



Guia de preparação de dados de pesquisa

Agosto de 2021

A adoção de boas práticas na preparação de dados para depósito é um aspecto de fundamental importância para o acesso, compartilhamento e reuso dos dados de pesquisa, além de contribuir para que eles seja *as FAIR as possible*.¹

Para melhor entender os Princípios FAIR e como tornar os seus dados o mais FAIR possível recomendamos o uso da ferramenta FAIR-Aware: <https://fairaware.dans.knaw.nl/>.

1. Nomeação dos arquivos

A adoção de boas práticas para nomeação de arquivos evita que estes sejam sobrescritos e facilita a localização e o reuso por outros pesquisadores.

Boas práticas:

- Use nomes significativos / descritivos, com no máximo 100 caracteres;
- Não use acentuação;
- Use caracteres alfanuméricos, underline (meus_dados) ou hífen (meus-dados);
- Evite espaços (meus dados), pontos (meus.dados), iniciais maiúsculas (MeusDados) e caracteres especiais (tais como \ / ? : * " > < | : # % " { } | ^ ` ~ @ & ; ° æ Æ ø Ø å Å ä Ä ö Ö);
- Use o formato YYYY-MM-DD (meus_dados_2021-01-07) ou YYYYMMDD (meus_dados_20210107) para datas;
- Inclua o número da versão na nomeação quando apropriado (meus-dados001.csv, meus-dados002.csv, ..., meus-dados010.csv, ...);
- Use a mesma nomeação para arquivos com o mesmo conteúdo, mas em formatos diferentes (meus-dados.doc e meus-dados.txt).

2. Formato dos arquivos

O SciELO Data aceita qualquer tipo de arquivo, porém é recomendado o depósito de formatos:

- Não proprietários;
- Abertos, com padrões internacionais documentados;
- Com codificação de caracteres padrão, de preferência Unicode (por exemplo, UTF-8).

Tipo de documento	Formatos recomendados	Formatos não recomendados (mas aceitos)
Análises estatísticas	R (.r, .rdata) SPSS (.dat/.sps) STATA (.dat/.do)	SPSS Portable (.por) SPSS (.sav)

¹ Recomendamos o webinar "Cómo hacer que los datos sean FAIR? buenas prácticas para datos (abiertos) de investigación" disponível em: <https://www.youtube.com/watch?v=114SwZxIRHY>.

Arquivos compactados	.zip*	.rar
Dados tabulares	Comma Separated Values (.csv) Arquivo de texto (.txt)	Excel (.xlsx)
Dados textuais	Arquivo de texto (.txt) OpenOffice (.odt, .ods ou .odp) PDF (.pdf)	Word (.doc ou .docx)

*Arquivos compactados com a extensão .zip serão descompactados após upload dos arquivos.

3. Tamanho dos arquivos

O tamanho limite dos arquivos individuais é 2GB. Para adicionar arquivos acima deste limite entre em contato com data@scielo.org.

4. Descrição dos dados

Para ajudar a garantir que os dados depositados possam ser interpretados corretamente e reutilizados, tanto por você em uma data posterior quanto por outros pesquisadores, é essencial que eles sejam descritos de forma mais detalhada e compreensível possível.

Essa descrição detalhada deve ser fornecida através do preenchimento dos campos durante o depósito e de um arquivo README, que funciona como um guia para o usuário, e deverá ser depositado junto aos arquivos de dados.

O arquivo README deve ser escrito como um texto simples com codificação de caracteres Unicode UTF-8 (.txt) ou como PDF se for necessário ilustrar ou formatar a descrição dos dados.

O arquivo README deve conter pelo menos as seguintes informações:

- Título do conjunto de dados;
- Informações de contato (nome, instituição e e-mail) do pesquisador correspondente / principal ou da pessoa responsável pela coleta dos dados;
- Data de coleta dos dados (data única ou intervalo de tempo);
- Visão geral dos dados e arquivos (breve descrição dos dados que cada arquivo contém, data de criação de cada arquivo e como eles se relacionam entre si, etc);
- Descrição dos métodos de coleta ou geração dos dados;
- Descrição dos métodos usados para o processamento dos dados;
- Informações específicas dos dados (lista de variáveis, unidades de medida, definições de códigos ou símbolos, calibragem de equipamentos, etc).

Para exemplos de outras informações que podem ser acrescentadas ao arquivo README consultar:

- [Guide to writing "readme" style metadata](#)

Exemplos de templates para o arquivo README:

- <https://drive.google.com/file/d/167cJdaRy4sxQWA5qEEp2cgfWqKV-smZV/view>

- https://drive.google.com/file/d/1VIDF489DDr044Uta8z1G7EuLj-Wm_UtG/view
- <https://cornell.app.box.com/v/ReadmeTemplate>

Para mais informações sobre a preparação de dados consulte também:

Códigos e dados de computadores	<ul style="list-style-type: none"> • Experiences on reproducibility of paper experiments
Dados de Ciências Sociais	<ul style="list-style-type: none"> • Guide to Social Science Data Preparation and Archiving
Dados tabulares	<ul style="list-style-type: none"> • Preparing tabular data for description and archiving

Referências

Cornell University. Guide to writing "readme" style metadata. *Cornell University* [online]. [viewed 12 February 2021]. Available from: <https://data.research.cornell.edu/content/readme>.

DataverseNO. Prepare your data. *DataverseNO* [online]. [viewed 12 February 2021]. Available from: <https://site.uit.no/dataverseno/deposit/prepare/>.

Nanyang Technological University. DR-NTU (Data) User Guides and Policies. *Nanyang Technological University* [online]. [viewed 12 February 2021]. Available from: <https://libguides.ntu.edu.sg/drntudataguidespolicies/depositor#s-lg-box-21651979>.

UC Santa Barbara Library. The Dos and Don'ts of file naming. *UC Santa Barbara* [online]. [viewed 30 March 2021]. Available from: <https://www.library.ucsb.edu/sites/default/files/dls-n01-2021-filenaming.pdf>.

University of Illinois at Urbana-Champaign. Dataset Documentation. *Illinois Data Bank* [online]. [viewed 12 February 2021]. Available from: https://databank.illinois.edu/help#dataset_documentation.

Como citar este documento

SciELO. *Guia de preparação de dados de pesquisa* [online]. SciELO, 2021 [cited DD Month YYYY]. Available from: _____.