SciELO
Scientific Electronic Library Online

Guía de preparación de datos de investigación

Mayo de 2025
Versión 1.1

This is an Open Access document distributed under the terms of the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly.

1. Preparación de datos de pesquisa **2**
2. Anonimización de datos **4**
2.1 Anonimización de las transcripciones 5
2.2 Anonimización de imágenes 5
3. Formato de los archivos **6**
4. Nombre de archivos **8**
5. Descripción de los datos **8**
6. Archivo README **10**
7. Tamaño de los archivos **11**
8. Cambios en el manuscrito **11**
Referencias 12
Anexo 1. Fluxograma SciELO Data **13**
Anexo 2. Formatos adecuados para la conservación **14**

1. Preparación de datos de pesquisa

Adoptar buenas prácticas en la preparación de datos para su depósito es fundamental para el acceso, el intercambio y la reutilización de datos de investigación. Preparar los datos adecuadamente implica organizar y documentar la información de forma que promueva el acceso y la reutilización por parte de otros investigadores, lo que repercute significativamente en el progreso de la investigación y la transparencia científica, y ayuda a los investigadores a seguir los Principios FAIR y a ser lo más justos posible.[1]

Los datos FAIR cumplen con los estándares de facilidad de búsqueda, accesibilidad, interoperabilidad y reutilización. La curación de datos, basada en los principios FAIR, busca promover la fácil búsqueda de datos y hacerlos más accesibles, interoperables y reutilizables. Para lograr estos objetivos, es fundamental que los datos se almacenen en repositorios adecuados, con metadatos claros y consistentes, lo que facilita su descubrimiento y acceso..

Para comprender mejor los Principios FAIR y cómo hacer que sus datos sean lo más FAIR posible, le recomendamos utilizar la herramienta FAIR-Aware: https://fairaware.dans.knaw.nl/.

Además de los principios FAIR, es importante considerar las pautas éticas y culturales relacionadas con los datos de investigación, en particular cuando se trata de información que involucra a comunidades, como pueblos indígenas o poblaciones vulnerables y sus derechos. Fue en este contexto que el *CARE*[2] (*Collective Benefit, Access, Respect, Ethics* - Beneficio colectivo, acceso, respeto, ética) que buscan garantizar que el uso de los datos respete las

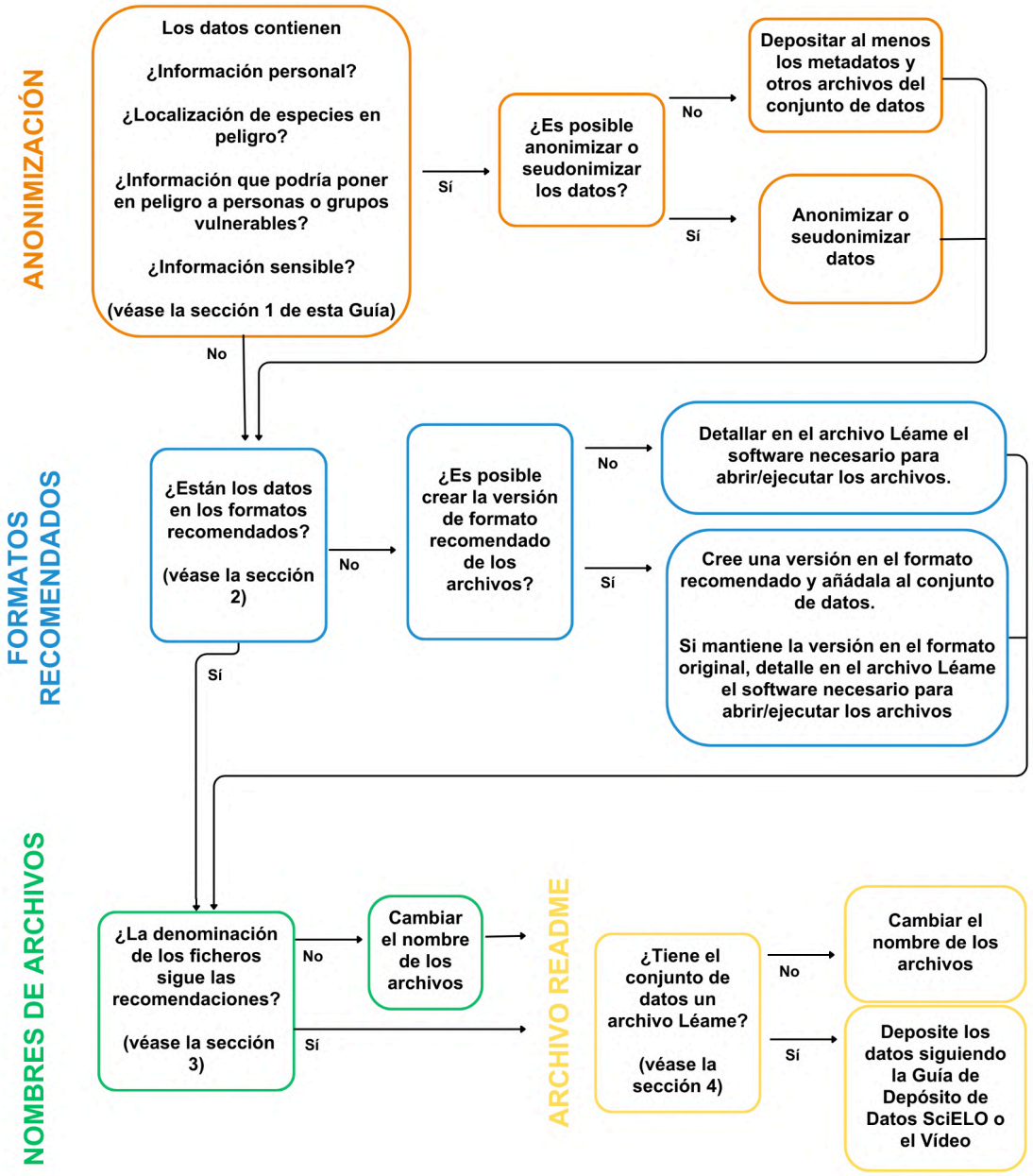
[1] Recomendamos el webinar "¿Cómo hacer los datos sean FAIR? Buenas prácticas para datos (abiertos) de investigación" disponible en: https://www.youtube.com/watch?v=l14SwZxIRHY.

[2] VIDOTTI, Silvana Aparecida Borsetti Gregorio; TORINO, Emanuelle; CONEGLIAN, Caio Saraiva. #SejaJUSTOeCUIDADOSO: princípios FAIR e CARE na gestão de dados de pesquisa. In: SALES, Luana Farias; VEIGA, Viviane dos Santos; HENNING, Patrícia; SAYÃO, Luís Fernando (org.). Princípios FAIR aplicados à gestão de dados de pesquisa. Rio de Janeiro: Ibict, 2021. p. 201 - 214. DOI: 10.22477/9786589167242.cap15

normas éticas y culturales.

La aplicación de los principios CARE junto con los principios FAIR ayuda a crear un entorno de investigación más justo, responsable y transparente, promoviendo no sólo la reutilización de los datos, sino también su uso ético y responsable.

Al preparar sus datos de investigación para depositarlos en un repositorio, se deben tomar una serie de decisiones y procedimientos considerando la información contenida en los datos, así como el formato de sus archivos y la documentación relacionada. A continuación, se presenta un diagrama de flujo de decisiones que puede ayudar en el proceso de adaptación de los datos de investigación:



La versión en video de esta guía está disponible en YouTube: Guía de preparación de datos de investigación (Sólo en portugués). Para obtener información sobre el flujo de datos de SciELO Data, consulte el Anexo 1.

2. Anonimización de datos

Se debe anonimizar lo siguiente: datos personales, ya sean sensibles o no[3], información que exceda el derecho a la privacidad de las personas involucradas o que los ponga en riesgo, así como coordenadas de áreas protegidas, en peligro de extinción, o información que viole acuerdos comerciales, patentes o que pertenezca a terceros.

Reducir la presencia de identificadores directos en los archivos que conforman el conjunto de datos para reducir la precisión y el detalle de personas, ubicaciones o información que no pueda identificarse mediante agregaciones, como:

- Año o década de nacimiento en lugar de la fecha precisa de nacimiento;
- Rango de edad en lugar de edad específica;
- Región en lugar de ciudad;
- Ubicación urbana/rural o general (ej: Zona Norte, Región Sur del municipio, edificio comercial en el centro de la ciudad, etc.) en lugar del nombre de los lugares;
- Ocupación o área de especialización en lugar del título del puesto específico;
- Período de tiempo en lugar de fecha u hora específica.

Ejemplo de anonimización de datos[4]:

Información sin anonimización	Responder sin anonimizar
Nombre	Juan Pérez
País de nacimiento	Argentina
Edad	54
Años de experiencia	15
Modelo de aviones	Boeing 777 \ 747
Fecha de último vuelo	05/01/2022

Información anonimizada	Respuesta anonimizada
-	-
Continente	América del Sur
Rango de Edad	50-60
Rango de años de experiencia	10-20
Modelo de aviones	Comerciales
Fecha de último vuelo	01/2022

[3]Datos personales: Se consideran datos personales: nombre y apellidos; domicilio; correo electrónico (si contiene elementos que permitan identificar al titular, como nombre y apellidos); género; fecha de nacimiento; número de documentos de registro, como DNI, CPF y tarjeta de trabajo; datos de geolocalización de un teléfono celular; número de teléfono personal. Disponible en: https://portal.fiocruz.br/noticia/entenda-melhor-lei-geral-de-protecao-de-dados-pessoais Consultado el 21 de marzo de 2023.
Datos personales sensibles: "datos personales sobre origen racial o étnico, creencias religiosas, opiniones políticas, afiliación a un sindicato u organización de carácter religioso, filosófico o político, datos relativos a la salud o la vida sexual, datos genéticos o biométricos, cuando estén vinculados a una persona física". https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Consultado el 30 de enero de 2023.
[4] Ejemplo tomado de: Gestión de Datos de Investigación - Parte I. Disponible en: https://www.youtube.com/watch?v=BM-lZ2XCCN0

2.1 Anonimización de las transcripciones

Para anonimizar las transcripciones, no basta con eliminar el nombre de la persona entrevistada. Analice cuidadosamente las respuestas de la entrevista y elimine la información que permita su identificación directa o indirecta, como:

- Nombres de lugares mencionados que permitan identificar dónde nació, vive o trabaja la persona;
- Número de teléfono, número de documento, fecha de nacimiento o dirección;
- Cargo desempeñado o trabajo específico realizado que pueda identificar a la persona o afiliación institucional;
- Citas de nombres de otras personas (por ejemplo, nombres de profesores o compañeros de trabajo).

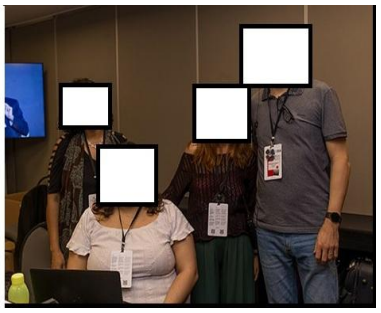
Evite eliminar o reemplazar elementos sin indicar que se realizó una edición. Use seudónimos o reemplazos entre corchetes (p. ej.: "Ana vive en Brasilia" -> "[E1] vive en [ciudad del Medio Oeste brasileño]") para indicar el lugar donde se realizó la edición.

Ana vive en **Brasília** → [**E1**] vive en [**ciudad del Medio Oeste brasileño**]

2.2 Anonimización de imágenes

Las imágenes y los videos también deben anonimizarse. Difumine o cubra los videos para evitar la identificación de personas u otra información que no pueda compartirse abiertamente (nombres de calles, matrículas, números de documentos, etc.).

Ejemplo de imágenes anonimizadas:



Fuente: https://25.scielo.org/fotos/. Foto: Carla Formanek

Se debe aplicar el mismo tratamiento a los archivos de audio, prestando atención a la posibilidad de identificación por voz, acento o defectos lingüísticos.

En los casos en que la anonimización sea imposible, se debe intentar usar seudónimos o, si los participantes aún son identificables, no se debe depositar el archivo en cuestión, sino solo los demás archivos que conformarán el conjunto de datos (como las preguntas formuladas y el análisis de las respuestas), así como un archivo Léame u otro archivo de documentación del

conjunto de datos (véase la sección 4), explicando que no todos los archivos están disponibles y el motivo por el cual se impide compartirlos en acceso abierto.

Algunas herramientas que pueden ayudar con la anonimización son: Amnesia de Openaire, anonymoUUs de la Universidad de Utrecht y la herramienta de ayuda para la anonimización de texto de UK Data Service.

3. Formato de los archivos

SciELO Data acepta cualquier tipo de archivo, pero se recomienda depositar los siguientes formatos:

- No propietarios;
- Abiertos, con padrones internacionales documentados;
- Con codificación de caracteres estándar, preferiblemente Unicode (por ejemplo. UTF-8).

Los archivos en formatos comunes no requieren software específico para su apertura. Por lo tanto, priorice los formatos recomendados a continuación para facilitar el acceso a los archivos por parte de otras personas, así como la conservación de datos y el acceso a largo plazo.

Tabla 1 - Formatos recomendados[5]

Tipo de documento	**Formatos recomendados**	**Formatos no recomendados** (pero aceptados)
Análisis estadístico	R (.r, .rdata) SPSS (.dat/.sps) STATA (.dat/.do)	SPSS Portable (.por) SPSS (.sav)
Archivos comprimidos	.zip*	.rar
Datos tabulares	Comma Separated Values (.csv) Arquivo de texto (.txt) Excel (.xlsx)	Excel (.xls)
Datos textuales	Microsoft Word (docx) Arquivo de texto (.txt) HTML, XHTML ou XML, que contiene referencia al DTD o esquema de validación de formato (.html, .xhtml, .xml) OpenOffice (.odt, .ods ou .odp) PDF (.pdf)	Microsoft Word (.doc) Postscript (.ps) PDF cifrado (.pdf) WordPerfect (.wpd) DVI (.dvi)

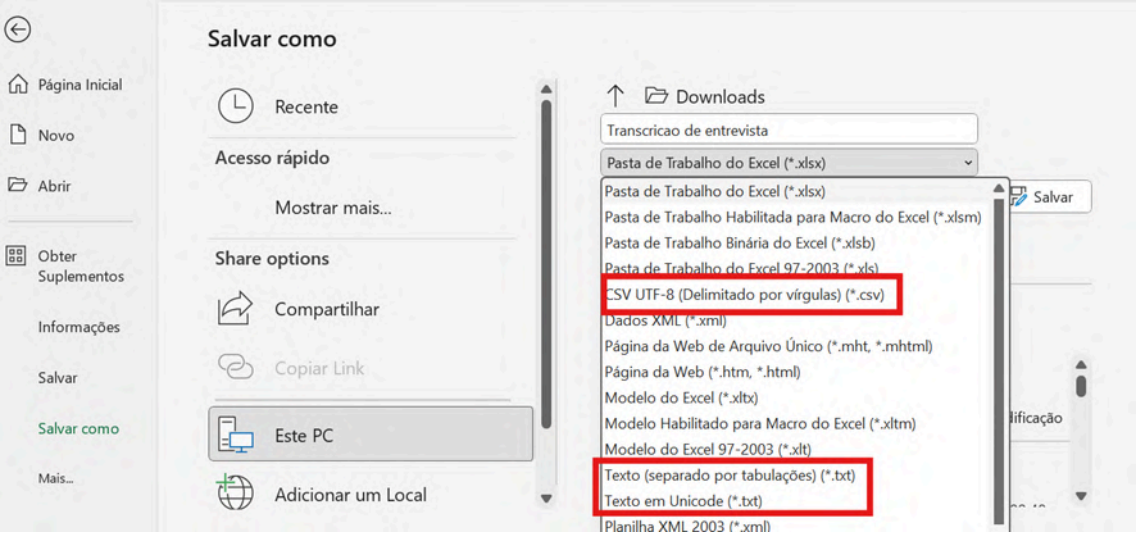
[5] Formatos adequados para preservação. Arquivo.pt. Sin fecha. Disponible en: https://sobre.arquivo.pt/pt/colabore/recomendacoes/formatos-adequados-para-preservacao/. Acceso: 30 de abril de 2025.

Imágenes	PNG (.png) JPEG2000 compresión sin pérdida de información (.jp2) TIFF sin comprimir SVG (.svg)	Macromedia Flash (*.swf) PhotoShop (.psd) JPEG 2000 Part 2 (.jpf, .jpx) MrSID (.sid) TIFF en formato Planar (.tiff) FlashPix (.fpx) RAW Encapsulated Postscript (.eps)
Audio	AIFF con Pulse-code modulation (.aif, .aiff) WAV con Pulse-code modulation (.wav, bwf) Ogg Vorbis (.ogg, .oga)	RealNetworks 'Real Audio' (.ra, .rm, .ram) Windows Media Audio (.wma) WAV con compresión (.wav) AIFC con compresión (.aifc) NeXT SND (.snd)
Video	QuickTime Movie sin compresión (.mov) AVI sin compresión (.avi) Motion JPEG 2000 (ISO/IEC 15444-4) (mj2) Motion JPEG (.avi, .mov)	Windows Media Video (.wmv) AVI con compresión (.avi) QuickTime Movie con compresión (.mov) RealNetworks 'Real Video' (.rv, .rm)

*Los archivos comprimidos con la extensión .zip se descomprimirán una vez cargados.

Una vez que los archivos se depositan en formatos no recomendados o que no están clasificados como de alta o media idoneidad para la conservación según el Anexo 1, no se garantiza su correcta conservación a largo plazo.

La conversión de formato de archivo suele realizarse mediante el programa utilizado para crear los documentos. Por ejemplo, para archivos creados con Office, simplemente abra el archivo, haga clic en "Guardar como" y seleccione el formato de apertura correspondiente, como los ejemplos resaltados en la imagen a continuación.



Si convierte archivos de un formato a otro, infórmelo en el archivo Readme o en la documentación complementaria depositada con el conjunto de datos.

4. Nombre de archivos

Adoptar buenas prácticas para nombrar archivos evita que se sobrescriban y facilita que otros investigadores los encuentren y reutilicen.

- Utilice nombres autoexplicativos/descriptivos, con un máximo de 100 caracteres (por ejemplo: “entrevista-nino-1”);
- No utilice acentuación;
- Utilice caracteres alfanuméricos, guiones bajos o guiones (por ejemplo: entrevista-nino_1);
- Evite los espacios por ejemplo: entrevista nino 1), puntos (entrevista.nino.1), letras mayúsculas (EntrevistaNino1) y caracteres especiales (como \ /?: * ”> <|: #%” {} | ^ `~ @ & ; ° æÆ øØ åÅ äÄ öÖ);
- Utilice el formato AAAA-MM-DD (entrevista_nino_1_2021-01-07) o AAAAMMDD (entrevista_nino_1_20210107) para las fechas;
- Incluya el número de versión en el nombre cuando corresponda (entrevista-niño-1-001.csv, entrevista-nino-1-002.csv, …, entrevista-nino-1-010.csv, …);
- Utilice el mismo nombre para archivos con el mismo contenido pero en diferentes formatos (entrevista-nino-1-data.doc y entrevista-nino-1.txt).

5. Descripción de los datos

Para que los datos depositados puedan ser interpretados y reutilizados correctamente, tanto por usted en el futuro como por otros investigadores, es esencial que se describan de la forma más detallada y comprensible posible.

Esta descripción detallada deberá proporcionarse rellenando los campos durante el depósito y en un archivo README, que sirve de guía para el usuario y deberá depositarse junto con los archivos de datos.

Como alternativa, se puede enviar otra documentación, como cuadernos de laboratorio y diccionarios de datos (que proporcionan información detallada sobre tablas, columnas, tipos de datos, restricciones y otras características de los datos), que contienen información descriptiva y contextual sobre los datos depositados. Data Dictionary Creator es una aplicación que permite crear un diccionario de datos a partir de los archivos de investigación.

Ejemplo de diccionario de datos

Tabela: Nome da tabela no banco de dados (ex: Tratamentos, Estudos).
Nome da Coluna: Nome da coluna que está na tabela (ex: ID_Paciente, ID_Tratamento).
Tipo de Dado: Tipo de dado que a coluna irá armazenar (ex: INT para números inteiros, VARCHAR(255) para textos com até 255 caracteres, DATE para datas).
Descrição: Uma breve explicação sobre o que a coluna representa (ex: Faixa etária do paciente, Duração do tratamento em dias).
Restrições: Regras que se aplicam à coluna. A restrição pode incluir "Não nulo", "Chave primária", "Chave estrangeira" e "Opcional".

Tabela,Nome da Coluna,Tipo de Dado,Descrição,Restrições
Pacientes,ID_Paciente,INT,Identificador único do paciente,Chave primária, Não nulo
Pacientes,Nome,VARCHAR(255),Nome completo do paciente,Não nulo
Pacientes,Idade,INT,Idade do paciente,Não nulo
Pacientes,Sexo,VARCHAR(10),Sexo do paciente (masculino ou feminino),Não nulo
Tratamentos,ID_Tratamento,INT,Identificador único do tratamento,Chave primária, Não nulo
Tratamentos,Nome_Tratamento,VARCHAR(255),Nome do tratamento administrado,Não nulo
Tratamentos,Tipo_Tratamento,VARCHAR(100),Tipo de tratamento (ex: farmacológico, fisioterápico),Não nulo
Tratamentos,Duração,INT,Duração do tratamento em dias,Não nulo
Estudos,ID_Estudio,INT,Identificador único do estudo,Chave primária, Não nulo

Ejemplo de un cuaderno de laboratorio

Exemplo de Caderno de Laboratório (Anônimo)

Página 1: Dados do Estudo

Data: 10/2023
Responsável: Dra. Maria Oliveira
Título do Estudo: Avaliação da eficácia de tratamento cognitivo-comportamental em pacientes com transtorno de ansiedade generalizada
Objetivo: Avaliar a redução dos sintomas de ansiedade após 5 semanas de tratamento cognitivo-comportamental em pacientes diagnosticados com transtorno de ansiedade generalizada.

Página 2: Participantes do Estudo

Número de Participantes: 5
Códigos dos Participantes:

- Paciente 01
- Paciente 02
- Paciente 03
- Paciente 04
- Paciente 05

Critérios de Inclusão:

- Diagnóstico de transtorno de ansiedade generalizada (conforme DSM-5).
- Idade entre 18 e 55 anos.
- Aceitação do termo de consentimento informado.

Critérios de Exclusão:

- História de transtornos psicóticos.
- Uso recente de medicamentos psicotrópicos (menos de 3 meses).

Página 3: Métodos

Tratamento Aplicado:

- **Tipo de Tratamento**: Terapia cognitivo-comportamental (TCC), 1 sessão semanal de 60 minutos durante 8 semanas.
- **Objetivos do Tratamento**: Redução dos sintomas de ansiedade, melhoria nas estratégias de enfrentamento e alteração de padrões cognitivos disfuncionais.

Instrumentos de Avaliação:

1. **Escala de Ansiedade**: Utilizada para medir a gravidade da ansiedade antes e após o tratamento.
2. **Questionário de Qualidade**: Avalia a qualidade de vida dos participantes.
3. **Entrevistas Semiestruturadas**: Conduzidas no início e no final do estudo para avaliar a experiência do paciente com o tratamento.

Página 4: Cronograma de Sessões

Semana 1: Introdução à TCC, identificação de pensamentos disfuncionais.
Semana 2: Estratégias de enfrentamento do estresse.
Semana 3: Modificação de pensamentos catastróficos.
Semana 4: Técnicas de relaxamento e controle da ansiedade.
Semana 5: Exposição gradual aos medos.

6. Archivo README

El archivo README, en particular, es **obligatorio para los conjuntos de datos depositados en SciELO Data**. Debe estar escrito como texto plano con codificación de caracteres Unicode UTF-8 (.txt) o como PDF si es necesario para ilustrar o dar formato a la descripción de los datos.

El archivo README debe contener al menos la siguiente información:

- Título del conjunto de datos;
- Información de contacto (nombre, institución y correo electrónico) del investigador principal / corresponsal o responsable de la recopilación de datos;
- Fecha de recopilación de datos (fecha única o intervalo de tiempo);
- Resumen de datos y archivos (breve descripción de los datos que contiene cada archivo, fecha de creación de cada archivo y cómo se relacionan entre sí, etc.);
- Descripción de los métodos de recopilación o generación de datos;
- Descripción de los métodos utilizados para procesar los datos;
- Información de datos específicos (lista de variables, unidades de medida, definiciones de códigos o símbolos, calibración de equipos, etc.).

SciELO Data cuenta con su propia plantilla README, que abarca los temas mencionados y puede servir de base para los autores. Está disponible para su descarga.

Al crear su conjunto de datos, inserte la mayor cantidad posible de información y descripciones de los datos, como un archivo README según la plantilla de SciELO Data, junto con un diccionario de datos que especifique las columnas, filas y variables de las tablas presentes.

Ejemplos de plantillas para el archivo README:

General	• https://drive.google.com/drive/folders/1LNxx4YKxN1a2DF10bbVoRW-ioZwNz-9M • https://drive.google.com/file/d/167cJdaRy4sxQWA5qEEp2cgfWqKV-smZV/view • https://cornell.app.box.com/v/ReadmeTemplate
Ciencias Sociales	• https://social-science-data-editors.github.io/template_README/template-README.html
Código de software	• https://drive.google.com/file/d/1VIDF489DDr044Uta8z1G7EuLj-Wm_UtG/view

Para ver ejemplos de otra información que se puede agregar al archivo README, consulte:

- Guide to writing "readme" style metadata

Para obtener más información sobre la preparación de datos, consulte también:

Códigos y datos de computadores	• Experiences on reproducibility of paper experiments • Research Code

Datos de Ciencias Sociales	• Guide to Social Science Data Preparation and Archiving
Datos tabulares	• Preparing tabular data for description and archiving

7. Tamaño de los archivos

El tamaño límite de los archivos individuales es de 2 GB. Para agregar archivos por encima de este límite, comuníquese con data@scielo.org.

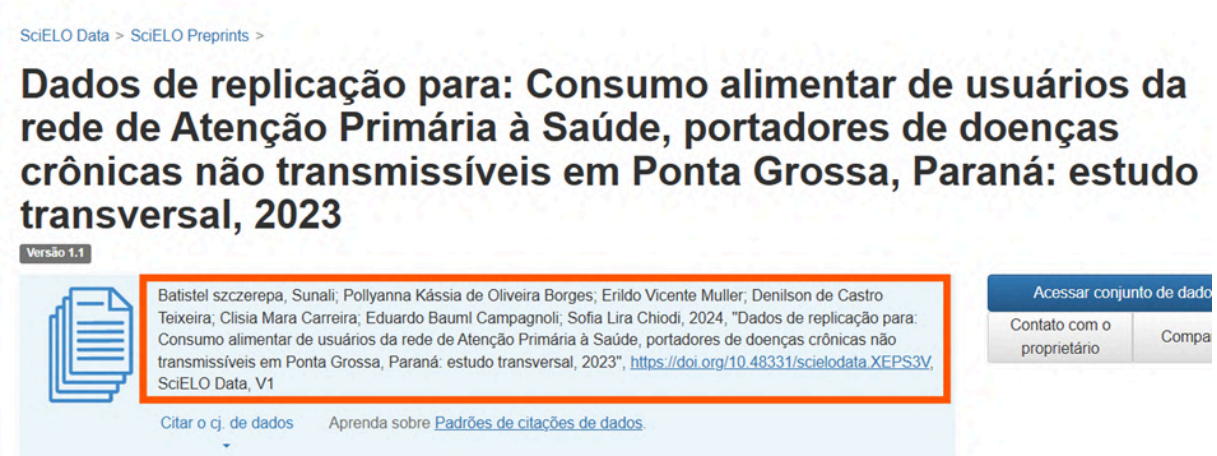
8. Cambios en el manuscrito

Al depositar datos de investigación relacionados con un artículo, el DOI del conjunto de datos debe estar presente en la sección Disponibilidad de datos, o sección equivalente, según la norma de la revista, o inserte la sección en su preprint.

Como ejemplo de texto, sugerimos:
"Todo el conjunto de datos anónimos que apoya los resultados de este estudio fue publicado en SciELO Data y se puede acceder a él en [**URL o DOI del conjunto de datos**]."

También es posible citar su propio conjunto de datos en el cuerpo del texto (ejemplos de citas estándar de conjuntos de datos están disponibles en la Guía de citación de datos de investigación (solo en Portugués), lo que implica insertar el conjunto de datos en la lista de referencias del artículo siguiendo el estándar normativo de referencias determinado por su revista.

Para acceder a la citación completa y DOI del dataset, después de depositarlo en SciELO Data, vaya a la página inicial del dataset. La información estará en el recuadro azul debajo del título del conjunto de datos. El DOI sólo estará activo una vez que el conjunto de datos haya sido curado por el equipo editorial y publicado.



Referencias

Arquivo.PT. Formatos adequados para preservação. [S.l.]: Arquivo.pt, [s.d.]. Available from: https://sobre.arquivo.pt/pt/colabore/recomendacoes/formatos-adequados-para-preservacao/.

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.

Syracuse University. De-Identification. *Qualitative Data Repository*, 2023. Available from: https://qdr.syr.edu/guidance/human-participants/deidentification.

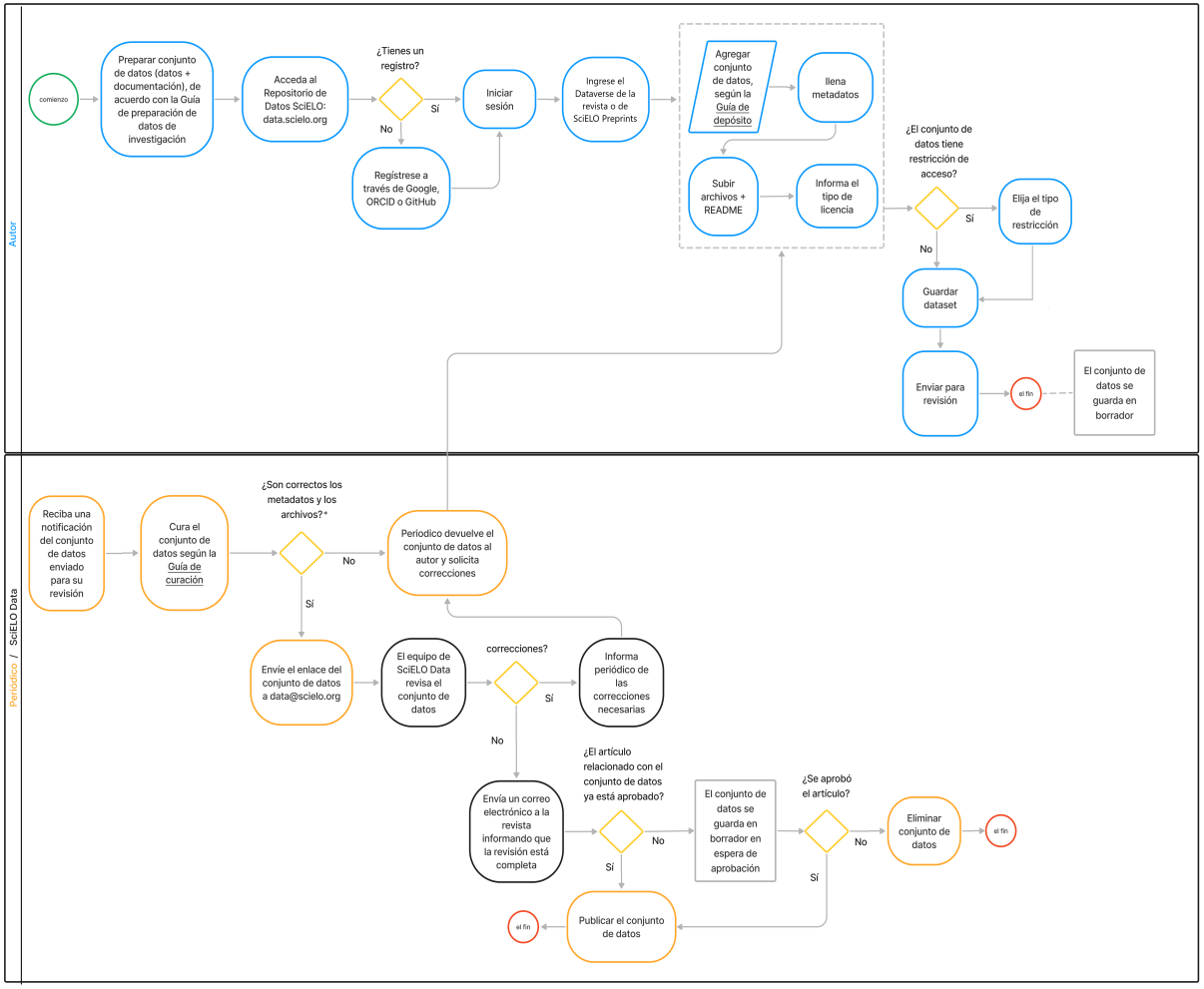
UC Santa Barbara Library. The Dos and Don'ts of file naming. *UC Santa Barbara* [online]. [viewed 30 March 2021]. Available from: https://www.library.ucsb.edu/sites/default/files/dls-n01-2021-filenaming.pdf.

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

Cómo citar este documento

SciELO. *Guía de preparación de datos de investigación* [online]. SciELO, 2025 [cited DD Month YYYY]. Available from: _____.

Anexo 1. Diagrama de flujo SciELO Data



***Solo para los conjuntos de datos depositados en el Dataverse de SciELO Preprints:** la curaduría está a cargo del Equipo SciELO Data. Si se requieren ediciones o correcciones, el equipo de SciELO se comunicará con el autor depositante por correo electrónico para solicitarlas. El conjunto de datos solo se publicará después de la aprobación del preprint y la realización de las correcciones.

Anexo 2. Formatos adecuados para la conservación

Algunos formatos de archivo se consideran más adecuados para la conservación a largo plazo porque no requieren software específico para abrirlos y acceder a ellos. Para saber si sus formatos de archivo se clasifican como de **alta**, **media** o **baja** idoneidad para la conservación, consulte la lista a continuación. Siempre que sea posible, guarde sus archivos en formatos de alta idoneidad para la conservación.

Texto

Alta idoneidad para la conservación

- HTML, XHTML o XML, que contiene referencia al DTD o esquema de validación de formato (.html, .xhtml, .xml)
- Textos simples que utilizan codificación de caracteres UTF-8, USASCII o UTF-16 con marca de orden de bytes (.txt)
- PDF/A-1 según ISO 19005-1 (.pdf)
- Abrir texto del documento (.odt)

Media para la conservación

- HTML, XHTML o XML, sin referencia al DTD o esquema de validación de formato (.html, .xhtml, .xml)
- Hojas de estilo en cascada (.css)
- Texto sin formato con codificación de caracteres ISO-8859-1 (.txt)
- PDF con fuentes incrustadas (.pdf)
- Formato de texto enriquecido 1.x (.rtf)
- HTML 4.x incluida la declaración DOCTYPE (.html)
- Documento de texto de Open Office (.sxw)
- Office Open XML (.docx)
- DTD (.dtd)
- SGML (.sgml)

Baja idoneidad para la conservación

- Microsoft Word (.doc)
- Posdata (.ps)
- PDF cifrado (.pdf)
- WordPerfect (.wpd)
- DVI (.dvi)

Imagen

Alta idoneidad para la conservación

- PNG (.png)
- JPEG2000 con compresión sin pérdida (.jp2)
- TIFF sin comprimir
- SVG (.svg)

Media para la conservación

- JPEG2000 con compresión con pérdida (.jp2)
- GIF (.gif)
- JPEG/JFIF (.jpg)
- TIFF con compresión
- BMP (.bmp)
- Negativo digital (.dng)
- Metarchivo de gráficos de computadora y WebCGM (.cgm)

Baja idoneidad para la conservación

- Macromedia Flash (*.swf)
- Photoshop (.psd)
- JPEG 2000 Parte 2 (.jpf, .jpx)
- MrSID (.sid)
- TIFF en formato planar (.tiff)
- FlashPix (.fpx)
- CRUDO
- PostScript encapsulado (.eps)

Audio

Alta idoneidad para la conservación

- AIFF con modulación por código de pulsos (.aif, .aiff)
- WAV con modulación por código de pulsos (.wav, bwf)
- Ogg Vorbis (.ogg, .oga)

Media idoneidad para la conservación

- MP3 (MPEG-1/2, capa 3) (.mp3)
- Códec de audio sin pérdida gratuito (.flac)
- SUN Audio sin comprimir (.au)
- MIDI estándar (.mid, midi)
- Codificación de audio avanzada (.mp4, .m4a, .aac)

Baja idoneidad para la conservación

- RealNetworks 'Audio real' (.ra, .rm, .ram)
- Audio de Windows Media (.wma)
- WAV comprimido (.wav)
- AIFC con compresión (.aifc)
- Siguiente SND(.snd)

Video

Alta idoneidad para la conservación

- Película QuickTime sin comprimir (.mov)
- AVI sin comprimir (.avi)
- Motion JPEG 2000 (ISO/IEC 15444-4) (mj2)
- JPEG en movimiento (.avi, .mov)

Media idoneidad para la conservación

- MPEG-1, MPEG-2 (.mpg, .mpeg)
- MPEG-4 (.mp4)
- Ogg Theora (.ogg, .ogm, .ogv)

Baja idoneidad para la conservación

- Vídeo de Windows Media (.wmv)
- AVI comprimido (.avi)
- Película QuickTime con compresión (.mov)
- Vídeo real de RealNetworks (.rv, .rm)

Otros formatos

Alta idoneidad para la conservación

- Valores separados por comas (.csv)
- DDL de SQL

Media idoneidad para la conservación

- OpenOffice (.sxc/.ods, .sxi/.odp))
- OOXML según el estándar ISO/IEC DIS 29500 (.xlsx, .pptx)

Baja idoneidad para la conservación

- Microsoft Excel (.xls)
- Microsoft PowerPoint (.ppt)

- Microsoft Access (.mdb)
- Microsoft Visio (.vsd)

Fuente: Formatos adequados para preservação. Arquivo.pt. Sin fecha. Disponible en: https://sobre.arquivo.pt/pt/colabore/recomendacoes/formatos-adequados-para-preservacao/. Acceso: 30 de abril de 2025.