

REPOSITÓRIOS DE DADOS DE PESQUISA: CRITÉRIOS *CORETRUSTSEAL*

Lyvia Rocha de Jesus Araujo¹, Eloísa Príncipe², Maria Simone de Menezes Alencar³

^{1, 2} Programa de Pós-Graduação em Ciência da Informação, convênio Ibict/UFRJ

³ Programa de Pós-Graduação em Enfermagem e Biociências (PPGENFBIO/UNIRIO)

Introdução

- Ciência Aberta:
 - ▶ Novas ações: práticas abertas, transparentes, colaborativas e inclusivas
 - ▶ Destacam-se o compartilhamento e o reuso de dados de pesquisa.

Introdução

- Compartilhamento e reuso de dados:
 - ▶ Surgimento dos repositórios de dados e *data journals*;
 - ▶ *Data papers/journals* raramente solicitam ou fornecem aos autores informações sobre os repositórios em que os dados podem ser compartilhados.

Kim (2020)

Introdução

- Repositórios de dados:
 - ▶ Mecanismos para certificar a qualidade de repositórios;



Objetivo

- Investigar a dimensão da compatibilidade dos critérios definidos pela atualização 2020-2022 do certificado com a descrição e as políticas dos repositórios de dados brasileiros, indexados no Re3data.



Materiais e Métodos

- Repositórios selecionados em março de 2022;
- Dados organizados no *Google Sheets*;
- Síntese comparativa dos repositórios.

- Avaliação feita de acordo com os 16 critérios do *CoreTrustSeal* 2020-2022:

- 1) Missão e escopo;
- 2) Licenças;
- 3) Continuidade do acesso;
- 4) Ética e confiabilidade;
- 5) Infraestrutura;
- 6) Orientação de especialistas;
- 7) Integridade e confidencialidade dos dados;
- 8) Avaliações/Feedbacks;
- 9) Procedimentos documentados de armazenamento;
- 10) Plano de preservação;
- 11) Qualidade dos dados;
- 12) Fluxo de trabalho;
- 13) Descoberta e identificação dos dados;
- 14) Reuso dos dados;
- 15) Infraestrutura técnica;
- 16) Segurança.



- Níveis de conformidade com os critérios

Níveis	Significado adotado nesta pesquisa
0	o critério não é citado na descrição/documentação do repositório
1	o critério é citado genericamente na documentação, porém não é observado no repositório
2	o repositório aponta políticas e definições relacionadas, mas não declara o procedimento específico adotado
3	o repositório declara que irá implementar o critério
4	o repositório especifica e aplica o critério na descrição do site/documentação fornecida

Principais resultados

13

Repositórios com vínculo institucional com uma ou mais instituições brasileiras

9

Repositórios avaliados

Exclusões: Programas ou Grupos de colaboração, que não se classificam exclusivamente como repositórios de dados (2); repositórios com links inválidos(1); e repositórios ainda em construção(1)

Principais resultados

Repositórios Disciplinares

Programa de Pesquisa em Biodiversidade

CENBAM PPBio

PORTUGUÊS
ENGLISH

HOME ABOUT US SURVEY SITES REPOSITORY PUBLICATIONS EXTENSION METHODS CENBAM DOCUMENTS BIOTECAS USEFUL LINKS

Data Imagens Vídeos

Home

Repository

The Repository contains data on PPBio Biological Surveys as well as Pictures and Videos. Use of the Biological Surveys data follow the same policy as use of the PPBio database. There are three Repositories, containing data, images and videos & sound files. All have been generated by research conducted in the Program study sites or by researchers associated with the program. All files are available for consultations and download, and can be used for educational purposes provided that the original authors are credited.

- [How to cite](#)
- [Data repository](#)
- [Image repository](#)
- [Video & sounds repository](#)
- [Data use policies](#)

COVID-19 Data Sharing/BR

Banco de dados compartilhados FAPESP

FAPESP USP Grupo Fleury SÍRIO-LIBANÊS ALBERT EINSTEIN HC FMUSP bp

Página inicial

FAPESP COVID-19 DataSharing/BR

INTRODUÇÃO

O repositório COVID-19 Data Sharing/BR é uma iniciativa da FAPESP em cooperação com a Universidade de São Paulo, e, inicialmente, participação do Instituto Fleury, Hospital Sírio-Libanês e Hospital Israelita Albert Einstein, com o objetivo de disponibilizar dados relacionadas à COVID-19 que possam contribuir para pesquisas desta temática.

Os dados disponibilizados apresentam duas categorias de informações: dados demográficos (gênero, ano de nascimento e região em que o paciente reside), dados de exames clínicos e/ou laboratoriais, além de informações, quando disponíveis, sobre a movimentação do paciente, como internações, por exemplo, e desfecho dos casos, como recuperação e óbito.

Para mais informações, por favor, veja o artigo <https://doi.org/10.5281/zenodo.3966427>, que descreve como o repositório foi concebido e criado.

O repositório é atualizado periodicamente, devendo ser verificado de tempos em tempos para baixar novos dados. Seus dados são fornecidos já pseudonimizados pelas instituições de saúde participantes, que não se responsabilizam pelo uso indevido dos dados.

DISPONIBILIDADE DOS DADOS E REGRAS DE CITAÇÃO

Todos os dados do repositório são abertos e disponíveis para download, sem qualquer custo para quem baixá-los.

O repositório e seus dados adotam a licença CC-BY de dados abertos. Toda e qualquer publicação ou apresentação que usar total ou parcialmente os dados do repositório deve citar a página web do FAPESP COVID-19 Data Sharing/BR. No caso de textos, a página deve ser citada como uma referência bibliográfica, conforme definido em: <https://repositoriodatasharingfapesp.uspdigital.usp.br/static/images/refPT.html>

Buscar no Repositório

Navegar

Todo o repositório

- [Comunidades e Coleções](#)
- [Por data do documento](#)
- [Autores](#)
- [Títulos](#)
- [Assuntos](#)

Discover

Autor

- [Albert Einstein, Hospital Israelita \(1\)](#)
- [BP, A Beneficência Portuguesa de São Paulo \(1\)](#)
- [Faculdade de Medicina da Universidade de São Paulo, Hospital](#)



Principais resultados

Repositórios Institucionais



Principais resultados

- Repositórios com escopo e missão bem definidos;
- Sobre o financiamento estável de instituições de ensino e pesquisa confiáveis;
- Sistema de gestão de dados que assegura sua integridade;
- Exigem informações suficientes para comprovar a qualidade dos dados armazenados;
- Desenvolvidos sobre uma infraestrutura estável que garante sua permanência online a longo prazo;
- Protegidos pelas políticas de segurança dos sistemas *Dspace*, *Dataverse* ou *Ckan*.

Principais resultados

Repositórios	Nível de conformidade
<u>SciELO Data</u>	60
<u>FAPESP COVID-19 Data Sharing/BR</u>	54
<u>Open Research Data @PUC-Rio</u>	50
<u>PPBio Data Repository</u>	46
<u>IBICT Cariniana Dataverse Network</u>	42
<u>UNESP Institutional Repository</u>	40
<u>Base de Dados Científicos da UFPR</u>	35
<u>CEDAP Research Data Repository</u>	34
<u>Dados Abertos De Pesquisas</u>	32

CONSIDERAÇÕES FINAIS

- Limitações: Materiais bibliográficos e descritivos;
- Nenhum dos repositórios consultados possui o selo *CoreTrustSeal*;
- Desde 2018, vem sendo cobrada uma taxa administrativa de € 1.000,00, equivalente à R\$5.017,17 para certificação;
- Sugere-se o aprofundamento desta pesquisa, a consulta direta às instituições responsáveis pelos repositórios;
- Ressalta-se a ideia de que a descrição mais esclarecida dos repositórios de dados pode aumentar a garantia da qualidade dos dados de pesquisa publicados em revistas híbridas ou *data journals*.

REFERÊNCIAS

- ▶ **CORETRUSTSEAL TRUSTWORTHY DATA REPOSITORIES REQUIREMENTS 2020–2022.** Zenodo, 2019. Disponível em: <https://zenodo.org/record/3638211#.YltNEXXMK00>. Acesso em: 20 set. 2022.
- ▶ KIM, Jihyun. An Analysis of Data Paper Templates and Guidelines: Types of Contextual Information Described by Data Journals. **Science Editing**, vol. 7, n. 1, p. 16–23, 2020. Disponível em: <https://doi.org/10.6087/kcse.185>. Acesso em: 20 set. 2022.
- ▶ LIN, D. *et al.* The TRUST Principles for digital repositories. **Sci Data**, vol. 7, n. 144, 2020. Disponível em: <https://doi.org/10.1038/s41597-020-0486-7>. Acesso em: 20 set. 2022.
- ▶ **MAPPING THE GLOBAL REPOSITORY LANDSCAPE.** 2022. Disponível em: <https://coref.project.re3data.org/blog/mapping-the-global-repository-landscape>. Acesso em: 20 set. 2022.

OBRIGADA!

