diferentes línguas. Esse controle terminológico ocorre a partir da normalização gramatical, isto é, a padronização quanto ao gênero,

número e grau, bem como a partir da normalização semântica que, de acordo com Cintra *et al.* (2002, p. 43), “procura garantir a univocidade na representação dos conceitos de áreas de especialidade, por meio das relações lógico-semânticas.”

Para garantir o controle via linguagem e viabilizar a recuperação da informação, destaca-se a importância das notas de escopo, notas de uso e definições presentes nos sistemas de organização do conhecimento.

As notas de escopo “têm por finalidade explicitar a amplitude ou o entendimento atribuído ao conceito”, conforme demarcam Smit e Kobashi (2003, p. 31). As autoras alertam que as notas de escopo não podem ser confundidas com definições de dicionário, pois

não se pode perder de vista que os vocabulários controlados têm objetivos concretos e “industriais”: como a organização do acesso à informação contida nos documentos. Desse modo, as definições contidas nas notas de escopo têm objetivos pragmáticos (“organizar o acesso à informação”) e não teóricos (SMIT; KOBASHI, 2003, p. 31).

Apesar de serem construídos para “organizar o acesso à informação”, há relatos na literatura revelando a estruturação dos sistemas de organização do conhecimento influencia a forma como a informação é percebida pelo usuário, bem como pode fomentar impactos negativos na autoimagem de usuários que se valerem de sistemas de organização do conhecimento com vieses culturais ou preconceituosos para chegarem até informações e documentos que tratavam de assuntos relacionados a sua identidade, características ou necessidades informacionais.

Outro recurso importante no âmbito dos sistemas de organização do conhecimento são as notas de uso, as quais, como o próprio nome diz, trazem diretrizes para o uso do termo. O objetivo das notas de uso não é explicitar a amplitude ou o entendimento atribuído ao conceito, mas como usá-lo naquele sistema, como ensinam Smit e Kobashi (2003).

O léxico é composto por termos autorizados e não autorizados e a delimitação dos termos autorizados é feita por meio de definições elaboradas pelos próprios gestores ou extraídas de outras fontes de informação.

Cintra *et al.* (2002, p. 73) esclarecem que o objetivo da definição é “propor uma expressão (sintagma ou palavra) semanticamente equivalente à unidade a ser definida”.

A definição caracteriza um conceito e o distingue de outros conceitos, assim “fazer uma definição equivale a estabelecer uma ‘equação de sentido’ [como propõe Ingetraut Dahlberg], limitá-lo de algum modo, para fixar os limites de um conceito ou ideia” (LARA, 2004, p. 93).

As notas de escopo, notas de uso e definições podem viabilizar a criação de “espaços” nos sistemas de

organização do conhecimento e um “convite” aos usuários a participarem, com o intuito de tornar os “limites dos sistemas permeáveis”, como já recomendava Olson (2002).

Isso torna o entendimento pragmático e das limitações dos sistemas de organização do conhecimento *sine qua non*, assim como o reconhecimento dos esforços das instituições que os gerenciam.

Por exemplo: esforços para consertar a sexualidade gay sob a categoria de *Vida sexual* ao invés de *Desvio sexual* não asseguram a verdade, mas simplesmente revelam os processos pelos quais essas categorias e conhecimentos sobre eles são produzidos. A mudança da categoria não revela a emergência por uma verdade eterna e imutável sobre a sexualidade gay, mas descreve um arco discursivo por meio da própria história da estrutura da organização do conhecimento (DRABINSKI, 2013, p. 102).

A autora acredita que estimular os usuários a se comprometerem criticamente com as fontes de informação, incluindo os catálogos de biblioteca, é uma parte crítica do trabalho contemporâneo dos bibliotecários. Nesse sentido, os usuários seriam convidados ao trabalho discursivo de usar e resistir aos sistemas padronizados desenvolvendo a capacidade da reflexão crítica também sobre os sistemas de organização do conhecimento.

Olson (2000, p. 70) afirma que a Lista de Cabeçalhos de Assuntos da Biblioteca do Congresso, por exemplo, é um recurso a ser usado enquanto agente de mudanças, pois é um instrumento dominante em nossa cultura e “podemos fazê-lo permeando os limites da autoridade cultural”.

A autoridade cultural advém das áreas do conhecimento relacionados ao espaço que se quer construir. Para isso, é necessário realizar um diagnóstico sobre quais são as áreas do conhecimento e as comunidades de usuários excluídas ou representadas de forma inapropriada ou preconceituosa no catálogo. Em seguida, é possível promover ações relacionadas a questões de linguagem envolvendo essas comunidades de usuários e os sistemas de organização do conhecimento usados para viabilizar o acesso e a apropriação de informações por elas.

3 Percorso metodológico

Essa pesquisa qualitativa está inserida no contexto das Ciências Sociais Aplicadas, mais especificamente na área de Ciência da Informação. A principal fonte de informação utilizada foi o conjunto de respostas obtido a partir do instrumento de coleta de dados, isto é, do questionário respondido pelos gestores dos seguintes sistemas de organização do conhecimento brasileiros: Lista de Cabeçalho de Assuntos da Fundação Biblioteca Nacional, Vocabulário Controlado Básico e Vocabulário Controlado USP.

O questionário foi composto por perguntas abertas, como recomendado por Günther (2003) quando se realiza uma pesquisa em que não se dimensiona a abrangência ou a variabilidade das respostas. Essas perguntas foram especializadas buscando apresentar os contextos de tempo e espaço.

Com o intuito de avaliar a clareza das perguntas e sua consistência em relação aos objetivos da pesquisa, um pré-teste foi realizado. Para tanto, um questionário foi aplicado à bibliotecária-chefe de uma das bibliotecas universitárias integrantes do Sistema de Bibliotecas da Universidade Federal Fluminense (UFF).

Após contato com a bibliotecária-sujeito do pré-teste, o questionário lhe foi remetido, no formato Word e por e-mail, juntamente com um documento, o Termo de Consentimento Livre e Esclarecido, contendo informações sobre a pesquisa. Consta ainda do Termo o objetivo estritamente acadêmico do estudo, bem como a razão para a aplicação do questionário pré-teste, isto é, mensurar a compreensão das perguntas e sua coerência e adequação quanto às proposições da pesquisa.

O questionário em seu formato final foi, então, enviado por e-mail, após contato inicial com os responsáveis pela gestão dos três sistemas de organização do conhecimento. Cabe ressaltar que, no esforço para contatar os responsáveis pelo Vocabulário Controlado USP foi realizado um contato com a sua Assessoria Acadêmica, o qual resultou no recebimento de mais um questionário respondido.

Os questionários, respondidos no fim do primeiro semestre de 2020, foram analisados como textos produzidos por especialistas em um contexto de reflexão proposto pela pesquisa. Devido à natureza da pesquisa e à amostra reduzida de respondentes, nenhum tratamento estatístico foi aplicado aos dados.

Para análise dos textos foi utilizada a técnica do *close reading*, cujo objetivo é a observação dos fatos e detalhes do texto ao “lê-lo mais de perto.” Em um primeiro momento, foram observadas “todas as características marcantes do texto, incluindo os recursos retóricos, os elementos estruturais, as referências culturais” (KAIN, 1998).

Em um segundo momento, as observações foram interpretadas utilizando o raciocínio indutivo, passando, conforme explica Kain (1998), “da observação de fatos particulares e de detalhes para uma conclusão ou interpretação baseada nessas observações.” Após análises, os dados foram discutidos a partir de premissas identificadas na literatura e encaminhamentos para pesquisas futuras foram delineados.

4 Resultados

Os sistemas de organização do conhecimento Lista de Cabeçalho de Assuntos da Fundação Biblioteca Nacional, Vocabulário Controlado Básico e Vocabulário Controlado USP foram escolhidos como *corpus* de análise desta reflexão por serem sistemas de organização do conhecimento, também denominados linguagens documentárias alfabéticas, consolidados no Brasil, terem diferentes escopos, bem como possuírem uma instituição responsável por sua gestão e atualização.

4.1 Caracterização dos sistemas de organização do conhecimento analisados

A Lista de Cabeçalho de Assuntos da Fundação Biblioteca Nacional trata-se de uma lista de cabeçalhos de assuntos gerais e especializados utilizados como pontos de acesso no catálogo da Fundação Biblioteca Nacional. Os cabeçalhos são apresentados como conceitos, cabeçalhos frase ou conceitos seguidos de explicação entre parênteses; subdivisões também estão disponíveis. Estes cabeçalhos são originários da Lista de Cabeçalho de Assuntos da Library of Congress norte-americana, usada como guia e modelo desde o Projeto CAU, para elaboração do Catálogo de Assunto Unificado da Biblioteca Nacional, na década de 1980 (GOMES, 1983).

A Fundação Biblioteca Nacional é a representante do depósito legal no Brasil e seu acervo abrange todas as áreas do conhecimento e, segundo os seus gestores, tem um público vasto incluindo pesquisadores, profissionais, público em geral e outras bibliotecas.

O Vocabulário Controlado Básico é adotado pela Rede Virtual de Bibliotecas ligada à Administração Pública Federal, Governo do Distrito Federal, Poderes Legislativo, Executivo e Judiciário. Esse sistema de organização do conhecimento está sob coordenação direta da Biblioteca do Senado Federal, e representa recursos informacionais que atendem às demandas de informações dos órgãos gestores das bibliotecas participantes.

Desenvolvido na década de 1980, o Vocabulário Controlado Básico possui a estrutura de um tesouro e a sua versão de 2017 dispõe de 9925 descritores incluindo descritores de assunto, geográficos e gêneros/formas literárias e especificadores gerais, cronológicos e de forma. O Vocabulário Controlado Básico abrange, principalmente, as áreas do conhecimento das Ciências Sociais e Humanas, com ênfase no Direito, expressando as áreas de atuação das bibliotecas integrantes da Rede.

O Vocabulário Controlado USP é utilizado para a representação temática de recursos informacionais contidos no Banco de Dados Bibliográficos da Universidade de São Paulo (USP), isto é, abarca todas as áreas do conhecimento no que diz respeito às atividades de ensino, pesquisa e extensão da

Universidade. Pode ser acessado online, embora seja vedada sua exportação para sistemas de informação fora da USP.

Integrando o conjunto de processos sistêmicos da Agência USP de Gestão da Informação Acadêmica, o vocabulário é gerenciado por representantes das bibliotecas da Universidade por área do conhecimento, além de representantes da coordenadoria da Agência e representantes acadêmicos das áreas de Biblioteconomia e de Ciência da Informação. O Vocabulário Controlado USP possui cerca de 47000 termos, sendo constituído por termos autorizados e não autorizados para a indexação, os quais operam como remissivas de sinonímia.

4.2 Análise dos resultados

As respostas obtidas via questionários suscitam algumas observações, as quais serão aqui abordadas caso a caso.

A primeira questão do questionário indagou sobre a frequência de uso das notas de escopo e definições no sistema de organização do conhecimento. A respeito da Terminologia de Assuntos da Fundação Biblioteca Nacional, a instituição sinalizou sobre a pouca frequência no uso das definições. Já as notas de escopo e de uso são de aplicação contínua nela.

A partir da criação e adoção do Vocabulário Controlado Básico sob a forma de tesouro estruturado na década de 1980, dado o número expressivo de inconsistências limitantes que foram sendo identificadas, foi criado um Grupo de Estudo do Tesouro da Rede Virtual de Bibliotecas, cujo trabalho deu origem, em 2007, ao documento “Diretrizes para construção do tesouro da Rede Virtual de Bibliotecas”, e que permanece como parâmetro até os dias atuais. Para a manutenção e atualização, uma Base de Sugestões serve aos bibliotecários como instrumento, tanto para proposição de novos termos, quanto para a inclusão ou alteração de classificações de termos já existentes no tesouro. Visando tais resultados, a Gerência faz uso de definições e de notas explicativas, sem que se subdividam em escopo ou uso, o que permite que os dois tipos de recursos possam ser encontrados no sistema de organização do conhecimento.

A Agência USP de Gestão da Informação Acadêmica informa que o Vocabulário Controlado USP conta com grupo gestor do vocabulário controlado atuante e uma metodologia de atualização de termos materializada em uma política de desenvolvimento, como explicado na resposta dada ao item do questionário. Mediante as constantes atualizações perpetradas pela manutenção, o órgão gestor do sistema de organização do conhecimento tornou obrigatório o preenchimento das notas de escopo para sugestão de inclusão ou alteração de termos a partir de um dado período, e segue mantendo tal posicionamento. Nos outros sistemas de organização

do conhecimento não há sinalização de tal obrigatoriedade.

Sobre orientações para elaboração das definições e notas de escopo, objeto da segunda questão, todos os sistemas de organização do conhecimento possuem gestões atuantes. Dos três sistemas pesquisados, contudo, apenas o Vocabulário Controlado USP possui uma equipe que se dedica especificamente, até os dias de hoje, ao gerenciamento da manutenção e atualização do vocabulário controlado que promoveu a obrigatoriedade de preenchimento de notas de escopo, oportunizando constantes alterações, em documento formal criado para tal fim, e contando as diretrizes que conduzem o trabalho dos bibliotecários. A particularidade destacada acima parece sinalizar para uma característica em comum entre sistemas de organização do conhecimento gerenciados por bibliotecas universitárias, pois na fase do pré-teste realizado em uma das bibliotecas universitárias da UFF, a bibliotecária respondente sinalizou para a existência de uma Assessoria Acadêmica, onde um grupo trataria da avaliação do sistema de organização do conhecimento da Universidade.

Embora a Terminologia de Assuntos da Fundação Biblioteca Nacional se apoie na Library of Congress e em dicionários especializados e o Vocabulário Controlado Básico possua uma Base de Sugestões interna onde atuam bibliotecários, não há ainda uma normatização específica e constante que norteie o preenchimento das notas de escopo e definições. Os dois sistemas de organização do conhecimento citados fazem atualizações sob demanda do próprio serviço que realizam, mas sem uma obrigação formal e documentada quanto a estes aspectos.

A respeito da frequência do recebimento de sugestões para a inserção ou revisão das notas de escopo e definições, objeto da terceira questão, os gestores informam que há inserção e atualização tanto de notas de escopo quanto de definições, contudo estes pedidos teriam origem muito mais a nível interno, ou de bibliotecários de bibliotecas de suas Redes.

Apesar destas inserções e modificações buscando atualização serem parte do trabalho, não parece haver, ainda, interação com as comunidades de usuários dos sistemas de informação em que os sistemas de organização do conhecimento são usados, fato este reforçado pelo comentário final da questão acerca da Terminologia de Assuntos da Fundação Biblioteca Nacional, de que não há quase mudanças solicitadas por fontes externas.

A comunicação entre os gestores dos sistemas de organização do conhecimento e seus usuários ocorre por e-mail, como nos casos do Vocabulário Controlado Básico e do Vocabulário Controlado USP, esta última aliás disponibiliza na página inicial do seu website um relatório de atualização com alterações, exclusões e

inclusões feitas no sistema de organização do conhecimento.

Em relação aos catálogos das bibliotecas das instituições responsáveis pelos sistemas de organização do conhecimento avaliados, não foi possível encontrar algum campo específico ou indicativo que estimule e oportunize os usuários a fazerem alguma sugestão quanto a atualizações ou destaques no sistema de organização do conhecimento de modo direto, evidenciando um certo distanciamento entre os sistemas de organização do conhecimento e as suas comunidades de usuários no quesito de uma participação ativa e colaborativa na atualização dos termos, por exemplo.

Em relação à avaliação das relações de equivalência e associativas entre os elementos descritores do sistema de organização do conhecimento, os gestores da Terminologia de Assuntos da Fundação Biblioteca Nacional explicaram que ao inserir um termo novo, são empreendidas pesquisas na lista de cabeçalhos de assuntos da Library of Congress e outras fontes de informação seguras e validadas. Reforçaram que os termos e suas relações são avaliados e as atualizações são de fluxo contínuo.

No Vocabulário Controlado USP, a inclusão de termos novos ou a revisão de termos ocorre a partir da demanda específica dos bibliotecários de cada área do conhecimento atendida via Base de Sugestões, promovendo alterações quando necessário. Esse vocabulário está constituído por termos autorizados e termos não autorizados os quais operam como remissivas de sinonímia. O sistema de organização do conhecimento usa como recursos os qualificadores, notas de escopo, e para as relações de equivalência, o operador “ver”. Todavia, conforme asseverado na resposta dada no questionário, o Vocabulário Controlado USP não utiliza relações associativas, e isto fica claro, posto não haver indicativo do uso da remissiva “ver também”.

Partindo de um exemplo extraído do Tesouro do Supremo Tribunal Federal, foi perguntado se as notas de escopo, definições e relações entre os descritores são usadas de forma a viabilizar que o catálogo atenda às diferentes necessidades de informação e representatividade das comunidades de usuários de forma mais equitativa e inclusiva, isto é, mais justa. Por exemplo, no item dois da nota abaixo, a definição do termo “homossexualidade” que consta do Tesouro do Supremo Tribunal Federal foi extraída de um manual especializado produzido por pessoas diretamente ligadas ao movimento LGBTQIA+, abrindo um espaço para representar esse conceito de forma defensável, equitativa e inclusiva.

Figura 1. *Termo Homossexualidade*

HOMOSSEXUALIDADE	
NOTA	1. Expressão que se refere à atração ou à manutenção de relações sexuais entre indivíduos do mesmo sexo. (Silva, De Plácido. Vocabulário Jurídico. 28ª ed. Rio de Janeiro: Forense, 2010) 2. "O termo se refere à orientação sexual do indivíduo, indicando 'modo de ser e sentir'". (Manual de Comunicação LGBTI+)
UP	HOMOSSEXUALISMO
TR	DISCRIMINAÇÃO SEXUAL HOMOSSEXUAL UNIÃO ESTÁVEL HOMOAFETIVA UNIÃO HOMOAFETIVA
CAT	NOC NOMES COMUNS

Fonte: Tesouro do Supremo Tribunal Federal.

Na Terminologia de Assuntos da Fundação Biblioteca Nacional, a inclusão de novos termos se dá mediante demanda dos assuntos utilizados nas obras de seu catálogo, isto é, a partir da garantia literária. A atualização ocorre quando um cabeçalho se torna obsoleto ou necessita de adequações. Os gestores reforçam que têm consciência de sua responsabilidade e acreditam que as notas de escopo, juntamente com os termos, restringem ou ampliam a aplicação de um termo e podem auxiliar o bibliotecário na escolha mais apropriada para a representação de um termo selecionado na etapa de análise da Análise Documentária.

Os gestores do Vocabulário Controlado Básico sinalizaram que concordam com a premissa levantada pela questão, assim como os gestores do Vocabulário Controlado USP que completaram a resposta explicando que as inclusões e alterações de termos preveem que a fonte consultada seja indicada, sendo que essa fonte pode ser bibliográfica ou via consulta a especialistas da área de conhecimento em que o termo se insere.

Ao se pensar sobre políticas de inclusão, é preciso considerar que um quantitativo expressivo e diverso de comunidades de usuários é excluído ou se encontra às margens dos fluxos de informação.

Fluxo da informação é o caminho que a informação percorre desde o momento em que é produzida até o momento em que é acessada e alguém se apropria dela para geração de novo conhecimento e/ou tomada de decisão. Na Biblioteconomia e na Ciência da Informação esse fluxo envolve o armazenamento e o tratamento para a recuperação da informação, sendo assim a exclusão dos fluxos de informação provoca a exclusão de outros fluxos, como do poder, da riqueza e dos direitos sociais. Uma das consequências de tal exclusão é a pouca ou nenhuma representatividade e/ou acesso das pessoas a contextos políticos, econômicos, educacionais e culturais, ficando às margens dos sistemas, incluindo os sistemas de organização do conhecimento.

Assim, no que concerne aos três sistemas de organização do conhecimento analisados, os seus gestores estão atentos à necessária política de atualização, bem como conscientes de sua responsabilidade junto às comunidades de usuários dos seus sistemas de informação. Contudo, esforços podem ser empreendidos de modo a atender à garantia cultural.

É importante salientar que os três sistemas de organização do conhecimento fazem uso de recursos apropriados para potencializar a representatividade das comunidades de usuários, quais sejam notas de escopo e definições. A missão de promover espaços que possam ser “percorridos” pelas comunidades de usuários deve inspirar também a política de indexação que oferece diretrizes para os processos de representação temática, incluindo a aplicação dos sistemas de organização do conhecimento.

No que compete às fontes de informação consultadas para a inserção ou atualização de termos, sua seleção e exame para elaboração de uma definição, por exemplo, deve-se considerar aquelas que incluem comunidades de usuários invisibilizadas. Este movimento contemplaria grupos de atuação marcante, como os coletivos, incluindo os coletivos universitários, que podem contribuir para discussões entre comunidades de usuários e gestores, estes últimos quase sempre os únicos responsáveis pela manutenção do sistema de organização do conhecimento.

Percebeu-se uma abertura dos gestores dos sistemas de organização do conhecimento no tocante à reflexão sobre a ampliação da visão sobre representatividade, e conseqüente diminuição de arbitrariedade, em relação às comunidades que possuem crenças e valores morais divergentes daquele estabelecido como padrão, isto é, do contexto dominante nos sistemas de organização do conhecimento.

A literatura de Organização do Conhecimento aponta que visões e preconceitos nos sistemas de organização do conhecimento podem envolver “os contextos de gênero, sexualidade, raça, idade, habilidade, etnicidade, língua e religião enquanto limites para a representação da diversidade e [limites para fornecer] serviços biblioteconômicos eficazes para populações diversas” (OLSON; SCHLEGL, 2001, p. 62).

A promoção de catálogos de bibliotecas mais equitativos, inclusivos e justos será viável a partir do momento que as comunidades de usuários que se encontram às margens estiverem presentes.

Para tanto, estudos de usuários devem ser empreendidos a partir de métodos e técnicas já sedimentados, como explicado por Cunha, Amaral e Dantas (2015), bem como a partir de teorias, métodos e técnicas desenvolvidos no interior das áreas ou comunidades de usuários analisados. Exemplos de instrumentos de coleta de dados para esse fim incluem: questionários,

entrevistas semiestruturadas ou estruturadas, grupo focal, tal como exemplificado por Caswell (2018).

Enquanto os estudos de usuários da informação avaliam características e comportamentos informacionais de grupos de usuários no que diz respeito à produção e ao uso da informação, abordagens como a analítica de domínio valem-se de critérios epistemológicos e de relevância compartilhados por comunidades discursivas a partir de práticas baseadas em evidência ou posturas epistêmicas, por exemplo. A combinação dessas abordagens tem sido usada para reflexões sobre a organização e representação do Conhecimento? Espera-se que esse questionamento estimule pesquisas futuras sobre o tema.

As comunidades de usuários também podem participar de ações relacionadas à folksonomia. Segundo Moura (2009, p. 34), há “a necessidade de compreender e dimensionar os desdobramentos da participação dos usuários na constituição de linguagens de referência adotadas na organização e recuperação da informação em ambientes digitais.” A autora já sinalizava para a necessidade de se pensar em planos de formação humana voltados a esse contexto.

O termo folksonomia, também denominado *social tagging*, foi cunhado por Wander Wal (2007), que explica que “[o] valor dessa marcação externa advém do fato de as pessoas usarem o seu próprio vocabulário e adicionarem significado explícito, o qual pode vir de inferências sobre aquela informação/objeto.” Exercícios desse tipo podem atuar como um incentivo para a manutenção dos sistemas de organização do conhecimento, em razão do feedback que os gestores obteriam dos usuários, porém alguns princípios norteadores devem ser disponibilizados, pois o exercício é contextual. Cumpre ressaltar que a folksonomia traria elementos para a reflexão acerca da adequação dos termos a partir da perspectiva da garantia cultural, porém não figuraria como uma alternativa para manutenções instantâneas.

Ao refletir sobre esse “movimento”, Mai (2011, p. 120) desenha a seguinte ação coletiva:

As *tags* escolhidas para os objetos seriam escolhidas pelos membros de um grupo a partir de uma discussão e interação entre os seus membros, talvez com base em uma votação entre um grupo maior de pessoas. Esse tipo de sistema permitiria uma verdadeira produção colaborativa de folksonomias e seria fiel aos princípios democráticos.

No que diz respeito a comunidades de usuários cujos conhecimentos nem sempre são registrados em suportes, estudos devem ser conduzidos com a presença e colaboração de membros da comunidade estudada. Por exemplo, no caso do conhecimento indígena, Sandy e Bossaller (2017, p. 146) argumentam que:

[...] abordagens publicadas e não universais que remontam aos “tempos mais simples” da organização local são

fundamentais para fornecer acesso e que esses métodos devem ser combinados com abordagens universais e tecnologias baseadas na web para fornecer acesso confiável e cognitivamente justo aos recursos indígenas.

Por fim, destaca-se que outros movimentos também podem ser pensados em parceria com programas de educação de usuários, visando contribuir para a promoção de catálogos mais equitativos e inclusivos, bem como para uma aproximação dos usuários com as formas pelas quais o conhecimento é formalmente organizado e representado nos sistemas de informação que utilizam.

5 Conclusão

O objetivo desta pesquisa foi oferecer elementos para uma reflexão sobre posturas e ações que podem ser tomadas pelos bibliotecários no tocante à manutenção e uso de sistemas de organização do conhecimento com vistas à inclusão das diferentes comunidades de usuários nos sistemas de informação.

Observa-se uma tradição de pesquisa relacionada às perspectivas sociais e culturais no campo da Organização e Representação do Conhecimento, que se manifesta em temas de eventos científicos e acadêmicos, números especiais de revistas e livros temáticos de Biblioteconomia, Arquivologia, Ciência da Informação e Organização do Conhecimento. Há também bibliotecários defendendo e aplicando técnicas e tecnologias para inclusão de comunidades de usuários que ainda se encontram às margens nos sistemas.

No que diz respeito ao uso e manutenção dos sistemas de organização do conhecimento estudados, destacam-se: o uso das notas de escopo e as orientações para sua elaboração; a importância da política para sugestões de inclusões e alterações de termos justificadas a partir de fontes de informação defensáveis; os procedimentos para o recebimento de sugestões de inserção ou revisão de termos, notas de escopo e definições; a atenção dos gestores à manutenção e à sua responsabilidade em relação aos usuários; a necessidade do mapeamento e caracterização das comunidades de usuários que não têm as suas necessidades de informação contempladas pelo sistema de informação.

No âmbito da formação profissional, conteúdos acerca da construção, manutenção e uso de sistemas de organização do conhecimento trazem a segurança necessária para que os seus recursos possam ser aplicados, e outros criados, de forma consistente, ética e crítica pelos bibliotecários.

É fundamental incentivar os bibliotecários e todos os envolvidos e interessados na área de Organização do Conhecimento, por meio dos saberes de que dispõem, a se valerem das notas de escopo, notas de uso e definições dos sistemas de organização do

conhecimento, com vistas a promoção de catálogos equitativos e inclusivos nos sistemas de informação. Os bibliotecários éticos e comprometidos com a sua responsabilidade social compreendem e valorizam a diversidade nas instituições em que trabalham e nos produtos que geram e disponibilizam para a sociedade.

Agradecimento

A autora agradece a bolsista Luciana Maria Gomes de Alcantara pelo desenvolvimento de pesquisa nesta temática no âmbito do PIBIC-UFF (2020-2021).

Referências

- ADLER, M. **For Sexual perversions see Paraphilias:** disciplining sexual deviance at the Library of Congress. 2012. 260 f. Tese (Doutorado em Library and Information Studies) – University of Wisconsin-Madison, Madison, 2012.
- ADLER, M.; TENNIS, J. T. Toward a taxonomy of harm in knowledge organization systems. **Knowledge Organization**, Würzburg, v. 40, n. 4, p. 266-272, 2013.
- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K. (org.). **Educação, universidade e pesquisa**. Marília: UNESP; São Paulo: FAPESP, 2001. p. 35-60.
- BEGHTOL, C. A proposed ethical warrant for global knowledge representation and organization systems. **Journal of Documentation**, London, v. 58, n. 5, p. 507-532, 2002.
- BEGHTOL, C. Ethical decision-making for knowledge representation and organization systems for global use. **Journal of the American Society for Information Science and Technology**, New York, v. 56, n. 9, p. 903-912, 2005.
- BELL, H. K. Bias in indexing and loaded language. **The indexer**, London, v. 17, n. 3, p. 173-177, 1991.
- BERMAN, S. **Prejudices and antipathies: A Tract on the LC Subject Heads Concerning People**. 2. ed. Jefferson; North Carolina; London: McFarland & Company, Inc., 1993.
- CASWELL, M. **Assessing the Affective Impact of Community Archives: A Toolkit**. Los Angeles: University of California/ UCLA Community Archives Lab, 2018. Disponível em: <https://communityarchiveslab.ucla.edu/toolkit>. Acesso em: 09 mar. 2023.
- CINTRA, A. M. M.; TÁLAMO, M. de F. G. M.; GINEZ DE LARA, M. L. KOBASHI; N.Y. **Para entender as linguagens documentárias**. 2. ed. São Paulo: Polis, 2002.
- CUNHA, M. B. da; AMARAL, S. A. do; DANTAS, E. B. **Manual de estudo de usuários da informação**. São Paulo: Atlas, 2015.
- DRABINSKI, E. Queering the catalog: queer theory and the politics of correction. **Library Quarterly**, Chicago, v. 83, n. 2, p. 94-111, 2013.

- ESTEBAN NAVARRO, M. A. Principios, reglas y técnica para la gestión del vocabulario y la estructura de los lenguajes documentales. In: GARCÍA MARCO, F. J. (ed.). **Organización del conocimiento en sistemas de información y documentación**. Zaragoza: Universidad de Zaragoza, 1997. p. 127-147.
- FISCHER, K. S. Critical views of LCSH, 1990–2001: the third bibliographic essay. **Cataloging & Classification Quarterly**, New York, v. 41, n. 1, p. 63-109, 2005.
- FOX, M. J.; MARTÍNEZ-ÁVILA, D.; MILANI, S. O. A Interseccionalidade e o Respeito às Pessoas na Organização do Conhecimento. In: ALVES, M. A; GRACIO, M. C. C; MARTÍNEZ-ÁVILA, D. (org.). **Informação, conhecimento e modelos**. Campinas: UNICAMP, 2017. p. 239-253. (Coleção CLE, 78).
- FOX, M. J. Subjects in doubt: the ontogeny of Intersex in the Dewey Decimal Classification. **Knowledge Organization**, Würzburg, v. 43, n. 8, p. 581-593, 2016.
- FURNER, J.; DUNBAR, A. W. The treatment of topics relating to people of mixed race in bibliographic classification schemes: a critical race-theoretic approach. In: McILWAINE, I. C. (ed.). **Knowledge Organization and the Global Information Society**: Proceedings of the Eighth International ISKO Conference, London, 13-16 July 2004. Würzburg: Ergon, 2004. p. 115-120. (Advances in Knowledge Organization, 9).
- GOMES, H. E. Cabeçalhos de assuntos unificados da Biblioteca Nacional. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 16, n. 3/4, p. 90-101, 1983.
- GÜNTHER, H. **Como elaborar um questionário**. Brasília: UnB, 2003. (Planejamento e pesquisa para as Ciências Sociais, 1). Disponível em: www.ic.unicamp.br/~wainer/cursos/2s2006/epistemico/01Questionario.pdf. Acesso em: 15 maio 2023.
- HJØRLAND, B. Deliberate bias in knowledge organization? In: ARSENAULT, C.; TENNIS, J. T. (ed.). **Culture and Identity in Knowledge Organization**: Proceedings of the Tenth International ISKO Conference, Montréal, 5-8 August 2008. Würzburg: Ergon, 2008. p. 256-261. (Advances in Knowledge Organization, 11).
- HOMOSSEXUALIDADE. In: Tesauro do Supremo Tribunal Federal. Brasília: Supremo Tribunal Federal, 2020. Disponível em: <https://portal.stf.jus.br/jurisprudencia/tesauro/pesquisa.asp>. Acesso em: 09 mar. 2023.
- JACOBS, C. Ethical places, ethical spaces: stopping to listen. **The indexer**, London, v. 25, n. 3, p. 161-166, 2007.
- KAIN, P. **How to do a close reading**. 1998. Disponível em: <http://writingcenter.fas.harvard.edu/pages/how-do-close-reading>. Acesso em: 08 mar. 2023.
- KNOWLTON, S. A. Three decades since *Prejudices and Antipathies*: a study of changes in the Library of Congress Subject Headings. **Cataloging & Classification Quarterly**, New York, v. 40, n. 2, p. 123-145, 2005.
- LARA, M. L. G. de. Diferenças conceituais sobre termos e definições e implicações na organização da linguagem documentária. **Ciência da Informação**, Brasília, v. 33, n. 2, p. 91-96, 2004.
- LEE, W.-C. Culture and classification: an introduction to thinking about ethical issues of adopting global classification standards to local environments. **Knowledge Organization**, Würzburg, v. 42, n. 5, p. 302-307, 2015.
- LÓPEZ-HUERTAS, M. J.; LÓPEZ-PÉREZ, M. J. Epistemological dynamics in scientific domains and their influence in knowledge organization. In: GNOLI, C.; MAZZOCCHI, F. (ed.). **Paradigms and conceptual systems in knowledge organization**: Proceedings of the Eleventh International ISKO Conference, Rome, 23-26 February 2010. Würzburg: Ergon, 2010. p. 91-97. (Advances in Knowledge Organization, 12).
- MAI, J.-E. Classification in a social world: bias and trust. **Journal of Documentation**, London, v. 66, n. 5, p. 627-642, 2010.
- MAI, J.-E. Ethics, values and morality in contemporary library classifications. **Knowledge Organization**, Würzburg, v. 40, n. 4, p. 242-253, 2013.
- MAI, J.-E. Folksonomies and the New Order: Authority in the Digital Disorder. **Knowledge Organization**, Würzburg, v. 8, n. 2, p. 114-122, 2011.
- MARTÍNEZ ÁVILA, D.; GUIMARÃES, J. A. C. Críticas de las clasificaciones bibliográficas: universalidad, postestructuralismo y ética. **Scire**, Zaragoza, v. 19, n. 2, p. 21-26, 2013.
- MARTÍNEZ-ÁVILA, D.; SEMIDÃO, R.; FERREIRA, M. Methodological aspects of critical theories in knowledge organization. **Knowledge Organization**, Würzburg, v. 43, n. 2, p. 118-125, 2016.
- McTAVISH, J. The ethics of querying and permeating Canadian everyday life nutritional classification technologies and processes. **Knowledge Organization**, Würzburg, v. 42, n. 5, p. 308-315, 2015.
- MILANI, S. O.; GUIMARÃES, J. A. C. Problemas relacionados a biases em sistemas de organização do conhecimento: Perspectivas para a representação de assunto. In: PINHO, F. A.; GUIMARÃES, J. A. C. **Memória, tecnologia e cultura na organização do conhecimento**. Recife: Ed. UFPE, 2017. p. 399-408. (Estudos Avançados em Organização do Conhecimento, 4).
- MIRANDA, M. L. C. de. A organização do etnoconhecimento: a representação do conhecimento afrodescendente em religião na CDD. **Revista África e Africanidades**, Rio de Janeiro, v. 1, n. 4, 2009.
- MOURA, M. A. Folksonomias, redes sociais e a formação para o tagging literacy: Desafios para a Organização da Informação em ambientes colaborativos virtuais. **Informação & Informação**, Londrina, v. 14, n. esp., p. 25-45, 2009.
- OLSON, H. A. Difference, culture and change: the untapped potential of LCSH. **Cataloging & Classification Quarterly**, New York, v. 29, n. 1/2, p. 53-71, 2000.
- OLSON, H. A. Exclusivity, teleology and hierarchy: our Aristotelian legacy. **Knowledge Organization**, Würzburg, v. 26, n. 2, p. 65-73, 1999.
- OLSON, H. A.; SCHLEGL, R. Standardization, objectivity, and user focus: a meta-analysis of subject access critiques.

Cataloging & Classification Quarterly, New York, v. 32, n. 2, p. 61-80, 2001.

OLSON, H. A. **The power to name**: locating the limits or subject representation in libraries. Dordrecht: Kluwer Academic Publisher, 2002.

PINHO, F. A. Percurso investigativo para contextualização de metáforas relativas à gênero e sexualidade em linguagens documentais. **Informação&Informação**, Londrina, v. 22, n. 2, p. 117-143, 2017.

SANDY, H. M.; BOSSALLER, J. Providing Cognitively Just Subject Access to Indigenous Knowledge through Knowledge Organization Systems, **Cataloging&Classification Quarterly**, New York, v. 55, n. 3, p. 129-152, 2017.

SMIT, J. W.; KOBASHI, N. Y. **Como elaborar vocabulário controlado para aplicação em arquivos**. São Paulo: Arquivo do Estado, Imprensa Oficial do Estado de São Paulo, 2003. (Projeto Como fazer, 10).

STROTTMAN, T. A. Some of our fifty are missing: Library of Congress Subject Headings for southwestern cultures and history. **Cataloging & Classification Quarterly**, New York, v. 45, n. 2, p. 41-64, 2007.

WANDER WAL, T. Folksonomy Coinage and Definition. *In*: WANDER WAL, T. **Room to Stretch for Thomas Vander Wal**. Bethesda, 02 fev. 2007. Disponível em: www.vanderwal.net/folksonomy.html. Acesso em: 20 mar. 2023.

Terminologia da Documentação Sanitária na literatura científica

Terminology of Sanitary Documentation in scientific literature

Virgínia Bentes Pinto (1), Raquel Ellen Gomes Pessoa (2), Raniane Alves Dias (3)
Eulália de Sousa Fernandes (4), Gabriela Belmont de Farias (5)

(1) Universidade Federal do Ceará, Av. da Universidade, 2762 - Centro, Fortaleza - CE, 60020-181, vbentes@ufc.br.
(2) raquel.ellen@alu.ufc.br. (3) ranianealves@alu.ufc.br. (4) eulaliafer01@alu.ufc.br. (5) gabriela_belmont@ufc.br.

Resumo

Embora que o sintagma Documentação Sanitária não esteja inserido no contexto das organizações de saúde brasileiras, ainda assim, considera-se tratar-se de uma temática de grande interesse na pesquisa, visto que, um de seus componentes é o prontuário do paciente. A Documentação Sanitária é constituída da Documentação Clínica (Prontuário do Paciente) e Documentação Não Clínica que tem influência nas ações de cuidados dos pacientes. Observando essa realidade, desenvolvemos a pesquisa que tem como **objetivo** investigar a temática Documentação Sanitária, presente na literatura das áreas da Ciências da Informação e da Saúde, no período de 1999-2022, visando a elaboração de uma ficha terminológica. **Metodologia:** pesquisa bibliográfica, cuja empiria foi feita no Portal de Periódicos da CAPES, SCIELO, BRAPCi e BVS. Estratégia de busca: “documentación sanitaria”, “documentação médica”, “prontuário do paciente”, “competência informacional”, “competência em informação” e “mediação”, nos idiomas português e espanhol. **Resultados:** Os achados da pesquisa evidenciaram um total de 154 termos coletados em 55 artigos, sendo os mais repetidos: Prontuário do Paciente (11), Ciência da Informação (8) e Prontuário Eletrônico do Paciente (7). Referente a Documentação Sanitária, foi identificado somente em dois trabalhos. Ficha terminológica: termo, classificação, termos sinônimos, definição, contextos, termos associados e notas. **Conclusão:** Compreendemos que o termo Documentação Sanitária, por ser de origem espanhola, pode ter contribuído para a redução do seu uso em contextos da língua portuguesa. No que diz respeito à terminologia, o Prontuário do Paciente, considerado documento fundamental na constituição do termo Documentação Sanitária, se destacou, bem como os sinônimos.

Palavras-chave: documentação sanitária. terminologia. prontuário do paciente.

Abstract

Even though the expression Sanitary Documentation is not yet inserted in the context of Brazilian health organizations, it is still considered to be a topic of great interest in research, given that one of its components is the patient's medical record. The Sanitary Documentation consists of the Clinical Documentation (Patient's Records) and Non-Clinical Documentation that influences the actions towards patient care. Analyzing this reality, we developed a research that has the **purpose** to study the Sanitary Documentation, that appears in the literature of the areas on Information and Health Sciences, in the period between 1999-2022, intending to create a terminological record. **Methodology:** bibliographical research, whose empirical work was carried out in the Portal de Periódicos da CAPES, SCIELO, BRAPCi and BVS. Search strategy: "sanitary documentation", "medical documentation", "patient's record", "informational competence", "information competence" and "mediation", in Portuguese and Spanish. **Results:** The research findings showed a total of 154 terms collected in 55 articles, the most repeated being: Patient Record (11), Information Science (8) and Electronic Patient Record (7). Referring to Health Documentation, it was identified only in two works. Terminology sheet: term, classification, synonymous terms, definition, contexts, associated terms and notes. **Conclusion:** We understand that the term Sanitary Documentation, being of Spanish origin, may have contributed to the reduction of its use in contexts of the Portuguese language. When it comes to terminology, the Patient Record, considered a fundamental document in the constitution of the term Sanitary Documentation, stood out, as well as the synonyms.

Keywords: sanitary documentation. terminology. patient record.

Resumen

Aunque el sintagma Documentación Sanitaria aún no se inserta en el contexto de las organizaciones de salud brasileñas, todavía se considera un tema de gran interés en la investigación, dado que uno de sus componentes es la historia clínica del paciente. La Documentación Sanitaria está compuesta por la Documentación Clínica (Historia clínica del Paciente) y la Documentación non Clínica que influye en las acciones de atención al paciente. Al observar esta realidad, desarrollamos una investigación que tiene como **objetivo** investigar el tema de la Documentación Sanitaria, presente en la literatura en las áreas de Ciencia de la Información y de la Salud, en el período 1999-2022, para elaborar una ficha terminológica. **Metodología:** investigación bibliográfica, cuyo empiria fue realizado en el Portal de Periódicos da CAPES, SCIELO, BRAPCi y BVS. Estrategia de búsqueda: "documentación sanitaria", "documentación médica", "historial del paciente", "competencia informacional", "competencia informativa" y "mediación", en

português y español. **Resultados:** Los hallazgos de la investigación mostraron un total de 154 términos recopilados en 55 artículos, siendo los más repetidos: Historia Clínica del paciente (11), Ciencias de la Información (8) y Historia Clínica Electrónica (7). Referente a la Documentación Sanitaria, se identificó solamente en dos trabajos. Ficha terminológica: término, clasificación, sinónimos, definición, contextos, términos asociados y notas. **Conclusión:** Entendemos que el término Documentación Sanitaria, siendo de origen español, puede haber contribuido a la reducción de su uso en contextos de la lengua portuguesa a la terminología, el Registro del Paciente, considerado un documento en la constitución del término Documentación Sanitaria, se destacó, así como las sinonimias.

Palabras clave: documentación sanitaria, terminología. História clinica del paciente.

1 Introdução

O termo Documentação Sanitária, embora ainda não esteja inserido no contexto das organizações de saúde brasileiras, tratar-se de uma temática de grande interesse na pesquisa, haja vista que um de seus componentes é o prontuário do paciente. Essa documentação é constituída pela Documentação Clínica (Prontuário do Paciente) e Documentação Não Clínica (todos os outros tipos de documentos que são condicionantes de saúde), porém que tem influência nas ações de cuidado com pacientes.

Conforme o artigo 1º da Resolução nº 1.638/2002 do Conselho Federal de Medicina (CFM), o prontuário do paciente é definido como:

O documento único constituído de um conjunto de informações, sinais e imagens registradas, geradas a partir de fatos, acontecimentos e situações sobre a saúde do paciente e a assistência a ele prestada, de caráter legal, sigiloso e científico, que possibilita a comunicação entre membros da equipe multiprofissional e a continuidade da assistência prestada ao indivíduo. (CFM, 2002).

Trata-se de um documento ímpar do domínio da saúde sendo redigido por uma terminologia de especialidade, que no contexto da redação do prontuário do paciente, é de valor inestimável para a práxis sanitária. Talvez, por esse fato é que já se percebe, na literatura, que outras áreas de conhecimento se interessam por pesquisas contemplando esse documento. É, pois, nesse âmbito que desenvolvemos esta pesquisa embasada no seguinte *problema*: quais foram os termos mais utilizados na literatura científica das áreas da Ciência da Informação e da Saúde relacionadas à Documentação Sanitária, particularmente ao Prontuário?

Visando encontrar respostas a esse questionamento, definimos como *objetivo geral*: analisar a terminologia relacionada a Documentação Sanitária presente na literatura científica dos campos de Ciência da Informação e da Saúde que podem contribuir para a elaboração de uma ficha terminológica.

Tendo como *objetivos específicos*: a) levantar a literatura científica referente a documentação sanitária publicada de 1999 a 2022; b) mapear a terminologia das palavras-chave encontradas nos artigos selecionados; c) construir a ficha terminológica relativa às temáticas encontradas. Destarte, esses são os

aspectos tratados neste artigo, e que assumimos que poderá incentivar outros trabalhos nessa temática.

2 Documentação Sanitária

Podemos considerar que documentos são representados de diversas formas e se encontram em vários ambientes, que podem ser utilizados para fins de prova, comunicação, registro, análise, entre outras. Briet (2016, p. 1) os define como “todo indício, concreto ou simbólico, conservado ou registrado, com a finalidade de representar, reconstituir ou provar um fenômeno físico ou intelectual”. Além disso, tomando como base os princípios arquivísticos e os seus estágios de evolução, um documento pode ter diferentes graus de formalidade e características a depender do suporte e da função que ele vá cumprir juntamente com a importância que ele tem no meio onde foi criado.

Segundo Sprague Jr (1995, p.15) “documentação, como atividade, significa coleta e arquivamento sistemático dos registros de informações, a fim de preservá-los para uso futuro”. Dessa maneira, a documentação é importante para todas as áreas, e, especificamente para a saúde, é fundamental, tanto para a equipe multiprofissional, como para o pacientes, seus familiares e para a pesquisa. Deste modo, considerando o conceito de documentação discutido, adentrarmos em uma documentação mais específica, a Documentação Sanitária. Casabona (2006, p. 140) explica que

o adjetivo "sanitário" denota pertencimento ou relação com a sanidade e, em consequência, o dado sanitário são os dados de saúde que se obtém e se utiliza em instruções sanitárias com fim de preservar a saúde dos cidadãos. (CASABONA, 2006, p.140, tradução nossa).

O autor diz ainda que as organizações de saúde lidam com dados relativos à saúde, dados sanitários e dados médicos. Documentação Sanitária, portanto, é aquela produzida nas organizações de saúde referentes aos atos de cuidados do paciente. Conforme Casabona (2006) integram essa documentação os registros de dados, informações e conhecimentos clínicos e não clínicos. No primeiro caso, dados, informações e conhecimentos clínicos são concernentes à atenção primária, especializada ou em outros níveis, de cuidados dispensados aos pacientes. No segundo, também são produzidos outros registros de dados, informação e conhecimentos, porém, considerados como não clínicos. Entretanto, são necessários ao

processo assistencial de um cidadão para o seu melhor atendimento nas organizações de saúde. Todavia, essa pesquisa irá se conter na Documentação Clínica, devido nosso objeto de pesquisa ser o Prontuário do Paciente.

Segundo o Conselho Regional de Medicina do Distrito Federal (2006, p. 11), o prontuário

[...] serve como instrumento de consulta, avaliações, ensino, pesquisa, auditoria, estatística médico-hospitalar, sindicâncias, prova de que o doente foi ou está sendo tratado convenientemente, investigação epidemiológica, processos éticos e legais, comunicação entre os profissionais de assistência ao paciente, defesa e acusação. (CONSELHO REGIONAL DE MEDICINA DO DISTRITO FEDERAL, 2006, p. 11)

Por sua vez, Oliveira (2007, p.161), conceitua como sendo

aquela também referida como documentação de saúde, documentação clínica ou documentação médica (e, acrescente-se, odontológica). Centra-se, em suma, num documento específico chamado prontuário do paciente. (OLIVEIRA, 2017, p. 161).

Na área da saúde, a Documentação Sanitária, seja Clínica ou Não Clínica, contém registros tanto do estado de saúde da pessoa enferma quanto dos condicionantes de saúde e outras. Documentos produzidos desde a recepção com o registro de dados pessoais, ficha de atendimento, prontuários médicos, fichas de evolução clínica, laudos e exames, entre outros. Esse tipo de documentação é fundamental tanto para a segurança do indivíduo perante os cuidados de saúde que estão sendo oferecidos, quanto para resguardá-lo, com a finalidade de uma boa comunicação com a equipe multiprofissional de saúde e entre elas e os pacientes. A documentação supramencionada, serve também como fonte de pesquisa, monitoramento de doenças e planejamento de políticas públicas para a população.

Isto posto, sabemos que todos os registros referentes a Documentação Sanitária relacionados às informações de saúde, são de extrema importância para o levantamento de dados da população em geral, assim como também a precisão e confiabilidade das informações. Portanto, o cuidado tanto na produção, no manuseio e guarda dessas informações clínicas exigem cautelas e providências que devem ser seguidas de acordo com suas leis regulamentares. Desse modo, tanto as instituições de saúde, quanto os profissionais da área estão cientes das obrigações relacionadas ao registro dessa documentação.

As informações registradas nos prontuários seguem uma terminologia bem particular da área da saúde de modo que os profissionais envolvidos tenham o mínimo possível de interferências, no processo de comunicação.

3 Terminologia

O avanço do desenvolvimento científico e tecnológico trouxe mudanças de paradigmas que afetam de modo geral, todos os campos de conhecimento. Tal fato, naturalmente, resultou, entre outras coisas, na grande produção de conhecimentos e, conseqüentemente, o uso das palavras do léxico comum, ganhou outras proporções, e com isso, foram percebidas interferências no processo de comunicação. Tal fato veio a exigir a padronização do conhecimento produzido, de modo que aquela informação não ficasse "perdida", sem recuperação, sem acesso, sem utilidade. É então, diante disso, que foi percebido a iminência de padronização dos termos, e entra em cena a Terminologia da especialidade, tendo em Wuster, o seu precursor ao publicar a Teoria Geral da Terminologia, nos anos 30 do século passado.

Para alguns pesquisadores, a proposta de Wuster parecia muito positiva, enquanto que para outros, trazia empecilhos devido a sua linearidade. Nesse sentido, entram em cena outras teorias terminológicas a saber: Socioterminologia, Teoria Comunicativa da Terminologia (TCT), Teoria Sociocognitiva da Terminologia (TST) e a Terminologia Cultural.

3.1 Correntes Teóricas da Terminologia

Surge então, a Socioterminologia buscando analisar o termo a partir de uma perspectiva sociolinguística, dando ênfase ao significado. Gaudin (2005, p. 81) aponta que a Socioterminologia estuda a "circulação dos termos em sincronia e diacronia, incluindo análise e modelagem de significados e conceituações". A partir de um olhar sociocrítico, aborda a relação entre a produção de sentido dos termos às condições de sua origem, sob a ótica da diversidade dos seus usos sociais. Faulstich (2006, p. 29) reitera que essa corrente abre espaço para os diferentes discursos especializados, entre eles, os contextos orais, pois entende que os termos variam, e essas variantes devem ser levadas em consideração no momento de suas elaborações. Auxiliando assim, na mediação entre terminologia e sociedade.

Outra proposta teórica criada foi a Teoria Comunicativa da Terminologia (TCT), sistematizada por Maria Teresa Cabré. A TCT, como o próprio nome diz, destaca o aspecto comunicativo da terminologia. Barros (2006, p. 23) afirma que os termos, da perspectiva da TCT, podem ser vistos de diferentes ângulos e analisados segundo os modelos das ciências linguísticas, cognitivas e da comunicação, como uma entidade poliédrica. Cabré *et al.* (1998, p. 4) reitera que os conceitos são influenciados por fatores socioculturais e linguísticos de uma comunidade, bem como, os canais de comunicação pelos quais eles circulam e são disseminados. Sales e Café (2009, p. 104) reiteram que, a partir da TCT, o termo é visto como "unidade de conhecimento, composto por uma

forma e um conteúdo”, onde a forma é a unidade lexical que denomina um determinado conceito, que é o conteúdo.

Já a Teoria Sociocognitiva da Terminologia (TST), proposta por Temmerman (1997, p. 55) sugere o uso da semântica cognitiva na terminologia, pois esta se utiliza da interação “entre o mundo, a linguagem e a mente humana” e da percepção destes em funcionamento no ambiente social. De acordo com Brangel (2010, p. 6), a TST trata os termos como unidades de entendimento, pois estes, são produtos da cognição humana, que podem ser conduzidos por diversos fatores. E por isso, é impossível reduzir um termo a um conceito padronizado. A delimitação do conteúdo, portanto, toma como base o texto no qual o termo está inserido, dessa forma, adotando assim, uma nova concepção ao termo.

Outro pesquisador interessado pelos estudos da terminologia, é Marcel Diki-Kidiri, que trouxe o enfoque da Terminologia Cultural. Diki-Kidiri (2009, p. 1) afirma que a terminologia acaba sendo uma disciplina tanto para a construção do conhecimento, como para apropriação de uma determinada cultura, e por isso, em seus fundamentos teóricos e métodos, devem ser considerados aspectos socioculturais, históricos, fenomenológicos e psicológicos. Ainda segundo o autor, essa abordagem cultural

[...] contribui para o recente fluxo sociocognitivo de pesquisa terminológica e que leve cada vez mais em conta a diversidade cultural no próprio processo de elaboração do terminologia como ciência da linguagem. Isso nos obriga a estabelecer uma metodologia mais adequada para o procedimento utilizado, para garantir a produção ideal dos termos como produtos de línguas culturalmente integradas. (DIKI-KIDIRI, 2009, p.1, tradução nossa).

Reis e Ferreira (2017, p. 198) afirmam que na perspectiva da Terminologia Cultural, “os termos ganham uma dimensão social, sobretudo, através da cultura, considerada o pilar dessa abordagem”, onde é valorizado a apropriação dos saberes para a criação de representações simbólicas. Trazendo assim, uma outra concepção da Terminologia.

Já no Brasil, os estudos de Terminologia tiveram bastante destaque no ano 1995, com a iniciativa do Projeto Banco de Dados Terminológico do Brasil (BrasilTerm), uma parceria do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e o LIV/UnB, que visava a implantação e difusão de Terminologia Científica e Técnica no Brasil, bem como, a capacitação de especialistas neste domínio.

Em suma, percebe-se que os desenvolvimentos e as evoluções da Terminologia durante o decorrer do tempo, são direcionados para a normalização e padronização de termos em uma área de especialidade, tendo como objetivo a comunicação, não apenas entre

profissionais, mas também com a sociedade, e por isso, busca abranger o conhecimento como todo e não apenas o científico. A terminologia então, por ser uma ferramenta comum a todas as áreas, passa a ser explorada por vários campos de conhecimentos, a exemplo da saúde, como citado anteriormente, não se mantendo apenas na documentação ou na Ciência da Informação.

Neste trabalho, portanto, optou-se pela abordagem da Teoria Comunicativa da Terminologia (TCT), tendo em vista a relevância comunicativa da terminologia no contexto do prontuário do paciente.

4 Metodologia

Trata-se de uma pesquisa bibliográfica, de natureza exploratória, descritiva tendo-se enfoque quantitativo, cuja a empiria se efetivou no portal de periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Biblioteca Eletrônica Científica Online (SCIELO), Base de Dados de Ciência da Informação (BRAPCI) e Biblioteca Virtual de Saúde (BVS).

A estratégia utilizada foi a busca de termos como “documentação médica”, “prontuário do paciente”, “*documentación sanitaria*”, “competência informacional”, “competência em informação” e “mediação”, tendo como período coberto de 1999-2022, nas áreas da Ciência da Informação e da Saúde e nos idiomas português e espanhol. A pesquisa foi feita utilizando-se as palavras-chave presentes nos resumos nas línguas escolhidas para esta pesquisa.

Os *corpora* foram constituídos de todos os artigos referentes às temáticas obtidos no levantamento bibliográfico nas bases de dados anteriormente mencionados.

De posse da terminologia mapeada nas palavras-chave dos resumos, *abstracts* e *resumén*, ressaltamos que conforme Cabré (1993, p. 181) as fichas terminológicas

são materiais estruturados que devem conter todas as informações relevantes sobre cada termo. As informações que apresentam, são extraídas [...] da documentação de referência, e é representada segundo critérios previamente estabelecidos. Existem diversos modelos de fichas terminológicas, conforme os objetivos de cada função e as necessidades de cada organização. (CABRÉ, 1993, p.181, tradução nossa)

Para a estruturação da ficha terminológica neste artigo, adotamos o modelo proposto por Barros (2004) considerando os seguintes elementos: termo, classificação, termos sinônimos, definição, contextos, termos associados e notas.

4 Análise dos Dados e Discussão dos Resultados

Os resultados da pesquisa evidenciaram um total de 724 artigos, sendo estes publicados na área da Ciência da Informação e da Saúde, nos idiomas português e espanhol. Entretanto, após análise dos resumos, somente 55 artigos tratavam da temática em questão. Após a seleção dos artigos, foi mapeada a terminologia das palavras-chave utilizadas, e do total de 212 termos, após exclusão devido às repetições, foram totalizados 154 termos coletados.

Os termos com mais repetições são: Prontuário do Paciente (11), Ciência da Informação (8), Prontuário Eletrônico do Paciente (7), Sistemas de Informação (5), Registros Eletrônicos de Saúde (4), Sistemas Computadorizados de Registros Médicos (4), Brasil (3) e Documentação Clínica (3). Os termos com menor frequência de repetição são: Atenção Primária à Saúde, Auxílio à decisão, Codificação, Documentação Sanitária, Gestão Documental, Historia Clínica, Indexação, Instalações de Saúde, Interoperabilidade da Informação em Saúde, Percepção do Paciente, Representação da informação, Representação indexal, Revisão de Literatura, Saúde, Saúde Pública, Segurança do Paciente, Sistemas de Informação Hospitalar. Todos os termos supracitados tiveram 2 repetições cada. Os demais termos que não estão evidenciados aqui, não tiveram repetições. Por conseguinte, foi elaborada uma nuvem de palavras com todos os termos - Imagem 1.

Imagem 1. Nuvem de palavras com os termos



Fonte: Dados da Pesquisa.

Tabela 1. Termos com maior número de repetições

TERMOS	QTD
Prontuário do Paciente	11
Ciência da Informação	8
Prontuário Eletrônico do Paciente	7
Sistemas de Informação	5
Registros Eletrônicos de Saúde	4
Sistemas Computadorizados de Registros Médicos	4
Documentação Clínica	3
Brasil	3

Fonte: Dados da Pesquisa.

Em um segundo momento, construímos a *ficha terminológica - Apêndice A*. De acordo com a ISO (1990), “uma ficha terminológica é um conjunto estruturado de dados terminológicos relativos a uma noção.” A ficha terminológica foi elaborada a partir do resultado da pesquisa. Estando constituída com o total dos termos (154), foi feita a classificação destes em três grupos: *Termo Principal* (1), aquele que encabeça a ficha terminológica e é o qual norteará a temática da ficha. O escolhido foi Documentação Sanitária, devido ser a temática central desta pesquisa. *Termos Sinônimos* (14), são aqueles que fazem referência ao termo principal, porém em outras palavras. Sendo eles: Arquivo hospitalares; Documentação clínica; Documentação de saúde; Documentação hospitalar; Documentação Técnica; Prontuário do Paciente; Prontuário do paciente baseado em computador; Prontuário Eletrônico; Prontuário Eletrônico do Paciente (PEP); Registros; Registros de Enfermagem; Registros Eletrônicos de Saúde; Registros Médicos; Sistemas Computadorizados de Registros Médicos. E por fim, os *Termos Associados* (139), que são termos que estavam presentes nas áreas investigadas porém não se enquadram nos termos sinônimos, entretanto, estão vinculados ao termo principal de alguma maneira. Sendo eles: Abreviaturas; Acessibilidade; Acesso legal; Acrônimos; Administração da justiça; Agir do bibliotecário; Anamnese; Aplicações de informática médica; Arquitetura da informação; Arquivos; Atenção Primária à Saúde; Atividade assistencial; Atuação do Bibliotecário na Área da Saúde; Atuação profissional; Auxílio à decisão; Avaliação; Bases de dados; Bibliotecário Clínico; Bibliotecário em Hospitais; Bibliotecários médicos canadenses; Bibliotecas médicas; Biomedicina; BioNER; BioNLP; Brasil; Cidadania; CIE-9-MC; Ciência da Informação; Ciência de Dados; Classificações sanitárias; Classificação Internacional de Doenças; Codificação; Codificação de

Prontuários do Paciente; Comitê de Ética em Pesquisa; Competência Informacional; Competências; Competências e Habilidades; Comunicação; Comunicação social; Consentimento; Controle de Formulários; Curadoria digital; Diagnóstico primário; Diagnóstico secundário; Difusão de inovações; Disseminação de informações; E-saúde médica; Enfermagem; Estratégias de eSaúde; Estruturação da Informação; Ética da Informação; Evolução Clínica; Fonte de Informação; Gestão da Informação Orgânica; Gestão clínica; Gestão da Informação; Gestão da Informação em Saúde; Gestão da Qualidade; Gestão de Bibliotecas; Gestão Documental; Gestão dos Fluxos de Informação; Gestores de Informação em Saúde; GRD; História clínica; Indexação; Infecções por Coronavírus; Informação; Informação clínica; Informações do consumidor; Informática em saúde; Informática Médica; Informática na doença; Inovação em Gestão da Informação e do Conhecimento; Instituições de saúde; Internet; Interoperabilidade sintática; Interoperabilidade da Informação em Saúde; Interoperabilidade semântica; Jornalismo médico; Legislação Médica; Lei de Acesso à Informação; Linguística textual; Mediação informacional; Medicina Baseada em Evidência; Meios de comunicação; Memória social da saúde; NLP; Ontologias de saúde; Organização da Informação; Organização do conhecimento; Paciente; Pacientes hospitalizados; Padrões de interoperabilidade; Padrões de Referência; Pandemias; Percepção do paciente; Perfis profissionais; Pessoal de Saúde; Pneumonia Viral; Política de indexação; Práticas do Profissional da Informação; Preservação digital; Processo de linguagem natural; Produção científica; Profissional da Informação na área de saúde; Recuperação de informações; Redes de comunicação de computadores; Redução Léxica em Resumo de Alta; Relação médico-paciente; Representação da informação; Representação da informação – Indexação; Representação indexal; Representação Temática da Informação; Resumo de Alta; Revisão de literatura; Revisão Sistemática; Saúde; Saúde Pública; Segurança Computacional; Segurança do paciente; Semântica; Serviços de Saúde; Serviços Médicos de Emergência; Sigilo informacional; Siglas; Símbolos; Sistema Único de Saúde; Sistemas avançados integrados de gerenciamento de informações; Sistemas de Informação; Sistemas de Informação em Saúde; Sistemas de Informação Hospitalar; Sistemas de Saúde; Taxonomias de saúde; Tecnologia da Informação; Telemedicina; Terminologia em saúde; Terminologias; Text-mining; Unidade de dor.

Ressaltamos que concernente ao termo competência ainda foi pouco evidenciado (1) e mediação (1), cuja ocorrência deu-se somente em um documento. Talvez esse fato seja decorrente da pouca inserção de pesquisas envolvendo essas temáticas, em uma perspectiva interdisciplinar com a Informação para a Saúde, que ainda estejam debutando.

3 Conclusão

A finalização deste artigo nos remete ao retorno de onde partimos, a fim de verificar se a questão-problema foi respondida satisfatoriamente. Nesse contexto, saltam aos olhos o sintagma prontuário do paciente, ciência da informação e prontuário eletrônico do paciente.

Inferimos também grande importância do diálogo entre a Terminologia e a Ciência da Informação para contribuir com a pesquisa, particularmente nesta, cujo objetivo geral é analisar a terminologia relacionada a Documentação Sanitária presente na literatura científica dos campos de Ciência da Informação e da Saúde.

Assim, foram identificados vários termos relacionados à documentação sanitária, embora que no Brasil, essa terminologia esteja sendo adotada, ainda, de forma elementar.

Além disso, constata-se também as contribuições da terminologia para o melhor entendimento da temática deste artigo, que a ficha terminológica apresenta o termo, definição, sinônimos, contextos, termos associados e notas relacionadas à documentação sanitária que aqui no Brasil, refere-se ao prontuário do paciente.

Notas

As autoras agradecem ao CNPQ, o financiamento desta pesquisa, por meio do Projeto Universal - 2021.

Referências

- ALMEIDA, M. B.; ANDRADE, A. Q. Organização da informação em prontuários de pacientes: uma abordagem Popperiana. **Informação & Tecnologia**, [S. l.], v. 1, n. 1, p. 29–41, 2014. Disponível em: <https://periodicos.ufpb.br/index.php/itec/article/view/19195>. Acesso em: 31 mar. 2023.
- BARROS, Lídia Almeida. Aspectos epistemológicos e perspectivas científicas da terminologia. **Ciência e Cultura**, v. 58, n. 2, p. 22-26, 2006. Disponível em: http://cienciaecultura.bvs.br/scielo.php?pid=s00096725200600200011&script=sci_arttext&tlng=pt. Acesso em: 14 mar. 2023.
- BARROS, L. A. **Curso básico de Terminologia**. São Paulo: Editora da Universidade de São Paulo, 2004. 285 p.
- BENTES PINTO, V. Prontuário eletrônico do paciente: documento técnico de informação e comunicação do domínio da saúde. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 11, n. 21, p. 34–48, 2007. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2006v11n21p34>. Acesso em: 27 mar. 2023.
- BRANGEL, L. M. . Quando uma cor é um termo: em busca de uma perspectiva de abordagem terminológica. *In: Semana*

- de Letras da PUCRS, 10., 2010, Porto Alegre. **Anais....** Porto Alegre: Edipucrs, 2010. Disponível em: <https://ebooks.pucrs.br/edipucrs/anais/Xsemanadeletras/comunicacoes/Larissa-Moreira-Brangel.pdf>. Acesso em: 20 mar. 2023.
- BRIET, Suzanne. **O que é a documentação**. Trad. Maria Nazareth Fendt. Brasília, DF: Briquet de Lemos, 2016. Disponível em: https://edisciplinas.usp.br/pluginfile.php/5389052/mod_resouce/content/1/O_que_%C3%A9_a_documenta%C3%A7%C3%A3o_Parapublicar.pdf. Acesso em: 10 abr. 2023.
- CABRÉ, M. T. C.; FREIXA, J.; LORENTE, M.; TEBÉ, C. La terminología hoy: replanteamiento o diversificación. **Organon**, Porto Alegre, v. 12, n. 26, 1998. DOI: 10.22456/2238-8915.29557. Disponível em: <https://www.seer.ufrgs.br/index.php/organon/article/view/2957>. Acesso em: 15 mar. 2023.
- CABRÉ, M. T. C. **La terminología: teoría, metodología, aplicaciones**. Barcelona: Editorial Antártida/Empúries, 1993. p. 281-282.
- CASABONA, Carlos María Romeo. **La ética y el derecho ante la biomedicina del futuro**. Bilbao: Universidad de Deusto, 2006. 356 p. v. 86. ISBN 8498300614.
- CONSELHO FEDERAL DE MEDICINA (Brasil). (2002). **Resolução nº 1.638, de 10 de julho de 2002**. Dispõe sobre o prontuário médico e torna obrigatória a criação da Comissão de Revisão de Prontuários nas instituições de saúde. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília.
- CONSELHO REGIONAL DE MEDICINA DO DISTRITO FEDERAL. **Prontuário médico do paciente: guia para uso prático**. Brasília, 2006. Disponível em: <https://www.saudedireta.com.br/docsupload/1370271458PEP.pdf>. Acesso em: 12 abr. 2023.
- DIKI-KIDIRI, Marcel; BIGOUNDU, Rodrigue. Un enfoque cultural de la terminología. **Debate Terminológico**, n. 05, 2009. ISSN: 1813-1867. Disponível em: <https://www.seer.ufrgs.br/riterm/article/view/23955>. Acesso em: 12 abr. 2023.
- FAULSTICH, Enilde. A socioterminologia na comunicação científica e técnica. **Ciência e Cultura**, v. 58, n. 2, p. 27-31, 2006. Disponível em: http://cienciaecultura.bvs.br/scielo.php?pid=S0009-67252006000200012&script=sci_arttext. Acesso em: 16 mar. 2023.
- GAUDIN, François. La socioterminologie. *In: Langages*, 39^e année, nº157, 2005. La terminologie: nature et enjeux. p. 80-92. Disponível em: https://www.persee.fr/doc/lgge_0458-726x_2005_num_39_157_976. Acesso em: 10 abr. 2023.
- GOMES, H. E. Terminologia e estrutura conceitual. **PontodeAcesso**, [S. l.], v. 15, n. 3, 2021. DOI: 10.9771/rpa.v15i3.47464. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/47464>. Acesso em: 31 mar. 2023.
- INTERNATIONAL STANDARD ORGANIZATION, Geneva. **ISO 1087-1990: Terminology / Vocabulary**. Geneva, 1990. Disponível em: <https://www.iso.org/standard/5591.html>. Acesso em: 05 abr. 2023.
- LARA, Marilda Lopes Ginez de. **Elementos de terminologia: (apostila para uso didático)**. São Paulo: ECA-USP, 2005. Disponível em: <https://bibliotextos.files.wordpress.com/2012/03/elementos-de-terminologia.pdf>. Acesso em: 20 mar. 2023.
- OLIVEIRA, Juliana Buse de. Conservação preventiva de documentação sanitária em papel. *In: PINTO, Virginia Bentes; CAMPOS, Henry de Holanda; SALES, Odete Mayra Mesquita; OLIVEIRA, Henry Poncio Cruz de. (Orgs.). Múltiplas abordagens sobre informação e tecnologias emergentes para a saúde*. João Pessoa: Editora UFPB, 2017, p. 161-179.
- EIS, J. de A.; FERREIRA, M. de N. de O. Pajé, conhecimento cultural e terminologia de plantas medicinais em parkatêjê. **Espaço Ameríndio**, Porto Alegre, v. 11, n. 2, p. 186, 2017. DOI: 10.22456/1982-6524.70276. Disponível em: <https://www.seer.ufrgs.br/index.php/EspacoAmerindio/article/view/70276>. Acesso em: 11 abr. 2023.
- SALES, Rodrigo de; CAFÉ, Lígia. Diferenças entre tesouros e ontologias. **Perspectivas em Ciência da Informação**, v. 14, p. 99-116, 2009. Disponível em: <https://www.scielo.br/j/pci/a/qr7hthZ93SnkHfRybCBLBxd/a/bstract/?lang=pt>. Acesso em: 10 abr. 2023.
- SCHAEFER, Fernanda. **Proteção de dados de saúde na sociedade de informação: a busca pelo equilíbrio entre privacidade e interesse social**. Curitiba: Juruá, 2010.
- SPRAGUE JR, Ralph H. **Electronic document management: Challenges and opportunities for information systems managers**. https://www.getty.edu/conservation/publications_resources/pdf_publications/recordim.html. Acesso em: 12 abr. 2023.
- TEMMERMAN, R. Questioning the univocity ideal: the difference between socio-cognitive Terminology and traditional Terminology. **HERMES - Journal of Language and Communication in Business**, [S. l.], v. 10, n. 18, p. 51-90, 1997. DOI: 10.7146/hjlc.v10i18.25412. Disponível em: <https://tidsskrift.dk/her/article/view/25412>. Acesso em: 12 apr. 2023.

Apêndice A – Ficha Terminológica Geral

FICHA TERMINOLÓGICA GERAL	
Termo	Documentação Sanitária
Classificação	1 - Termo Principal
Termo Sinônimo	Arquivo hospitalares; Documentação clínica; Documentação de saúde; Documentação hospitalar; Documentação Técnica; Prontuário do Paciente; Prontuário do paciente baseado em computador; Prontuário Eletrônico; Prontuário Eletrônico do Paciente (PEP); Registros; Registros de Enfermagem; Registros Eletrônicos de Saúde; Registros Médicos; Sistemas Computadorizados de Registros Médicos.
Definição	O artigo 1º da Resolução nº 1.638/2002 do Conselho Federal de Medicina do Brasil define o prontuário do paciente como “o documento único constituído de um conjunto de informações, sinais e imagens registradas, geradas a partir de fatos, acontecimentos e situações sobre a saúde do paciente e a assistência a ele prestada, de caráter legal, sigiloso e científico, que possibilita a comunicação entre membros da equipe multiprofissional e a continuidade da assistência prestada ao indivíduo”.
Contexto	Ciência da Informação: Bentes Pinto (2006, p.39): é um documento que contém registradas todas as informações concernentes a um paciente, sejam elas de caráter de identificação, socioeconômico, de saúde (as observações dos profissionais da saúde, as radiografias, as receitas, os resultados dos exames, o diagnóstico dos especialistas, as notas de evolução redigidas pelo pessoal da enfermagem com relação ao progresso observado) ou administrativo, dentre outros. [...] Trata-se da memória escrita da história da pessoa doente, sendo, portanto, indispensável, para a comunicação intra e entre a equipe de saúde e o paciente, a continuidade, a segurança, a eficácia e a qualidade de seu tratamento, bem como da gestão das organizações hospitalares. Almeida e Andrade (2014, p. 29) afirma que “os prontuários são, na verdade, tipos documentais produzidos e utilizados no âmbito das atividades de cuidados à saúde relevantes para a gestão da informação nas instituições de saúde”.
Alguns termos associados nas áreas investigadas	Abreviaturas; Acessibilidade; Acesso legal; Acrônimos; Administração da justiça; Agir do bibliotecário; Anamnese; Aplicações de informática médica; Arquitetura da informação; Arquivos; Atenção Primária à Saúde; Atividade assistencial; Atuação do Bibliotecário na Área da Saúde; Atuação profissional; Auxílio à decisão; Avaliação; Bases de dados; Bibliotecário Clínico; Bibliotecário em Hospitais; Bibliotecários médicos canadenses; Bibliotecas médicas; Biomedicina; BioNER; BioNLP; Brasil; Cidadania; CIE-9-MC; Ciência da Informação; Ciência de Dados; Classificações sanitárias; Classificação Internacional de Doenças; Codificação; Codificação de Prontuários do Paciente; Comitê de Ética em Pesquisa; Competência Informacional; Competências; Competências e Habilidades; Comunicação; Comunicação social; Consentimento; Controle de Formulários; Curadoria digital; Diagnóstico primário; Diagnóstico secundário; Difusão de inovações; Disseminação de informações; E-saúde médica; Enfermagem; Estratégias de eSaúde; Estruturação da Informação; Ética da Informação; Evolução Clínica; Fonte de Informação; Gestão da Informação Orgânica; Gestão clínica; Gestão da Informação; Gestão da Informação em Saúde; Gestão da Qualidade; Gestão de Bibliotecas; Gestão Documental; Gestão dos Fluxos de Informação; Gestores de Informação em Saúde; GRD; História clínica; Indexação; Infecções por Coronavírus; Informação; Informação clínica; Informações do consumidor; Informática em saúde; Informática Médica; Informática na doença; Inovação em Gestão da Informação e do Conhecimento; Instituições de saúde; Internet; Interoperabilidade sintática; Interoperabilidade da Informação em Saúde; Interoperabilidade semântica; Jornalismo médico; Legislação Médica; Lei de Acesso à Informação; Linguística textual; Mediação informacional; Medicina Baseada em Evidência; Meios de comunicação; Memória social da saúde; NLP; Ontologias de saúde; Organização da Informação; Organização do conhecimento; Paciente; Pacientes hospitalizados; Padrões de interoperabilidade; Padrões de Referência; Pandemias; Percepção do paciente; Perfis profissionais; Pessoal de Saúde; Pneumonia Viral; Política de indexação; Práticas do Profissional da Informação; Preservação digital; Processo de linguagem natural; Produção científica; Profissional da Informação na área de saúde; Recuperação de informações; Redes de comunicação de computadores; Redução Léxica em Resumo de Alta; Relação médico-paciente; Representação da informação; Representação da informação – Indexação; Representação indexal; Representação Temática da Informação; Resumo de Alta; Revisão de literatura; Revisão Sistemática; Saúde; Saúde Pública; Segurança Computacional; Segurança do paciente; Semântica; Serviços de Saúde; Serviços Médicos de Emergência; Sigilo informacional; Siglas; Símbolos; Sistema Único de Saúde; Sistemas avançados integrados de gerenciamento de informações; Sistemas de Informação; Sistemas de Informação em Saúde; Sistemas de Informação Hospitalar; Sistemas de Saúde; Taxonomias de saúde; Tecnologia da Informação; Telemedicina; Terminologia em saúde; Terminologias; Text-mining; Unidade de dor.
Notas	Percebe-se que ocorre a substituição do Termo Principal pelo termo “Prontuário do Paciente” no Brasil. O termo Prontuário do Paciente sofre muitas variações em sua redação, particularmente em línguas neolatinas: português, francês, espanhol.

Fonte: Dados da Pesquisa.

A utilização de dados como moedas digitais através da coleta por *cookies*

The use of data as digital currency through its collection by cookies

Bárbara Souza da Silva (1), Deise Maria Antonio Sabbag (2)

(1) Universidade Estadual Paulista (UNESP), Marília - SP, bs.silva@unesp.br. (2) Universidade de São Paulo, (USP/FFCLRP), Campus de Ribeirão Preto; Pós-Graduação em Ciência da Informação (UNESP/Campus de Marília), deise.sabbag@unesp.br.

Resumo

Configura-se um trabalho onde explora um debate sobre datificação e o uso dos dados como um novo formato de valor monetário dentro de um ambiente virtual, juntamente com um aprofundamento acerca do dataísmo em uma sociedade informacional. Designou-se uma busca acerca das definições dos conceitos norteadores do texto, individual e como se dá suas relações. O trabalho se baseia em recortes sobre o papel ao qual se dá os *cookies* e o serviços personalizados respaldados nos perfis dos usuários enquanto se utilizam algum serviço online, no qual reflete na mentalidade do dataísmo, onde encara os dados resultados de usuários como máxima na sociedade a qual se insere. A pesquisa advém com um caráter bibliográfico substancial para o levantamento do referencial teórico o qual, permitiu que houvesse um aprofundamento sobre como as percepções sobre os dados dos usuários são levemente convertidas em moedas digitais dissimuladamente vendidas a desculpa de personalização de serviços quanto há um armazenamento e análise do que se obtém a fim de capitalizar as informações cedidas ao navegar na web; com o provedor de informações inconsciente do seu produto.

Palavras-chave: Datificação. Dataísmo. Sociedade Informacional.

Abstract

This work explores the debate on datification and the use of data as a new format of monetary value inside a virtual environment, together with a deepening on dataism in an informational society. A search was made for the definitions of the guiding concepts of the text, individual and how their relations take place. It is based on clippings about the role played by *cookies* and personalised services based on users' profiles when using online services, which reflects the mentality of dataism, where data is seen as a maxim of the society in which it is inserted. The research comes with a substantial bibliographic character for the survey of the theoretical framework which, allowed there to be a deepening on how the perceptions of user data are lightly converted into digital currencies covertly sold the excuse of personalization of services when there is a storage and analysis of what is obtained in order to capitalize on the information given when browsing the web; with the information provider unaware of your product.

Keywords: Datification. Informational Society. Dataism.

Resumen

Esta obra explora el debate sobre la datificación y el uso de los datos como nuevo formato de valor monetario dentro de un entorno virtual, junto con una profundización sobre el dataísmo en una sociedad informacional. Se han buscado definiciones de los conceptos que guían el texto, así como conceptos individuales y sus relaciones. El estudio se basa en recortes sobre el papel que desempeñan las *cookies* y los servicios personalizados basados en los perfiles de los usuarios cuando utilizan servicios en línea, lo que refleja la mentalidad del dataísmo, que ve los datos resultantes de los usuarios como una máxima en la sociedad en la que se inserta. La investigación viene con un carácter bibliográfico sustancial para el levantamiento del marco teórico que, permitió que hubiera una profundización sobre cómo las percepciones de los datos de los usuarios se convierten a la ligera en monedas digitales vendidas encubiertamente la excusa de la personalización de los servicios cuando hay un almacenamiento y análisis de lo que se obtiene con el fin de capitalizar la información dada a navegar por la web; con el proveedor de información ajeno a su producto.

Palabras clave: Datificación. Sociedad de la Información. Dataísm.

1 Introdução

A realização de censo populacional, a pedido do Estado, é uma prática comum, visando coletar dados sobre a população, a fim de compreender a realidade do país e propor melhorias ou novos empreendimentos

que atendam às demandas da população. Esse tipo de coleta não se trata de um esquema desconhecido a humanidade, pois de acordo com Mayer-Schönberger e Cukier (2013), o registro mais antigo referente a esse assunto se trata de uma tabulação de bens comerciáveis por sumérios 8000 a.C. Ao longo do tempo, o

aprimoramento dessa prática esteve relacionado ao interesse dos Estados em coletar, organizar e manter informações sobre sua população.

Empresas como *Facebook*, *Google*, *Twitter* e *LinkedIn* tem datificado rede de amigos, busca e recuperação de informações, notícias e rede de contatos profissionais. De forma com que cada empresa de relacionamento busca aproveitar essa abundância de dados não estruturados (*big data*) para armazenar, analisar e rentabilizar a informação em torno das suas atividades e negócios (Antoniutti, 2015, p. 54).

Ao realizar o cadastro em uma plataforma provedora de serviços web, é necessário o cadastramento de informações para identificação do novo perfil que ali será criado, tais como nome, e-mail, data de nascimento e uma senha recém-criada para a entrada na plataforma. Além da concordância com os termos de uso e de privacidade daquele site. Contudo, a natureza do que se é dado o consentimento por vezes é deixada de lado a sua compreensão. Embora a privacidade seja associada à esfera íntima e a documentos pessoais, como a identidade ou passaporte, no contexto do aumento da presença na web e da utilização das redes sociais, a privacidade adquire contornos diferentes daqueles a quais normalmente é relacionada.

Ignorar os termos de adesão é uma prática comum realizada por novos usuários das redes sociais. Esse comportamento implica em conscientemente desconsiderar o impacto das políticas e termos de utilização e desconhecer as possíveis consequências decorrentes disso. Salientamos que a inserção de informações em uma plataforma digital, consiste em uma troca informativa para a utilização dos recursos online, os quais são indispensáveis e valiosos para os indivíduos e, para aqueles que cedem o seu uso.

O processo de coleta de dados por parte das empresas fornecedoras de tecnologias de informação e comunicação, é um procedimento natural de funcionalidade; a recolha de informações produzidas pelos usuários é voltada ao aperfeiçoamento dos produtos disponibilizados para posteriores usos. Essa característica de funcionamento é comum de ser encontrada em *cookies*.

O entendimento de que há uma possível troca de informações se deve ao fato do aproveitamento que se dá perante os dados colhidos. O que segundo Cavalcanti (2021) configura como uma forma de escambo digital perante a utilização do produto. Conforme mencionado, ocasionalmente quando se inscreve em rede social disponibilizamos informações para a nossa identificação, o contrário desse processo, ou seja, a rede social nos identificar, também, como parte do armazenamento, monitoramento e análise dos dados que ela vai atribuindo após a coleta. Não se trata apenas de quem se segue, ou curtidas, segundo Antoniutti (2015) apenas a pequena biografia

disponível em perfis como Twitter ou Instagram, serve como um meio de datificação para a rede social compreender o seu usuário e moldar o conteúdo a ser apresentado.

No presente trabalho pretende explorar os conceitos de datificação, e a sua relação com a organização e representação da informação em uma sociedade informacional onde existe um capitalismo de dados. Entendendo que para a ocorrência de uma relação no meio digital, é necessária uma moeda para essa troca, em razão disso compreendemos a existência dos *cookies* como expressão dessas moedas, que se constitui mediante a transformação dos dados vindos dos usuários em produto rentável para os fornecedores de uma plataforma na web ou em redes sociais, como se é observado no caso dos denominados “*Influencers*”, surgidos primariamente na rede social *Instagram* – cujo é parte do foco de estudo – onde o aprendizado do significado das métricas disponibilizadas para perfis profissionais, representa uma maneira de extração de conhecimento de seu público e não-público, onde se resulta em escolhas que interferem impessoalmente nas escolhas dos indivíduos. Aplicou-se uma leitura em textos acerca do que é o conceito de datificação após uma pesquisa na base de dados em ciência da informação (BRAPCI) e no Google Acadêmico afim de compreensão do termo e de seu encaixe com a proposta de análise nos termos de uso e de privacidade da rede social *Instagram*, em razão de permear-se uma base fixa para a visualização desse conceito entorno das circunstâncias que envolvem essa rede social e no papel que se atribui a ela em caráter de interferência na vida pessoal e profissional de seus consumidores.

2 Desenvolvimento

O conceito proposto por Manuel Castells nos anos finais da década de 90, acerca da sociedade informacional, compreende que as práticas realizadas no cotidiano da humanidade ganham representações em meios virtuais em razão das tecnologias de informação e comunicação.

[...] que o processamento da informação é focalizado na melhoria da tecnologia do processamento da informação como fonte de produtividade, em um círculo virtuoso de interação entre as fontes de conhecimentos tecnológicos e a aplicação da tecnologia para melhorar a geração de conhecimentos e o processamento da informação [...] (Castells, 2002, p. 54).

Para Antoniutti (2015), a sociedade informacional é uma sociedade marcada pela produção, circulação, armazenamento e controle de uma quantidade massiva de dados. Há uma desenfreada disseminação e utilização referente a informação.

A influência das redes baseadas na Internet vai além do número de seus usuários: diz respeito também à qualidade do uso. Atividades econômicas, sociais, políticas, e

culturais essenciais por todo o planeta, estão sendo estruturadas pela Internet e em torno dela, como por outras redes de computadores. (Castells, 2003, p. 8).

É possível observar a ocorrência de uma datificação das informações, ao qual é caracterizada pela produção em massa de dados por parte dos usuários que navegam na internet. Aspecto ocorrendo inconscientemente, por meio da coleta, armazenamento e análise de informações acerca dos padrões de uso dos dados associados à utilização de redes. Eventualmente, o caráter desse processo em uma forma de organização das informações, com o intuito de torná-las compreensíveis para empresas conhecidas como Big Techs, ou gigantes corporações fornecedoras de tecnologia de informação e comunicação ao redor do planeta, tais como *Amazon, Apple, Google e Meta (Facebook, Instagram e WhatsApp)*.

A informação na sociedade informacional é tratada em duas formas distintas: como matéria-prima ou como commodity. É parte de seu cerne, e é manifestado através do processo de troca ao qual pode ser caracterizada como um escambo digital.

O sistema de trocas diretas de mercadorias, conhecido como escambo, por não possuir um valor atribuído. Tal sistema foi criado em épocas em que não havia dinheiro ou moedas para serem utilizados como meio de troca. Nesse contexto, as trocas ocorriam baseadas na necessidade dos indivíduos. No entanto, é importante ressaltar que também ocorriam situações de engodo, como o ocorrido entre os colonizadores portugueses e os povos originários. Ao necessitarem extrair o Pau Brasil para a Europa, foi realizado o troca de trabalho braçal pelo território até o litoral. Segundo Cavalcanti (2021, p. 99),

[...] o escambo digital mais se afigura ao segundo exemplo que ao primeiro, isto é, suas lógicas e práticas parecem se galgar mais no engodo e na oportunidade do que na autodeterminação e no consentimento livre, pleno e desimpedido dos cidadãos, mais nos critérios de descumprimento e desconhecimento, do que no binômio necessidade/possibilidade.

Interpreta-se nesse estudo a informação como commodity/mercadoria na qual é produzida, coletada, armazenada e analisada para posteriormente, ser identificada como um padrão de usuário, sem que este tenha pleno conhecimento de sua utilização, e posteriormente tendo-lhe atribuído valor. Ao iniciar um navegador, ao pesquisar algum conteúdo, ou adentrar em algum site consegue-se resultar em uma representação do usuário a partir de sua atividade. De acordo com Agência Nacional de Proteção de Dados (2022), os *cookies* são arquivos instalados no dispositivo de um usuário que permitem a coleta de determinadas informações, inclusive de dados pessoais, em algumas situações, visando ao atendimento de finalidades diversas. É comum que quando aparece as caixas relacionadas a permissão de *cookies*, é dito que

aquele site poderá não funcionar corretamente caso não seja permitido a coleta de *cookies*. É possível que haja até diversas categorias de *cookies*, para as mais diferentes funções, como para a funcionalidade do site, publicidade de desempenho e *cookies* não necessários que são utilizados para rastrear o comportamento.

Os *cookies* são apresentados ao usuário assim que se dá o acesso a uma página na web, são arquivos de texto codificados que são gerados através das atividades do usuário perante o navegador ou site que acessa, contendo informações resultantes do modo de navegação. Alguns dados pessoais, como nome ou e-mail, ficam salvos em sites onde o cookie permitido identificará qual usuário é e quais eram as suas preferências na última navegação. Os *cookies* se tornaram nos últimos anos ferramentas indispensáveis para marketing e propaganda ao saber direcionar produtos de desejo, ao armazenar buscas de navegadores e distribuí-los em razão de publicidade. De acordo com Castelluccia (2012, p. 23-4) os *cookies* podem pertencer à própria página visitada ou a outras entidades (*cookies* de terceiros), além de serem temporariamente apagáveis (*cookies* de sessão) ou persistentes (*cookies* permanentes).

Os *cookies* de sessão são aqueles que após a navegação e fechamento do navegador pelo usuário são desfeitos normalmente, o problema envolvendo privacidade se dá entorno *cookies* de terceiros e *cookies* permanentes. Conforme o que um dos maiores softwares de proteção do mundo, Kaspersky, os *cookies* permanentes são divididos e duas categorias:

1. **Autenticação:** esses *cookies* controlam se o usuário está conectado e qual conta ele está usando. Eles também simplificam as informações de login para que os usuários não tenham de lembrar das senhas do site.
2. **Rastreamento:** esses *cookies* acompanham as várias visitas ao mesmo site ao longo do tempo. Alguns lojistas online, por exemplo, usam *cookies* para controlar as visitas de usuários específicos, incluindo as páginas e produtos vistos. As informações que eles obtêm permitem sugerir outros itens que possam interessar aos visitantes. Gradualmente, é criado um perfil com base no histórico de navegação do usuário nesse site.

Os *cookies* de terceiros são aqueles que são gerados por outros sites que tem conteúdo apresentado o site em que está acessando. Sem que o usuário tenha o conhecimento, podem ser gerados mais de dez (10) *cookies* de terceiros apenas por uma visita ou clique em um site o qual já é acostumado a utilizar. Segundo Kaspersky, esse tipo de *cookie* permite que os anunciantes ou as empresas de análise acompanhem o histórico de navegação de um indivíduo em toda a *Web*, nos sites que contêm seus anúncios.

É elencável os *cookies* zumbis, consistentes na impossibilidade ou rara exclusão dos arquivos do computador ou smartphone, são *cookies* nonde os usuários não escolheram não instalar e são ligeiramente infiltrados nos seus dispositivos e que acabam

tornando-se um meio de rastreamento por parte de pessoas ou corporações. Além dessa definição, a Agência Nacional de Proteção de Dados (2022), criou um guia acerca dos *cookies* onde os categoriza em 1) entidade responsável pela sua gestão; 2) necessidade; 3) finalidade e 4) período de retenção de informação.

Tabela 1. *Compilação das características de cookies consoantes com o guia orientativo da Agência Nacional de Proteção de Dados.*

<i>Cookies de acordo com a entidade responsável pela sua gestão</i>	<p><i>Cookies próprios ou primários:</i> Os <i>cookies</i> primários geralmente não podem ser usados para a estrear atividade em outro site que não seja o site original em que foi colocado. Esses tipos de <i>cookies</i> podem incluir informações como credenciais de login, itens do carrinho de compras ou Idioma preferido.</p>
<i>Cookies de acordo com a necessidade</i>	<p><i>Cookies de terceiros:</i> são criados por um domínio diferente daquele em que o usuário está visitando. Decorrem de funcionalidades de outros domínios que são incorporadas a uma página eletrônica como, por exemplo, exibição de anúncios.</p>
<i>Cookies de acordo com a necessidade</i>	<p><i>Cookies necessários:</i> são utilizados para que o site ou aplicativo realize funções básicas de operação. A sua coleta é essencial para assegurar o funcionamento da página eletrônica ou para adequada prestação de serviço.</p>
<i>Cookies de acordo com a necessidade</i>	<p><i>Cookies não necessários:</i> são quantos aqueles que não se encaixam na definição de <i>cookies</i> necessários e cuja desabilitação não impede o funcionamento do site ou aplicação ou a utilizando os serviços pelo usuário; exemplos desses <i>cookies</i> são os utilizados para rastrear comportamentos, medir o desempenho da página ou serviço além de exibir anúncios ou outros conteúdos incorporados.</p>
<i>Cookies de acordo com a necessidade</i>	<p><i>Cookies de funcionalidade:</i> os atos para fornecimento dos serviços básicos solicitados pelo usuário e possibilitam lembrar preferências do site ou aplicativo como o nome do usuário, região e idioma. Dentre esses <i>cookies</i> é possível incluir <i>cookies</i> próprios, de terceiros, persistentes ou de sessão.</p>
<i>Cookies de acordo com a necessidade</i>	<p><i>Cookies de publicidade:</i> são usados para coletar informações do titular para exibição de anúncios. São coletados a partir de informações relativas aos hábitos de navegação do usuário, eles então permitem a identificação e a construção de perfis para a Futura exibição de anúncios personalizados.</p>
<i>Cookies de acordo com o período de retenção das informações</i>	<p><i>Cookies de sessão ou temporários:</i> projetados para coletar, armazenar informações enquanto os usuários acessam o site. Costumo usar descartados após o exame da sessão após o usuário fechar o</p>

navegador.

Cookies persistentes: por meio desses *cookies* ficam imaginados podem ser acessados e processados por um período definido pelo controlador, que pode variar de alguns minutos a vários anos.

Fonte: Autoras.

Os *cookies* não representam malefícios aos desempenhos na web ou do próprio computador/smartphone, ou conseguem criar uma invasão por meio de vírus, ou malwares. O problema, reside na segurança e na capacidade de rastreamento de comportamento dos usuários, onde transformasse em um “Big Brother” com um nome inofensivo.

2.1 Para onde vai os meus dados?

O *Big Brother*, ou Grande Irmão em tradução literal, foi um personagem criado por George Orwell para o livro em 1984, onde a sua onipresença na vida da população é absoluta. Sua presença é predominante na vigilância constante de sua população, tornando-se uma ferramenta de controle da ordem social. Famoso no Brasil e em países como Argentina (*Gran Hermano*) e México (*La Casa dos Famosos*), o programa de reality show *Big Brother*, criado no ano de mil novecentos e noventa e nove (1999) nos Países Baixos, tem a sua proposta e nome derivados do personagem do livro de Orwell no qual, durante uma parcela de tempo, pessoas reais serão postas em confinamento para serem assistidas a todo momento com a finalidade de serem julgadas pelos telespectadores se são merecedoras do prêmio final do programa, onde no Brasil facilmente ultrapassa a marca de um milhão de reais.

Em uma perspectiva paralela, nossos aparelhos de informação e comunicação, como os smartphones realizam um trabalho semelhante com o do personagem literário.

Em 2020, despontou entre os usuários das redes sociais o documentário O Dilema das Redes, lançado na plataforma de streaming Netflix. A discussão se vê presente no grande impacto nas vidas de seus usuários, além de tingindo níveis de interferências em governos pelo globo. No documentário, é dita uma frase pelo escrito americano Andrew Lewis, “se você não paga pelo produto, o produto é você”, disseminada em larga escala essa frase não é inédita restrita ao documentário, ela foi dita por Lewis no ano de 2010 na rede social *Reddit*. Apesar de produzir e disponibilizar o documentário, o mais famoso streaming é um dos que mais utiliza a tecnologia de coleta de dados para a personalização do conteúdo a ser disponibilizado.

À medida que empresas fornecedoras de tecnologias de comunicação foram modificando seus sistemas para atender as preferências e gostos de seus consumidores, também surge a interferência contrária através dessas

mudanças visando absorver aspectos da vida social para ser quantificado. Van Dijck (2017) diz que amizades, interesses, conversações casuais, buscas por informação, expressão de gostos, respostas emocionais, são um dos meios que são buscados para serem datificados pelas empresas. Dentre as mais famosas a utilizarem o processo de datificação, está o *Twitter* e o *Instagram*.

A ideologia por detrás da datificação, proposta por Mayer-Schoenberger e Cukier (2013), é o que Van Dijck (2017) explica como *dataísmo*, o qual acaba sendo a crença generalizada na quantificação objetiva e no potencial monitoramento de todos os tipos de comportamento humano e de seus círculos sociais através da sua atividade online. O autor ainda define que dentro do *dataísmo* é essencial que ocorra um processo de confiança em quem irá coletar, interpretar e armazenar os dados produzidos nas redes sociais ou em websites. Ou seja, existe uma tênue confiança nas redes em que está meus perfis e onde está as informações fornecidas como condição de uso.

A rede social do passarinho, *Twitter*, é conhecida por ser uma rede do tempo real, onde assuntos que são uma “febre do momento” podem ser comentados e engajados ao mesmo tempo. Contudo, essa contagem de pessoas é passível de ser dribladas, nessa rede social a criação de perfis falsos ou bots (perfis apenas para a criação da sensação de um falso engajamento), e era deixada a possibilidade infinita de publicações que os perfis dos usuários podiam realizar ou participar na discussão ali existente. Entretanto, no período do último ano essa rede social foi adquirida pelo bilionário da tecnologia Elon Musk, um conhecido defensor da ‘liberdade de expressão’, realizou modificações onde reduziu o número de publicações possíveis pelos perfis gratuitos.

Outra rede social que utiliza o modelo de personalização do seu conteúdo com base no que se obtém do seu usuário é a pertencente ao grupo Meta, *Instagram* do bilionário Mark Zuckerberg. Inicialmente uma rede de fotos, ocorre uma hibridação a partir do momento no qual incorpora os chamados “stories” – derivados da rede social *Snapchat*, no qual permita que se colocasse na rede foto ou vídeos por um curto período - e posteriormente a inclusão dos “reels” uma ferramenta dentro do aplicativo apenas para vídeos. Esse último, é uma clara tentativa de competição para com uma gigante asiática, o *TikTok*. O *Instagram* hoje é uma rede altamente personalizável conforme os gostos e futuros gostos de seus usuários. Uma parte se deve a sua transformação como uma rede de trabalho e comercio. Segundo diz Van Dijck (2017, p. 45),

seu uso como moeda para pagar pelos serviços online e por segurança tornou os metadados uma espécie de ativos invisível, processados, na maioria das vezes, fora de seu contexto original e sem que as pessoas tenham consciência. As companhias de mídia social monetizam

os metadados ao reprocessá-los e vendê-los para anunciantes ou companhias de dados.

A ocorrência de um achismo em que o algoritmo está nos servindo ao permitir a sua personalização, mas ao mesmo o deixando nos moldar é uma demonstração dos sintomas aos quais os usuários estão sujeitos. Se a cada like, eu modifico um conteúdo na minha linha do tempo (*timeline*), hoje os dados coletados pelas redes ou através dos *cookies* oferecem a oportunidade de saber quanto tempo o conteúdo teve a atenção do usuário a ponto daquilo se tornar algo comerciável e desejável.

É comum a venda de estilos de vidas nas redes sociais, inclusive foi isso que se tornou possível a cunhagem do termo *influencer*, que são pessoas comuns no qual constitui em seus perfis dentro da rede uma grande massa de seguidores partindo desde mil a milhões. O dentro núcleo haverá uma contínua ligação dentre o *influencer* e seus seguidores através de postagens. O que acontece é uma pequena problemática acerca do quão real e em qual quantidade aquilo que se é postado na rede interfere na vida social e pessoal de quem está consumindo aquilo. Perfis de influências dedicados a certos assuntos como temos para estilo vida saudável, gurus econômicos e até perfis dedicados a mães e bebês (muito conhecido por *momsfluencer*), interferem na vida dos seguidores ao ponto de modificarem aspectos de seu cotidiano para a tentarem alcançar aquilo que se está sendo vendido.

3 Conclusão

Com o avanço das tecnologias de coleta e análise de dados, os usuários se veem cada vez mais obrigados a ceder informações pessoais para permanecerem conectados e inseridos em seus círculos sociais virtuais. Tal prática leva a uma crescente datificação da sociedade, onde as preferências e comportamentos dos usuários são interpretados e moldados a partir da leitura proveniente das suas ações. Embora possa ser útil para fins de personalização de serviços e publicidade direcionada, também traz à tona preocupações em relação à privacidade e à autonomia individual. É importante que os usuários estejam cientes das implicações dessa prática.

No início exemplificamos que a coleta e análise de dados é corriqueira como ferramentas é encontrada nas redes sociais, em particular o *Instagram*, onde ocorre uma influência sentida em âmbito real, através do uso e interpretação dos dados produzidos pelos usuários, onde se transformam em fornecedores e consumidores; pois como elencado, uma coleta não necessariamente é explicada ao usuário em linguagem acessível, ela pode estar disponível para a sua compreensão e consentimento, no entanto, em uma sociedade com uma produção massificada de informação e um culto

aos dados por parte daqueles que contém o poder de coletá-los e armazená-los, dificilmente usuários irão se interessar para ler o termo de uso ou a política de *cookies*. Pois como explicado por Van Dijck (2017), a ocorrência dentro da ideologia do dataísmo é que ocorra uma confiança já existente no processo de troca. Contudo, ao se propor uma troca de serviços pela permissão do uso dos dados é um formato bem barato e lucrativo de se produzir anúncios nos quais o objetivo é alcançar não só uma grande parcela de pessoas, mas também aqueles nos quais inconscientemente produziram material para a sua criação. O escambo digital, continua semelhante ao que se dá origem no termo, não permanecendo uma troca muito justa de mercadorias quando seus consumidores são sua fonte de produção.

Referências

- AGÊNCIA NACIONAL DE PROTEÇÃO DE DADOS (ANPD). **Guia orientativo - Cookies e proteção de dados pessoais**. 2022. Disponível em: <https://www.gov.br/anpd/pt-br/assuntos/noticias-periodo-eleitoral/anpd-lanca-guia-orientativo-201cookies-e-protecao-de-dados-pessoais201d>. Acesso em: 03 abr. 2023.
- ANTONIUTTI, Cleide Luciane. **Usos do big data em campanhas eleitorais**. 2015. 270 f. Tese (Doutorado) - Curso de Ciência da Informação, Universidade Estadual do Rio de Janeiro, Rio de Janeiro, 2015. Disponível em: https://bdtd.ibict.br/vufind/Record/IBICT_30edf379ab0daa0d84d6d58da44a03e0. Acesso em: 02 abr. 2023.
- CASTELLUCCIA, Claude. Behavioural Tracking on the Internet: a technical perspective. **European Data Protection: In Good Health?**, [S.L.], p. 21-33, 2012. Springer Netherlands. http://dx.doi.org/10.1007/978-94-007-2903-2_2.
- CASTELLS, Manuel. **A sociedade em rede, a era da informação: economia, sociedade e cultura**. 6 ed. São Paulo: Paz e Terra, 2002. Vol. 1.
- CASTELLS, Manuel. **A galáxia da internet: reflexões sobre a internet, os negócios e a sociedade**. Rio de Janeiro: Zahar, 2003.
- CAVALCANTI, Mario Filipe. *Cookies para quem? Entre o escambo digital e os direitos à privacidade e proteção de dados*. **Revista Acadêmica da Faculdade de Direito do Recife**, [S.l.], v. 93, n. 2, p. 96 - 115, out. 2021. ISSN 2448-2307. Disponível em: <https://periodicos.ufpe.br/revistas/ACADEMICA/articloe/view/249887>. Acesso em: 18 abr. 2023. doi:<https://doi.org/10.51359/2448-2307.2021.249887>.
- KASPERSKY. **O que são cookies?** Disponível em: <https://www.kaspersky.com.br/resource-center/definicoes/cookies>. Acesso em: 01 abr. 2023.
- MAYER-Schopenhauer, Viktor; CUKIER, Kenneth. **Big data: a revolution that will transform how we live, work, and think**. Nova York: Houghton Mifflin Harcourt, 2013.
- OLIVEIRA, J. V. de; SILVA, L. A. da. *Cookies de computador e história da internet: desafios à lei brasileira de proteção de dados pessoais*. **Revista de Estudos Jurídicos da UNESP**, [S. l.], v. 22, n. 36, 2021. DOI: 10.22171/rej.v22i36.2767. Disponível em: <https://ojs.franca.unesp.br/index.php/estudosjuridicosunesp/article/view/2767>. Acesso em: 10 abr. 2023.
- SANTAELLA, Lucia; KAUFMAN, Dora. Os dados estão nos engolindo? **Civitas - Revista de Ciências Sociais**, [S.L.], v. 21, n. 2, p. 214-223, 24 ago. 2021. EDIPUCRS. <http://dx.doi.org/10.15448/1984-7289.2021.2.39640>. Disponível em: <https://www.scielo.br/j/civitas/a/bKChbysdyGw5bX9Q9GgCWNH/#>. Acesso em: 04 abr. 2023.
- VAN DIJCK, J. Confiamos nos dados? As implicações da datificação para o monitoramento social. **MATRIZES**, [S. l.], v. 11, n. 1, p. 39-59, 2017. DOI: 10.11606/issn.1982-8160.v11i1p39-59. Disponível em: <https://www.revistas.usp.br/matrizes/article/view/131620>. Acesso em: 19 abr. 2023.



**RESUMO DOS TRABALHOS
APRESENTADOS NO FÓRUM DE DISCENTES**

*Papers' abstracts presented
at the Student's Forum*

Bibliografias em Museus: um estudo em museus brasileiros do início da república a meados do século XX (1893-1949)

Mestrando: Carlos Eduardo da Silva Carvalho (1)

Orientadora: Ana Cristina Albuquerque (2)

(1) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação (PPGCI UEL), eduardo.carvalho98@uel.br; (2) albuana@uel.br

Introdução

Este trabalho é uma parte teórica de pesquisa de Mestrado em desenvolvimento no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (PPGCI UEL), enquadrado na linha de pesquisa Organização e Representação da Informação e do Conhecimento e associa-se aos estudos históricos da informação e aos estudos de bibliografias.

Pode-se dizer que a Ciência da Informação ocupa-se, dentre outras questões, da recuperação e comunicação da informação, bem como de suas propriedades, atores, meio de produção e circulação, nos espaços de origem, coleção, organização, armazenamento, interpretação, uso e compartilhamento (BORKO, 1968; FREIRE, 2006). Nesse sentido, e compreendendo como de interesse para a Ciência da Informação, a cultura material, em sua expressão física (BUCKLAND, 1991), contorna-se como objeto de estudo as unidades de informação museais, mais especificamente no que se refere à organização da informação e do conhecimento em seus interiores, destacando-se o leque de possibilidades no estudo de seu desenvolvimento através do tempo. Portanto, pretende-se estudar as práticas museológicas referentes aos processos de organização do conhecimento nos finais do século XIX e início do século XX no Brasil.

Desenvolvimento

O caminho percorrido por um objeto, desde a existência com sua função “original” e prática no mundo até sua transformação em objeto de museu (a ser estudado, admirado, tratado como algo valioso ou “exemplar”, digno de contemplação e preservação), é atravessado por uma série de etapas de seleção, reflexão, separação e julgamento realizados por profissionais da instituição (GIOVANAZ, 1999; LARA FILHO, 2009; ALBUQUERQUE, 2012; 2015). Um objeto não nasce num museu, ele é selecionado para lá estar. Este processo de seleção e tratamento está

associado e é marcado, dentre outras etapas, pela atividade de classificação realizada nas unidades de informação - no caso, no museu. Tratando-se da atividade onde em determinado domínio do conhecimento se realizam agrupamentos com base em semelhanças e diferenças, com fim a organizar de forma determinada por uma lógica específica seres, saberes ou documentos, de maneira inteligível (ALBUQUERQUE, 2012), é importante ressaltar a natureza social e humana desta atividade. Da mesma forma como as ações sociais dos seres humanos de forma ampla constituem-se em atos de classificação, organizando de forma pragmática relações sociais ou objetos com vistas ao seu entendimento, observação ou recuperação, devemos entender as atividades de classificação em espaços ligados ao conhecimento como atravessadas por dinâmicas sociais e culturais (COSTA, 1997; ALBUQUERQUE, 2012).

Desta forma, não pode-se posicionar diante destes sistemas como estruturas puramente técnicas e desvinculadas de aspectos marcantes da vida social, seja por vias objetivas ou subjetivas, já que sistemas de classificação são, antes de tudo, humanos e históricos, não universais, espécies de códigos ordenadores da cultura - basta atentar para o estranhamento experimentado quando nos relacionamos com sistemas de organização e classificação baseados em códigos de cultura com as quais não estamos habituados (POMBO, 1998; BURKE, 2003).

No sentido proposto por Duarte e Chagas (2020), entendemos como possível e frutífera a aproximação entre os estudos sobre museus e estudos da bibliografia, a se observar as práticas relacionadas aos inventários, catálogos e registros de objetos. Entendida tanto como uma disciplina como um instrumento de organização do conhecimento, a bibliografia está associada a práticas de prospecção, seleção, avaliação e organização do conhecimento há, no mínimo, séculos, com sua prática podendo ser compreendida como a matriz disciplinar que permitiu o desenvolvimento de ciências que pensam a organização do conhecimento humano, como a própria Ciência da Informação

(CAPACCIONI, 2006; ARAUJO; CRIPPA; SALDANHA, 2015; LARA, 2018).

Para desenvolver os estudos no sentido proposto, foram selecionados cinco museus públicos a serem estudados, um representando cada região do país, estabelecendo como recorte temporal da pesquisa a ser desenvolvida o período decorrido entre a Proclamação da República (1889) e abarcando a primeira metade do século XX: Museu Paulista (SP), Museu de Numismática Bernardo Ramos (AM), Museu Julio de Castilhos (RS), Museu de Arte da Bahia (BA) e Museu das Bandeiras (GO). Tal demarcação se deu em razão da ampliação do número de museus no Brasil a partir do final do século XIX.

A metodologia de análise se baseia na Análise Dialógica do Discurso sob a perspectiva bakhtiniana, entendendo a vinculação dos enunciados linguísticos em, por exemplo, documentos institucionais, com o seu entorno sócio-cultural, assim possibilitando atribuição de sentido e significado a estes itens (FIORIN, 2011; ROSA, 2021; ROHLING, 2014). Pretendemos utilizar a aplicação do percurso metodológico desenvolvido e descrito por Rohling (2014), objetivado em nossa pesquisa pelas seguintes etapas: (1) estudar a relação dialógica nos museus selecionados; (2) descrever os papéis e relações presentes na prática bibliográfica entre os sujeitos nestes museus; (3) estudar as práticas bibliográficas e museológicas na contemporaneidade de nossos objetos; (4) estudar os horizontes temáticos-valorativos de seus enunciados; (5) identificar e analisar os tipos de discursos na sua materialidade do contexto sócio-cultural observado.

Resultados Parciais

Compreendendo sistemas de organização da informação e do conhecimento - neste caso, museus - como estruturas atravessadas pela ação humana e por aspectos ordenados pela cultura e a vida em sociedade, torna-se possível encará-los como passíveis de uma série de problemas - não apenas em sua eficiência técnica, mas também em suas noções categoriais, fundamentação teórica e filosófica e, em última instância, na concepção de mundo que os norteiam e ordenam. Nesse sentido, apoiamos-nos em Turner (2015; 2016; 2017) e suas reflexões sobre o caráter mantenedor de relações de expressão colonial a nível do conhecimento nos instrumentos e práticas vinculadas à classificação e catalogação realizadas pelas instituições museais. A autora (TURNER, 2016) contribui para que se observe as práticas documentais de um museu não apenas como parte de sua história, mas como construtoras ativas da história dos museus.

Conclusão

Espera-se no decorrer do desenvolvimento da pesquisa contribuir com os estudos que relacionam a organização da informação e do conhecimento e os objetos de museu, bem como com os estudos vinculados à bibliografia e suas práticas, atentando para problemas contemporâneos da Ciência da Informação, procurando compreender o desenvolvimento histórico de determinados sistemas de organização da informação e do conhecimento.

Referências

- ALBUQUERQUE, A. C. de. **A classificação de documentos fotográficos**: um estudo em arquivos, bibliotecas e museus. 2012. 287 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2012.
- ALBUQUERQUE, A. C. Tratamento temático da informação e a documentação museológica: aspectos e reflexões referentes à classificação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 16., João Pessoa, 2015. **Anais [...]**. João Pessoa: UFPB/ANCIB.
- ARAUJO, A. V. F.; CRIPPA, G.; SALDANHA, G. S. Em busca da bibliografia: sobre o i seminário internacional "a arte da bibliografia". **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 11, n. especial, p. 495-512, 2015.
- BORKO, H. **Information Science**: What is it? *American Documentation*, [S. l.], v. 19, n. 1, p. 3-5, 1968.
- BUCKLAND, M. K. Information as thing. **Journal of the American Society for Information Science**. v. 42, n. 5, p. 351-60, Jun. 1991.
- BURKE, P. **Uma história social do conhecimento de Gutenberg a Diderot**. Rio de Janeiro: Jorge Zahar, 2003
- CAPACCIONI, A. Mapas y memorias: apostillas a una historia de la Bibliografía. **Documentación de las Ciencias de la Información**, Madrid, v. 29, p. 9-24, maio 2006.
- COSTA, A. F. da. Classificações sociais. **Leitura**. Lisboa, v. 3, n. 2, p. 65-75, out. 1997.
- DUARTE, B. T.; CHAGAS, M. de S. Documentação, museu e memória: A Coleção Getúlio Vargas do Museu Histórico Nacional. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 25, p. 01-18. 2020.
- FIORIN, J. L. **Introdução ao pensamento de Bakhtin**. São Paulo: Ática, 2011.
- FREIRE, G. H. Ciência da Informação: temática, histórias e fundamentos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 1, p. 6-19, jan./abr. 2006.
- GIOVANAZ, M. M. Práticas de Coleção: seleção e classificação dos restos do passado. **Anos 90**: Revista do Programa de Pós-graduação em História. Porto Alegre, v. 11, p. 162-171, 1999

LARA, M. L. G. de. Conceito de bibliografia, ou conceitos de bibliografia?. **Informação & Informação**, Londrina, v. 23, n. 2, p. 127-151, set. 2018.

LARA FILHO, D. Museu, objeto e informação. **TransInformação**, Campinas, v. 21, n. 2, p. 163-169, maio/ago., 2009.

POMBO, O. Da classificação dos seres à classificação dos saberes. **Leituras**: Revista da Biblioteca Nacional de Lisboa, Lisboa, n. 2, 1998.

ROSA, A. P. C. **Arquivologia, Biblioteconomia e Museologia**: a noção de sociocultural a partir das concepções dialógicas. 2021. 125 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual de Londrina, Londrina, 2021.

ROHLING, N. A pesquisa qualitativa e análise dialógica do discurso: caminhos possíveis. **Cadernos de Linguagem & Sociedade**, Brasília, v. 15, n. 2, p. 44-60, 2014.

TURNER, H. Researching the care of collections information. **Practicing Anthropology**, [S. l.], v. 37, n. 3, 2015, p. 46.

TURNER, H. Critical histories of museum catalogues. **Museum Anthropology**, v. 39, n. 2, p. 102–10, 2016.

TURNER, H. “Organizing Knowledge in Museums: A Review of Concepts and Concerns.” **Knowledge Organization**, [S. l.], v. 44, n. 7, p. 472-484, 2017.

Enfoques sobre Sistemas de Organização do Conhecimento (SOC) nos currículos dos programas de pós-graduação em Ciência da Informação do Brasil

Mestranda: Clarice Luzia Rodrigues Casoni (1)
Orientadora: Raimunda Fernanda dos Santos (2)

(1) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação, clarice.luzia.casoni@uel.br, (2) fernanda@facc.ufrj.br

RESUMO

O projeto de pesquisa intitulado “Enfoques sobre Sistemas de Organização do Conhecimento nos currículos dos Programas de Pós-Graduação em Ciência da Informação do Brasil” estuda questões conceituais, teóricas, metodológicas e instrumentais a respeito dos Sistemas de Organização do Conhecimento. Entende-se que há a necessidade latente em investigar se os currículos dos Programas de Pós-Graduação em Ciência da Informação contemplam questões epistemológicas e/ou pragmáticas relacionadas aos conceitos, teorias, métodos e instrumentos de representação do conhecimento em ambientes tradicionais e digitais. Seguindo essa linha de raciocínio, tem-se como objetivo geral: Analisar os componentes curriculares sobre Sistemas de Organização do Conhecimento nos programas de Pós-Graduação em Ciência da Informação no Brasil. Objetiva-se especificamente: mapear as disciplinas e ementas voltadas para os Sistemas de Organização do Conhecimento; analisar os componentes curriculares sobre Sistemas de Organização do Conhecimento nos programas de Pós-Graduação em Ciência da Informação no Brasil; demonstrar quais os tipos de Sistemas de Organização do Conhecimento são contemplados nos currículos; investigar se as novas possibilidades de representação do conhecimento por meio da Folksonomia encontram-se contempladas nas estruturas curriculares com vistas a formação de pesquisadores em Ciência da Informação no Brasil.

Palavras-chave: Sistemas de Organização do Conhecimento; Programas de Pós-Graduação; Ciência da Informação Brasil.

Introdução

A organização e a representação do conhecimento (ORC) “tem como objeto de estudo o conhecimento e suas atividades” (CARLAN, 2010, p.28), contemplando questões epistemológicas e pragmáticas voltadas para a representação do conhecimento.

A representação do conhecimento, por sua vez, contempla teorias, métodos e técnicas aplicadas ao tratamento dos documentos. Tais documentos passam a ser descritos por um “esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, conseqüentemente, facilitar a recuperação das informações [...]” (CARLAN, 2010, p.28). Essa representação é feita por meio dos diferentes tipos de Sistemas de Organização do Conhecimento (tesauros, ontologias, taxonomias, mapas conceituais, Folksonomia etc.) que apresentam características e tipologias que variam de acordo com o grau de expressividade semântica.

Considerando que a pesquisa tem como foco os Sistemas de Organização do Conhecimento nos Currículos dos cursos de Pós-Graduação em Ciência da Informação do Brasil, faz-se necessário caracterizar os

procedimentos metodológicos empreendidos na pesquisa.

Procedimentos Metodológicos

Este estudo utiliza as pesquisas bibliográfica, documental, exploratória e descritiva com abordagem quali-quantitativa utilizando a técnica de análise de conteúdo. O percurso metodológico empreendido na proposta de estudo baseia-se em autores como Minayo (1994), Gil (2008), Bardin (2016), Richardson (2012), Creswell (2007) e Leitão (2005). A partir dessa metodologia serão analisados os currículos dos cursos de Pós-Graduação (*Stricto e lato sensu*) da Ciência da Informação no Brasil.

Para este estudo, a pesquisa bibliográfica será realizada em bases de dados nacionais e internacionais como: Base de Dados em Ciência da Informação (BRAPCI); Biblioteca Digital Brasileira de Teses e Dissertações (BDTD); Portal de Periódicos CAPES; Web of Science e ScieLo. Serão empregados os seguintes termos nos campos de busca: Organização do Conhecimento; Representação do Conhecimento; Sistemas de Organização do Conhecimento; Organização da Informação; Representação da Informação; Ciência da Informação; Pós-Graduação. Desse modo, a pesquisa

bibliográfica contribuirá para o desenvolvimento do referencial teórico, bem como para as análises qualitativas da presente pesquisa.

Em seguida, será realizada a pesquisa documental, a partir dos sites institucionais dos programas de pós-graduação em Ciência da Informação do Brasil (PPGCI-BRASIL) vinculados às Universidades Federais e Estaduais, para a realização do mapeamento e coleta de dados. Com a finalidade de identificar os programas de pós-graduações, inicialmente foi realizada uma busca na Sucupira (PLATAFORMA, [2023]), no intuito de validar todos os registros ativos de pós-graduações em CI, a serem mapeados individualmente nesta proposta de estudo.

Posteriormente a coleta dos dados se dará a análise das Propostas Curriculares, desta futura pesquisa. Nesse sentido, serão analisadas as propostas curriculares dos Programas de Pós-Graduação em Ciência da Informação do Brasil, com o intuito de identificar disciplinas e ementas voltadas para os Sistemas de Organização do Conhecimento, a partir da análise desses documentos nos sites das respectivas instituições que oferecem cursos de Mestrado e Doutorado (acadêmico e/ou profissional) na Ciência da Informação brasileira.

A pesquisa exploratória, por sua vez, servirá como base para a exploração e análise da produção científica nacional e internacional para fins de construção do referencial teórico do presente estudo, além de fornecer subsídios para a exploração e identificação dos Programas de Pós-Graduação e cursos que fornecem disciplinas relacionadas à temática Sistemas de Organização do Conhecimento. De maneira complementar, a pesquisa descritiva contribuirá para a descrição dos nomes e dos conteúdos presentes nas ementas das disciplinas, bem como dos enfoques temáticos contemplados por estas.

Além disso, a pesquisa utilizará as abordagens quantitativa e qualitativa, em que no primeiro momento as análises serão pautadas na abordagem quantitativa. Nesse sentido, a análise quantitativa possibilitará elaborar estatísticas, que visam dar apoio as inferências percebidas durante a análise dos dados, ou seja, quantificar as categorias estudadas e/ou dados concretos, que poderão ser demonstrados por meio de gráficos e tabelas. De outro modo disposto, a pesquisa qualitativa contribuirá tanto no processo de compreensão dos assuntos enfatizados nas produções científicas recuperadas, como para as análises e discussões dos resultados com base nos dados a serem coletados para o estudo proposto.

Considerando que uma abordagem complementa a outra, esta proposta de estudo possui abordagem quali-quantitativa.

Para fins de identificação dos conteúdos contemplados nos programas e ementas das disciplinas será realizada a análise do conteúdo proposta por Bardin (2011, p. 15), a qual permite a “sequência de tarefas e atividades a serem seguidas para fazer a análise dos dados qualitativos”.

Tais pesquisas e abordagens darão suporte para atender o objetivo geral da pesquisa que consiste em analisar os componentes curriculares sobre Sistemas de Organização do Conhecimento nos currículos dos programas de Pós-Graduação em Ciência da Informação das Universidades Federais e Estaduais do Brasil.

Resultados Preliminares

Como resultados preliminares, em março de 2023 foi realizado um mapeamento dos Programas de Pós-Graduação da área de avaliação “Comunicação e Informação” e da área do Conhecimento “Ciência da Informação” a partir de uma busca na plataforma Sucupira (PLATAFORMA, [2023]), em que foram identificadas 21 (vinte e uma) Instituições de Ensino Superior Federais e Estaduais, com mestrado acadêmico e profissional e doutorado acadêmico. Não foram identificados registro de Doutorado profissional na plataforma em questão.

Dentre as Instituições de Ensino Superior Federais e Estaduais citadas anteriormente, verificou-se a existência de 27 (vinte e sete) cursos de Pós-Graduação, sendo eles: 06 (seis) mestrados acadêmicos; 08 (oito) mestrados profissionais e 13 (treze) doutorados acadêmicos. Os dados provenientes dos currículos formulados pelos Programas de Pós-Graduação em questão proporcionarão valiosos elementos a serem analisados na proposta de pesquisa.

Como etapa seguinte, serão acessados os sites institucionais dos Programas de Pós-graduação em Ciência da Informação do Brasil que estão listados no resultado da busca na plataforma Sucupira, no intuito de analisar a grade curricular de cada programa e recuperar as ementas dos componentes curriculares cujo foco esteja voltado para Sistema(s) de Organização do Conhecimento. Nos casos em que não seja possível acessar a ementa das disciplinas diretamente nos sites institucionais dos programas, estas serão solicitadas via e-mail aos coordenadores dos PPGCI, individualmente, a fim de identificar elementos relacionados à temática estudada.

Com relação às informações a serem analisadas nos currículos dos cursos, serão considerados as seguintes categorias criadas previamente na proposta de estudo:

- a) Instituição e nome do Programa;
- b) Nome do curso de Pós-Graduação;

- d) Modalidade do curso de Pós-Graduação (acadêmico ou profissional);
- e) Área de concentração do Programa e linhas de pesquisa;
- f) Nomenclatura da disciplina;
- g) Caracterização da disciplina (obrigatória ou eletiva);
- g) Carga horária;
- i) Tipos de Sistemas de Organização do Conhecimento citados nas ementas das disciplinas;
- j) Principais teorias contempladas nesses componentes curriculares.

Considerações Finais

Com isso, espera-se obter um panorama do ensino sobre Sistemas de Organização do Conhecimento na Ciência da Informação Brasileira, considerando a importância de se promover no âmbito da Pós-Graduação *stricto e lato sensu*, um conjunto de reflexões e ações que priorizem essas metalinguagens não só como instrumentos que são construídos a partir de um conjunto de métodos e teorias da Organização do Conhecimento com vistas a recuperação e acesso à informação em diferentes contextos, mas como aparatos discursivos que possuem natureza enviesada e que podem viabilizar a proeminência de discursos hegemônicos e regulatórios. A formação educacional e científica nesse campo se configura como um potente instrumento de mudança e transformação social, considerando as novas possibilidades de representação

do conhecimento, bem como as novas tecnologias que dão suporte à sua estruturação.

Referências

- BARDIN, L. **Análise de Conteúdo**. São Paulo: Edições 70, 2016.
- CARLAN, E. **Sistemas de Organização do Conhecimento: uma reflexão no contexto da Ciência da Informação**. 2010. Dissertação (Mestrado em Ciência da Informação). Universidade de Brasília, 2010.
- COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Plataforma Sucupira**. 2023. Disponível em: <https://sucupira.capes.gov.br/sucupira/>. Acesso em: 27 mar. 2023.
- COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR (CAPES). **Plataforma Sucupira: cursos avaliados e re-conhecidos**. [2023] Disponível em: <https://sucupira.capes.gov.br/sucupira/public/consultas/coleta/programa/quantitativos/quantitativoAreaConhecimento.jsf?reaAvaliacao=31>. Acesso em: 28 abr. 2023.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. - São Paulo: Atlas, 2008.
- LEITÃO, B. J. M. **Avaliação qualitativa e quantitativa numa Biblioteca Universitária**. Niterói-RJ. Intertexto. 2005.
- CRESWELL, J. W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. Tradutor: Luciana de Oliveira da Rocha. Porto Alegre: Artmed, 2007.
- MINAYO, M.C. de. (Org.) **Pesquisa Social: teoria, método e criatividade**. Petrópolis: Vozes, 1994.
- RICHARDSON, R. J. **Pesquisa social: métodos e técnicas**. São Paulo: Atlas, 2012.

Organização e Representação do Conhecimento aplicado na área de “controle externo” sob a luz de Foskett: mapeamento dos processos para aperfeiçoamento do tesouro de contas nacional

Mestranda: Ilvana Sousa do Amaral (1)

Orientador: Marcos Luiz Cavalcanti de Miranda (2)

(1) Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Programa de Pós-Graduação em Biblioteconomia (PPGB), ilvana.amaral@edu.unirio.br, (2) Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Professor Associado – PPGB/UNIRIO; Doutor em Ciência da Informação – PPGCI IBICT UFRJ, Líder do Grupo de Pesquisa Organização do Conhecimento para Recuperação da Informação – GP/OCRI, mlmiranda@unirio.br

Introdução

A construção do conhecimento é um processo que ocorre por diferentes meios, dentre eles, a interação entre os sujeitos e as informações, com acervos de documentos para acesso e compreensão dos conteúdos informacionais.

Na perspectiva da Organização do Conhecimento, Hjørland (2008, p. 86) indica a relação “[...] com atividades como: descrição de documentos, indexação e classificação em bibliotecas, bases de dados bibliográficas, arquivos e outros tipos de ‘memória institucional’”. Para o autor, a Organização do Conhecimento, dialoga com a Biblioteconomia e a Ciência da Informação, pois dizem respeito à organização dos conceitos e a construção de sistemas de organização do conhecimento.

Segundo Souza (2006, p. 29), “a área da Ciência da Informação possui uma vasta literatura sobre questões relacionadas à classificação e indexação da produção científica publicada em áreas de C&T”. Nesse sentido, podemos destacar, particularmente, a produção literária na Inglaterra, entre as décadas de 1950 a 1970, com destaque especial para a obra de Foskett “*Classification and indexing in the social sciences*” (1963 e 1974, 1ª e 2ª ed., respectivamente), sobre classificação e indexação, dentro da área do conhecimento em Ciências Sociais, relacionando-se com o presente projeto, a estrutura de esquema de classificação de documentos (FOSKETT, 1974, cap. 7 e 8).

Para fundamentar este estudo, quanto a organização do conhecimento em biblioteca temática especializada, foi recuperado a dissertação de Assunção (1972), orientada por Foskett, que desenvolveu uma estrutura básica para o desenvolvimento de um sistema de classificação bibliográfica da biblioteca do Instituto de Ciências Biológicas/UFMG, incluindo todos os

assuntos abrangidos pelos Departamentos de Pesquisa, visando atender aos problemas de indexação de documentos especializados, em quaisquer suportes.

Assunção (1972) ainda argumentou que não pretendia fazer um levantamento exaustivo de toda terminologia especializada e, tão pouco, considerou estar com o sistema completo para a classificação de publicações em ciências biológicas, mas apresentar uma estrutura e o método ideal para o sistema de classificação. Complementa, dizendo que, os termos em cada classe demonstram como os mesmos devem ser enumerados e desenvolvidos, segundo os princípios dos sistemas analítico-sintéticos ou facetados. Finalizando, a autora desenvolveu um vocabulário, visando completar a sistemática animal e vegetal, eliminando detalhes desnecessários, e preparação de índice alfabético em forma de tesouro, tornando o sistema um instrumento completo de indexação e recuperação da informação na área de Biologia.

Correlacionando com a dissertação apresentada por Assunção (1972), o projeto objetiva aperfeiçoar o Tesouro de Contas Nacional (TCN), hospedado no site do Instituto Rui Barbosa (IRB), que é uma associação civil criada pelos Tribunais de Contas do Brasil, conhecido como “braço acadêmico” do Sistema de Controle Externo, auxiliando no desenvolvimento/aperfeiçoamento das atividades dos Tribunais de Contas brasileiros. O Vocabulário Controlado, criado em 2015, por meio de Comissão de Estudo formada por membros do TC-DF e TCU, em colaboração com bibliotecários/as das unidades informacionais dos TCs, buscou atender uma demanda de controle e uniformização das linguagens. Os regionalismos, por exemplo, são fatores que envolve os processos de organização de informações jurisprudenciais, e a implementação de um sistema unificado permitirá minimizar a subjetividade. Contudo, desde sua criação, não houve uma revisão

contínua devido a diversos fatores, principalmente pela falta de pessoal, e como consequência, alguns TCs adotaram seu próprio Vocabulário Controlado.

Desenvolvimento

Compreende revisar e atualizar o TCN, que atualmente utiliza o *software* TemaTres, sob a perspectiva da sua utilização, pelos profissionais bibliotecários das unidades informacionais, bem como pelos usuários/servidores dos Tribunais de Contas no Brasil, que são associados ao IRB. Quanto ao *software* TemaTres, Gonzales-Aguilar *et al.* (2012, p. 320) aponta ser uma ferramenta de “aplicação web para gestão de representações linguísticas formais do conhecimento: vocabulários controlados, taxonomias, tesouros, listas de cabeçalhos de assunto etc”.

Como procedimento metodológico será empregada a técnica de observação. Como membro-titular do Subcomitê do Tesouro, do Comitê Técnico de Gestão da Informação do IRB (CTGI/IRB), observou-se como fraqueza a ausência de plano de ação e procedimentos visando a continuidade para manutenção e atualização do sistema. Pretende-se, para organização deste trabalho, realizar um mapeamento dos processos e definir diretrizes com descrição dos procedimentos a serem seguidos pelos bibliotecários, tais como: estudar a Teoria do Conceito, Teoria da Classificação Facetada, Teoria da Terminologia (questões semânticas, sintáticas e conceituais), contemplando, sobretudo, os termos regionais dos TCs, para validação e atualização de termos linguísticos da área de “Controle Externo”.

A abrangência temática do TCN, atualmente, está organizada em 17 subcategorias, o qual podemos chamar de “Árvore do Conhecimento”, conceito denominado por Burke (2003, p. 83), dentro do Controle Externo. O processo de representação é dependente da etapa de análise de assunto por meio da identificação e seleção de conceitos, com vistas à “tradução” desses conceitos identificados e selecionados por meio de termos constituintes de uma linguagem documentária.

Na descrição das etapas serão consideradas: (i) incorporação constante de novos termos, visando à atualização da linguagem que se fizer necessária, por meio de coleta em fontes de informação formais (dicionários especializados, glossários técnico-científicos, diretórios, entre outros) e informais (formulários de sugestões de assuntos preenchidos pelos usuários, catálogo e listas de assuntos locais elaborados pela biblioteca); (ii) eleição de termos expressivos, visando à clareza na designação do assunto; (iii) revisão da tradução de termos existentes na linguagem, tendo em vista a devida correspondência conceitual que deve ocorrer em relação à terminologia

das áreas científicas nacionais; (iv) eliminação das ambiguidades causadas pela homonímia e polissemia advindas da linguagem natural; (v) incorporação de termos específicos, entre outras

Resultados Parciais

Pretende-se redigir um manual de procedimentos, e propor elaboração de uma norma técnica por meio do IRB, que deverá ser seguido pelos futuros bibliotecários-indexadores que possam vir integrar o Subcomitê do Tesouro, visando manter o vocabulário atualizado, de alta especificidade e representativo das necessidades de indexação e recuperação da informação, com vistas ao usuário final.

Conclusão

A elaboração do “Manual de Procedimentos” versará um recurso norteador e facilitador aos bibliotecários-indexadores visando dar continuidade ao trabalho de manutenção, aperfeiçoamento, adequação e atualização do TemaTres, tornando esse instrumento benéfico e confiável, não somente para os bibliotecários dos TCs, mas sobretudo sua usabilidade pelos servidores/usuários nos Tribunais de Contas no país.

Referências

- ASSUNÇÃO, Jandira Batista de. **Projeto de um sistema de classificação bibliográfica analítico-sintético (ou facetado) para a indexação e recuperação de informações em Biologia**. Orientador: Douglas John Foskett. 1972. 145 f. Dissertação (Mestrado em Biblioteconomia e Documentação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, 1972.
- BURKE, Peter. **Uma história social do conhecimento**: de Gutenberg a Diderot. Rio de Janeiro, RJ: Jorge Zahar, 2003.
- CAMPOS, Maria Luiza de Almeida. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói, RJ: EdUFF, 2001.
- FOSKETT, D. J. **Classification and indexing in the social sciences**. 2nd. ed. London: Butterworths, 1974.
- GONZALES-AGUILAR, Audilio; RAMÍREZ POSADA, María; FERREYRA, Diego *Tematres: software* para gestionar tesouros. **El profesional de la información**, v. 21, n. 3, p. 319-325, mayo/jun. 2012. Disponível em: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2012.may.14>. Acesso em: 19 abr. 2023.
- HJORLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, Bonn, Germany, v. 35, n. 2/3, p. 86-101, 2008.
- MELO, M. A. F.; BRÄSCHER, M. Termo, conceito e relações conceituais: um estudo das propostas de Dahlberg e Hjørland. **Ciência da Informação**, Brasília/DF, v. 43, n. 1, p. 67-80, jan./abr., 2014. Disponível em:

<https://revista.ibict.br/ciinf/article/view/1419>. Acesso em: 18 abr. 2023.

MIRANDA, Marcos Luiz Cavalcanti de.; CAMPOS, Maria Luiza de Almeida. A organização do conhecimento no domínio das ciências sociais e humanidades na perspectiva de Douglas John Foskett e Derek Wilton Langridge: uma abordagem aos Estudos Culturais e inovações disruptivas.

P2P & INOVAÇÃO, Rio de Janeiro, v. 8, n. 2, p. 160-177, 2022.

SOUZA, Rosali Fernandez de. Organização e representação de áreas do conhecimento em ciência e tecnologia: princípios de agregação em grandes áreas segundo diferentes contextos de produção e uso de informação. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 11, n. 1, p. 27-41, 1º sem. 2006.

Proposta de formalização de política de indexação de biblioteca universitária

Mestranda: Lilian Morais Brum (1)

Orientadora: Benildes Coura Moreira dos Santos Maculan

(1) Universidade Federal de Minas Gerais, PPGGOC - ECI, lilianmoraibrum@gmail.com

Introdução

A literatura brasileira sobre política de indexação vem crescendo desde o estudo pioneiro de Carneiro (1985), como resultado das investigações de Isidoro Gil Leiva, Mariângela S. L. Fujita e Milena Polsinelli Rubi. Nesses estudos nota-se a ligação que as políticas de indexação estabelecem com a gestão de unidade de informação, fato registrado em Fujita (2012). Ramos (1996, p. 4) ressalta que a gestão da informação em unidades de informação “compreende a definição da estrutura interna de operação, [...] dos processos de produção e dos produtos a serem gerados, a montagem de grupos de trabalho, divisão do poder com a atribuição de responsabilidades e tarefas”, assim como deve se preocupar com “a obtenção e a distribuição de recursos físicos e humanos compatíveis com o trabalho a ser realizado, propiciando a montagem de uma infraestrutura de trabalho”. Nonato (2022, p. 107) determina que a gestão da informação compreende três elementos básicos: o ciclo de vida da informação, que compreende todo o percurso que a informação faz, a gestão organizacional, que determina os produtos e serviços adequados ao público-alvo, e a tecnologia como um meio para manter a eficiência e eficácia na recuperação e desenvolvimento dos produtos e serviços de informação.

Desenvolvimento

Considera-se que todos os elementos para a gestão da informação devem estar presentes na política de indexação, pois, conforme Carneiro (1985), ela é um guia de tomada de decisões, definindo quais instrumentos e quais ações serão necessários para alcançar o propósito do serviço de informação. A essência dela é determinar “variáveis, estabelecer princípios e critérios que servirão de guia na tomada de decisões para otimização do serviço, racionalização dos processos e consistência das operações nele envolvidas” (CARNEIRO, 1985, p. 222).

Oliveira (2017, p. 43) destaca que ela deve ser formalizada, visto que uma “política formalmente estabelecida garante uniformidade e padronização na

análise e atribuição de descritores de assunto, o que irá favorecer diretamente o tratamento temático e a recuperação da informação”, devendo ser constantemente revista e atualizada.

Nunes (2004, p. 57 *apud* OLIVEIRA, 2017, p. 49) reforça esta posição dizendo que “em primeiro lugar a política de indexação deve ser enunciada formalmente num documento oficial da biblioteca ou serviço de informação”. Depois é que se deve analisar “o tratamento que será dado aos diferentes domínios disciplinares cobertos pelo acervo da biblioteca.” (NUNES, 2004, p. 58 *apud* OLIVEIRA, 2017, p. 49).

O problema da ausência de políticas formalmente registradas prejudica a gestão da informação, e seu desenvolvimento se justifica porque se pressupõe que traga mais vantagens que desvantagens. Como medida de racionalização das atividades biblioteconômicas, defende-se que a formalização resultará em economia de recursos humanos, materiais e financeiros, e contribuir para a satisfação do usuário, assegurando-se padronização, que dá maior consistência aos produtos e serviços de informação.

Objetivos

Tendo em vista o problema evidenciado até este ponto, este projeto de pesquisa tem como objetivo formalizar o conhecimento empírico das práticas de gestão da equipe da Biblioteca Universitária Professora Etelvina Lima, a partir de discussões e reflexões sobre as políticas de indexação e classificação em bibliotecas universitárias.

Metodologia

Este estudo se caracteriza como de natureza aplicada, pois visa resultados com consequências práticas, tratando de problemas específicos; quanto aos objetivos é descritiva e utiliza uma abordagem qualitativa de análise, interpretando questões relacionadas ao fenômeno estudado (GIL, 2008).

Como procedimento técnico, o estudo utiliza a observação participante, que se caracteriza como a

inserção do pesquisador no ambiente do fenômeno analisado, interagindo e compartilhando com os atores envolvidos o cotidiano daquela situação. As três etapas da observação participante seguem os passos descritos por Queiroz *et al.* (2007, p. 279), a saber:

1. primeira etapa: aproximação com o fenômeno em estudo, quando o pesquisador deve ser aceito para realizar o estudo juntamente com o grupo de atores;
2. Segunda etapa: visão da comunidade objeto de estudo, com a análise de documentos oficiais, da instituição, levantamento de autores, etc.;
3. terceira etapa: sistematização e organização dos dados, registrando a situação real e a percepção pessoal do pesquisador.

Na implantação da política de indexação será aplicada a proposta de Rubi (2012), em três fases:

preparação, desenvolvimento e avaliação, conforme descrito a seguir.

1. Preparação: fazer um diagnóstico inicial sobre a situação real;
2. Desenvolvimento: determinar diretrizes e decisões alinhadas aos objetivos e à filosofia da biblioteca, da organização e do público-alvo;
3. Avaliação: avaliar a política de indexação implantada tendo em vista ajustes e melhorias.

Resultados Parciais

Este projeto de estudo foi iniciado em fevereiro de 2023, e já cumpriu a primeira etapa da técnica da observação participante, de aproximação do pesquisador com o fenômeno. A proposta deste projeto foi apresentada e aceita pela equipe da biblioteca universitária da Escola de Ciência da Informação da Universidade Federal de Minas Gerais (UFMG), que é uma das 25 unidades setoriais do sistema de bibliotecas da UFMG. A equipe parece motivada com o empreendimento, visto que estão passando por processo de adaptação no processamento técnico, pois a catalogadora responsável pela atividade se aposentou recentemente. Durante o curto período em que as bibliotecárias da equipe estão assumindo as atividades do processamento técnico, estão sendo encontrados alguns obstáculos e algumas inconsistências de indexação e classificação que precisam passar por ajustes e formalização. A rotina das atividades de gestão da biblioteca como um todo já existe, alguns estão registrados de modo não formalizado, porém, já é possível perceber que há questões que ainda não foram refletidas pelo grupo. Sobre a segunda etapa da técnica da observação participante, alguns elementos também já foram coletados no site oficial da biblioteca.

Considerações Parciais

A formalização da política de indexação em unidades de informação tem como função primordial servir de apoio à gestão da informação, orientando a tomada de decisão durante as atividades da equipe. No caso específico deste estudo, já houve uma aproximação do pesquisador com o fenômeno a ser estudado, e a equipe da biblioteca já aceitou o desenvolvimento da proposta, considerando que trará ganhos consideráveis, fato já registrado na literatura (OLIVEIRA, 2017).

Referências

- CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 221-241, 1985.
- GIL LEIVA, Isidoro; RUBI, Milena P.; FUJITA, Mariângela S. L. Consistência na indexação em bibliotecas universitárias brasileiras. **TransInformação**, Campinas, v. 20, n. 3, p. 233-253, set./dez. 2008.
- FUJITA, Mariângela S. L.; LACRUZ, Maria Del C. A.; DÍAZ, Raquel G. A situação atual da indexação nas tarefas bibliotecárias. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 17, n. 1, p. 94-109, jan./mar. 2012.
- NONATO, Rafael dos Santos. **Gestão da informação e sistemas de gestão de segurança da informação: modelo para a garantia de disponibilidade em processos de contratação**. 2022. 304f. Tese (Doutorado na área de concentração da Ciência da Informação) – Programa de Pós-Graduação em Gestão e Organização do Conhecimento, Universidade Federal de Minas Gerais, Escola de Ciência da Informação. Belo Horizonte: PPGOC/UFMG, 2022.
- NUNES, C. O. I. Algumas considerações acerca da ausência de políticas de indexação em bibliotecas brasileiras. **BIBLOS - Revista do Instituto de Ciências Humanas e da Informação**, v. 16, p. 55-61, 2004. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/22625>. Acesso em: 2 maio 2023.
- OLIVEIRA, L. P. Política de indexação: concepções acerca do conceito e percepções em torno de sua elaboração. **Ciência da Informação em Revista**, v. 4, n. 2, p. 39-58, 2017. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/36466>. Acesso em: 2 maio 2023.
- QUEIROZ, Danielle Teixeira *et al.* Observação participante na pesquisa qualitativa: conceitos e aplicações na área da saúde. **Revista Enfermagem UERJ**, Rio de Janeiro, v. 15, n. 2, p. 276-283, abr./jun. 2007.
- RUBI, Milena Polsinelli. Política de indexação. In: LEIVA, Isidoro Gil; FUJITA, Mariângela Spotti Lopes (eds.). **Política de indexação**. Marília: Cultura Acadêmica, 2012. cap. 3, p. 107-120.
- RUBI, Milena Polsinelli. Proposta para implantação de política de indexação em bibliotecas. In: LEIVA, Isidoro Gil; FUJITA, Mariângela Spotti Lopes (eds.). **Política de indexação**. Marília: Cultura Acadêmica, 2012. cap. 7, p. 171-186.

Mapeamento de padrões de metadados para dados de pesquisa alinhados aos princípios FAIR

Mestrando: Anderson Araujo (1)

Orientadores: Viviane Santos de Oliveira Veiga, Carlos Henrique Marcondes

(1) FIOCRUZ, Programa de Pós-Graduação em Informação e Comunicação em Saúde, moranderson0182@gmail.com

Introdução

Diversas instituições, as agências de fomento e pesquisadores da área da ciência da informação compreendem que o compartilhamento de dados de pesquisa pode constituir uma fonte riquíssima de recursos para a pesquisa científica, para isso, é indispensável dispor de ambientes informacionais, de plataformas, de sistemas, que assegurem o armazenamento dos dados, a fim de disponibilizar informações seguras e confiáveis aos diversos tipos de usuários (especialistas e não especialistas).

Nesse contexto os padrões ou esquemas de metadados são indispensáveis. São eles que determinarão a descrição, a representação, o acesso, a confiabilidade e a persistência do recurso/objeto digital no ambiente informacional, além de definir a interoperabilidade entre sistemas.

Os princípios FAIR (Findable, Accessible, Interoperable e Reusable) foram desenvolvidos para orientar as boas práticas na pesquisa científica, de modo a facilitar a localização, o acesso, a interoperabilidade e o reuso de dados de pesquisa, para que conjuntos de dados sejam FAIR, dados e metadados precisam estar alinhados a estes princípios. Desta forma, metadados e padrões de metadados são importantes para garantir que as plataformas digitais disponibilizem dados encontráveis, acessíveis, interoperáveis e reutilizáveis (FORMENTON et al, 2017).

Desenvolvimento

Objetivo: O presente estudo teve como objetivo o mapeamento dos padrões de metadados para dados de pesquisa alinhados aos princípios FAIR.

A metodologia consistiu na pesquisa bibliográfica em fontes de informação nacionais e internacionais que forneceu as bases teóricas da pesquisa e foi realizado um mapeamento das diretrizes internacionais para padrão de metadados para dados de pesquisa nas

iniciativas internacionais Fairsharing (<https://fairsharing.org/>), Digital Curation Centre (DCC) (<https://www.dcc.ac.uk/>) e Research Data Alliance (RDA) (<https://www.rd-alliance.org/>), permitindo identificar os padrões de metadados utilizados e a mapear as principais diretrizes.

Resultados Parciais ou Finais

O Fairsharing é um recurso com curadoria, informativo e educacional sobre padrões de dados e metadados, inter-relacionados a bancos de dados e políticas de dados (<https://fairsharing.org/communities#governance>).

Para alcançar o objetivo de levantar diretrizes, políticas ou guidelines para padrões de metadados de pesquisa, criamos a seguinte estratégia de busca: “metadata standard” “research data” policy.

Ao aplicarmos a estratégia no campo de busca, para mapearmos as diretrizes em relação aos metadados de pesquisa, encontramos dois resultados: Genomic Science Program (GSP) Information and Data Sharing Policy, (<https://genomicscience.energy.gov/datasharing/>) e RDA Covid-19 WG Resources (<https://www.rd-alliance.org/groups/rda-covid19>).

O Digital Curation Center (DCC) no que se refere a metadados disciplinares, reuniu um conjunto de disciplinas com links para informação sobre padrões de metadados disciplinares, incluindo perfis, ferramentas para implementar os padrões e casos de uso de repositórios de dados.

Nos diferentes padrões disciplinares encontramos um conjunto de lista de Padrões de metadados gerais, conforme descritos a seguir:

CERIF - É o padrão que a União Européia recomenda aos seus estados membros para registrar informações sobre atividades de pesquisa.

Data Package - A especificação Data Package é um formato wrapper genérico (é um programa que extrai o conteúdo de uma fonte de informação específica e o

converte em um formato relacional) para troca de dados, consistindo em uma pasta contendo arquivos de dados e um arquivo descritor.

DataCite Metadata Schema - É uma lista de propriedades principais de metadados escolhidas para uma identificação consistente de um recurso para fins de citação e recuperação, juntamente com instruções de uso recomendado (definições e usos dos metadados).

DCAT - é um vocabulário RDF projetado para facilitar a interoperabilidade entre catálogos de dados publicados na Web. Serve para descrever catálogos e conjuntos de dados – datasets.

Dublin Core é um padrão de metadados, composto por 15 elementos, planejado para facilitar a descrição de recursos eletrônicos, foi feito inicialmente para descrever documentos digitais. É um dos padrões de metadados mais conhecidos e mais amplamente usados.

Open Archives Initiative - desenvolve e promove padrões que permitem a distribuição de conteúdo.

Observations and Measurements - O padrão O&M foi desenvolvido no contexto de sistemas de informações geográficas.

PREMIS (Preservation Metadata: Implementation Strategies) define um conjunto de metadados que a maioria dos repositórios de objetos digitais precisaria registrar e usar para preservar esses objetos a longo prazo.

PROV define um modelo, serializações correspondentes e outras definições de suporte para permitir o intercâmbio interoperável de informações de proveniência em ambientes heterogêneos como a Web.

RDF Data Cube Vocabulary - O vocabulário Data Cube é uma base fundamental que suporta vocabulários de extensão para permitir a publicação de outros aspectos de fluxos de dados estatísticos ou outros conjuntos de dados multidimensionais.

Alguns repositórios decidiram que os padrões atuais não atendem às suas necessidades de metadados e, portanto, criaram seus próprios requisitos. São denominados de “Repository-Developed Metadata Schemas”.

A RDA é uma organização internacional, que engloba participantes de todas as áreas e de diferentes nacionalidades. A RDA apresenta uma lista significativa de recomendações fornecidas pelos grupos de trabalho ou grupos de interesse da RDA. As

recomendações são os resultados oficiais e endossados da RDA e considerados os seus principais resultados (<https://www.rd-alliance.org/recommendations-outputs/standards>).

Na RDA são quatro grupos que foram criados para apoiar os estudos sobre padrões de metadados para dados de pesquisa. São eles: MIG (Metadata Interest Group), MSDWG (Metadata Standards Directory WG); DICIG (Data in Context IG) e RDPIG (Research Data Provenance IG). Eles se uniram com a 'tarefa' de produzir um diretório de metadados, o Metadata Standards Directory WG.

Fairsharing, DCC e RDA trabalham em conjunto para a criação de padrões de metadados alinhados ao FAIR.

Com a criação do Metadata Standards Catalog o pesquisador ou a instituição pode pesquisar qual o padrão da lista que aparece no catálogo se adequa melhor a sua área ou campo de conhecimento.

Conclusão

Nos padrões visto no levantamento encontramos com bastante frequência o padrão para catálogo de Dados, o DCAT, que é recomendado pelo W3C (World Wide Web Consortium), a principal organização de padronização da World Wide Web. Este foi o padrão de metadados escolhido pelo Projeto VODAN (<https://portal.fiocruz.br/en/vodan-brazil>) para o tratamento de dados em Covid-19, mas verificou-se que ele não é suficiente para descrever os metadados de proveniência para dados de pesquisa em Covid-19. Existem iniciativas que propõem extensões ao padrão DCAT como a iniciativa COVID-19 Data Portal da União Europeia, que usa uma extensão do DCAT chamada COVID-19 DCAT-AP (<https://data.europa.eu/data/datasets/dfa6ef4a-6d10-44ad-b51a-42bd8dda4476?locale=en>). Para preencher esta lacuna, é necessário conhecer a percepção dos pesquisadores dos metadados de proveniência necessários para o reuso dos dados de pesquisa.

Referências

FORMENTON, D.; GRACIOSO, L. S.; CASTRO, F. F. Revisitando a preservação digital na perspectiva da ciência da informação: aproximações conceituais. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, SP, v. 13, n. 1, p. 170-191, jan./abr. 2015. Disponível em: <http://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1587/1571>. Acesso em: 21 jul. 2022.

Análise do domínio *Computational Archival Science*: contribuições para a Arquivologia e a Organização do Conhecimento Arquivístico

Doutoranda: Raquel Torrão Valentim (1)

Orientadora: Natália Bolfarini Tognoli

(1) Universidade Federal Fluminense, Programa de Pós-Graduação em Ciência da Informação, raquel_valentim@id.uff.br

Introdução

A constante evolução e uso das tecnologias digitais trouxe transformações significativas para a Ciência da Informação, demandando estudos teóricos e uma base conceitual para auxiliar seus profissionais em suas atividades (SMIT; BARRETO, 2002). Entretanto, essas tecnologias também significam novas soluções para o campo e as atividades de tratamento, organização e representação do conhecimento.

Propostas inter e transdisciplinares que prometem esforços colaborativos entre os campos da Ciência da Informação e suas áreas afins, da Organização do Conhecimento e da Tecnologia da Informação, vêm ganhando espaço nos debates acadêmicos.

Em especial no que tange à Arquivologia e a Organização do Conhecimento Arquivístico (OCA), observa-se a emergência de um novo domínio chamado *Computational Archival Science* (CAS) que promete ser um campo destinado a encontrar soluções para as atividades arquivísticas com o uso de recursos computacionais. Sua caracterização enquanto novo domínio deu-se a partir de uma análise de domínio que apontou a necessidade de um aprofundamento teórico e epistemológico.

Desenvolvimento

A CAS propõe-se a construir um novo campo que reúne e aplica os conhecimentos teóricos e metodológicos provenientes da integração do pensamento computacional, isto é, um conjunto de práticas agrupadas em quatro categorias: práticas de dados, práticas de modelagem e simulação, práticas de resolução de problemas computacionais e práticas de pensamento sistêmico (WEINTROP et al., 2016; UNDERWOOD; MARCIANO, 2019), e do pensamento arquivístico, ou seja, o conjunto teórico e metodológico da Arquivologia, onde se incluem suas atividades e aplicações práticas.

Na universidade de Maryland nos Estados Unidos, onde estão concentrados a maioria de seus pesquisadores, esse campo já é reconhecido e nele desenvolvem-se ações educativas que oferecem soluções práticas para aplicação de métodos e recursos computacionais para o tratamento e preservação de documentos de arquivo a fim de melhorar as atividades arquivísticas e permitir engajamento e acesso a esses materiais (ADVANCED INFORMATION COLLABORATORY, 2022; PAYNE, 2018).

Como exemplo, podemos citar o Programa Legado da Escravidão (*Legacy of Slavery Program*), desenvolvido no Arquivo do Estado de Maryland (EUA), que provou ser possível o uso de recursos e ferramentas computacionais para reunir de forma compreensível as relações e contextos que formavam a realidade da escravidão em Maryland. No projeto, os pesquisadores lograram êxito em demonstrar as complexas relações comuns entre os indivíduos a série de registros onde seus nomes foram encontrados, através da aplicação dos conceitos de pensamento computacional por meio de processos de análise e engenharia de dados. É certo que sem a automação das atividades por meio de um sistema de softwares desenvolvido e aplicado, banco de dados, ferramentas de visualização para organizar o conhecimento de centenas de milhares de documentos a serem transcritos, os profissionais demandariam um tempo muito maior do que o disposto (COX et al., 2018).

Nesse sentido, afim de compreender melhor a CAS, desenvolveu-se uma pesquisa em nível de Mestrado, no Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense (PPGCI/UFF) com apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) no âmbito do Grupo de Pesquisa Arquivos, Bibliotecas e Organização do Conhecimento (GPABOC/UFF), entre os anos de 2021 e 2023, explorando-a enquanto um domínio emergente à luz da Organização do Conhecimento e da metodologia de análise de domínio, a partir do mapeamento de sua

comunidade discursiva e da identificação das temáticas mais abordadas pelos seus atores e de suas influências teóricas. A partir da análise de domínio desenvolvida sob as perspectivas histórica, epistemológica, bibliométrica e terminológica, constatou-se a carência de um material teórico consistente sobre a CAS que sistematize suas bases teóricas e principais ideias, sobretudo em âmbito nacional.

A escolha metodológica, justifica-se, pois, a análise de domínio possibilita múltiplas perspectivas (HJØRLAND, 2016) e a verificação do que é importante ou significativo sobre um campo, proporcionando a análise de elementos como tendências, padrões, agentes, processos e relacionamentos existentes (DANUELLO, 2007), além de possibilitar a identificação das condições que permitem a construção e socialização do conhecimento científico (GUIMARÃES, 2015).

Resultados Parciais ou Finais

Os resultados da pesquisa inicial demonstram que a CAS configura um domínio emergente (SMIRAGLIA, 2012; BARITÉ, 2020), com um corpo de conhecimento único pautado na integração do pensamento computacional e arquivístico com uma comunidade discursiva ativa com ideias convergentes e conceitos compartilhados, mas que ainda carece de aprofundamento teórico sobre sua natureza disciplinar, suas bases teóricas e a extensão de sua comunidade discursiva.

Portanto, durante a nova fase da pesquisa, busca-se aprofundar-se sobre as questões pendentes na primeira etapa. Em especial, questiona-se a natureza da CAS enquanto uma transdisciplina, um campo de pesquisa ou uma nova abordagem na Arquivologia, bem como sua real contribuição para a Arquivologia e a OCA. Tem-se, então, como objetivo geral compreendê-la a partir de sua construção teórico-conceitual e, objetivos específicos o desenvolvimento de uma revisão teórico-epistemológica a partir da literatura produzida pelos seus marcos teóricos e o aprofundamento sobre as bases teóricas que a subsidiam a partir da análise de sua comunidade discursiva.

Deste modo, será realizada uma nova análise de domínio sobre a bibliografia já existente da CAS e sua comunidade discursiva a partir de artigos, capítulos de livros e livros sobre o tema, encontrados em bases de dados internacionais, bem como no *Advanced Information Collaboratory website*. A análise da bibliografia coletada se dará por meio de uma perspectiva epistemológica e histórica, fundamentais a toda pesquisa (HJØRLAND; 2002), pela perspectiva bibliométrica, que permitirá a verificação da extensão da comunidade discursiva, e pela terminológica, que

contemplará a identificação dos termos e conceitos utilizados por tal comunidade.

Além disso, essa etapa da pesquisa prevê um contato direto com os pesquisadores e cofundadores da CAS a partir da realização de entrevistas estruturadas, contribuindo para a complementação dos estudos epistemológicos. Espera-se, então, em resultados futuros, identificar a natureza disciplinar da CAS, assim como aprofundar-se em sua epistemologia, suas bases teórico-conceituais e os projetos que vêm sendo desenvolvidos em seu escopo.

Conclusão

Conclui-se que, a CAS enquanto um domínio emergente possui uma comunidade discursiva ainda em desenvolvimento e crescimento, concentrada na Universidade de Maryland, cuja literatura ainda requer uma expansão para além do desenvolvimento teórico de seus cofundadores e suas bases teóricas enfatizam os conhecimentos da Computação. Apesar de apresentar inconsistências no que tange a sua natureza disciplinar, seus pesquisadores demonstram esforços em desenvolver materiais educacionais a fim de promover o ensino dos estudos teóricos e metodológicos que vêm sendo construídos, sobretudo a partir dos projetos no âmbito da CAS.

Além disso, enfatiza-se que a CAS não propõe romper com os conhecimentos já produzidos sobre as tecnologias digitais na Arquivologia e na OCA, mas reuni-los junto a novas iniciativas em um campo. Nesse sentido, percebe-se que a CAS configura uma iniciativa promissora no cenário arquivístico, devendo ainda ser mais bem avaliada por seus pares no que tange a sua aplicabilidade, mas que já possibilita novos horizontes para a realização das atividades arquivísticas.

Referências

- ADVANCED INFORMATION COLLABORATORY, 2022. Disponível em: <https://ai-collaboratory.net/cas/>. Acesso em: 4 jul. 2022.
- BARITÉ, M. Terminología de urgencia y garantías para la representación temática: elementos para el análisis de dominios de emergencia súbita (des). **Informação & Informação**, v. 25, n. 3, p. 239-265, 2020.
- COX, R. *et al.* A case study in creating transparency in using cultural Big Data: the legacy of slavery project. *In*: IEEE BIG DATA CONFERENCE, 10-13 dec. 2018, Seattle, WA. **Proceedings** [...]. [S.l.]: IEEE, 2018.
- DANUELLO, J. C. **Produção científica docente em tratamento temático da informação no Brasil: uma abordagem métrica como subsídio para análise do domínio**. 2007. 115 f. Dissertação (Mestrado em Ciência da

Informação) – Programa de Pós-Graduação em Ciência da Informação, UNESP, Marília, SP, 2007.

GUIMARÃES, J. A. C. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, Brasília, DF, v. 43, n. 1, 2015. Disponível em: <https://revista.ibict.br/ciinf/article/view/1415>. Acesso em: 5 jul. 2022.

HJØRLAND, B. Domain analysis in information science: eleven approaches: traditional as well as innovative. **Journal of Documentation**. v. 58, p. 422-462, 2002.

HJØRLAND, Birger. Knowledge Organization. In: **ISKO Encyclopedia of Knowledge Organization**, 2016. Disponível em: www.isko.org/cyclo/knowledge_organization. Acesso em: 13 set. 2022.

PAYNE, N. Stirring the Cauldron: Redefining Computational Archival Science (CAS) for the Big Data Domain. In: IEEE INTERNATIONAL CONFERENCE ON BIG DATA,, 2018, Seattle. Proceedings [...]. Seattle: IEEE, 2018.

SMIRAGLIA, R. P. Epistemology of domain analysis. In: SMIRAGLIA, R. P.; LEE, H.-L. (eds.). **Cultural frames of knowledge**. Würzburg: Ergon, 2012, p. 111-124.

SMIT, J. W.; BARRETO, A. A. Ciência da Informação: base conceitual para a formação do profissional. In: VALENTIM, M. (org.). **Formação do profissional da informação**. São Paulo: Polis, 2002. p. 9-23. Disponível em: <https://abecin.org.br/wpcontent/uploads/2021/03/Formacao-do-profissional.pdf>. Acesso em: 10 ago. 2022.

UNDERWOOD, W.; MARCIANO, R. Computational thinking in Archival Science research and education. In: IEEE BIG DATA CONFERENCE, 9-12 dec. 2019, Los Angeles, CA. **Proceedings** [...]. Los Angeles, CA: IEEE, 2019.

WEINTROP, D. *et al.* Defining computational thinking for mathematics and science classrooms. **Journal of Science Education and Technology**, v. 25, n. 1, p. 127-147, 2016.

Aparecimento e descontinuidade de termos: um resgate histórico em periódicos de Ciência da Informação

Doutoranda: Patrícia Ofélia Pereira de Almeida (1)

Orientadora: Brígida Maria Nogueira Cervantes

(1) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação, pereira@uel.br

RESUMO

A presente pesquisa tem como foco analisar aspectos da abordagem temática da informação, principalmente no que se refere aos termos utilizados para a descrição de assuntos de objetos informacionais. Apresenta como objetivo geral investigar o aparecimento e a descontinuidade de termos em Ciência da Informação, utilizando como ponto de partida as palavras-chave dos artigos de periódicos da Área. Trata-se de uma pesquisa básica, documental, com abordagem qualitativa e descritiva. Serão analisados os artigos publicados na Revista Ciência da Informação, pois trata-se do periódico brasileiro mais tradicional na Área, publicado desde junho de 1972, quase ininterruptamente.

Palavras-chave: representação da informação; conceitos; periódicos científicos; terminologia; descritores.

Introdução

A representação da informação pode ser concebida a partir de diversos suportes, como um objeto, uma imagem, uma música, ou qualquer outra coisa que remete a algum significado, a algum conceito, com a finalidade de extrair elementos que identifiquem o objeto em meio a uma massa documental, de maneira que possa ser agrupado junto a outros com os mesmos atributos físicos e temáticos (TAYLOR; JOUDREY, 2009). Na Ciência da Informação, representar é uma atividade que consiste em traduzir a informação contida em um documento para uma versão resumida, de forma que possa condensar em poucos termos aquilo que é essencial para descrever sua estrutura, forma e conteúdo (NOVELLINO, 1998).

Ao refletir a respeito da representação da informação em bases de dados, e em como as alterações ocorrem e os termos são remodelados ao longo do tempo, suscitou a seguinte inquietação: é possível identificar o aparecimento e a descontinuidade de termos da Ciência da Informação com base na indexação dos artigos de periódicos da Área? Por conseguinte, manifestou-se a necessidade de conhecer quais os fatores que influenciam nessa dinâmica e o contexto em que essas mudanças estão inseridas. Algumas hipóteses que podem ser associadas ao contexto são:

- a) incidência do assunto: em que uma temática pode ter maior recorrência de interesses em um período devido a uma demanda por acontecimentos específicos, como questões políticas, econômicas, pandemias ou desastres naturais;

- b) alterações gramaticais e ortográficas: seguindo as regras vigentes na língua portuguesa;
- c) adequação do termo ao conceito: em que o entendimento acerca de um objeto ou um tema, tem seu conceito alterado à medida em que se discute seus aspectos, como a etimologia da palavra, ou origem/função do objeto,
- d) questões socioculturais: em que termos que possam causar desconforto ou constrangimento em determinados contextos, como questões raciais, de sexualidade, peso corporal ou desenvolvimento cognitivo – são substituídos por outros que mais se adaptam ao assunto representado.

Apresenta como objetivo geral investigar o aparecimento e a descontinuidade de termos em Ciência da Informação, utilizando como ponto de partida as palavras-chave dos artigos de periódicos da Área. De forma mais específica, pretende:

- a) levantar os as palavras-chave utilizadas nos artigos publicados pela Revista Ciência da Informação;
- b) identificar termos que surgiram, atingiram um ápice e posteriormente caíram em desuso,
- c) verificar fatores que influenciaram para que os termos tenham aparecido e descontinuado na indexação dos artigos.

Trata-se de uma pesquisa básica, documental, com abordagem qualitativa e descritiva. Serão analisados os artigos publicados na Revista Ciência da Informação,

pois trata-se do periódico brasileiro mais tradicional na Área, publicado desde junho de 1972, quase ininterruptamente.

Desenvolvimento

Em bancos de dados, a informação é representada por palavras e números, que são caracteres legíveis por computador e compreensíveis por humanos. Dessa forma, é registrado um conjunto de dados que descrevem os atributos físicos e temáticos do documento, ou seja, os metadados (TAYLOR; JOUDREY, 2009; DIAS; NAVES, 2013), o que permite a interação homem-máquina, e viabiliza a representação do objeto informacional, bem como sua recuperação por meio dos mecanismos de busca.

Conforme Carneiro (1985), ao selecionar os termos que irão representar tematicamente um documento, deve-se evitar ser demasiadamente específico, com a finalidade de não restringir os descritores a um assunto exclusivo, pois pode prejudicar a revocação adequada nos resultados de busca; ou demasiadamente exaustivo, incluindo descritores que não tem relevância significativa dentro do contexto do documento, e que assim tornam os resultados de busca numerosos e contraproducentes.

Nos periódicos científicos, que “[...] constituem o meio mais importante para a comunicação científica.” (GIL, 2008, p. 62), os termos são os primeiros, e talvez os mais importantes, elementos de indexação dos assuntos tratados nos artigos. Ao longo da história das publicações científicas houve diversos estágios nas normas de indexação, que variam desde a ausência absoluta de orientações aos autores, passam pela indicação de fontes, e até mesmo a obrigatoriedade de utilização de tesouros ou outras linguagens documentárias específicas da área.

As mudanças terminológicas que ocorrem ao longo dos anos podem influenciar nos resultados de buscas nas bases de dados, bem como na Gestão da Terminologia (MAGARZO ARIAS; FRANKLIN, 2018), o que torna relevante investigar as transformações e variações na representação de assuntos, e quais seriam os fatores determinantes para que termos comecem a ser utilizados, ou mesmo caiam em desuso por não serem mais adequados.

Resultados Parciais

Atualmente a estrutura da Revista Ciência da Informação está dividida em apenas quatro seções: artigos, relatos de experiência, opinion paper e revisão de literatura. No período de 1972 a 2023 foram publicados 1.829 artigos, incluindo editoriais (71), expedientes (30), resenhas (7) e entrevistas (3). Em

edições comemorativas ou temáticas, 29 artigos foram publicados duas vezes, e 14 artigos foram publicados três vezes. O presente estudo está em fase inicial, e deverão ser selecionados os artigos com palavras-chave em português, e a grafia dos termos será padronizada (plural, singular, acentos etc.) para que se possa agrupá-los adequadamente.

Serão selecionados alguns termos que se destacarem pela recorrência ou pela extensão do período de utilização, e então serão contextualizados alguns aspectos do período em que surgiram, entraram em ápice, e desapareceram, com a expectativa de identificar quais as circunstâncias em que isso ocorreu – se por algum aspecto relacionado à estrutura das palavras, ou se pela obsolescência do assunto relacionado. Nesse sentido, o contexto em que os conceitos surgem pode influenciar em seu significado (MAGARZO ARIAS; FRANKLIN, 2018), então, se não houver constância, coerência ou fundamentação teórica que sustente a aplicação de um conceito em um âmbito geral, esse pode se enquadrar a apenas uma única área ou campo, e ainda assim não prolongar sua utilização.

Os Editoriais dos fascículos (quando houver) poderão servir como uma das formas de associar os termos utilizados ao contexto histórico. Também os dois números especiais publicados em comemoração aos 45 anos do periódico – em 2017, uma homenagem aos autores mais produtivos, e em 2018, um panorama dos assuntos mais abordados ao longo de sua trajetória – poderão auxiliar nas análises.

Embora a pesquisa esteja em fase inicial, ao averiguar os dados obtidos foi possível observar que a Revista Ciência da Informação apresenta um longo e sólido percurso, e com isso agrega uma quantidade significativa de publicação de artigos, de respeitáveis autores da Área, o que a torna uma preciosa fonte de informação para o estudo proposto, para a fundamentação de pesquisas, e para a memória da Ciência da Informação.

Conclusão

Ressalta-se que o desenvolvimento e utilização de linguagens documentárias é uma preocupação constante de estudiosos da Ciência da Informação, assim como é relevante pesquisar as metamorfoses terminológicas da Área. Nisso, inclui-se a identificação da compatibilidade de representação dos conceitos, no intuito de aprofundar o conhecimento a respeito das consequências que a representação temática detém na recuperação e no acesso de documentos em bases de dados.

Referências

- CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, 1985. Disponível em: <https://brapci.inf.br/index.php/res/v/73170>. Acesso em: 04 dez. 2022.
- DIAS, Eduardo Wense; NAVES, Madalena Martins Lopes. **Análise de assunto: teoria e prática**. 2. ed. Brasília, DF: Briquet de Lemos Livros, 2013.
- GIL, Antonio C. **Métodos e técnicas de pesquisa social**. 6.ed. São Paulo: Atlas, 2008.
- MAGARZO ARIAS, Miguel Ivan; FRANKLIN, Benjamin Luiz. O ciclo de vida terminológico em periódicos científicos eletrônicos na Ciência da Informação do Brasil. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 19, Londrina, 2018. **Anais...** Londrina: UEL, 2018. Disponível em: <https://brapci.inf.br/index.php/res/v/102617>. Acesso em: 14 nov. 2022.
- NOVELLINO, M. S. F. A linguagem como meio de representação ou de comunicação da informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 3, n. 2, p. 137-146, jul./dez. 1998. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/38505>. Acesso em: 22 dez. 2022.
- TAYLOR, Arlene G.; JOUDREY, Daniel N. Organization of recorded information *In*: TAYLOR, Arlene G.; JOUDREY, Daniel N. **The organization of information**. 3. Ed. Westport: Libraries Unlimited, 2009.

O contexto social de criação dos álbuns fotográficos: uma reflexão no âmbito da Organização do Conhecimento

Doutoranda: Cristina Ribeiro dos Santos

Orientadora: Ana Cristina de Albuquerque

Universidade Estadual de Londrina

Introdução

Considerando as ideias de Brascher Café (2008), compreende-se que o conhecimento e a informação são produtos culturais, conteúdos da experiência humana, que estão presentes nos diferentes tempos históricos, ou seja, são uma construção humana que se utiliza de diferentes técnicas, tecnologias e linguagens.

A Informação, nesse sentido, é uma forma material da existência do conhecimento. (BRASCHER; CAFÉ, 2008). Hjørland (2008) destaca que a Organização do Conhecimento (OC) tem entre seus processos a descrição, indexação e classificação, que são desenvolvidas em vários ambientes informacionais. A OC, busca a qualidade dos processos de Organização do Conhecimento e dos Sistemas de Organização do Conhecimento. (HJØRLAND, 2008, p.23). Assim, estão entre as suas preocupações a análise, coleta, organização, classificação, manipulação, acúmulo, armazenamento, representação, recuperação, propriedades, acesso e a disseminação da informação registrada. Neste contexto, há o pressuposto, de construir uma base teórica e metodológica que se instala nas abordagens, técnicas, nos instrumentos, bem como nos procedimentos utilizados por aqueles que trabalham com a informação em seu dia a dia, entre estes; mas não limitados a estes; os bibliotecários, arquivistas, documentalistas. Seguindo com esta reflexão, que envolve o álbum fotográfico e a Organização do Conhecimento e a qualidade desse processo observa-se a necessidade de uma análise a respeito do contexto de criação deste recurso informacional e a sua proximidade elementar com o domínio que lhe deu origem.

A partir desses preceitos, busca-se como objetivo principal, refletir a respeito dos pontos de convergência entre a perspectiva de Domínio de Hjørland (2002), o princípio da proveniência e o contexto social de criação dos álbuns fotográficos. Desta forma pergunta-se quais contribuições para a OC que podem ser encontradas na intersecção entre a perspectiva de Domínio de Hjørland (2002), o princípio da proveniência e o contexto social

de criação dos álbuns fotográficos? Esta reflexão tem como justificativa, colaborar com discussões que destacam os elementos socioculturais de álbuns fotográficos. Neste sentido, essa investigação utiliza como metodologia a pesquisa bibliográfica, tendo ainda como pressuposto ser exploratória.

Desenvolvimento

O álbum fotográfico, visto pelo contexto informacional, mais especificamente pelo olhar da Organização do Conhecimento, se destaca pelo seu universo de questões, que percorrem desde a sua gênese (ALBUQUERQUE, SANTOS, FREITAS, 2016), passando pela perspectiva de ser entendido como um recurso informacional (SANTOS, ALBUQUERQUE, 2019), ou ainda em sua especificidade como os álbuns fotográficos familiares e a sua transição do âmbito privado para o público (SANTOS, ALBUQUERQUE, 2014), ou mesmo os elementos constitutivos dos álbuns de fotografias escolares e as potencialidades da relação com as humanidades digitais (ALBUQUERQUE, SANTOS, 2021).

A busca pela compreensão deste recurso informacional, possui ainda alguns desafios, que são proporcionais a importância deste material e o seu potencial informacional, bem como a sua crescente presença nos centros de informação. Observa-se que há discussões direcionadas ao domínio (HJØRLAND, 2002), que podem ser frutíferas para a busca do princípio da proveniência e do contexto social de criação dos álbuns fotográficos. Um olhar transversal que busca mapear, agentes, produtores, processos e seus relacionamentos, ou seja identificar os elementos fundamentais que servirão de base para as estruturas e processos da OC.

Resultados Parciais

Evidencia-se como parte dos resultados obtidos desta pesquisa as perspectivas de Domínio de Hjørland

(2002), como um aporte para mapear o princípio da proveniência e o contexto social de criação dos álbuns fotográficos. Contribuições e enriquecimento para a compreensão e análise de forma integrada deste recurso informacional.

Há nesta intersecção, a emersão de informações enriquecedoras a respeito do álbum fotográfico, tais como a origem e a classe social, do produtor, as relações inter-pessoais, acadêmicas e sociocultural do seu produtor, ou mesmo a construção de narrativas visuais (narrativa esta que pode ter por base a vivência prática, a experiência real e a narrativa criada a partir das memórias repassadas de terceiros e não vivenciadas.).(SANTOS, ALBUQUERQUE, 2017). Hjørland (2002), Barité (2001), são alguns exemplos de autores que são referência para a OC e que percorrem em suas discussões, a importância deste caráter social para a OC. Hjørland, tem o devido destaque nessas reflexões, de início pela Análise de Domínio, que neste texto é percebida pela perspectiva de ferramenta.

Destaca-se que, ao se mapear a comunidade discursiva, que está vinculada com a geração, materialização, uso, transmissão e disseminação de determinado recurso informacional, possibilita-se a compreensão, uma visão contextual ampla de criação/ gênese documental dos álbuns fotográficos. Nesse sentido, as características sociais do álbum fotográfico, sua incorporação e interpretação vistas pelo prisma do seu produtor ganha destaque.

Conclusão

A realidade social, a diagramação das páginas, as escolhas e materialização do elemento essencial dos álbuns, a fotografia possui uma simbiose com o contexto social do seu produtor, tal afirmação, advém do fato, que por mais tecnológico que se materialize o álbum, este é permeado de escolhas e decisões que refletem o seu contexto de produção, percepções de mundo, tecnologia disponível uma ligação inequívoca com o seu tempo e espaço, ou seja, o álbum fotográfico, não decorre de uma ação mecânica, pura e sem interferência do seu contexto social.

Assim, há um recurso informacional impregnado de elementos do seu contexto sociocultural de criação, o crivo das questões como o colecionismo de imagens fotográficas, o contexto sociocultural do produtor, a criação de uma narrativa visual, classificações entre outras questões. Uma simbiose de elementos que ao serem identificados oferecem subsídios informacionais para os processos da OC. E está presente nesse escopo, informações que estão registradas e indicadas de forma clara, explícita e conteúdo que estão presentes de forma subjetiva que apenas são evidenciados quando

entra a perspectiva de uma OC com viés social, que busca respeitar tais questões.

A identificação da presença destas questões é relevante para a compreensão, de forma ampla do álbum fotográfico ou mesmo para a implementação dos processos relacionados a Organização do Conhecimento, pois percebe-se que a OC, não se preocupa apenas com a recuperação da informação pelo usuário, mas sim, com o conjunto, que envolve a suas práticas e a qualidade desses processos.

Referências

- BRÄSCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais [...]**. São Paulo: ANCIB, 2008. Disponível em: <https://brapci.inf.br/index.php/res/v/176535>. Acesso em: 27 fev. 2023.
- HJORLAND, B. What is knowledge organization (KO)? **Knowledge Organization**, [S. l.], 2003. Disponível em: https://www.researchgate.net/publication/277803483_What_is_Knowledge_Organization_KO. Acesso em: 5 fev. 2023.
- HJORLAND, B. Domain analysis in information Science: eleven approaches traditional as well as innov. **Journal of Documentation**, [S. l.], n. 58, v. 4, p. 422-462, 2002. Disponível em: https://www.researchgate.net/publication/249366184_Do_main_analysis_in_information_science_Eleven_approaches_-_Traditional_as_well_as_innovative. Acesso em: 27.fev.2023.
- SANTOS, C. R. D.; ALBUQUERQUE, A. C. Álbuns fotográficos familiares: reflexões sobre a transição do âmbito privado para o público. **Informação Arquivística**, Rio de Janeiro, v. 3, n. 1, 2014. DOI: 10.18377/2316-7300/informacaoarquivistica.v3n1p% Acesso em: 21 fev. 2023.
- SANTOS, C. R.; FREITAS, L. M.; ALBUQUERQUE, A. C. de. Um Estudo da Gênese Documental de Álbuns Fotográficos. In: SEMINÁRIO DE PESQUISA EM CIÊNCIAS HUMANAS, 11., São Paulo, 2016. **Anais [...]**. **Blucher Social Sciences Proceedings**, v. 2, n. 4, p. 53-60, 2016.
- SANTOS, C. R. dos. **Gênese documental de álbuns fotográficos escolares: entre avanços e ausências**. 2017. 123f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual de Londrina, Londrina, 2017. Disponível em: <http://www.bibliotecadigital.uel.br/document/?code=vtls000214285>
- ALBUQUERQUE, A. C.; SANTOS, C. R. dos. Os elementos constitutivos de álbuns fotográficos escolares e as potencialidades a partir da relação com as Humanidades Digitais: primeiras interlocuções. **ARCIC [online]**, v. 10, n. 25, p. 3-28. 01-Abr-2021. ISSN 2411-9970.

Contribuições da Linguística Sistêmico-Funcional (SFL) para o aprimoramento da construção de Sistemas de Organização do Conhecimento (SOCs)

Doutorando: José Mario de Oliveira Mendes (1)

Orientadora: Vania Mara Alves Lima

(1) Escola de Comunicações e Artes da Universidade de São Paulo (ECA/USP), Programa de Pós-Graduação em Ciência da Informação, jose.mario.oliveira.mendes@gmail.com

Introdução

Esta pesquisa tem como objetivo geral identificar na Linguística Sistêmico-Funcional (LSF) os conceitos que podem contribuir para a elaboração de um Sistema de Organização do Conhecimento (SOC), no nosso caso os Tesouros, a partir das relações já estabelecidas entre a Linguística e a Organização do Conhecimento (OC) no âmbito da Ciência da Informação (CI). Como objetivos específicos temos: a) definir e contextualizar os SOCs no âmbito da OC; b) descrever os conceitos da Linguística Documentária (LD), da Teoria do Conceito e da Terminologia para o desenvolvimento de um SOC do tipo Tesouro; c) descrever os conceitos da LSF e; d) identificar as relações entre conceitos da LSF, da Teoria do Conceito e da Terminologia apontando as contribuições para o desenvolvimento, aprimoramento, construção e uso de SOCs.

Como um fenômeno social, é através e pela linguagem que interagimos com o mundo e pessoas, aprendemos sobre cultura, valores, identidades e experiências de mundo que passam, com a prensa tipográfica (Gutenberg, 1450) e a revolução industrial (século XVIII), a ser registradas, compartilhadas e disseminadas em grande escala. Com o conhecimento produzido passando a ser compilado e organizado, desde a Enciclopédia de Diderot (1750), passando pelos sistemas de classificação bibliográfica até os SOCs, a linguagem se faz instrumento essencial para que esse conhecimento produzido pela sociedade ou por comunidades de domínios específicos possa ser organizado, representado e recuperado a partir das demandas existentes.

A teoria da Linguística Sistêmico-Funcional (Halliday, 1994/2004/2014) aborda a conexão entre sociedade e conhecimento num contexto de cultura e situação, a partir de um sistema de convenções e transmissão de significado. No âmbito da CI, na linha de pesquisa Organização da Informação e Conhecimento, são os SOCs que têm por objetivo facilitar o gerenciamento e

a recuperação desse conhecimento e informação em determinado domínio (MAZZOCCHI, 2018). São considerados SOCs os glossários, dicionários, listas de cabeçalhos de assunto, esquemas de classificação, taxonomias, ontologias e tesouros (também denominados de Linguagens Documentárias (LD)), nosso foco de estudo.

Desenvolvimento

Iniciamos com um passeio histórico desde a Documentação até os SOCs, contando um pouco da história da Documentação, passando pela Grupo TEMMA e chegando até a construção do capítulo da ISKO-Brasil, num *state of the art* sobre a área da OC atualmente no país. O Grupo TEMMA, parte do Diretório de Grupos de Pesquisa do CNPq, e a produção de seus integrantes é marco para a reflexão e o ensino da Organização e Representação do Conhecimento no Brasil, principalmente no que diz respeito à integração de referenciais de linguagem de pesquisas, preocupando-se em estabelecer relações entre a Documentação e a Linguística, a Lógica, a Comunicação e a Terminologia, pensando na “organização de conceitos, na categorização e em parâmetros para a definição de hipóteses de organização”, considerados essenciais para o desenvolvimento de tesouros, ontologias, *topic maps*, bem como para a arquitetura da informação” (LARA, 2012, p.4). “No Brasil, onde os estudos de análise documental são fortemente influenciados pela abordagem francesa, identifica-se nas pesquisas desenvolvidas pelo Grupo TEMMA (ECA/USP) o espaço acadêmico responsável pela germinação e crescimento dos fundamentos teóricos e metodológicos da Análise documental neste país. Reconhecendo os estudos de Anna Maria Marques Cintra, Mariângela Spotti Lopes Fujita, João Batista Ernesto de Moraes, Marilda Lopes Ginez de Lara, Johanna Wilhelmina Smit, Nair Yumiko Kobashi, José Augusto Chaves

Guimarães, Vânia Mara Alves Lima, Maria de Fátima Gonçalves Moreira Tálamo entre outros, vem diversificando seu espectro investigativo entre estudos de terminologia aplicada à organização e transferência da informação, nos processos de leitura em análise documental e linguagem de especialidade da Ciência da Informação.

Em seguida, tratamos da Linguística Documentária, que parte do pressuposto que os problemas relacionados à representação e recuperação da informação são problemas de linguagem. Passamos pela Teoria do Conceito (Dahlberg, 1978), explicando que é por meio da linguagem que se formulam enunciados, quando todo enunciado verdadeiro representa um elemento de seu respectivo conceito. E trazemos, da Terminologia, a Teoria Geral da Terminologia (TGT, Wüster, 1930, 1960) e a Teoria Comunicativa da Terminologia (TCT, Cabré, 1993, 1999), abordagem de base linguística que zela pelo caráter comunicativo dos termos efetivamente utilizados em domínios específicos.

No capítulo seguinte, trazemos os conceitos dos SOCs, seus conceitos e funções, no âmbito da OC, quais os aspectos linguísticos, documentários e terminológicos envolvidos no seu desenvolvimento, fundamentados na LD, na Teoria do Conceito e na Terminologia. Destacando que em 2005, no Brasil, quando da criação do capítulo ISKO Brasil, durante o VI ENANCIB o termo Organização do Conhecimento passa a ser identificado, na literatura nacional, como subdomínio da Ciência da Informação, abordando os processos que incluem a organização, representação e recuperação do conhecimento e da informação.

A seguir, apresentamos a teoria Linguística Sistêmico-Funcional (Halliday (1985, 1994 e 2004) e Halliday e Matthiessen (2004, 2014)), trazendo os conceitos de Contexto de Cultura, de Situação e as Metafunções da Linguagem, focando no Sistema de Tema-Rema, inferindo as conexões desta teoria com a LD e seu uso para a construção de SOCs, levando em conta que “o termo texto se refere a qualquer instância da linguagem, produzida em qualquer meio, que faz sentido para alguém que conhece tal linguagem/língua” (HALLIDAY; MATTHIESSEN, 2004, p. 03) focando, na Metafunção Textual, o Sistema Tema e Rema, que se realiza pela estrutura temática que dá à sentença seu caráter de mensagem (Halliday, 1994, p. 37) e que usa a linguagem para organizar nossos significados experienciais e interpessoais de uma maneira coerente, coesa e linear. As diferentes escolhas do Tema, do ponto de partida da mensagem, é fator determinante na produção de sentido de fala, chamando a atenção para o que virá a seguir e para todo o significado textual da sentença. O Tema é o elemento que funciona como ponto de partida, aquele que determina sobre qual será

o assunto da oração (Halliday, 1994, p. 39; Eggins, 1994, p. 275; Thompson, 1996, p. 72), a primeira parte da informação, podendo trazer uma informação dada ou uma informação nova, o Rema que é a parte da oração na qual o tema será desenvolvido. Ou seja, tudo o que não é tema, é rema (Eggins, 1994, p. 275).

Por fim, discutimos nossa hipótese inicial já que acreditamos que a LSF fornece subsídios para o desenvolvimento dos SOCs, pois analisa a língua e a linguagem em uso e em um contexto de cultura e situação específicos, ajudando a criar indexadores e formas de organizar cada área do conhecimento por meio de investigações prévias de como a área X ou Y registra, comunica e armazena seus estudos e/ou discursos, gerando classificação, organização e recuperação mais precisa e pertinente por meios de SOC em cada área X ou Y.

Resultados Parciais

Ainda em fase de construção, já apresentamos algumas contribuições da LSF para a OC, como a) a LSF se localiza como uma área de relevante contribuição para estudos em Linguística Aplicada, o que pode nos levar a pensar que bases da LD também podem se encontrar com bases da LSF; b) a importância e o papel da mensagem durante a comunicação, destacando o ponto de partida da mensagem, a identificação do Tema e do Rema num texto; c) a terminologia de um domínio específico é parte da linguagem real compilada para dar conta de uma denominação especializada, como um recurso utilizado para se criar e representar significados, quando o conceito “pode ser concebido independentemente da denominação ou termo que o designa” (CABRÉ, 1995, p.06) já que para a Terminologia e a LSF os termos interessam por si sós, quando o que importa é o significado que aquele termo realiza dentro daquele contexto específico; d) a diferença fundamental pensando em palavra e em termo, e que, segundo Cabré (1995, p.06), “uma palavra é uma unidade descrita por um conjunto de características linguísticas sistemáticas e dotada da propriedade de se referir a um elemento da realidade [enquanto] um termo é uma unidade de características linguísticas similares, utilizada em um domínio de especialidade [portanto] uma palavra que faz parte de um campo especializado seria um termo”; e) podemos traçar paralelos entre a TGT, a LSF e a TCT; f) quando Halliday (1985, 1994, p.xvii) diz que a “perspectiva funcional dos estudos linguísticos é diferente da perspectiva chamada de tradicional pois enxerga a linguagem vinculada a um contexto de uso e de situação”, o que ocorre com uma LD no momento da indexação e da recuperação da informação.

Referências

- CABRÉ, Maria Teresa. **La terminología: representación y comunicación: elementos para una teoría de base comunicativa y outros artículos**. Barcelona: Institut Universitari de Lingüística Aplicada, 1999.
- CABRÉ, Marisa Teresa. La terminología hoy: concepciones, tendencias y aplicaciones. **Ciência da Informação**, Rio de Janeiro, v.24, n.3, 1995. Disponível em: <http://revista.ibict.br/ciinf/article/view/567>. Acesso em: 06 dez. 2019.
- DAHLBERG, Ingetraut. **Teoria do Conceito**. Ciência da Informação, Rio de Janeiro, v. 7, n. 2, 1978. p. 101-107. Disponível em: <http://revista.ibict.br/ciinf/article/view/115>. Acesso: 12 dez. 2019.
- EGGINS, S. **An Introduction to Systemic Functional Linguistics**. Grammar. Londres: Continuum Intl. Publishing Group, 2nd ed. 2004.
- GRUPO Temma. *In*: CNPq. Diretório dos Grupos de Pesquisa no Brasil. <http://dgp.cnpq.br/buscaoperacional/detalhegrupo.jsp?grupo=0067607UVP2I76>.
- HALLIDAY, Michael Alexander Kirkwood. **An Introduction to Functional Grammar**. 2. ed. Londres: Edward Arnold, 1994.
- HALLIDAY, Michael Alexander Kirkwood; HASAN, Ruqaiya. **Language, context and text: aspects of language in a social-semiotic perspective**. Oxford: Oxford Press University, 2. ed., 1989.
- HALLIDAY, Michael Alexander Kirkwood; MATTHIESSEN, Christian Matthias Ingemar Martin. **An introduction to functional grammar**. 3. ed. Londres: Edward Arnold, 2004.
- HALLIDAY, Michael Alexander Kirkwood; MATTHIESSEN, Christian Matthias Ingemar Martin. **An introduction to functional grammar**. 4. ed. Londres: Edward. Arnold, 2014.
- HALLIDAY, Michael Alexander Kirkwood; WEBSTER, Johnathan (ed.). **The language of science**. Londres: Continuum Intl. Publishing Group, v. 5, 2004. (Collected Works of M. A. K. Halliday).
- LARA, Marilda Lopes Ginez de. **O Grupo Temma na ECA-USP, 2001-2011**. Relatório de pesquisa em iniciação científica apresentado ao Departamento de Biblioteconomia e Documentação da Escola de Comunicações da Universidade de São Paulo. Projeto PIBIC/CNPq. Bolsista: Davilson Antonio, 2012.
- MAZZOCCHI, Fulvio. Knowledge organization system (KOS): an introductory critical account. **Knowledge Organization**, [s.l.], v. 45, n.1, p. 54-78, 2018. Disponível em: <http://www.isko.org/cyclo/kos>. Acesso em: 15 jul. 2021.
- THOMPSON, G. **Introducing functional grammar**. London: Arnold, 1996.
- WÜSTER, Eugen. **Introducción a la teoría general de la terminología y a la lexicografía terminológica**. Tradução Anne-Cécile Nokerman. Barcelona: IULA, 1998.

As concepções de Michel Foucault para a compreensão da informação como poder na Ciência da Informação

Doutoranda: Liliane Cristina Soares Sousa (1)

Orientadora: Ana Cristina Albuquerque (2)

(1) UEL (Universidade Estadual de Londrina), Programa de Pós-Graduação em Ciência da Informação, lilianeli.sousa@uel.br, (2) albuana@uel.br

Introdução

A Ciência da Informação é um campo científico que se dedica ao estudo dos processos de produção, organização, representação, disseminação e uso da informação. Nesse sentido, preocupa-se com a geração, processamento e disseminação de informações por meio de diversos canais e meios de comunicação. Capurro (2003) argumenta que para identificar o papel de uma Ciência da Informação autônoma, é necessário ir além das questões práticas e técnicas e pensar nos conceitos fundamentais e nas questões mais abrangentes que norteiam a disciplina. Ele defende que a Ciência da Informação lida com questões filosóficas, sociológicas, psicológicas e tecnológicas relacionadas ao uso e ao significado da informação na sociedade, e que uma reflexão epistemológica é fundamental para compreender a relação entre o conceito de informação na disciplina e seu uso em outras áreas do conhecimento e em diferentes contextos culturais e históricos. Segundo Gómez (2003) há a necessidade em estudar as conexões entre diferentes esferas de produção de conhecimento e sua relação com a sociedade em geral, como um caminho para uma pesquisa mais aprofundada em Ciência da Informação. Sendo assim, sugere que a compreensão dessas relações pode levar a uma melhor compreensão da produção e circulação de conhecimento na sociedade contemporânea. A partir do prisma de Saracevic (1996), a Ciência da Informação tem uma forte dimensão social e humana, que ultrapassa a tecnologia. O autor argumenta que essa disciplina não deve ser vista apenas como técnica, mas também como uma ciência social preocupada em compreender os processos de informação que ocorrem em contextos sociais e culturais. Além disso, ele destaca a importância da Ciência da Informação na sociedade por estudar todos os processos de informação que afetam os indivíduos e as comunidades. Hjørland (2002) entende que a gestão da informação é uma atividade que está inserida em uma sociedade complexa e que, portanto, ela é moldada pelas relações

sociais e políticas que permeiam essa sociedade. Visto que, a informação não é um produto neutro, mas é influenciada pelas relações de poder e controle que existem dentro da sociedade.

Desenvolvimento

Correntes de pensamento contribuem para refletirmos a respeito da epistemologia da Ciência da Informação, no sentido de propor novos olhares críticos para o campo da organização e representação do conhecimento. Buschman (2010) destaca a importância de Foucault na estruturação de um projeto para uma teoria crítica viável para a Ciência da Informação, enfatizando a importância de suas ideias para a compreensão das relações entre a informação, conhecimento, poder e controle. A concepção de Foucault é vista como uma fonte importante de discussão de teoria crítica na Ciência da Informação, que busca examinar criticamente as estruturas sociais, políticas e econômicas que moldam a gestão da informação e suas implicações para a sociedade. Por isso, a Ciência da Informação é um campo dinâmico e em constante transformação, no qual diferentes atores lutam para manter ou modificar suas “formações discursivas.” Segundo Foucault (2008, p. 133) a formação discursiva é “um conjunto de regras anônimas, históricas, sempre determinadas no tempo e no espaço, que definiram em uma época dada, e para uma área social, econômica, geográfica ou linguística dada, as condições de exercício da função enunciativa.” Essas formações discursivas podem ser vistas como conjuntos de ideias, valores e práticas que são compartilhados pelos atores envolvidos no campo da Ciência da Informação. Elas são marcadas pelo exercício da “ordem do discurso”, que de acordo com a observação de Foucault (1996, p. 9) “a produção do discurso é ao mesmo tempo controlada, selecionada, organizada e redistribuída por certo número de procedimentos que têm por função conjurar seus poderes e perigos, dominar seu acontecimento aleatório, esquivar sua pesada e temível materialidade.” O autor ressalta que a produção do

discurso é um processo controlado e organizado por meio de diversos procedimentos, que têm como objetivo evitar que a disseminação da informação se torne um perigo ou ameaça para aqueles que detêm o poder. Ele sugere que a produção e disseminação dos discursos não são um processo neutro e imparcial, mas sim um processo que se tenta controlar e organizar.

Resultados Parciais ou Finais

Diante do exposto, a presença de Foucault é imprescindível em uma agenda epistemológica da ciência, em especial na Ciência da Informação, pelo fato de que sua obra apresenta importantes contribuições para a compreensão das relações entre o poder, o conhecimento e a produção do discurso. Seus estudos demonstram de que maneira as relações de poder estão presentes em todas as esferas da vida social, inclusive na organização e representação do conhecimento. Ele argumenta que o poder não é algo que está nas mãos de indivíduos ou instituições, mas sim algo que permeia todas as relações sociais, e que se manifesta de maneira sutil e difusa através de mecanismos disciplinares e de controle. Sendo assim, a questão que norteia nossa pesquisa é: em que medida as teorias e conceitos de Michel Foucault podem contribuir para a construção de uma análise crítica na organização e representação do conhecimento em Ciência da Informação?

Conclusão

Nesse sentido, as contribuições da leitura de Foucault para a Ciência da Informação são importantes por fornecerem uma análise crítica das estruturas de poder e controle na gestão da informação, uma compreensão da relação entre o conhecimento e o poder e uma análise das práticas de documentação e arquivo. Essas

contribuições continuam a influenciar a teoria e a prática da Ciência da Informação, permitindo uma compreensão mais profunda de como a informação é usada para exercer poder sobre as pessoas e como a análise crítica pode ajudar a desafiar essas estruturas de poder.

Referências

- BUSCHMAN, John. Transgredir ou estagnar? desafiando Foucault na teoria da Biblioteconomia e Ciência da Informação (BCI). **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 1, n.1, p. 3 - 31, 2010.
- CAPURRO, Rafael. Epistemologia e ciência da informação. *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO -ENANCIB*, 5., 2003. Belo Horizonte. **Anais...**Belo Horizonte: UFMG, 2003. Disponível em:<http://www.capurro.de/enancib_p.htm>. Acesso em: 25 mar. 2023.
- FOUCAULT, Michel. **A ordem do discurso**. Aula inaugural no College de France. Pronunciada em 2 de dezembro de 1970. Tradução de Laura Fraga de Almeida Sampaio. São Paulo: Loyola: 1996.
- FOUCAULT, Michel. **A arqueologia do saber**. 6. ed. Rio de Janeiro: Forense - Universitária, 2008.
- GÓMEZ, M. N. G. As relações entre ciência, estado e sociedade: um domínio de visibilidade para as questões da informação. **Ciência da Informação**, v.32, n.1, 2003. DOI: <https://doi.org/10.18225/ci.inf.v32i1.1020>. Acesso em: 20 mar. 2023.
- HJORLAND, B. (2002). Epistemology and the socio-cognitive perspective in information science. **Journal of the American Society for Information Science and Technology**, 53(4), 257-270.
- SARACEVIC, T. Interdisciplinary nature of information science. **Ciência da Informação**, Brasília, v. 24, n. 1, p. 36-41, 1996.

A decolonialidade na Organização e Representação do Conhecimento para a construção de Sistemas de Organização do Conhecimento Decoloniais

Doutoranda: Ana Flávia Dias Zammataro (1)

Orientadora: Ana Cristina de Albuquerque (2)

(1) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação, ana.flavia.zammataro@uel.br, (2) albuati@uel.br

Introdução

Essa pesquisa é parte de uma tese de doutorado, em andamento, pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de Londrina (PPGCI – UEL) e se insere na Linha de Pesquisa 1 - Organização e Representação da Informação do Conhecimento.

Com o propósito de trazer fundamentos epistemológicos decoloniais à Organização e Representação do Conhecimento (ORC), parte-se de uma abordagem pragmática da área, sob a fundamentação de Hjørland (2003), considerando, dessa forma, as influências sociais nos processos e Sistemas de Organização do Conhecimento (SOC) e, “a intermediação cultural na relação das pessoas com a informação.” (HJØRLAND, 2003, p. 99). Evidencia-se, nesse sentido, uma ORC preocupada com questões relativas à realidade social, demonstrando que seus sistemas de organização do conhecimento devem ser pensados para desconstruir e colocar em evidência saberes marginalizados, elaborando, assim, instrumentos, processos e produtos que representem a pluralidade de saberes e de vivências, além de estarem alinhados às suas respectivas unidades de informação.

As bases epistemológicas decoloniais da ORC, fundamentais no trabalho, foram identificadas a partir de procedimentos metodológicos da pesquisa bibliográfica de caráter qualitativo. Foi realizado levantamento e análise de trabalhos acadêmicos em repositórios nacionais, como a Brapci, as edições dos livros da série *Estudos Avançados em Organização do Conhecimento*, da ISKO Brasil, e também as bases do Benancib, buscando por trabalhos no âmbito do GT2 – Organização e Representação do Conhecimento que discutem os fundamentos teóricos e epistemológicos da Organização e Representação do Conhecimento, além do conceito de decolonialidade no âmbito da ORC, constituindo-se como conceito base para a elaboração

de sistemas de organização do conhecimento com caráter decolonial.

A partir da identificação de trabalhos que trazem abordagens epistemológicas decoloniais à ORC, foi aplicada a análise de natureza qualitativa, buscando compreender de que maneira esses trabalhos abordaram o conceito de decolonialidade e sua relação com a ORC. De acordo com Richardson (1999): “[...] Os estudos que empregam uma metodologia qualitativa podem descrever a complexidade de determinado problema, analisar a interação de certas variáveis, [...] contribuir no processo de mudança de determinado grupo [...]. (Richardson, 1999, p. 80). Destaca-se, nesse sentido, o caráter subjetivo da pesquisa, que nos permite descrever e analisar fenômenos envolvidos, identificando, nesse caso, as abordagens dadas ao conceito de decolonialidade nos respectivos trabalhos.

Os resultados parciais da pesquisa demonstram a importância dos debates decoloniais no âmbito da ORC para que os sistemas de organização do conhecimento sejam elaborados em consonância com as necessidades sociais, priorizando as culturas e os saberes marginalizados e refletindo a função social e política da ORC e por extensão dos próprios Sistemas de Organização do Conhecimento.

Desenvolvimento

Entre os autores que trazem bases epistemológicas decoloniais para a ORC estão Lima e Almeida (2019) e Moura (2018). Lima e Almeida (2019) destacam a importância de se construir SOC multiculturais, que evidenciem saberes marginalizados e que, nesse sentido, tenham a própria sociedade que se deseja representar como participante do processo de constituição desses SOC. Na perspectiva dos autores: “[...] a teoria decolonial busca a emancipação de qualquer tipo de opressão [...]” (LIMA; ALMEIDA, 2019, p. 527).

Os autores, além disso, revelam o caráter universalista e imperialista que ainda fundamenta os estudos na ORC, cujas “[...] metodologias, diretrizes, disciplinas, processos de organizar e representar o conhecimento e suas práticas, são ainda de ordem positivistas e com perspectivas eurocêntricas [...]. (LIMA; ALMEIDA, 2019, p. 525), e por isso, questionam as limitações impostas pelos sistemas europeus dominantes às representações culturais dos países colonizados.

A respeito das relações de poder envolvidas na elaboração dos sistemas de organização do conhecimento, Moura (2018), que os trata como “dispositivos de organização do conhecimento”, esclarece que muitas vezes esses dispositivos são usados para expressar sistemas de dominação e “[...] viabilizar a circulação e a proeminência de discursos hegemônicos, autoritários e regulatórios.” (MOURA, 2018, p. 12). Coloca-se como urgente, nesse sentido, trazer à ORC embasamento epistemológico decolonial, conferindo à área a possibilidade de desconstruir discursos hegemônicos e, a partir disso, ampliar a maneira como a diversidade de saberes e de culturas é organizada e representada nos sistemas de organização do conhecimento.

Resultados Parciais

Até o presente momento de desenvolvimento da pesquisa, foi possível perceber a importância dos debates decoloniais no âmbito da ORC, trazendo para a área fundamentos teóricos e epistemológicos que, primeiramente, questionem a neutralidade dos processos e sistemas de organização do conhecimento para, a partir disso, ampliar essas discussões no intuito de contribuir com a desconstrução do pensamento epistêmico imperialista (LIMA; ALMEIDA, 2019) que ainda se faz presente nessas discussões.

Demonstra-se que esses fundamentos podem ser usados como base para transformar a maneira como a diversidade de saberes e de culturas são organizadas e representadas nos sistemas de organização do conhecimento, contribuindo para dar voz à populações historicamente marginalizadas e, ao mesmo tempo, estando alinhados com as necessidades de seu domínio institucional, sejam eles arquivos, museus ou bibliotecas.

Demonstra-se também que a Ciência da Informação e a Organização e Representação do Conhecimento são potenciais enquanto áreas do conhecimento humano para transformar a realidade social por meio de suas pesquisas, ampliando discussões de bases decoloniais e, a partir disso, produzindo sistemas de classificação, ontologias, vocabulários controlados, taxonomias com caráter decolonial e que busquem “[...] promover o

acesso à informação dentro de um contexto e que dá voz às comunidades de usuários que dele farão uso.” (MILANI; GUIMARÃES, 2017, p. 85).

Conclusão

Além de demonstrar a relevância dos fundamentos epistemológicos decoloniais nos debates da ORC, essa pesquisa corrobora a importância de ampliar e aprofundar essas discussões no âmbito da ORC e da própria Ciência da Informação para que processos, instrumentos e produtos sejam elaborados com bases decoloniais. É necessário, nesse sentido, que as maneiras de se produzir, armazenar, gerir, buscar e distribuir informação e conhecimento sejam repensados, no intuito de que todos os trâmites envolvidos nos fluxos informacionais estejam alinhados não somente com suas respectivas unidades de informação, mas também com aqueles que são representados nesses processos, considerando a diversidade de gênero, a diversidade étnica, social e cultural. A partir disso, será possível questionar narrativas dominantes e assim representar o conhecimento de grupos diversos, buscando assim, dar bases decoloniais para os estudos da ORC.

Referências

- ALMEIDA, Carlos Cândido; LIMA, Graziela dos Santos. Perspectiva pós-colonial e decolonial no campo da Organização do Conhecimento: reflexões para a construção de SOCs multiculturais. *ISKO Brasil*, v.6, 2019; 524-530. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/125408>. Acesso em: 8 maio 2023.
- HJØRLAND, Birger. Fundamentals of Knowledge Organization. *Knowledge Organization*, 30(2). 87-111, 2003. Disponível em: <http://ppggoc.eci.ufmg.br/downloads/bibliografia/Hjorland2003.pdf>. Acesso em: 8 maio 2023.
- MILANI, Suellen Oliveira; GUIMARÃES, José Augusto Chaves. Problemas relacionados a Biases em Sistemas de Organização do Conhecimento: perspectivas para a Representação de Assunto. *Iris – Informação, Memória e Tecnologia*, Recife, v. 3, n. especial. Disponível em: <https://periodicos.ufpe.br/revistas/IRIS/article/view/236189/29018>. Acesso em: 8 maio 2023.
- MOURA, Maria Aparecida. Organização social do conhecimento e performatividade de gênero: dispositivos, regimes de saber e relações de poder. *Liinc em Revista*, Rio de Janeiro, v.14, n.2, p. 118-135, 2018. Disponível em: <http://dx.doi.org/10.18617/liinc.v14i2.4472>. Acesso em: 8 maio 2023.
- Richardson, R. J. *Pesquisa social: métodos e técnicas*. São Paulo: Atlas, 1999.

O domínio dos patrimônios culturais na Ciência da Informação pela perspectiva da Organização e Representação do Conhecimento em plataformas digitais

Doutoranda: Maria Lígia Triques (1)

Orientadora: Ana Cristina Albuquerque (2)

(1) Universidade Estadual de Londrina (UEL), Programa de Pós-Graduação em Ciência da Informação, mligia.triques@uel.br, (2) albuanati@uel.br

Introdução

O termo “patrimônio cultural”, tem sido cada vez mais encontrado em publicações científicas de diversas áreas do conhecimento, sobretudo, quando enquadra-se o leque de interesses de estudos da Ciência da Informação, notadamente, na relação com a organização e representação em ambiência digital. Ao longo da história, os principais agentes na organização e representação do patrimônio cultural institucionalizado têm sido os arquivos, as bibliotecas e os museus, entre outras instituições que podem ser reunidas sob o acrônimo LAM (da expressão inglesa “*Libraries, Archives and Museums*”), desempenhando diferentes atividades em relação à documentação, à preservação e à comunicação dos registros culturais. Seus acervos e coleções podem incluir tanto os objetos do patrimônio tangível, que são armazenados e documentados, quanto os objetos do patrimônio intangível e do patrimônio natural, que são registrados e documentados a partir de fontes textuais, fotográficas, audiovisuais, entre outras. Para todos os casos, tais objetos encontram-se cada vez mais em formatos digitais.

Diante disso, o problema de pesquisa que norteia essa proposição apresenta-se da seguinte forma: como se dão os processos de organização e representação da informação e do conhecimento que viabilizam o diálogo entre os campos de estudo da Ciência da Informação e do Patrimônio Cultural? Objetiva-se, assim, mapear as atividades concernentes ao domínio do patrimônio cultural, incluindo os contextos de tratamento e utilização pública que são disponibilizados em plataformas digitais de modo a investigar como o campo da Ciência da Informação reconhece e trabalha com essas formas de expressão culturais no que tange a organização e representação, sobretudo, em seus formatos digitais.

Desenvolvimento

Delimita-se a partir da abordagem teórico-metodológica da Análise de Domínio (AD) na Ciência da Informação (HJØRLAND; ALBRECHTSEN, 1995) uma pesquisa qualitativa e de caráter exploratório-descritivo. Enquanto método, a AD permite, por meio do emprego de suas abordagens, verificar o que é característico de um determinado domínio e delinear tais características de acordo com critérios definidos.

Assim, propõe-se trabalhar com três abordagens das 11 propostas por Hjørland (2002):

- a) Abordagem 6 - Estudos históricos: organizam tradições, paradigmas, documentos e formas de expressão e suas influências mútuas. Permitirá uma análise mais detalhada e coerente das perspectivas teóricas e relações multidisciplinares que influenciam as atividades do domínio;
- b) Abordagem 7 - Estudos de gêneros e tipologias documentais: revelam a organização e a estrutura de diferentes tipos de documentos em um domínio. Dará a noção de quais objetos informacionais são parte do domínio e como eles se apresentam,
- c) Abordagem 11 - Análise de domínio na cognição profissional e inteligência artificial: fornece modelos mentais de um domínio ou métodos para produzir sistemas especializados. Auxiliará na visualização dos fluxos de trabalho desempenhados no domínio, incluindo ferramentas e métodos empregados.

Ao objetivar mapear um domínio, é preciso identificar as características que lhe são próprias, o que implica em definir parâmetros da análise segundo critérios definidos. Considera-se, pois, seguindo os parâmetros de Tennis (2012), sob a nomenclatura Patrimônio Cultural na Ciência da Informação a cobertura da análise como sendo as atividades realizadas via

plataformas digitais e os processos de organização e representação da informação e do conhecimento, como os parâmetros dessa análise. Ainda, situa-se essa análise na intersecção com demais domínios que lidam com o objeto de estudo, a exemplo da Biblioteconomia, Arquivologia, Museologia.

Resultados Parciais

As tecnologias digitais protagonizam uma era de construção compartilhada de conhecimentos que se beneficiam das conexões em rede de diversificados ambientes informacionais, tais como repositórios, banco de dados e plataformas digitais que passam a configurar espaços de memória, e, assim, espaços que lidam com os registros do conhecimento humano. A investigação desses espaços, comunidades e de suas atividades enquanto domínio revela importantes mudanças na forma de tratar, preservar e comunicar os registros culturais.

Conclusões Parciais

Ao passo que mantém certa continuidade com atividades tradicionais, as plataformas digitais apresentam oportunidades de integração, interoperabilidade e sustentabilidade de dados, bem como de participação social, que confere a organização e representação do patrimônio cultural novos

horizontes e dá ao campo relacional do Patrimônio Cultural com a Ciência da Informação o caráter de um domínio do conhecimento, no qual encontra-se comunidades discursivas, termos e conceitos e implicações sociotécnicas de caráter singular.

Referências

- HJØRLAND, B. Domain analysis in information science: eleven approaches—traditional as well as innovative. **Journal of documentation**, v. 58, n. 4, p. 422-462, 2002. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/00220410210431136/full/html?fullSc=1>. Acesso em: 14 maio 2023.
- HJØRLAND, B.; ALBRECHTSEN, H. “Toward a new horizon in Information Science: domain-analysis”, **Journal of the American Society for Information Science**, v.46, n.6, p. 400-425, 1995. Disponível em: [https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/\(SICI\)1097-4571\(199507\)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y](https://asistdl.onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1097-4571(199507)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y). Acesso em: 14 maio 2023.
- TENNIS, J. T. Com o que uma análise de domínio se parece no tocante a sua forma, função e gênero?. **Brazilian Journal of Information Science**, v. 6, n. 1, p. 3-15, 2012. Disponível em: <https://dialnet.unirioja.es/servlet/articulo?codigo=4328123>. Acesso em: 14 maio 2023.

A importância da privacidade para a Organização do Conhecimento

Doutoranda: Lucilene Aparecida Francisco (1)

Orientadora: Terezinha Elisabeth da Silva (2)

(1) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação, lafrancisco08@gmail.com, (2) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação, Câmara dos Deputados

Introdução

A privacidade aparece no âmbito da Organização do Conhecimento (OC) como um valor ético que deve ser protegido e incentivado com vistas a atender às necessidades informacionais de uma comunidade usuária culturalmente diversificada e conectada às novas tecnologias. Atualmente, é uma preocupação global, formando agenda de políticas públicas no intuito de desenvolver meios para protegê-la, frente a um contexto de expansão do acesso à informação e reconfigurações do sistema capitalista, que expõe a necessidade de preservação dos direitos fundamentais, como intimidade, honra, direito de imagem e dignidade humana.

Em âmbito nacional, foi promulgada a Lei Geral de Proteção de Dados - (LGPD) com o objetivo de garantir a proteção dos direitos fundamentais de liberdade, privacidade e a preservação dos direitos de personalidade dos indivíduos. Dedicando especial atenção aos dados sensíveis, ou seja, aqueles que revelam características intrínsecas à personalidade das pessoas e que podem gerar algum tratamento discriminatório, relativo à origem racial ou étnica, convicção religiosa, opinião política, além dos dados relacionados à saúde do indivíduo. Importante compreender que a LGPD surge em um contexto de capitalismo de vigilância, compreendido como um novo modelo econômico fundamentado na exploração de dados pessoais que, uma vez capturados são convertidos em mercadorias, capazes de gerar grandes possibilidades de lucro (ZUBOF, 2019). Essa nova forma de exploração econômica vem explicitar a necessidade de políticas públicas, voltadas à regulação do acesso e uso de dados.

Este trabalho tem o objetivo de discutir as preocupações da Organização do Conhecimento (OC) com a preservação da privacidade. Para tal, realiza uma pesquisa bibliográfica, de carácter exploratório e delineamento qualitativo, buscando responder ao

seguinte questionamento: Como a OC tem tratado as questões relacionadas à privacidade? As respostas a tal indagação, decorrem de uma busca no Portal de Periódicos da Capes, mais especificamente na base *Library, Information Science & Technology Abstract (LISTA)* com o intuito de levantar na literatura da Ciência da Informação voltada à OC, as principais reflexões sobre este tema nos últimos 5 anos.

Resultados Parciais

O resultado, ainda preliminar, surpreende pela complexidade e variedade de questões envolvidas, uma vez que as bibliotecas e centros de documentação caracterizam-se como gestores de dados pessoais com especial obrigação de preservar a privacidade das pessoas. Como tais, devem lançar um olhar multidisciplinar para o tema e adotar políticas internas, treinamento de funcionários entre outras ações condizentes com o importante papel que essas instituições devem assumir na anomização e pseudomização dos dados de todo público atendido (KATULIC, 2018).

Importante notar que todos os setores das bibliotecas armazenam e processam grandes quantidades de dados pessoais, incluindo: identificação de usuários, registros de circulação e solicitações de reserva, representação dos documentos entre outros serviços. Além disso, as instalações da biblioteca, como computadores, contêm dados, como históricos de navegação, *caches* e *cookies*. Acrescenta-se a esse rol, as solicitações recebidas que revelam dados pessoais, tais como: sexualidade, opiniões políticas, crenças religiosas, condições de saúde, processos criminais, entre outros. Portanto, as bibliotecas estão repletas de dados pessoais (BAILEY (2018). Esse volume de dados vai muito além dos gerados pelos cadastros de usuários que realizam empréstimos ou utilizam os diferentes serviços da biblioteca, basta olhar para os registros de autores dos catálogos do acervo para perceber a proporção disso.

Práticas comuns nas bibliotecas, como registrar a data de nascimento para diferenciar autores homônimos, ou ainda identificar nominalmente autores que preferem pseudônimos, já mostram o quanto isso é frequente.

As próprias normativas de catalogação baseadas nos Requisitos Funcionais de Registros Bibliográficos (FRBR) e o padrão de catalogação Resources, Description and Access (RDA) orientam o registro do local e data de nascimento, sexo, local de residência dos autores para melhor identificá-los no acervo (VARELA - OROL; RODRIGUÉZ, 2018). Todas essas questões influenciam diretamente no cumprimento das legislações de privacidade pelas bibliotecas. Nesse aspecto, a responsabilidade desses espaços, sobretudo nas atividades ligadas à OC, com os dados vai muito além do cumprimento da legislação pertinente, alcança uma obrigação moral de manter a confidencialidade dessas informações, conforme definem os códigos de conduta declarações éticas da profissão de bibliotecário, publicados pela Federação Internacional de Associações e Instituições de Bibliotecas (IFLA) e Associação Americana de Bibliotecas (ALA) (BAILEY (2018)).

Cabe lembrar que a Declaração da IFLA Sobre as Bibliotecas e a Liberdade Intelectual, aprovada em 1999, já destacava o compromisso dessas instituições e com os princípios da liberdade intelectual, do livre acesso à informação e da liberdade de expressão e reconhecimento do direito à privacidade dos usuários. A declaração veta a divulgação da identidade dos usuários a terceiros, assim como a identificação dos documentos por eles utilizados (IFLA, 1999). Desse modo, as bibliotecas e suas atividades de OC têm um compromisso histórico com a privacidade, o que demanda o desenvolvimento de estratégias com vistas a preservação da confidencialidade dos dados dos usuários, incluindo a avaliação crítica acerca do uso das tecnologias disponíveis, principalmente as que se mostrarem capazes de traçar o perfil dos hábitos de leitura ou disponibilizar inadvertidamente dados dos usuários (KRITIKOS; ZIMMER, 2017). Isso porque a privacidade é um valor necessário para se alcançar a liberdade intelectual, o que pressupõe assegurar aos sujeitos que suas opções de leitura não estejam sujeitas a vigilância ou represálias e que possa usufruir da liberdade de explorar ideias, pesar argumentos e decidir por si mesmo no que acreditar e defender (RICHARDS, 2013).

Conclusão

Conclui-se de forma preliminar que as atividades de avaliação, classificação, representação e indexação, próprias da OC, devem desenvolver um olhar crítico e atento às questões de privacidade, observando sempre os riscos de informações confidenciais serem

disseminadas inadvertidamente. Fica evidente, que o uso indiscriminado de dados pessoais pelas atividades de OC, impacta na privacidade dos usuários, autores, organizadores e demais pessoas citadas nas diferentes nas obras. Podendo refletir de maneira especial na privacidade de grupos marginalizados ou sub-representados que disponibilizam seus trabalhos em bibliotecas/centros de informação ou que usam esses serviços para pesquisar assuntos pessoais e privados importantes para sua emancipação.

Referências

- BAILEY, Josephine. Data protection in UK library and information services: Are we ready for GDPR?. **Legal Information Management**, v. 18, n. 1, p. 28-34, 2018. Disponível em: <https://scihub.se/https://doi.org/10.1017/S1472669618000063>. Acesso em: 5 mai. 2023.
- INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). Declaração da IFLA sobre as bibliotecas e a liberdade intelectual. Haia, IFLA, 1999. Disponível em: https://www.ifla.org/wp-content/uploads/2019/05/assets/faiife/statements/iflastat_pt.pdf. Acesso em: 8 mai. 2023.
- KATULIC, Anita. Obveze knjižnica prema Općoj uredbi o zaštiti podataka: izazovi, pristupi i moguća rješenja. [Obrigações das bibliotecas ao abrigo do Regulamento Geral de Proteção de Dados: desafios, abordagens e possíveis soluções]. **Vjesnik bibliotekara Hrvatske**, Zagreb, v. 61, n. 2, p. 343-365, 2018. Disponível em: <https://hrcak.srce.hr/clanak/318154>. Acesso em: 5 mai. 2023.
- KRITIKOS, Katie Chamberlain; ZIMMER, Michael. Privacy policies and practices with cloud-based services in public libraries: An exploratory case of bibliocommons. **Journal of Intellectual Freedom & Privacy**, Chicago, v. 2, n. 1, p. 23-37, 2017. Disponível em: <https://www.journals.ala.org/index.php/jifp/article/view/6252>. Acesso em: 10 mai. 2023.
- RICHARDS, Neil M. The dangers of surveillance. **Harvard Law Review**, Cambridge, v. 126, n. 7, p. 1934-1965, 2013. Disponível em: <https://harvardlawreview.org/print/vol-126/the-dangers-of-surveillance/#:~:text=Fourth%2C%20we%20must%20recognize%20that,harm%20in%20constitutional%20standing%20doctrine>. Acesso em: 10 mai. 2023.
- VARELA-OROL, Concha; AMEIREIROS RODRÍGEZ, Rocío. La protección de datos personales en las bibliotecas universitarias españolas en el entorno digital. *Revista General de Información y Documentación*, Madrid, v. 28, n. 2, p. 685-702, 2018. Disponível em: <https://revistas.ucm.es/index.php/RGID/article/view/62844/4564456549069>. Acesso em: 21 abr. 2023.
- ZUBOFF, Shoshana. **A era do capitalismo de vigilância: a luta por um futuro humano na nova fronteira de poder**. Rio de Janeiro: Intrínseca, 2019.

Organização do Conhecimento e inteligência artificial: interseções teórico-conceituais

Doutoranda: Raíssa Yuri Hamanaka (1)

Orientador: Rogério Aparecido Sá Ramalho (2)

(1) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação, raissa.hamanaka@uel.br, (2) Universidade Federal de São Carlos, Programa de Pós-Graduação em Ciência da Informação, ramalho@ufscar.br

Introdução

A Era da Informação é caracterizada por uma sociedade conectada em rede dentro de um universo digital, globalizado e competitivo, ambiente que exige constantemente inovações, tanto no uso de recursos computacionais, quanto na análise massiva de grandes volumes de dados para a tomada de decisões estratégicas e em perfil profissional de *long life learning*. A afirmação de Borko (1968) que a essência da Ciência da Informação (CI) é representar e processar a informação para sua posterior recuperação e disseminação, ainda é pertinente e necessária. O uso de Sistemas de Organização do Conhecimento (SOCs) e a construção de linguagens documentárias possibilita a representação da informação de maneira estruturada e padronizada para posterior recuperação. Nesse contexto, as tecnologias podem potencializar a organização e tratamento da informação em ambientes digitais.

Objetivos

O objetivo geral do estudo é a realização de discussão teórico-conceitual que busca identificar as relações entre os fundamentos da Organização do Conhecimento (OC), especificamente, de SOCs com as técnicas do campo da Inteligência Artificial (IA).

Para atingir o objetivo geral, foram traçados os objetivos específicos de: a) discutir os fundamentos teóricos da OC e da elaboração de SOCs; b) caracterizar as teorias, técnicas e principais definições do campo da IA com o intuito de facilitar o entendimento da comunidade de profissionais da informação sobre esta temática; c) mapear SOCs sobre IA que fundamentem as aproximações entre os dois campos, e d) delinear as relações existentes entre os dois campos.

O problema desta pesquisa gira em torno da identificação de aproximações entre a CI e as

inovações proporcionadas pelo uso de técnicas de IA. Este estudo busca responder à seguinte pergunta: quais são as interseções teóricas, metodológicas e conceituais entre a OC e o campo da IA?

Metodologia

Trata-se de pesquisa de natureza bibliográfica, abordagem qualitativa, com objetivos exploratório e descritivo e de cunho teórico-reflexivo.

Reflexões Iniciais

O campo da IA nasce com a premissa de simular o funcionamento do cérebro humano. Como permitir que computadores pudessem pensar? Como capacitar sistemas a realizarem tarefas características de seres humanos, como a aprendizagem e a tomada de decisões? Os sistemas que aplicam algoritmos de IA podem ser categorizados em: a) sistemas que raciocinam como humanos; b) sistemas que raciocinam de forma lógica; c) sistemas que atuam como seres humanos e, d) sistemas que atuam de forma racional (CONEGLIAN; SANTARÉM SEGUNDO, 2022).

É possível assinalar algumas interseções entre a CI e a IA, como, a nutrição em torno de dados, informação, conhecimento e pelo uso de tecnologias; suas características interdisciplinares e, os interesses mútuos entre tópicos comuns aos dois campos (MARTÍNEZ MUSIÑO, 2021). No contexto da CI brasileira, a IA vem sendo estudada sob uma perspectiva teórica que analisa potencialidades de seu uso para os profissionais da informação e sob uma perspectiva crítica estudando temas como vigilância cibernética e seus limites éticos. Também ocorrem estudos práticos, como, o uso de processamento de linguagem natural em estudos bibliométricos, a aplicação de técnicas automáticas de categorização de documentos em bases de dados para a melhoria da recuperação da informação, a

representação do conhecimento em sistemas especialistas, etc.

O uso da IA trouxe diversos avanços tecnológicos para diferentes áreas do conhecimento, mas também apresentou dilemas éticos que impactam os direitos humanos, a democracia e o mercado de trabalho, por exemplo. A falta de transparência no desenvolvimento de algoritmos, a falta de políticas de regulamentação no uso de códigos, o surgimento de “bolhas informacionais” e a violação frequente da privacidade de usuários são preocupações recorrentes deste universo de transformações digitais (FERREIRA; VARÃO; BOSELLI; SANTOS; MORET, 2022).

A OC pode se beneficiar do uso de técnicas de IA. As comunidades dos dois campos representam o conhecimento, e a OC pode fazer uso de ontologias e *Linked Data*, por exemplo, na construção de tesouros, sistemas de classificação, taxonomias ou outros tipos de vocabulários controlados. Torna-se necessário repensar os currículos dos profissionais da informação levando-se em consideração modelos algorítmicos e o imperativo do pensamento lógico matemático, dos modelos conceituais e ferramentas de representação do mundo real, sem ignorar a análise crítica preponderante nas humanidades (TEMESIO VIZOSO, 2022).

A informação está presente nos diversos campos do conhecimento, assim como, a necessidade do uso e entendimento de tecnologias. Nesse sentido, o diálogo entre a CI e as tecnologias da informação vem ocorrendo desde sua origem com características interdisciplinares. O uso de técnicas de IA nas pesquisas em CI poderia trazer benefícios na coleta, análise e interpretação dos dados obtidos pelos pesquisadores, o que permitiria trabalhar com conjuntos de dados maiores na aplicação de teorias, métodos e técnicas desenvolvidas por cientistas da informação. O profissional da CI poderia se apropriar do uso de tecnologias semânticas tendo em vista o aperfeiçoamento de suas atividades em ambientes de representação do conhecimento e sistemas de recuperação da informação.

Algumas questões que suscitaram o recorte desta pesquisa foram: Qual seria o papel dos profissionais da informação no cenário de transformações digitais? Arquivos, bibliotecas e museus continuarão existindo no mundo digital? Como a IA afetará os serviços tradicionais de informação? Quais competências o profissional da informação terá de adquirir para atuar

em um mundo marcado pelo uso de técnicas de IA? Quais as contribuições do uso de IA na CI?

Lancaster (1994) alertava para a necessidade de bibliotecários se apropriarem das tecnologias da informação de forma a não apenas utilizá-las, se não também as aprimorando e propondo novos meios de incorporá-las às unidades de informação. As humanidades digitais podem ser entendidas como a interseção entre a tecnologia e as ciências humanas e sociais. A CI está desenvolvendo uma nova perspectiva de análise e interpretação tanto das humanidades digitais como da IA. O uso da IA permeia diversos campos de conhecimento e práticas, sendo matéria de estudo não apenas na CI, mas perpassando diversas áreas do conhecimento, surgindo questionamentos sobre quais serão as profissões do futuro e quais formações serão necessárias para exercê-las (TEMESIO VIZOSO, 2022).

Referências

- BORKO, Harold. Information Science what is it? **American Documentation**, p. 3-5, Jan. 1968.
- CONEGLIAN, Caio Saraiva; SANTARÉM SEGUNDO, José Eduardo. Inteligência artificial e ferramentas da Web Semântica aplicadas a recuperação da informação: um modelo conceitual com foco na linguagem natural. **Informação & Informação**, Londrina, v. 27, n. 1 p. 625-651, jan./mar. 2022. DOI: 10.5433/1981-8920.2022v27n1p625
- FERREIRA, Fernanda Vasques; VARÃO, Rafiza; BOSELLI, Marco Aurelio; SANTOS, Leandro Brito; MORET, Marcelo Albano. Uso de Python para detecção de fake news sobre a covid-19: desafios e possibilidades. **Reciis: Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, Rio de Janeiro, v. 16, n. 2, p. 266-280, abr./jun. 2022. DOI: <https://doi.org/10.29397/reciis.v16i2.3253>
- LANCASTER, Frederick Wilfrid. Ameaça ou oportunidade? O futuro dos serviços de biblioteca à luz das inovações tecnológicas. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 23, n. 1, p. 7-27, jan./jun. 1994.
- MARTÍNEZ MUSIÑO, Celso. La producción documental de la Inteligencia Artificial y la Ciencia de la Información en Scopus: análisis informétrico. **e-Ciencias de la Información**, San Pedro, v. 11, n. 1, p. 1-13, ene./jun. 2021. DOI: <http://dx.doi.org/10.15517/eci.v11i1.42252>
- TEMESIO VIZOSO, Silvana Grazia. Reflexiones sobre la inteligencia artificial y la bibliotecología. **Palabra clave**, La Plata, v. 11, n. 2, p. 159-159, abr./set. 2022. DOI: <https://doi.org/10.24215/18539912e159>

Organização e Representação do Conhecimento em Arquivos de associações político-partidárias: mapeamento de acervos documentais do Partido dos Trabalhadores no Estado de São Paulo

Doutorando: Wilson Roberto Veronez Júnior (1)

Orientadores: Sônia Maria Troitiño-Rodríguez (2), Daniel Martínez-Ávila(3)

(1) Universidade Estadual Paulista, Programa de Pós-Graduação em Ciência da Informação, wilson.veronez@unesp.br,

(2) sonia.troitiño@unesp.br, (3) Universidade de León, Espanha e Universidade Estadual Paulista, Programa de Pós-graduação em Ciência da Informação, dmarta@unileon.es

Introdução

Nos últimos 30 anos, de acordo com Tognoli, Rodrigues e Guimarães (2019), a relação científica, epistemológica e interdisciplinar entre Arquivologia e Organização do Conhecimento (OC) tem sido debatida com ênfase nos dois campos, o que em consequência dessa relação gerou discussões e pesquisas sobre inúmeras temáticas: Diplomática, Descrição, Classificação, Tipologia Documental, Identificação Documental, Fotografias, Vocabulários Controlados e Linguagens Documentárias/Documentais, entre outras. Segundo os autores, essas reflexões e produções científicas motivaram a Arquivologia a repensar seus principais métodos e práticas de organização e representação do conhecimento em arquivos.

Pesquisadores como Smit (2003), Tognoli (2009, 2022), Barros e Moraes (2012), Albuquerque e Souto (2013), Schmidt e Smit (2015), Alencar e Cervantes (2017), Fonseca e Troitiño-Rodríguez (2017), Troitiño-Rodríguez (2018), Barros e Sousa (2019), Tognoli, Rodrigues e Guimarães (2019), Corujo e Freitas (2021), Silva e Corujo (2022), Veronez Júnior, Martínez-Ávila e Troitiño-Rodríguez (2021a, 2022a, 2022b), Veronez Júnior *et al* (2021b), Veronez Júnior *et al* (2022c), Silva e Tognoli (2022), Tognoli, Schmidt e Guimarães (2022), Barros (2022a, 2022b) e Linden (2022), produziram pesquisas a respeito do tema.

A maior parte dessa produção científica foi apresentada e discutida no universo da *International Society for Knowledge Organization* (ISKO), tanto nos anais dos encontros dos capítulos brasileiros, internacionais, ibéricos e norte-americanos, bem como no Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), com foco no Grupo de Trabalho 2 - “Organização e Representação do Conhecimento”-, além de publicações em livros, revistas e periódicos nacionais e internacionais de grande fator de impacto.

1.1 Problema de Pesquisa

Na medida em que o arquivo tem o poder de restringir, ele propicia o acesso à informação e à memória, que para os partidos políticos, tem ampla significação no sentido de pertencimento, identidade, resgate histórico e reavivamento de si mesmo. Esse processo de significação não é apenas para as associações político-partidárias, mas também aos militantes, filiados e simpatizantes. Assim, questiona-se: o que seria uma organização político-partidária sem esses personagens? Qual o sentido de se ter uma agremiação sem a participação desses? Como se dá a organização e representação do conhecimento nos arquivos do Partido dos Trabalhadores no estado de São Paulo?

1.2 Objetivo Geral

Mapear os acervos documentais do Partido dos Trabalhadores (PT) em Diretórios Municipais, Arquivos, Centros de Documentação e Informação, Instituições Culturais e Tribunais Regionais Eleitorais localizados no estado de São Paulo.

1.3 Objetivos Específicos

- a) Realizar um estado da arte sobre os principais aspectos da Organização do Conhecimento e dos tópicos que relacionam a Organização do Conhecimento em Arquivos: Classificação Arquivística, Diplomática e Diplomática Arquivística Contemporânea;
- b) Realizar uma análise histórica sobre a fundação do Partido dos Trabalhadores e sua estrutura organizacional;
- c) Discutir e apresentar uma definição de Conhecimento em Arquivos com base na concepção de Conceito, Forma Documental, Contexto de Produção Documental, Classificação Arquivística e na Diplomática Arquivística Contemporânea;

- d) Discutir a elaboração de um guia de fontes para pesquisadores e militantes que trabalham com arquivos de associações político-partidárias.

1.4 Metodologia

A tese está sendo desenvolvida com base na pesquisa bibliográfica, etnografia documental e na abordagem qualitativa. Para a elaboração da parte teórica, foram recuperados e analisados materiais em bases de dados, como a Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI), *Service Electronic Library Online* (SCIELO), *Library and Information Science Abstracts* (LISA), *Scopus* (SCOPUS) e *Web of Science*, revistas e periódicos especializados em Arquivologia, Arquivo Permanente, Centro de Documentação e Memória, Associações Político-Partidárias e Organização do Conhecimento, além de consulta no Portal Capes de Teses e Dissertações e no *Google Scholar*.

Estão sendo analisados eventos científicos a nível nacional e internacional, como o ENANCIB, a ISKO (Internacional, Brasil, Ibérico e Nasko), o *Encuentros Internacionales sobre Sistemas de Información y Documentación* (IBERSID) e a *Asociación de Educación e Investigación en Ciencia de la Información de Iberoamérica y el Caribe* (EDICIC).

Na etapa prática, o objetivo é dar continuidade às visitas roteirizadas em Diretórios Municipais do Partido dos Trabalhadores no estado de São Paulo, além de instituições documentais e tribunais regionais eleitorais que custodiam acervos documentais do PT.

No decorrer desta etapa, foram analisados o Acervo do Diretório Nacional do Partido dos Trabalhadores (ADNPT), da Fundação Perseu Abramo (FPA), em São Paulo, e o Diretório Municipal de São José dos Campos, porém, o próximo passo é realizar visitas roteirizadas nos diretórios municipais do PT em Marília, Garça, Diadema, Carapicuíba, Osasco, Guarulhos, Ourinhos, Campinas, Taubaté e Jacaréi.

2 Organização e Representação do Conhecimento em Arquivos

Em meados da década de 1990, a discussão sobre a Organização e Representação do Conhecimento em Arquivos já havia sido aludida por Esteban Navarro (1995), que destacou a importância da inserção dos estudos sobre os arquivos no escopo da Organização do Conhecimento, especificamente por meio de publicações em eventos da ISKO.

Para Esteban Navarro (1995), os pesquisadores, as sociedades científicas, as reuniões e congressos que se situam no marco da Representação e Organização do

Conhecimento ignoram, habitualmente, a análise do substrato teórico, recursos metodológicos e dos instrumentos de gestão documental aplicados aos arquivos. Esteban Navarro (1995) e Garcia Marco (1997), advogam que a OC consolida-se como um campo científico interdisciplinar que mantém diálogo com os arquivos, que faz uso de elementos que contribuem para os aspectos teóricos e metodológicos dos processos de organização do conhecimento desenvolvidos em arquivos. Nesse sentido, havia a necessidade de se inserir o tema da OC no contexto dos arquivos, essa inserção se deu principalmente no âmbito da ISKO. Por outro lado, Barros e Sousa (2019) defendem que pesquisas nesse sentido já vinham sendo discutidas desde a década de 1980. Segundo Barros e Sousa (2019), é que no campo da Arquivologia, tem havido uma intensificação do trabalho dedicado aos temas que podemos relacionar com a OC, tais como esquemas de classificação e de sistemas de organização e representação do conhecimento. Com base em Barros e Sousa (2019), já havia indícios de uma expressiva aproximação da classificação arquivística no contexto da OC.

Resultados Parciais

Com base no mapeamento realizado no ADNPT e no Diretório do PT em São José dos Campos, é possível verificar que o processo de organização e representação do conhecimento em arquivos de associações político-partidárias é realizado com base em diretrizes e princípios arquivísticos, sobretudo no que trata sobre a descrição arquivística e a proveniência, no entanto, a partir do diagnóstico nos dois acervos, considera-se que o método de organização adotado pelas instituições são considerados frágeis, por outro lado, quando se analisa a preservação e a conservação dos materiais, entende-se que está de acordo com os padrões arquivísticos.

Além da falta de organização, observa-se que há uma fragmentação dos arquivos, e que essa massa documental, em sua maioria é constituída por documentos textuais, fotografias, bandeiras, faixas, panfletos e cartazes, ou seja, muitos desses objetos ficam em posse dos militantes, simpatizantes e gestores do Partido ou até mesmo depositados em porões e locais inadequados, o que dificulta com que esses arquivos sejam disponibilizados para tratamento, classificação e acesso.

Considera-se que a partir da fragmentação desses arquivos, existe uma dificuldade no processo de ressignificação, contextualização, reconstrução e na preservação da memória de associações político-partidárias.

Referências

- ALENCAR, M. F.; CERVANTES, B. M. N. A importância da organização do conhecimento arquivístico no acesso à informação: um olhar para tesouros funcionais. *In: XVIII Encontro Nacional de Pesquisa em Ciência da Informação (XVIII ENANCIB)*. Marília, São Paulo, 2017.
- BARROS, T. H. B.; MORAES, J. B. E. de. Archival classification and knowledge organization: theoretical possibilities for the archival field. *In: Categories, Contexts And Relations In Knowledge Organization*. Proceedings of the Twelfth International ISKO Conference. 6-9. (2012).
- BARROS, T. H. B.; TOGNOLI, N. B. Os processos de representação do conhecimento arquivístico: elementos históricos e conceituais da classificação e descrição. *In: GUIMARÃES, José Augusto Chaves; DODEBEI, Vera. Organização do conhecimento e diversidade cultural*. v. 3. Marília: Fundepe, 2015. p. 94-99.
- BARROS, T. H. B.; SOUSA, R. T. B. de. Organização do Conhecimento e Arquivologia: abordagens metodológicas. *Informação & Informação*, v. 24, n. 2, p. 76-92, 2019.
- CORUJO, L. M. N.; FREITAS, M. C. V de. Organização, gestão e representação da informação e do conhecimento na arquivística: interações e interlocuções. *Ponto de Acesso*, v. 15, n. 3, 2021.
- ESTEBAN NAVARRO, M. Á. La representación y la organización del conocimiento en los archivos: los lenguajes documentales ante los procesos de clasificación, ordenación y descripción. *In: Organización del conocimiento en sistemas de información y documentación: actas del I Encuentro de ISKO-España*, Madrid, 4 y 5 de noviembre de 1993. Universidad de Zaragoza, 1995. p. 65-90.
- FONSECA, G. F.; TROITIÑO RODRÍGUEZ, S. M. The contribution of archival identification to knowledge organization in personal archives. *Knowledge Organization and Cultural Diversity [recurso eletrônico] / José Augusto Chaves Guimarães, Vera Dodebei, organizadores*. -- Pernambuco: ISKO-Brasil; UFPE, 2017. 760 f.; 30 cm. **INTERNATIONAL SOCIETY KNOWLEDGE ORGANIZATION**, Marília – SP. 322-327.
- GARCÍA MARCO, F. J. Avances en organización del conocimiento en España: los II encuentros sobre organización del conocimiento en los sistemas de información y documentación. *In: GARCÍA MARCO, F.J. (ed.) Organización del conocimiento en sistemas de información y documentación*. Zaragoza: Librería General, p.7-12. 1997.
- GUIMARÃES, J. A. C.; TOGNOLI, N. B. Provenance as a domain analysis approach in archival knowledge organization. *Knowledge Organization*, v. 42, n. 8, p. 562-569, 2015.
- HJORLAND, B. Fundamentals of knowledge organization. *Knowledge Organization*, v. 30, n. 2, p. 87-111, 2003.
- LINDEN, L. L. Organização do conhecimento e Arquivologia: diálogos e prospecções. *In: Organização e representação do conhecimento em múltiplas abordagens / Organizadores Thiago Henrique Bragato Barros, Rita do Carmo Ferreira Laipelt*. – São Paulo: Pimenta Cultural, 2022.
- SILVA, A. M. S. da; TOGNOLI, N. A organização do conhecimento arquivístico: a emergência de uma comunidade discursiva brasileira. *Acervo*, v. 35, n. 2, p. 1-17, 5 abr. 2022.
- TOGNOLI, N. B.; GUIMARÃES, J. A. C. A diplomática contemporânea como base metodológica para a organização do conhecimento arquivístico: perspectivas de renovação a partir das ideias de Luciana Duranti. *In: Nuevas perspectivas para la difusión y organización del conocimiento: actas del congreso*. Servicio de Publicaciones, 2009. p. 38-47.
- TOGNOLI, N. B.; RODRIGUES, A. C.; GUIMARÃES, J. A. C. Definindo o conhecimento arquivístico: estruturas conceituais. *Informação & Informação*, v. 24, n. 2, p. 58-75, 2019.
- TOGNOLI, N. B.; SCHMIDT, C.; GUIMARÃES, J. A. C. Context as a Core Concept in Archival Knowledge Organization. *In: Knowledge Organization across Disciplines, Domains, Services and Technologies*. Ergon-Verlag, 2022. p. 273-284.
- VERONEZ JÚNIOR, W. R.; MARTÍNEZ-ÁVILA, D.; TROITIÑO RODRÍGUEZ, S. M. Arquivologia e Organização do Conhecimento: uma análise nos anais da Isko Brasil, Isko Internacional e Isko Ibérico. *In: Silva, Carlos Guardado da, Revez, Jorge & Corujo, Luís (Eds.). (2021). Organização do Conhecimento no Horizonte 2030: Desenvolvimento Sustentável e Saúde: Atas do V Congresso ISKO Espanha-Portugal*, Universidade de Lisboa. Faculdade de Letras, 25 e 26 de novembro de 2021. Lisboa: Centro de Estudos Clássicos, Colibri. (CA – Ciência Aberta; 1) ISBN 978-989-566-137-4 CDU 025.4(063).
- VERONEZ JÚNIOR, W. R., ALMEIDA JÚNIOR, O. F. de, TROITIÑO RODRÍGUEZ, S. M. T.; MARTÍNEZ-ÁVILA, D. (2022b). A Dimensão Política da Mediação da Informação em Arquivos de Partidos Políticos Brasileiros: o caso do Partido dos Trabalhadores. *Revista EDICIC*, 2(1). Disponível em: <https://ojs.edicic.org/index.php/revistaedicic/article/view/138> Acesso em 05 de mai. 2023.
- VERONEZ JÚNIOR, W. R.; MARTÍNEZ-ÁVILA, D.; TROITIÑO RODRÍGUEZ, S. M. Dimensões científicas e epistemológicas da Organização do Conhecimento Arquivístico: análise com base nas contribuições de Bunge e Japiassu. **XXII Encontro Nacional de Pesquisa em Ciência da Informação – XXII ENANCIB**, edição 2022c.

Paratextos: elementos que acompanham o texto

Doutorando: Hugo Figueiredo Mafra (1)

Orientadora: Rosa Inês de Novais Cordeiro (2)

(1) Universidade Federal Fluminense, Programa de Pós-Graduação em Ciência da Informação, hugom@id.uff.br,

(2) rosanovais@id.uff.br

Introdução

Um texto não é um livro. Esta é a ideia do teórico francês Gérard Genette (2009) que entende que para um texto se tornar livro é necessário estar rodeado de diversos elementos que o acompanham e o complementam, auxiliando sua recepção e consumo. Elementos estes intitulados, por Genette, de paratextos, conceito concebido no final do século XX e aprofundado em seu livro “Seuils” de 1987 que possui a versão traduzida para o português, Paratextos Editoriais (GENETTE, 2009).

Esta comunicação tem a intenção de apresentar os resultados preliminares da literatura levantada sobre paratextos, cuja pesquisa tem como problema inicial indagar quais contribuições o paratexto pode fornecer para a Organização de Conhecimento e seus processos técnicos, considerando o objetivo de trazer o conceito de paratexto para a Ciência da Informação, entendendo esses elementos como componentes do documento, que auxiliam sua recepção e consumo. O estudo do conceito de paratexto faz-se importante, conforme será evidenciado a seguir, devido a contribuição que exerce como complemento e informação do texto, garantindo que este seja recebido pelos indivíduos.

Desenvolvimento

O paratexto faz uma mediação para o texto principal. É a porta de entrada do texto, possibilitando ao leitor ir em frente na leitura ou voltar. Trata-se de um limite flexível entre o dentro e o fora do texto, pois o leitor ainda não chegou a ele, mas também não está fora do seu universo. O paratexto pode comunicar uma informação, expor uma intenção ou interpretação. Entretanto, está sempre subordinado ao texto que faz referência.

Na perspectiva de Genette (2009), o paratexto divide-se em peritexto e epitexto. Essa divisão fica explícita no esquema: paratexto = peritexto + epitexto. O peritexto é todos os elementos paratextuais que se encontram no limite do livro, ou seja, no interior:

título, prefácio, capa, agradecimentos etc. Já aqueles que estão localizados no exterior do limite do livro, recebem o nome de epitexto: entrevista, resenha, peças de divulgação etc. Todavia, ressaltamos que os paratextos podem aparecer ou desaparecer dependendo da edição do livro, uma vez que é uma escolha do editor ou autor mantê-los ou não. Para além disso, um epitexto pode vir a se tornar um peritexto e vice-versa.

O paratexto, compreende Genette (2009), influencia na forma como os indivíduos leem uma obra, impactando diretamente em sua recepção. Estendendo esta ideia, os próprios paratextos podem influenciar a forma como os leitores buscam as obras.

Skare (2020) compreende que o paratexto além de apresentar, também tem um poder de marketing/venda e de direcionamento da experiência de leitura. Uma peculiaridade é que sempre cabe ao leitor utilizar esses paratextos ou não, logo, o leitor possui um poder de escolha de consumi-los ou não. Wakarindi (2019) considera o paratexto como um fator crucial para qualquer livro, pois apresentam a substância do texto.

Le Coadic (1996) aponta o paratexto como tema que dialoga com a Ciência da Informação pela interdisciplinaridade. Percebemos que, já em 1996, havia uma preocupação, por parte da Ciência da Informação, em estabelecer a relação entre a área e o paratexto e entender como podemos nos beneficiar desse conceito. Todavia, Gross e Latham (2017) expõem que apesar de haver inúmeras pesquisas que o consideram, a maior parte dos estudos é fora da Ciência da Informação, pois a área talvez ainda não reconheça as contribuições da teoria do paratexto tanto para a pesquisa quanto para o trabalho profissional.

Resultados Parciais ou Finais

Como início da pesquisa, realizamos uma pesquisa bibliográfica, sem recorte temporal, em algumas bases de dados a fim de verificar se este conceito tem sido trabalhado na Ciência da Informação. Para tanto, selecionamos sete, sendo três de Ciência da Informação e quatro que abrangem outras áreas do conhecimento.

São elas: Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação, Base de Dados do ENANCIB, Library, Information Science and Technology Abstracts, Scientific Electronic Library Online, Web of Science, Repositórios Científicos de Acesso Aberto de Portugal e Biblioteca Digital Brasileira de Teses e Dissertações. Utilizaremos “Paratexto” e “Ciência da Informação” como palavras-chave.

Ao todo, tivemos 99 documentos como resultado. Todavia, havia 16 duplicados e dos 83 restantes, 15 estavam indisponíveis, sobrando, assim, 68 artigos para análise. Desses, 63 são artigos, quatro são apresentações de evento e uma dissertação, divididos nos idiomas Inglês (43), Português (14), Espanhol (6), Árabe (1), Eslovaco (1), Italiano (1), Japonês (1) e Língua Lituana (1).

Em um primeiro momento, decidimos observar em cada documento se os autores definiam Paratexto e se utilizavam as obras de Gerard Genette ou outra indicação de autoria. Percebemos que a maioria, totalizando 47 documentos, se valiam dos textos de Genette para apresentar o conceito. Já o restante, utilizou outro autor ou simplesmente não definiu. Houve apenas um caso que empregou Genette, porém a partir de um outro autor, ou seja, pelo recurso de citação de segunda mão (apud).

Conclusão

Constatou-se assim, um número baixo de resultados, dada a relevância que o assunto possui para os processos da área. Portanto, vemos como necessidade trazer o tema para a área de Ciência da Informação, bem como ampliar e aprofundar a sua discussão conceitual, especialmente por percebemos que os elementos que Genette (2009) elenca como paratextos apontam para a necessidade de revisão e atualização dos seus tipos, conceitos e esferas, levando em conta a gama de objetos digitais contemporâneos, os quais têm sido um desafio de estudo na área, devido, entre outros fatos, ao imbricado cenário de produção, circulação e recepção dessas fontes. Atualmente, há outros paratextos produzidos que o autor não abarca em sua obra, principalmente quando consideramos o ambiente digital.

Referências

- GENETTE, G. **Paratextos editoriais**. Cotia: Ateliê, 2009.
- GROSS, M.; LATHAM, D. The peritextual literacy framework: using the functions of peritext to support critical thinking. **Library & Information Science Research**, Amsterdam, v. 39, p. 116-123, 2017.
- LE COADIC, Y. F. **A ciência da informação**. Brasília: Briquet de Lemos, 1996.
- SKARE, R. Paratext. **Knowledge Organization**, Baden-Baden, v. 47, n. 6, p. 511-519, 2020.

O período de escravização no Brasil e a Organização e Representação da Informação e do Conhecimento: estudos nas instituições arquivísticas

Doutoranda: Roberta Albanielle Garcia (1)

Orientadora: Ana Cristina de Albuquerque (2)

(1) Universidade Estadual de Londrina, Programa de Pós Graduação em Ciência da Informação, robertagarcia@uel.br,
(2) albuati@uel.br

Introdução

Dentre os objetivos da Organização e Representação da Informação e do Conhecimento (ORIC) está compartilhar com público os objetos informacionais e os objetivos socioculturais, motivado em como os conhecimentos, crenças, valores e tradições de uma sociedade são transmitidos, armazenados e comunicados. Como disciplina a Organização do Conhecimento é responsável pelo desenvolvimento de técnicas de construção, gestão, uso e avaliações de classificações científicas, taxonomias, nomenclaturas e linguagem documentais, de acordo com Barité (2001), que determina como o objeto de estudo da Organização do Conhecimento o conhecimento socializado. Barité (2001) explica que a Organização do Conhecimento fornece subsídios teóricos para tudo aquilo que é relativo ao tratamento da informação e a gestão do uso social da informação.

O presente resumo tem por objetivo apresentar parte dos planos de pesquisa de doutorado que dão continuidade às pesquisas desenvolvidas pela pesquisadora na dissertação de mestrado acadêmico em Ciência da Informação, construindo sua trajetória acadêmica e tema de pesquisa. A temática selecionada para a problemática, análise e discussão no âmbito da ORIC representa um longo período da história brasileira: a escravização de pessoas traficadas do continente africano para o americano, e de seus descendentes; ocorrido no Brasil por mais de trezentos anos, submetendo-os a condições sub-humanas de sobrevivência.

O escravizado nesse período não era visto como um cidadão, possuía pouco ou nenhum direito, era mão de obra e mercadoria (ROBLES; QUEIROZ, 1987). Tais ações deixaram marcas na cultura e sociedade brasileiras e precisam ser objeto de estudo nos mais diversos ambientes de conhecimento, buscando ressignificar a identidade e a experiência afrodescendente no país. Ianni (1978) explica que estudos sobre as raízes históricas do racismo buscam

compreender o chamado “trauma da escravidão”, e a influência dele no período pós-abolicionista e no século XX, para o autor “Tanto em nível ideológico como no das relações econômicas, por exemplo, o negro liberto e os seus descendentes carregariam uma espécie de experiência negativa, a experiência escrava; [...]” (IANNI, 1978, p. 115), portanto o racismo estrutural (ALMEIDA, 2019) presente na sociedade brasileira pode ser resultado do desenvolvimento do negro dentro da sociedade sempre da perspectiva de um ex-escravo, descendente de escravos.

Essa é uma perspectiva lembrada por Ribeiro (2019, p.7) ao relatar: “Quando criança, fui ensinada que a população negra havia sido escrava e ponto, como se não tivesse existido uma vida anterior nas regiões de onde essas pessoas foram tiradas à força.” Esse cenário não foi alcançado, e mantido, por acaso, “[...] foi um ‘projeto nacional’ a produção de um discurso sobre o outro, tornando racional e emocionalmente aceitável a conquista e a destruição daqueles com os quais não se compartilha a mesma identidade.” (ALMEIDA, 2019, p.103), que se prolonga pelo período de centenas de anos de escravização e persegue os afrodescendentes na pós-abolição, de 1888 até os dias atuais.

Problematiza-se a pesquisa na consciência de que os documentos da escravização, fonte primária e essencial para estudos sobre o período, não foram criados com o objetivo de registrar a escravização como um período/acontecimento, os documentos de arquivo registram as funções administrativas, sendo salvaguardados por seu valor probatório público. No campo da ORIC ambienta-se a preocupação em criar sistemas de organização do conhecimento (SOCs) eficientes no tratamento e compartilhamento de produtos criados pelos processos humanos e artificiais de organizar conhecimento e informação nas instituições de custódia permanente. A pesquisa tem, portanto, o objetivo de analisar a temática e os aspectos da ORIC nos SOCs previamente selecionados.

Desenvolvimento

Metodologicamente a pesquisa se desenvolve em dois âmbitos: o teórico e a investigação e análise dos resultados obtidos. A pesquisa bibliográfica segundo Fachin (2001) é uma constante na vida de quem se propõe estudar, e reúne no trabalho as bases teóricas e conceituais sobre os assuntos que compõe o corpo teórico. No que diz respeito à coleta de fontes que integrarão o universo de estudo da pesquisa em sua totalidade a pesquisa se enquadra como documental, nesse sentido, Fachin (2001, p. 146) explica a multiplicidade das fontes que podem compor a pesquisa documental: “Para a pesquisa documental, considera-se documento qualquer informação sob a forma de textos, [...] entre outros.”. Delimitada as fontes e o universo de pesquisa com o método Análise de Conteúdo, de Lawrence Bardin (1977), para análise, exploração do material, o tratamento dos resultados, inferência e interpretação, será possível alcançar os objetivos propostos no projeto. Por fim, para reflexão das questões elaboradas pela pesquisa a abordagem em que se enquadra é a qualitativa.

Durante o desenvolvimento teórico da tese é imprescindível a reflexão sobre a relação tríplice entre a ORIC, o conhecimento custodiado nas instituições arquivísticas de caráter público e os SOCs. As definições daquilo que é público, atreladas ao conceito de Democracia que vivemos no Brasil, nos permitem entender como público aquilo que é do povo para o povo, o resultado de viver como pátria, amparados pela mesma Constituição. Por isso tratar a documentação pública com o objetivo de disponibilizá-la aos mais diversos usuários, pesquisadores, historiadores e cidadãos, é devolver para o povo aquilo que diz respeito ao povo. A organização, análise, interpretação e tratamento de informações e conhecimento criados pela função administrativa do Estado é de fundamental importância para o fazer administrativo, mas o tratamento específico da documentação histórica é igualmente importante e urgente: “Os arquivos públicos constituem um tipo de fonte de cultura entre muitos outros tipos como livros, manuscritos e tesouros de museus. São uma fonte tão importante como os parques, monumentos ou edifícios.” (SCHELLENBERG, 2006, p. 31)

Conclusão

Os resultados parciais demonstraram a importância dos processos, instrumentos e produtos da organização para o acesso e compartilhamento temático de aspectos históricos de tamanha relevância para o país, tais produções possibilitam encontrar nos documentos arquivísticos as vidas registradas, os nomes, as histórias, entre outros dados que permitem uma nova compreensão dos acontecimentos do passado. Para o desenvolvimento e conclusão da tese é esperado ainda vislumbrar novas relações conceituais nos documentos do período da escravização e nos SOCs onde estão inseridos.

Referências

- ALMEIDA, Silvio Luiz de. **Racismo Estrutural**. São Paulo: Sueli Carneiro; Pólen, 2019. 255p.
- BARDIN, Laurence. **Análise de conteúdo**. Lisboa: Edições 70, 1977. 229p.
- BARITÉ, Mario. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K. (Org.). **Educação, Universidade e Pesquisa**. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP, 2001. p.35-60.
- FACHIN, Odília. Pesquisa social, documental e de laboratório. In: FACHIN, Odília. **Fundamentos de metodologia**. 5. ed. rev. e atual. São Paulo: Saraiva, 2006. p. 130-150.
- IANNI, Octavio. **Escravidão e Racismo**. São Paulo: Editora Hucitec, 1978, 142p.
- RIBEIRO, Djamilia. **Pequeno Manual Antirracista**. São Paulo: Companhia das Letras, 2019. 135p.
- ROBLES, Suely; QUEIROZ, Reis de. **Escravidão negra no Brasil**. São Paulo: Ed. Ática, 1987. 85 p.
- SCHELLENBERG, Theodore Roosevelt. **Arquivos modernos: princípios e técnicas**. 6ª. ed. Rio de Janeiro: Editora FGV, 2006.

A produção científica dos líderes dos grupos de pesquisa em Organização e Representação da Informação e do Conhecimento: uma análise nos anais dos eventos da ISKO-Brasil

Doutoranda: Jéssica dos Santos Gonçalves (1)

Orientadores: Natália Bolfarini Tognoli (2), Daniel Martínez-Ávila (3)

(1) Universidade Federal Fluminense, Programa de Pós-Graduação em Ciência da Informação, jegoncalves@id.uff.br

Introdução

A Ciência não pode prescindir, nos tempos atuais, de grupos de pesquisadores trabalhando na resolução de problemas que afligem a humanidade. Os grupos de pesquisadores, que podem ou não formar grupos de pesquisa entre si, constituem o campo científico de determinado domínio do conhecimento.

Um grupo de pesquisa consiste em um conjunto de indivíduos com hierarquia estabelecida, que realizam tarefas interdependentes e compartilham responsabilidade na geração de resultados. Os grupos de pesquisa, cadastrados no Diretório dos Grupos de Pesquisa no Brasil - DGP do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq, possuem um ou dois pesquisadores líderes que definem as temáticas de pesquisa do grupo.

O aprimoramento teórico-conceitual do domínio da Organização do Conhecimento (OC) é uma das finalidades dos grupos de pesquisa deste domínio. Com isso, o objetivo geral da Tese é “analisar o capital científico da elite dos líderes dos grupos de pesquisa em Organização do Conhecimento no Brasil”. Para tanto, são desenhados os seguintes objetivos específicos: Discutir o capital científico dos atores ligados à OC à luz da Sociologia da Ciência; Identificar a elite de pesquisa em Organização do Conhecimento nos anais dos eventos da ISKO-Brasil; Identificar a frente de pesquisa nos anais dos eventos da ISKO-Brasil.

Tendo em vista os pressupostos que guiam a pesquisa, propõem-se as seguintes questões:

- 1) Qual a frente de pesquisa em Organização do Conhecimento no Brasil? Quais autores a compõem?;
- 2) Os autores da frente de pesquisa fazem parte dos grupos de pesquisa em OC como membros ou líderes?;

- 3) Se não participantes de grupos de pesquisa em OC, são membros ou líderes de grupos de pesquisa de outros domínios?;
- 4) Os líderes dos grupos de pesquisa em OC são a elite de pesquisa do domínio Organização do Conhecimento no Brasil?;
- 5) Os autores da elite de pesquisa, caso sejam membros ou líderes de grupos de pesquisa, são citados pelos demais membros destes grupos?;
- 6) Esses autores da elite e/ou da frente de pesquisa são filiados à ISKO-Brasil?;
- 7) Embora não sejam associados, tanto os autores da elite quanto da frente de pesquisa utilizam os eventos da ISKO-Brasil para divulgar suas pesquisas?

Desenvolvimento

O líder de grupo de pesquisa é definido no glossário do DGP/CNPq como o indivíduo que “[...] detém a liderança acadêmica e intelectual no seu ambiente de pesquisa. Normalmente, tem a responsabilidade de coordenação e planejamento dos trabalhos de pesquisa do grupo. Sua função aglutina os esforços dos demais pesquisadores e aponta horizontes e novas áreas de atuação dos trabalhos. Um grupo pode admitir até dois líderes, denominados 1º Líder e 2º Líder.” (CNPq, [202-]). Portanto, nessa definição já podemos notar a influência dos líderes dentro da hierarquia dos grupos de pesquisa e no direcionamento de temáticas que serão pesquisadas pelo grupo.

Denota-se a influência dos líderes dentro da hierarquia existente nos grupos de pesquisa. Essa influência torna o estudo da utilização dos anais de eventos como canal de divulgação pelos líderes de grupos de pesquisa primordial para compreender as dinâmicas comunicacionais dos campos científicos.

A produção científica dos grupos de pesquisa é divulgada através de diferentes fontes de informação, dentre elas os trabalhos apresentados em eventos, cuja publicação ocorre em anais, boletins de resumos, entre outros, atualmente em formato eletrônico. Essas publicações são importantes, pois nestes eventos as pesquisas passam a ser conhecidas, e podem ser devidamente criticadas e aprimoradas a partir dos comentários e refutações de pesquisadores especialistas. Posteriormente, servem de fontes a levantamentos bibliométricos.

Os indicadores bibliométricos, como por exemplo os indicadores de citação, possibilitam ressaltar os pesquisadores, temáticas investigadas, campos científicos – frentes de pesquisa, redes colaborativas, grupos de pesquisa, redes de citação ou cocitação – e os países cuja pesquisa está mais avançada ou que são mais produtivos. Isto posto, constata-se que a aplicação de indicadores de citação no domínio da Organização do Conhecimento contribuirá para ampliar a compreensão que se tem dele.

Quanto aos indicadores de citação, os autores dos artigos mais citados de um corpus de publicações em um período de tempo pré-determinado correspondem à frente de pesquisa. Relacionado a ele, o conceito de elite de pesquisa pode ser definido como a raiz quadrada de autores que mais produzem; portanto, a raiz quadrada da quantidade de autores que compõem a frente de pesquisa de um domínio (BRAGA, 1973; URBIZAGÁSTEGUI-ALVARADO, 2009).

Pretende-se estudar, com isso, os agentes e a estrutura do campo científico da OC. A fonte serão os anais da ISKO-Brasil no período de 2011 a 2021, o que evidencia a relevância dos anais de eventos científicos, pois neles ocorre a publicação dos primeiros resultados e com isso a confirmação (ou não) da validade das pesquisas científicas que estão sendo realizadas. A análise da produção se dará com base nos indicadores de citação dos líderes dos grupos de pesquisa nos trabalhos apresentados neste evento.

A representação dos grupos de pesquisa brasileiros em OC na ISKO-Brasil é fonte de interesse por ser um espaço importante de trocas e construção de conhecimento acerca da Organização do Conhecimento no cenário nacional.

Resultados Parciais

Como a intenção é verificar se os líderes (1º e 2º) de grupos de pesquisa fazem parte da elite e da frente de pesquisa de autores em Organização do Conhecimento, foi feito um levantamento prévio dos grupos de pesquisa em OC e seus líderes. Encontrou-se 51 (cinquenta e um) grupos de pesquisa utilizando o termo de busca “organização do conhecimento”, em “Todas

as palavras”, Consultar por “Grupo”, aplicando a busca nos campos “Nome do grupo”, “Nome da linha de pesquisa” e “Palavra-chave da linha de pesquisa”, filtrando por Área do Conhecimento “Ciências Sociais Aplicadas” e por Área “Ciência da Informação”, ambas marcadas como Predominante do grupo.

Em 5 de maio de 2023 a autora passou por exame de qualificação de Doutorado, no qual constatou-se que a estratégia de busca acima terá que ser modificada, a fim de contemplar grupos de pesquisa que contenham nos campos “Nome do grupo”, “Nome da linha de pesquisa” e “Palavra-chave da linha de pesquisa” termos/conceitos correlatos à ou contidos na organização do conhecimento, como organização da informação, representação temática, dentre outros. Isso se dá devido a esses conceitos estarem em voga em determinado período temporal, que corresponde a criação de alguns grupos de pesquisa cadastrados no DGP/CNPq.

Da mesma forma, ocorreu pré-teste com os 51 grupos encontrados, utilizando 10 trabalhos publicados no I Congresso da ISKO-Brasil em 2011, para identificar as autocitações e possíveis dificuldades na coleta de dados. Foram encontrados no pré-teste 9 autocitações pelos citantes e verificou-se que a maioria dos autores dos trabalhos não era 1º ou 2º líder dos grupos de pesquisa em OC listados. Os autores citados foram tabulados, mas ainda não foram analisados quanto à liderança de grupos de pesquisa e quantidade de citações.

Ainda será realizado o levantamento completo da produção de pesquisadores brasileiros nos anais dos eventos da ISKO-Brasil, a ser comparado com a listagem de líderes de grupos de pesquisa atualizada, para verificar se estes fazem parte da elite e/ou da frente de pesquisa em OC no Brasil encontrada nestes anais, confirmando ou não a relevância e produtividade dos líderes de grupos de pesquisa.

Conclusão

A relevância dos estudos que analisam a OC como domínio do conhecimento foi apontada por diversos autores, tais quais Hjørland (2008), Smiraglia (2012), Guimarães (2014), Sales (2016, 2017), dentre outros. O pensamento crítico acerca da Organização do Conhecimento é essencial, na medida em que contribui para o fortalecimento do corpus teórico que compõe o domínio em questão e, em consequência, do próprio domínio.

Da mesma forma, os hábitos e costumes das comunidades ou campos científicos influenciam no julgamento, publicação, recuperação e citação do conhecimento produzido neles. Esses ritos e práticas variam a depender do campo científico. Assim sendo,

entender a natureza e tradições de um campo é fundamental, pois só assim é possível compreender como influenciam sua produção, legitimação e uso do conhecimento.

Notas

Esse projeto de Tese compõe as discussões do grupo de pesquisa Arquivos, Bibliotecas e Organização do Conhecimento (GPABOC) do programa de pós-graduação do qual a autora faz parte, que tem como foco investigações práticas e epistemológicas em OC.

Referências

- BRAGA, G. M. Relações bibliométricas entre a frente de pesquisa (research front) e revisões de literatura: estudo aplicado a Ciência da Informação. **Ciência da Informação**, Rio de Janeiro, v. 2, n. 1, p. 9-26, 1973.
- CNPq. **Diretório dos grupos de pesquisa no Brasil**: glossário. Brasília, [202-]. Disponível em: <https://lattes.cnpq.br/web/dgp/glossario>. Acesso em: 11 abr. 2023.
- GUIMARÃES, J. A. C. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, Brasília, DF, v. 41, n. 1, p. 13-21, jan./abr. 2014.
- HJØRLAND, B. What is knowledge organization (KO)? **Knowledge Organization**, Baden-Baden, v. 35, n. 2/3, p. 86-101, 2008.
- SALES, R. Knowledge organization in the brazilian scientific community and its epistemological intersection with Information Science. *In*: INTERNATIONAL ISKO CONFERENCE, 14., Rio de Janeiro, 2016. **Anais eletrônicos [...]**. [Marília]: UNESP; ISKO-Brazil, 2016.
- SALES, R. Diferentes perspectivas nos contextos do GT2 da ANCIB e da ISKO-Brasil. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., Marília, 2017. **Anais eletrônicos [...]**. Marília: UNESP, 2017.
- SMIRAGLIA, Richard P. Chapter 6: epistemology of domain analysis. *In*: SMIRAGLIA, Richard P.; LEE, Hur-Li (eds.). **Cultural frames of knowledge**. Würzburg: Ergon Verlag, 2012. p. 111-124.
- URBIZAGÁSTEGUI-ALVARADO, R. A frente de pesquisa na literatura sobre a produtividade dos autores. **Encontros Bibli**, Florianópolis, v. 14, n. 28, p. 38-56, 2009.

Scientia Oblitus: o esquecer e o resgate dos saberes dentro da Organização e da Representação do Conhecimento

Doutorando: Pedro Henrique Cremonez Rosa (1)

Orientador: Miguel Luiz Contani

(1) Universidade Estadual de Londrina, Programa de Pós-Graduação em Ciência da Informação, pedro.cremonez@uel.br

Introdução

O conhecimento humano é fruto de um processo de construção e aprimoramento contínuo e cumulativo. A edificação do conhecimento se dá por dois movimentos, que ora se identificam como antagônicos, ora se complementam: a disseminação e a perda do conhecimento (BURKE, 2003, 2012). O primeiro movimento é o de pluralização daquilo que se conhece ou do que se descobre. O segundo é resultado de superações, desusos, negações, destruição, incapacidade de acesso e esquecimento. A perda do conhecimento, paradoxalmente, promove aspectos positivos, quando se trata da evolução do conhecimento resultante da busca de aprimoramento; o lado desfavorável se mostra em outras instâncias diretamente ligadas ao movimento de esquecimento e silenciamento de saberes. Dentro da dinâmica da ciência, conhecimentos passam por uma validação de acordo com critérios predeterminados, comumente resultantes de um pensamento cartesiano e empírico que, na busca de padronizar e legitimar as construções do conhecimento, restringem a produção dos saberes (SANTOS, 1999, 2007, 2010a, 2010b, 2019). Nessa dinâmica, a formulação de parâmetros e critérios destinados a certificar aquilo que é proposto, pode ser entendida como efeito do paradigma dominante, ou paradigma vigente, resultado de um conjunto de ideias, conceitos e práticas aceitas e amplamente utilizadas. Esse paradigma promove uma profunda articulação do pensamento científico, ao delinear e limitar aquilo que é ou não aceito como conhecimento científico.

Desenvolvimento

As formas de sustentação do paradigma dominante passam a manifestar-se enquanto tecnologias que buscam se aprimorar e manter as estruturas científicas tidas como legítimas. Neste cenário, o esquecimento passa a ser uma estratégia de manutenção de um

pensamento dominante e o silenciamento de outros conhecimentos tão pertinentes quanto os que estão sendo preservados. Dentro da Ciência da Informação, é possível observar que as metodologias e ferramentas que sustentam a Organização e a Representação do Conhecimento (OC e RC) (BRASCHER e CAFÉ, 2008) estão propensas a amparar práticas do paradigma dominante, como forma de excluir conhecimentos não enquadrados nos processos de classificação, catalogação e indexação, por exemplo (GÁRCIA GUTIÉRREZ, 2006, 2007, 2008, 2011).

Esta pesquisa enquadra-se no campo da reflexão sobre os processos de análise, pesquisa terminológica, tratamento documental associados à organização do conhecimento e sustenta-se na convicção de que uma mudança de olhar pode ser obtida na abordagem aqui denominada “fator esquecimento”. Esse novo olhar é necessário frente às lacunas encontradas na atividade de indexação.

O emprego da expressão esquecimento é associado a uma multiplicidade de acepções, explicitadas ao longo da fundamentação teórica (ALMEIDA, SANDANHA e RIBEIRO, 2016; ALVES e BUENO, 2017; BERGSON, 1999; HALBWACHS, 2013; HAUSSMAN, 2006; IZQUIERDO, 2002, 2004; MILNER, 2017; PEIRCE, 1995; PERGHER e STEIN, 2003; POLLAK, 1989; RICOEUR, 2007; SELIGMANN-SILVA, 2003; WEINRICH, 2001; YERUSHALMI, 2017) e da descrição dos procedimentos metodológicos. Tem-se, por objetivo, analisar a participação e a relevância do fator esquecimento nos processos de organização e representação do conhecimento, bem como o ajuste de uma abordagem de caráter normativo para a lógica da conduta dentro destes processos; busca-se oferecer uma contribuição com práticas inclusivas de conhecimentos alternativos dentro do processo de Organização e Representação do Conhecimento.

A pesquisa é básica, de natureza exploratória, em um método hipotético-dedutivo, com dados obtidos do levantamento em produções na área da Ciência da Informação e da Comunicação. Enquanto pesquisa bibliográfica, propõe-se a um levantamento narrativo bibliográfico como forma de investigar a temática, visando a um diálogo interdisciplinar. O conto de Jorge Luis Borges (Funes, o memorioso) fornece a metáfora pela qual se vislumbra o esquecimento como componente cognitivo evidenciado em perdas ocasionadas pelas lacunas terminológicas. A fundamentação principal é encontrada no conceito semiótico de objeto imediato e objeto dinâmico.

Resultados Parciais ou Finais

A tese a ser defendida é de que, dentro dos parâmetros atuais envolvendo os processos de OC e RC, observam-se características excludentes pelo emprego de tecnologia em serviço da manutenção do paradigma dominante. A alternativa que vem sendo explorada para o aspecto danoso da perda de conhecimento é a da inserção do conceito de tecnoinformação, proposto por Vilém Flusser (1983, 2008), referindo-se à relação entre tecnologia e imagem. As imagens são produtos da tecnologia e, ao mesmo tempo, moldam a percepção do indivíduo em relação ao mundo. A tecnoinformação entraria como uma ferramenta de Organização e Representação do Conhecimento que, associado às Inteligências Artificiais (IAs), cada vez mais incidentes no processo de construção do conhecimento, resultam em processo dinâmico que envolve a criação, transmissão e recepção das imagens a fim de estruturar todo o conhecimento. Nessa proposta, a construção de tecnoinformações por meio das IAs são entendidas como forma de auxiliar a OC e a RC.

Conclusão

As inteligências artificiais (IAs) podem produzir tecnoinformações por meio da análise de dados e do uso de algoritmos que permitem a criação de imagens a partir de padrões identificados em conjuntos de dados. Esse processo é conhecido como "aprendizado de máquina" (machine learning) e envolve a programação das IAs para reconhecer padrões em dados de entrada, e produzir saídas correspondentes. Por exemplo, uma IA pode ser treinada para reconhecer imagens de um determinado objeto a partir de uma grande quantidade de fotos desse objeto. Com base nesse treinamento, a IA pode produzir novas imagens do objeto a partir de dados de entrada, como descrições textuais ou outras imagens. No entanto, vale destacar que a produção de tecnoinformações por meio de IAs não se dá de forma completamente autônoma, mas sim por meio de processos que envolvem a intervenção humana, seja no treinamento das IAs ou na avaliação e seleção das

saídas geradas. Além disso, a produção de tecnoinformações, por meio delas, enseja questões éticas, como a responsabilidade por eventuais vieses ou distorções presentes nos dados utilizados no treinamento das IAs.

Referências

- ALMEIDA, T.; SALDANHA, G. S.; RIBEIRO, A. R. P.. Memória, esquecimento e recuperação da informação: contradição e dialética da práxis na organização do conhecimento. *Anais...17º Encontro Nacional de Pesquisa em Ciência*
- ALVES, M. V.; BUENO, O.. **Interferência Retroativa: O Esquecimento como uma Interrupção na Consolidação da Memória**. Temas em Psicologia, n. 25, 2017.
- BERGSON, H. **Matéria e memória**: ensaio sobre a relação do corpo com o espírito. São Paulo: Martins Fontes, 1999.
- BRASCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento?. **In**: Encontro Nacional de Pesquisa em Ciência da Informação, 9, 2008, São Paulo. *Anais...* São Paulo: Universidade Estadual de São Paulo, 2008.
- BURKE, P. **Uma história social do conhecimento 2: Da Enciclopédia ao Wikipédia**. 1ª ed. São Paulo: Zahar, 2012.
- BURKE, P. **Uma história social do conhecimento**: de Gutenberg a Diderot. Rio de Janeiro: Jorge Zahar Ed., 2003.
- da Informação: Bahia, ENANCIB, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/190734>. Acesso em: 18 jan. 2023.
- FLUSSER, V. **O universo das imagens técnicas**: Elogio da superficialidade. São Paulo, Annablume, 2008.
- FLUSSER, V. **Pós-história**: Vinte instantâneos e um modo de usar. São Paulo: Livraria Duas Cidades, 1983.
- GARCÍA GUTIÉRREZ, A. Cientificamente favelados: uma visão crítica do conhecimento a partir da epistemografia. **Transinformação**, Campinas, v.18, n.2, p.103-112, maio/ago. 2006.
- GARCÍA GUTIÉRREZ, A. **Desclasificados**: pluralismo lógico y violencia de la clasificación. Barcelona: Anthonopos, 2007.
- GARCÍA GUTIÉRREZ, A. **Epistemología de la documentación**. Barcelona: Stonberg, 2011.
- GARCÍA GUTIÉRREZ, A. **Outra memória é possível**: estratégias descolonizadoras do arquivo mundial. Petrópolis: Vozes, 2008.
- HALBWACHS, M. **A memória coletiva**. 2ª ed. São Paulo: Centauro, 2013.
- HAUSSMAN, C. R. Peirce's Semeiotic Applied to Perception: The Role of Dynamic Objects and Percepts in Perceptual Interpretation. **Cognitio**, São Paulo, v. 7, n. 2, p. 231-246, jul./dez., 2006.
- IZQUIERDO, I. **A arte de esquecer**. São Paulo: Vieira & Lent, 2004.

IZQUIERDO, I. **Memória**. Porto Alegre: ARTMED, 2002.

MILNER, JC. O material do esquecimento. In: YERUSHALMI, Y. et. al. **Usos do Esquecimento**: Conferências proferidas no colóquio de Royaumont. Campinas: Ed. Unicamp. 2017, p. 81-98.

PEIRCE, C. S. **Semiótica**. Trad. José Teixeira Coelho Neto. 2 ed. São Paulo: Perspectiva, 1995.

PERGHER, G. K; STEIN, L. M. Compreendendo o esquecimento: teorias clássicas e seus fundamentos experimentais. **Psicologia USP**, 14, 2003. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-65642003000100008&Ing=en&nrm=iso&tIng=pt. Acesso em: 14 ago. 2021.

POLLAK, M. Memória, esquecimento, silêncio. In: **Estudos Históricos**, 2, v.3. Rio de Janeiro, 1989.

RICOEUR, P. **A memória, a história, o esquecimento**. Campinas: Editora da Unicamp, 2007.

SANTOS, B. de S. Para além do pensamento abissal: das linhas globais a uma ecologia de saberes. **Revista Crítica de Ciências Sociais**, n. 78, p. 3-46, out. 2007. Disponível em: https://www.ces.uc.pt/myces/UserFiles/livros/147_Para%20alem%20do%20pensamento%20abissal_RCCS78.pdf. Acesso em: 05 fev. 2021

SANTOS, B. S. **A construção multicultural da igualdade e da diferença**. 1999. Disponível em: <https://estudogeral.uc.pt/bitstream/10316/11020/1/A%20Constru%C3%A7%C3%A3o%20Multicultural%20da%20Igualdade%20e%20da%20Diferen%C3%A7a.pdf>. Acesso em: 21 jan. 2022.

SANTOS, B. S. **O fim do império cognitivo**: a afirmação das epistemologias do Sul. Belo Horizonte: Autêntica Editora, 2019.

SANTOS, B. S. Para além do pensamento abissal: das linhas globais a uma ecologia dos saberes. In: SANTOS, B. S.; MENESES, M. P. **Epistemologias do Sul**. São Paulo: Cortez, 2010b, p. 31-83.

SANTOS, B. S. **Um discurso sobre as ciências**. 7. ed. São Paulo: Cortez, 2010a

SELIGMANN-SILVA, M. (org.). **História, Memória, Literatura**. O testemunho na era das catástrofes. Campinas, Editora da UNICAMP, 2003.

WEINRICH, H. **Lete**: Arte e crítica do esquecimento. Rio de Janeiro: Civilização Brasileira, 2001.

YERUSHALMI, Y. et. al. **Usos do Esquecimento**: Conferências proferidas no colóquio de Royaumont. Campinas: Ed. Unicamp. 2017.

Uma história da informação e seus efeitos éticos e epistemológicos

Doutorando: Duanne de Oliveira Ribeiro (1)

Orientador: Marivalde Moacir Francelin

(1) Universidade de São Paulo, PPGCI/USP, duanne.ribeiro@usp.br

Introdução

A partir da perspectiva e dos métodos da epistemologia histórica – corrente que abrange a história das ciências de Canguilhem e a arqueologia e a genealogia de Foucault, entre outros – este projeto pretende construir uma história do conceito ou dos conceitos de informação operantes na Ciência da Informação. Dadas as referências metodológicas a que aludimos, isso implica identificar tanto o momento em que surgem esses conceitos quanto as bases epistemológicas implicadas para que possam funcionar; por outro lado, se põe a tarefa de demonstrar possíveis pontos de ruptura com conceitos que atuavam de maneira análoga antes do seu estabelecimento. Essa análise conceitual, ademais, não se desvincula do social e ético, que condicionam o conceito e são condicionados por ele.

A visão histórico-epistemológica é múltipla, mas é possível defender que ela comporta a recusa de dois pressupostos: o de que existe uma contínua evolução do conhecimento, a qual procederia pela compreensão cada vez maior dos objetos visados; e o de que o objeto a conhecer, embora desconhecido, está dado, fixado, bastaria desvendá-lo. Pelo contrário, trata-se, por um lado, de procurar por linhas de corte no conhecimento, as quais separam formas de entender o mundo de maneira até incomensurável, e, por outro lado, de pensar a realidade como algo captável de modos distintos pelos conceitos. Nossa questão, assim, passa por não assumir a informação como algo estável, conhecido ou desconhecido desde sempre de uma forma ou outra, mas de investigar o *surgimento* de um conceito e de como ele dá acesso à realidade – uma disposição epistemológica e social inexistente antes.

Desenvolvimento

Para perseguir essa hipótese – lembrando, a ideia de que a informação e seu conceito têm data de nascimento e diferem estruturalmente de noções prévias que ocupavam o mesmo espaço de pensamento –, realizaremos uma pesquisa bibliográfica, com revisão de textos sobre a formação da Ciência da Informação e das histórias da informação publicadas

antes – entre elas, as de Capurro (2003), Day (2001) e Roszak (1988). Com base nisso, identificaremos as obras que possam testemunhar momentos de transição nas formas de pensamento, rompimentos entre conceitos. Como indicamos, esse procedimento terá de atentar às origens e consequências sociais do aparecimento do conceito e seus efeitos éticos, tendo esse termo aqui o sentido de formação de sujeitos, modo de se compreender e se portar.

Em específico quanto à Organização da Informação e do Conhecimento – que constituirá um tópico da tese –, as questões que emergem são: quando aquilo que se organiza passou a ser objetivado (ou seja, feito um objeto, disponível a um saber) como informação? Quais os desníveis entre essa concepção e as anteriores, isto é, o que muda quando passamos a organizar não *livros* ou *documentos* (conceitos com sua própria história e determinações) e sim conteúdos informacionais? No mesmo sentido, o que implica assumir uma gradação que vai da informação ao conhecimento, como isso altera a noção de conhecimento?

Pode-se dizer também que a organização do conhecimento surge como tema da pesquisa em outro nível, na medida em que a epistemologia histórica e, em particular, Foucault, se debruçam sobre as estruturas que regem a produção de conhecimentos, os limites do que se pode perceber e dizer em tal época, as regras que delineiam a criação de conceitos. Um diálogo entre essa metodologia e a perspectiva mais localizada da OIC pode ser realizada, como temática interna da tese ou mesmo em produções derivadas dela.

Resultados Parciais ou Finais

O projeto de pesquisa neste momento está sendo preparado para a banca de qualificação. O trabalho até aqui têm sido de estruturação das etapas e tópicos da tese e de aperfeiçoamento da definição do seu objeto e dos seus métodos. Os maiores avanços foram no debate sobre as relações entre Ciência da Informação e perspectiva histórica e sobre a tradição da epistemologia histórica, nisso incluindo uma leitura

detalhada dos procedimentos metodológicos de Michel Foucault. Com isso tudo, esboçamos este sumário preliminar:

- *Quadro teórico de referência* – bases teóricas da tese será desenvolvida; trata de epistemologia e Ciência da Informação, da perspectiva histórica nos estudos epistemológicos na área e introduz a epistemologia histórica e as práticas de Foucault.
- *As histórias da informação anteriores e suas questões em aberto* – estudo da obra de autores que trataram do conceito de informação com enfoque histórico e social.
- *Antepassados dos conceitos de informação da Ciência da Informação* – como o conceito de informação passa a funcionar em espaços antes ocupados por outros conceitos? Os efeitos disso serão analisados, por exemplo tendo como objeto a “troca” de *livro e documento* (usados pela Biblioteconomia e pela Documentação) por *informação*, como ela surge na CI.
- *Informação como aquilo que se organiza: efeitos de um objeto* – exercendo uma investigação semelhante à anterior, interroga-se a, por assim dizer, substituição de conceitos nas atividades de classificação, tratamento e disponibilização do saber, tendo como interesse nuclear a Organização da Informação e do Conhecimento.
- *Genealogia do sujeito informacional: subjetivação e informação* – neste passo, se exploram as correlações entre o conceito de informação e uma ética. Aqui se implica uma história do *sujeito informacional*, noção que demonstra como se conectam o conceito de informação e uma forma de entender ou de modelar a compreensão dos indivíduos.
- *Formas do saber: pensar os conceitos de informação com Foucault* – por fim, retomamos nossos achados e propomos debates sobre a *sociedade informacional*, em diálogo com Foucault, procurando trazer visões críticas do presente.

Conclusão

Acreditamos que esse trabalho possa trazer contribuições epistemológicas, metodológicas e históricas para a Ciência da Informação e, em especial, à Organização da Informação e do Conhecimento.

Contribuição epistemológica, pois analisa condições de possibilidade de conceitos em que essas áreas se sustentam. *Contribuição metodológica*, pois atualiza, no campo da CI, uma vertente da filosofia, da epistemologia, da história das ciências; isto é, a partir dos epistemólogos históricos, deve modelar um processo próprio. *Contribuição histórica*, pois tenta trazer uma nova ótica das trajetórias da CI e da OIC. Tais propósitos estão em construção, e serão ajustados de acordo com os subsídios trazidos pela banca de qualificação, que ocorre neste ano, e por eventos como o da Isko Brasil.

Referências

- CAPURRO, Rafael. The Concept of Information. **Annual Review of Information Science and Technology**, v. 37, cap. 8, p. 343-411, 2003. Disponível em: <http://www.capurro.de/infoconcept.html>. Acesso em: 29 jun. 2022.
- DAY, Ronald. **The modern invention of information: discourse, history, and power**. Carbondale: Southern Illinois University Press, 2001.
- FOUCAULT, Michel. **A arqueologia do saber**. Trad. Luiz Felipe Baeta Neves. Rio de Janeiro: Forense Universitária, 2012.
- HJØRLAND, Birger. Fundamentals of Knowledge Organization. **Knowledge Organization**, v. 30, n. 2, 2003.
- JAPIASSU, Hilton Ferreira. **Introdução ao pensamento epistemológico**. Rio de Janeiro: Francisco Alvez Editora, 1986.
- ROSZAK, Theodore. **O culto da informação**. São Paulo: Brasiliense, 1988.
- SILVA, Thiago José da Silva; FREIRE, Maria Isa. “Introdução”. In: SILVA, Thiago José da Silva; FREIRE, Maria Isa. **Epistemologia e historiografia na ciência da informação**. João Pessoa: Editora UFPB, 2020, p. 9-18. Disponível em: <http://www.editora.ufpb.br/sistema/press5/index.php/UFPB/catalog/view/740/871/6857-1>. Acesso em: 7 jun. 2022.
- SILVA, Thiago José da Silva; FREIRE, Maria Isa. “Tecendo a Rede Conceitual”. In: SILVA, Thiago José da Silva; FREIRE, Maria Isa. **Epistemologia e historiografia na ciência da informação**. João Pessoa: Editora UFPB, 2020, p. 19-62. Disponível em: <http://www.editora.ufpb.br/sistema/press5/index.php/UFPB/catalog/view/740/871/6857-1>. Acesso em: 7 jun. 2022.
- SOUTO, Caio. **O dever de um estilo: estudos de epistemologia histórica e filosofia contemporânea**. São Paulo: Editora Córrego, 2022.