



VIII ISKO BRASIL 2025



REPOSITÓRIOS DE DADOS DE PESQUISA BRASILEIROS À LUZ DA ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO

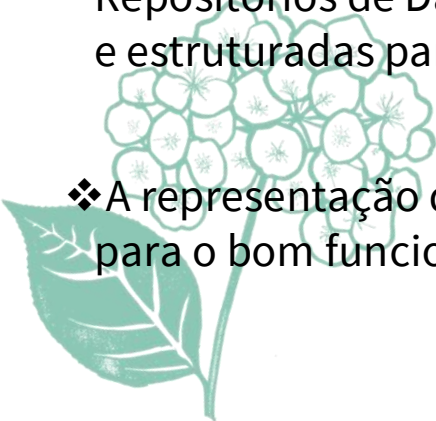


Raimunda Fernanda dos Santos *
Carla Beatriz Marques Felipe *

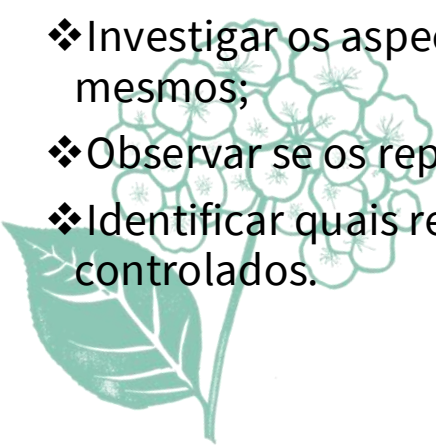
Universidade Federal do Rio Grande do Norte*; raimunda.fernanda@ufrn.br
Universidade Federal do Rio de Janeiro; felipecarla12@gmail.com



- ❖ Os dados de pesquisa desempenham um papel essencial no avanço da ciência, sendo insumos fundamentais para a produção do conhecimento;
- ❖ Os dados são armazenados em ambientes específicos, como os Repositórios de Dados de Pesquisa, que devem garantir informações claras e estruturadas para que o reuso dos dados seja eficaz.
- ❖ A representação da informação, nesse contexto, é um fator determinante para o bom funcionamento desses repositórios



- O presente estudo tem como objetivo geral analisar os Repositórios de Dados de Pesquisa brasileiros à luz das práticas de organização e representação da informação.
- Os objetivos específicos são:
 - ❖ Investigar os aspectos relacionados aos Metadados e o preenchimento dos mesmos;
 - ❖ Observar se os repositórios apresentam diretrizes para indexação;
 - ❖ Identificar quais repositórios Orientação para uso de vocabulários controlados.



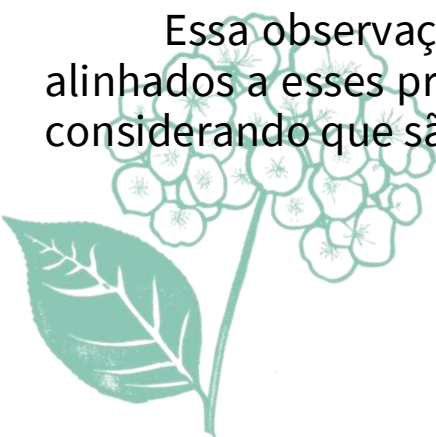
- Quadro1. *Indicadores analisados nos Repositórios de dados de pesquisa*

Indicadores gerais	Indicadores relacionados a representação da informação
Ano de criação	Padrões de metadados
Instituições as quais os repositórios estão vinculados	Diretrizes para preenchimento dos metadados
Temáticas contempladas	Diretrizes para indexação
	Orientação para uso de vocabulários controlados



Ainda como objeto de investigação foram analisados se os metadados presentes nos repositórios analisados estão em consonância com os princípios FAIR (do inglês, *Findable, Accessible, Interoperable, Reusable*).

Essa observação se justifica pelo fato de que, se os mesmos estiverem alinhados a esses princípios, poderão ser reaproveitados de maneira efetiva, considerando que são localizáveis, acessíveis, interoperáveis e reutilizáveis.



Repositórios de Dados de Pesquisa Brasileiros estão vinculados às seguintes instituições: Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT/UFRJ), Universidade Federal do Paraná (UFPR), Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Universidade Federal do Rio Grande do Sul (UFRGS), Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Universidade Estadual de Campinas (UNICAMP), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), Empresa Brasileira de Pesquisa Agropecuária (EMBRAPA) e Fundação Oswaldo Cruz (FIOCRUZ).

Ampla gama de áreas do conhecimento, como ciências naturais, ciências da vida, humanidades, ciências sociais, engenharias e ciências da computação.



Figura 1. *Padrões de metadados*

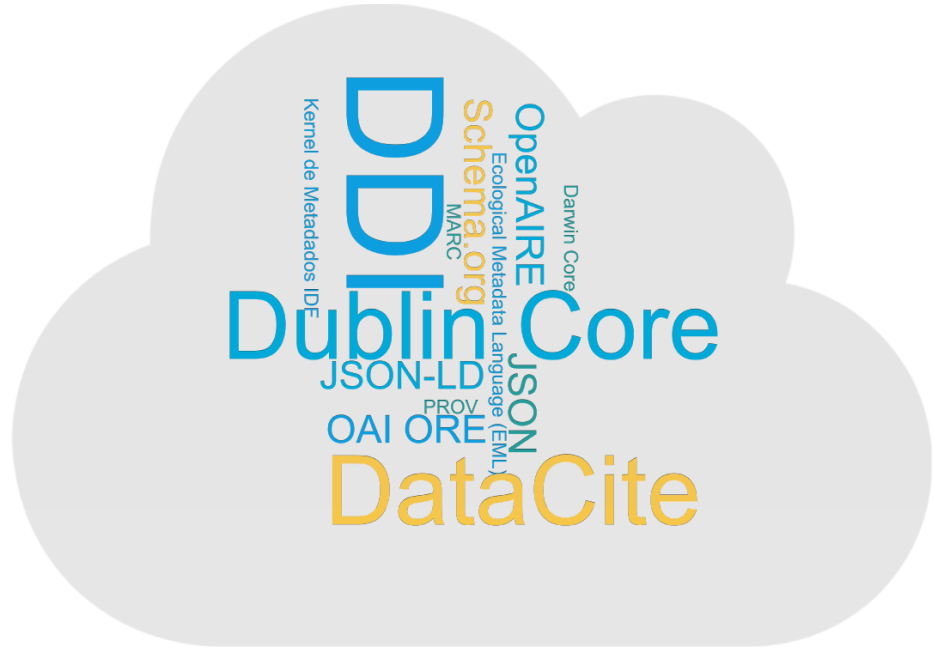
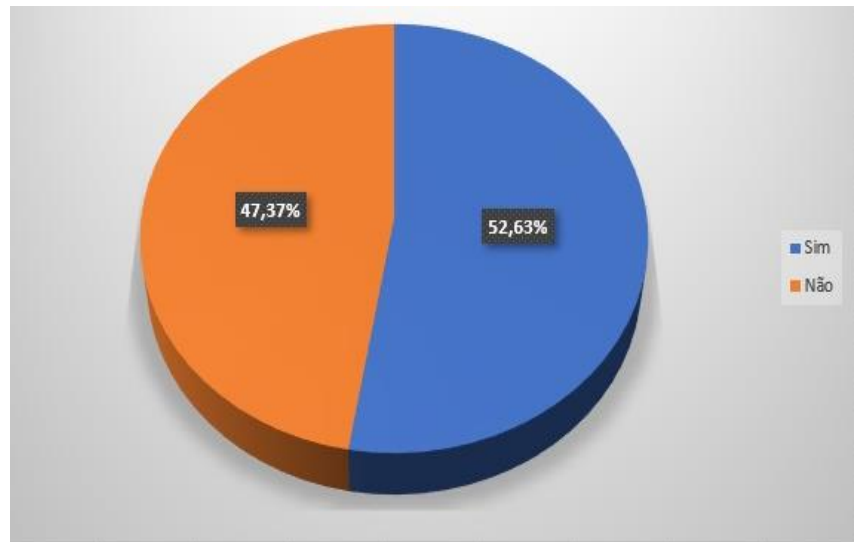


Gráfico 2. *Diretrizes para preenchimento dos metadados*

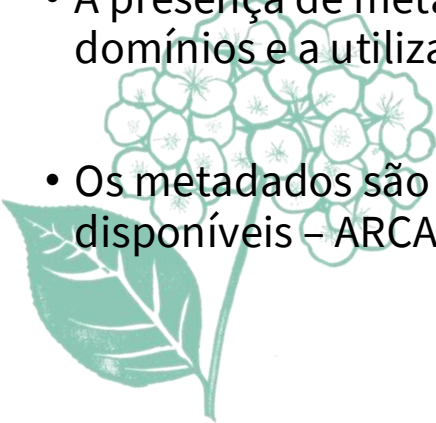


Apenas dois repositórios oferecem diretrizes para indexação: o WorldClim - Global Climate Data e o Repositório de Dados do PPBio. Juntamente com o Programa Internacional de Descoberta do Oceano, esses três repositórios são os únicos a fornecer orientações sobre o uso de vocabulários controlados.

Esses repositórios recomendam a utilização de quatro vocabulários controlados para a indexação: itution_id CMIP6, IODP Depth Scales Terminology, Thesaurus of Geographic Names (TGN) e o vocabulário do PPBio

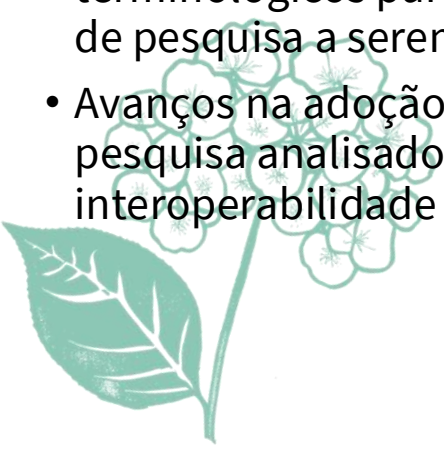


- A análise dos repositórios de dados de pesquisa evidencia avanços significativos na adoção dos princípios FAIR (Findable, Accessible, Interoperable, Reusable), fundamentais para a ciência aberta e a gestão eficiente de dados científicos.
- A presença de metadados bem estruturados e padrões específicos para os domínios e a utilização de identificadores persistentes;
- Os metadados são acessíveis, mesmo quando os dados não estão disponíveis – ARCA Fiocruz;



- Ferramentas de Inteligência Artificial (IA) podem automatizar processos de curadoria e padronização de metadados em repositórios de dados de pesquisa, aprimorando a indexação e a descoberta das informações.
- Permitem aplicar metadados de forma mais eficiente, analisando grandes volumes de dados e identificando padrões e categorias automaticamente.
- Ao empregar vocabulários controlados e ontologias específicas, a IA torna o processo de indexação mais rápido e preciso, facilitando a recuperação da informação e seu reuso por outros pesquisadores.
- Além disso, a IA pode sugerir classificações baseadas em padrões identificados em grandes volumes de dados, reduzindo o tempo de processamento e aumentando a padronização da informação.

- No que tange à representação da informação, a pesquisa identificou que a representação descritiva está bastante avançada, com boas práticas nos repositórios investigados. No entanto, aspectos relacionados à indexação ainda requerem maior atenção por parte dos administradores, sobretudo no que diz respeito ao uso e disponibilidade de instrumento de controle terminológicos para as práticas de representação e recuperação dos dados de pesquisa a serem armazenados em seus ambientes.
- Avanços na adoção dos princípios FAIR nos repositórios de dados de pesquisa analisados, persiste a necessidade de aprimorar a interoperabilidade e a reutilização dos dados;



Por fim, a gestão dos repositórios de dados não deve ser vista como uma responsabilidade exclusiva dos profissionais da Ciência da Computação ou ferramentas de Inteligência Artificial. Pelo contrário, deve envolver também os bibliotecários, que possuem conhecimento especializado em processos, produtos, instrumentos e serviços voltados à organização e representação da informação e do conhecimento, essenciais para garantir uma recuperação eficiente e eficaz da informação.

ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaços digitais. **Enc. Bibli.** R. Eletr. Bibliotecon. Ci. Inf., Florianópolis, v.15, n.1, 2003.

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.

OPENAIRE – 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: 2 ago. 2020.

PAMPEL, Heinz; KINDLING, Maxi. Informationsinfrastrukturangebote für digitale Forschungsdaten. In: SCHIRMBACHER, Peter. **E(hren)-Journal**. Berlin: Angaben gemäß § 5 TMG, 2017. p. 15-33. Disponível em: <https://edoc.hu-berlin.de/bitstream/handle/18452/2993/3.pdf?sequence=1&isAllowed=y>. Acesso em: 15 jun. 2020.

TAYLOR, A. G.; JOURDREY, D. N. **The organization of the information**. 3.ed. Santa Barbara, CA: Libraries Unlimited, 2008. 513p.

