

Aderência dos Conjuntos de Dados dos Repositórios Brasileiros de Dados de Pesquisa Indexados no Diretório re3data aos Princípios FAIR*

Rosane Teles Lins Castilho
Doutoranda PPGCI/IBICT UFRJ BR
ORCID <https://orcid.org/0000-0002-7142-6813>
Email: rosanetlcastilho@gmail.com

Introdução

A reutilização de dados de pesquisa tem recebido da comunidade científica estímulos para a prática e maior incentivo das agências de fomento. Para contribuir para que os dados sejam efetivamente reutilizados, foram propostos como iniciativa os Princípios FAIR, acrônimo de *Findable*, *Accessible*, *Interoperable* e *Reusable*. Adotando os Princípios, os dados de pesquisa gerados e, até então dispersos, passam a ser reunidos e disponibilizados em repositórios digitais, onde podem ser encontrados, coletados e reutilizados na produção de artigos científicos e de outras pesquisas (Sayão & Sales, 2016). E para ajudar o pesquisador a descobrir um repositório de dados compatível com o assunto da sua pesquisa, os diretórios internacionais de dados de pesquisa, como o *Registry of Research Data Repositories*, conhecido como *re3data*, tem sido considerado como fonte relevante para a descoberta.

Objetivo

A presente pesquisa teve por objetivo descobrir e relatar de que forma os conjuntos de dados dos repositórios digitais de dados científicos de instituições do Brasil se ajustam aos Princípios e qual o grau de *FAIRness* (Henning et al., 2021, Simões, Anjos & Dias (2021) aferido aos 12 repositórios que atenderam aos critérios de inclusão.

Metodologia

Pesquisa caracterizada como exploratória e descritiva, bem como qualitativa, qualitativa por analisar aderência dos conjuntos de dados aos Princípios FAIR, e quantitativa para obter a pontuação *FAIRness* de aderência aos princípios. A coleta de dados, realizada no mês de julho de 2023, selecionou 12 repositórios nacionais, cujo URL indicava o código de país BR. E para aferir o grau de *FAIRness* foi adotada ferramenta “*Self-Assessment Tool to Improve the FAIRness of Your Dataset*” – *SATIFYD*, disponível na forma de um questionário composto por 12 questões e dividido em 4 seções, que são as próprias letras do acrônimo FAIR, e permite uma autoavaliação prévia *on-line* dos dados pelo depositante, bem como uma pós-avaliação feita por outros.



Fonte: Template SATIFYD (2023)

Resultados

Os resultados da aferição da aderência dos 12 repositórios aos Princípios FAIR são apresentados nas tabelas 1 a 3, agrupados pelas pontuações do índice de *FAIRness*, classificados como Muito Alto, Alto e Mediano, conforme a semelhança das pontuações.

Tabela 1 – *FAIRness* Muito Alto

REPOSITÓRIO	Findable	Accessible	Interoperable	Reusable	FAIRNESS
1 ARCA Dados	100%	100%	100%	93%	98%
10 PPBIO INPA	100%	100%	92%	80%	98%
11 REDAPE EMBRAPA	100%	100%	100%	93%	98%
12 REDU UNICAMP	100%	100%	100%	93%	98%
14 SCIELO DATA	100%	100%	92%	93%	98%

Tabela 2 – *FAIRness* Alto

REPOSITÓRIO	FINDABLE	ACCESSIBLE	INTEROPERABLE	REUSEABLE	FAIRNESS
2 BDEP ANP	89%	50%	75%	61%	72%
5 DADOS I.F. GOIANO	67%	55%	92%	87%	78%
8 MAENDUAR LARH/UD/UFRJ	55%	55%	92%	72%	80%

Tabela 3 – *FAIRness* Mediano

REPOSITÓRIO	FINDABLE	ACCESSIBLE	INTEROPERABLE	REUSEABLE	FAIRNESS
4 CEDAP UFRGS	67%	55%	58%	56%	59%
6 FAPESP COVID-19 Data Sharing/BR	44%	55%	73%	74%	61%
9 OPEN RESEARCH DATA PUC-RIO	60%	55%	58%	77%	63%
13 INSTITUCIONAL UNESP	66%	55%	58%	74%	63%

Os cinco repositórios da Tabela 1 atingiram a pontuação de 100% em *Findable* e *Accessible*, princípios considerados essenciais para, já de início, o repositório conquistar a credibilidade do usuário. São pioneiros na implementação da filosofia e das práticas da Ciência Aberta, no Brasil, bem como dos princípios FAIR, além de difundirem a adoção de planos de gestão dados para prover repositórios de metadados ricos. Os repositórios da Tabela 2 demonstram preocupação quanto à curadoria dos dados, porém, carecem de informações e documentação ao usuário, por exemplo, sobre políticas e guias técnicos e de uso. Os repositórios da Tabela 3, contam com coleções de dados de pesquisa heterogêneos e de diversas áreas, o que pode explicar as dificuldades de atenderem às demandas de tratamento dos dados mais adequado e fornecimento de meta(dados) mais detalhados.

Referências

- Henning, P. et al. (2021) The FAIRness of data management plans: an assessment of some European DMPs. *Reciis – Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, Rio de Janeiro, v. 15 (3), 722-735, jul.-set. <https://www.arca.fiocruz.br/handle/icict/49019>
- Sayão, L. F. & Sales, L. F. (2016). Algumas considerações sobre os repositórios digitais de dados de pesquisa. *Informação & Informação*, v. 21 (2) 90–115, maio/ago.
- Simões, A. R. C., Anjos, R. L., & Dias, G. (2021) Análise dos conjuntos de dados disponíveis no repositório COVID-19 Data Sharing/BR à luz dos princípios FAIR. In: L. F. SALES, V. S., V. VEIGA, P. HENNING, P., L. F. SAYÃO, L. F. (Orgs.). *Princípios FAIR aplicados à gestão de dados de pesquisa*. IBICT. p. 91-102.

* Trabalho realizado como apoio da Coordenação de Aperfeiçoamento do Pessoal de Nível Superior – CAPES e produzido no âmbito do Grupo BRIET <https://grupobriet.com/>