

IDENTIFICANDO REPOSITÓRIOS DE DADOS NO BRASIL

Igor Yure Ramos Matos

Mestrando em Gestão da Informação PPGInfo/UDESC
Bibliotecário na Universidade Federal de Santa Catarina (UFSC)
<http://lattes.cnpq.br/4570351139474814>

Divino Ignácio Ribeiro Júnior

Professor do Programa de Pós-Graduação em Gestão da Informação da
Universidade do Estado de Santa Catarina (UDESC)
Doutor em Engenharia e Gestão do Conhecimento (EGC/UFSC)
<http://lattes.cnpq.br/1599118875018269>

INTRODUÇÃO

Os repositórios de dados são infraestruturas recentes utilizadas para armazenar dados científicos. Os dados de pesquisa são elementos desprovidos de significado e coletados ao longo da pesquisa científica. Silva (2019, p. 21) dados científicos são todas as evidências que um investigador necessita para validar suas conclusões após uma pesquisa. Para Sanchez, Vechiato e Vidotti (2019, p. 52) “[...] os repositórios dados, ambientes

informacionais digitais que buscam armazenar, organizar, representar, prover acesso, disseminar e preservar dados oriundos de pesquisas científicas”.

OBJETIVO

Identificar repositórios de dados no Brasil que estão em funcionamento.

METODOLOGIA

A metodologia consistiu em pesquisar no: 1) *Registry of Research Data Repositories Initiative* – re3data.org; 2) www.dataverse.org; 3) no metabuscador de dados de pesquisa – <https://metabuscador.uspdigital.usp.br/>. Identificou-se 15 (quinze) repositórios de dados no País. As buscas foram realizadas na segunda quinzena de fevereiro de 2021.

No re3data foi selecionado somente repositórios brasileiros e descartados repositórios internacionais que marcavam

parceria com o Brasil. No site do Dataverse encontrou-se seis instalações no Brasil, porém uma instalação o estava com link corrompido, assim não foi analisada. O Repositório do ITA mesmo fazendo parte do metabuscador da USP não possui itens armazenados e assim não tem link para o repositório, ficando de fora das análises.

ANÁLISE

| | Repositório de Dados | Software | Ano | Identificador persistente |
|----|---|--------------------------|---------|---------------------------|
| 1 | Base de Dados Científicos da Universidade Federal do Paraná | DSPACE | 2017 | DCI |
| 2 | Embrapa Informática Agropecuária (CNPTIA/Embrapa)* | DSPACE | [2019?] | Handle |
| 3 | Open Research Data @PUC-Rio Project | Maxwell (software Local) | 2015 | DCI |
| 4 | Repositório Cariniana IBICT | DATAVERSE | 2015 | DCI |
| 5 | Repositório COVID-19 Data Sharing/BR FAPESP | DSPACE | 2020 | Handle |
| 6 | Repositório de Dados Científicos da USP* | DSPACE | 2018 | Handle |
| 7 | Repositório de Dados da Unicamp (Redu)* | DATAVERSE | 2019 | DCI |
| 8 | Repositório de Dados da Universidade Federal do ABC* | DATAVERSE | 2018 | DCI |
| 9 | Repositório de Dados de Pesquisa do CEDAP - dados de pesquisa (UFRGS) | DSPACE | 2017 | Handle |
| 10 | Repositório de Dados de Pesquisas do Instituto Federal Goiano – Campus Urutai | DATAVERSE | 2020 | DCI |
| 11 | Repositório do CIDACS - Centro de Integração de Dados e Conhecimentos para Saúde / FIOCRUZ-BA | DATAVERSE | 2020 | Handle |
| 12 | Repositório Institucional da Universidade Estadual Paulista (Unesp) | DSPACE | 2019 | Handle |
| 13 | Repositório Institucional de Dados de Pesquisa da RNP (projeto Piloto) | DATAVERSE | 2019 | Handle |
| 14 | Universidade Federal de São Carlos (UFSCar)* | DSPACE | 2019 | Handle |
| 15 | Universidade Federal de São Paulo (Unifesp)* | DATAVERSE | 2019 | DCI |

Os Repositórios de Dados começaram a ser instalados no Brasil em 2015, tendo o ano de 2019 como o ano que mais teve iniciativas. A análise desses repositórios verificou-se que 100% dos repositórios de dados utilizam as licenças Creative Commons e identificadores persistentes DOI (47%) ou Handle (53%). Porém poucos possuem uma política de Gestão de Dados conforme Sayão e Sales (2015) orientam no Guia dos repositórios de dados. Em relação ao Software utilizado 47% utilizam o DSPACE; 46% utilizam o DATAVERSE e 7% (uma instituição) utiliza um software local. As iniciativas dos repositórios de dados estão 33% em centros de pesquisa ou institutos; 60% em universidades públicas e somente 7% em universidade particular.

CONSIDERAÇÕES FINAIS

Os repositórios no Brasil não estão bem disseminados e necessita de conhecimentos técnicos sobre Gestão de Dados, ciclo de vida dos dados e de como elaborar um Plano de Gestão de Dados (PGD). Aos bibliotecários é necessário

habilidades e capacitação para lidar com essa nova demanda que são os repositórios de dados. As iniciativas pertencem em sua maioria ao setor público.

REFERÊNCIAS

- FUNDAÇÃO DE AMPARO À PESQUISA DE SÃO PAULO. Metabuscador de dados de pesquisa. São Paulo: Fapesp, [2019?]. Disponível em: <https://metabuscador.uspdigital.usp.br/>. Acesso em: 15 fev. 2021.
- RE3DATA. Registry of Research Data Repositories. **Re3data.org Reaches a Milestone & Begins Offering Badges**. Apr. 2016. Disponível em: [https://www.re3data.org/search?query=&countries\[\]=BRA](https://www.re3data.org/search?query=&countries[]=BRA). Acesso em: 15 fev. 2021.
- SILVA, Fabiano Couto Corrêa da. **Gestão de dados científicos**. Rio de Janeiro: Interciência, 2019. 128 p.
- SANCHEZ, Fernanda Alves; VECHIATO, Fernando Luiz; VIDOTTI, Silvana Aparecida Borsetti Gregório. Encontrabilidade da informação em repositórios de dados: uma análise do DataONE. **Informação & Informação**, Londrina, v. 24, n. 1, p. 51-79, jan./abr. 2019. ISSN 1981-8920. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/30725>. Acesso em: 26 nov. 2020. DOI: <http://dx.doi.org/10.5433/1981-8920.2019v24n1p51>.