

# Metadados e Plataformas para dados de pesquisa em saúde

## uma análise dos repositórios cadastrados no Re3Data

# REDE BRASIL SAÚDE

## REDE DE BIBLIOTECAS FIOCRUZ

### **Autores\***

Adrienne Silva - COC/Fiocruz

Adriano Silva - ENSP/Fiocruz

Claudio Ribeiro - UNIRIO

Claudete Queiroz - ICICT/Fiocruz

Erick Penedo - ICICT/Fiocruz

Francijane Oliveira - INC

Gizele Rocha - ICICT/Fiocruz

Jefferson Lima - ICICT/Fiocruz

Marcus Silva - COC/Fiocruz

Maria Eduarda Teixeira Alves - ICICT/Fiocruz

Patrícia Henning - UNIRIO

Patrícia Mendes - ICICT/Fiocruz

Rejane Machado - ICICT/Fiocruz

Simone Dib - ICICT/Fiocruz

Solange Santos - Scielo

Viviane Veiga - ICICT/Fiocruz

**\*Membros do Grupo de Trabalho Metadados e Plataformas**



## Objetivo da pesquisa

---

Investigar padrões de metadados e plataformas para promover o compartilhamento e gestão de dados de pesquisa no campo da saúde alinhados aos princípios FAIR.

A network diagram with nodes and connecting lines, serving as a background for the slide.

# Objetivos específicos

---

- ✓ Identificar softwares e plataformas para dados de pesquisa genéricos e disciplinares;
- ✓ Identificar softwares e plataformas para dados de pesquisa genéricos e disciplinares;
- ✓ Identificar padrões de metadados do campo da saúde;
- ✓ Verificar se os padrões de metadados e as plataformas estão alinhadas aos princípios FAIR;
- ✓ Elaborar uma lista de padrões de metadados e plataformas alinhadas aos princípios FAIR.



# Identificar softwares e plataformas para dados de pesquisa genéricos e seus metadados

METADADOS	EUDAT		Dspace ARCA		Harvard Dataverse	
	<i>Ocorrência</i>	<i>Repetitivo</i>	<i>Ocorrência</i>	<i>Repetitivo</i>	<i>Ocorrência</i>	<i>Repetitivo</i>
Título	Obrigatório	Não	Obrigatório	Não	Obrigatório	Não
Autor	Obrigatório	Não	Obrigatório	Sim	Obrigatório	Sim
Afiliação	Obrigatório	Não	Obrigatório	Sim	Obrigatório	Sim
Licença	Restrito	Não	Obrigatório	Não	Opcional	Não
Data de Criação	Obrigatório	Sim	Ausente	x	Opcional	Não
Data de Publicação	Facultativo	Não	Obrigatório	Não	Opcional	Não
Financiador	Possui	Sim	Facultativo	Sim	Opcional	Sim
Se foi publicado	Facultativo	Não	Ausente	x	Opcional	Sim
Versionamento	Ausente	Não	Ausente	x	Obrigatório*	Não
Tipo de Dado	Ausente	Não	Ausente	x	Opcional	Não
Identificador Único Dado	Ausente	Não	Facultativo	Sim	Obrigatório*	Não
Identificador Único Autor	Ausente	Não	Ausente	x	Facultativo	Sim
Documento Relacionado	Facultativo	Não	Facultativo	Não	Facultativo	Sim
Tipo de Documento	Obrigatório	Sim	Obrigatório	Não	Facultativo	Sim
Upload de Documento	Obrigatório	Sim	Obrigatório	Sim	Opcional	Sim
Idioma	Facultativo	Sim	Obrigatório	Não	Opcional	Sim
Descrição	Obrigatório	Sim	Obrigatório	Sim	Obrigatório	Sim
Palavras-chave	Obrigatório	Sim	Obrigatório	Sim	Facultativo	Sim

# Identificar softwares e plataformas para dados de pesquisa genéricos e seus metadados

METADADOS	Zenodo		Figshare		Dendro	
	<i>Ocorrência</i>	<i>Repetitivo</i>	<i>Ocorrência</i>	<i>Repetitivo</i>	<i>Ocorrência</i>	<i>Repetitivo</i>
Título	Obrigatório	Não	Obrigatório	Não	Obrigatório	Não
Autor	Obrigatório	Sim	Obrigatório	Sim	Obrigatório	Sim
Afiliação	Facultativo	Não	Ausente	x	Ausente	x
Licença	Obrigatório	Não	Obrigatório	Não	Obrigatório	Não
Data de Criação	automatizado	automatizado	Ausente	x	Facultativo	Não
Data de Publicação	Obrigatório	Não	Ausente	x	Facultativo	Não
Financiador	Facultativo	Sim	Facultativo	Sim	Facultativo	Não
Se foi publicado	Obrigatório	Não	Facultativo	Não	Ausente	x
Versionamento	Facultativo	Não	Ausente	x	Obrigatório	Não
Tipo de Dado	Obrigatório	Não	Ausente	x	Ausente	x
Identificador Único Dado	Facultativo	Não	Facultativo	Não	Facultativo	Não
Identificador Único Autor	Facultativo	Sim	Ausente	x	Ausente	x
Documento Relacionado	Facultativo	Não	Facultativo	Sim	Facultativo	Não
Tipo de Documento	Ausente	x	Facultativo	Não	Ausente	x
Upload de Documento	Obrigatório	Sim	Facultativo	Sim	Obrigatório	Sim
Idioma	Facultativo	Não	Ausente	x	Facultativo	Não
Descrição	Obrigatório	Não	Obrigatório	Não	Obrigatório	Não
Palavras-chave	Facultativo	Não	Obrigatório	Sim	Ausente	x

**Identificar softwares e plataformas para dados de pesquisa disciplinar (em andamento)**

**Neuroscience**

<b>Bases</b>	<b>Endereço eletrônico</b>
<b>ModelDB</b>	<a href="http://senselab.med.yale.edu/modeldb">http://senselab.med.yale.edu/modeldb</a>
<b>OpenNeuro</b>	<a href="http://www.openneuro.org">http://www.openneuro.org</a>
<b>NeuroVault</b>	<a href="http://neurovault.org">http://neurovault.org</a>
<b>Open Connectome Project</b>	<a href="http://openconnecto.me">http://openconnecto.me</a>
<b>Donders Repository</b>	<a href="http://data.donders.ru.nl">http://data.donders.ru.nl</a>
<b>G-Node Data Infrastructure Services</b>	<a href="http://gin.g-node.org">http://gin.g-node.org</a>

**Health Science**

<b>Bases</b>	<b>Endereço eletrônico</b>
<b>Research Domain Criteria Database</b>	<a href="http://rdocdb.nimh.nih.gov">http://rdocdb.nimh.nih.gov</a>
<b>National Database for Clinical Trials related to Mental Illness</b>	<a href="https://data-archive.nimh.nih.gov/ndct">https://data-archive.nimh.nih.gov/ndct</a>
<b>Swedish National Data Service</b>	<a href="https://snd.gu.se/en">https://snd.gu.se/en</a>
<b>Ag Data Commons</b>	<a href="https://data.nal.usda.gov">https://data.nal.usda.gov</a>





# Identificar softwares e plataformas para dados de pesquisa **disciplinar**

Levantamento

**re3**data.org  
REGISTRY OF RESEARCH DATA REPOSITORIES

Total de  
repositórios

**2.423**

Áreas do  
Conhecimento

Public Health, Health Services Research,  
Social Medicine  
**124 repositórios**



# Identificar softwares e plataformas para dados de pesquisa

## Principais softwares utilizados nos repositórios

Software ↕	Repositórios ↕
unknown	75
-NA-	19
other	15
DataVerse	8
CKAN	4
EPrints	1
MySQL	1
Nesstar	1

A network diagram with nodes and connecting lines, serving as a background for the slide.

# Objetivos específicos

---

- Identificar softwares e plataformas para dados de pesquisa genéricos e disciplinares;
- Identificar padrões de metadados do campo da saúde;
- Verificar se os padrões de metadados e as plataformas estão alinhadas aos princípios FAIR;
- Elaborar uma lista de recomendação de padrões de metadados e plataformas alinhadas aos princípios FAIR.

# Identificar padrões de metadados do campo da saúde

## Avaliação Geral – Resultados Parciais

Padrão metadados ↕	Repositórios ↕
DDI - Data Documentation Initiative	20
Dublin Core	11
DataCite Metadata Schema	5
FGDC/CSDGM - Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata	5
ISO 19115	3
Repository-Developed Metadata Schemas	3
DCAT - Data Catalog Vocabulary	2
CF (Climate and Forecast) Metadata Conventions	1
SDMX - Statistical Data and Metadata Exchange	1
other	1

# Identificar padrões de metadados do campo da saúde

<b>Social Science &amp; Humanities - Metadata Standards</b>	
<b>DDI - Data Documentation Initiative</b>	Um padrão internacional para descrever dados das ciências sociais, comportamentais e econômicas. Expressa em XML, a especificação de metadados DDI suporta todo o ciclo de vida dos dados da pesquisa
<b>Encoded Archival Description (EAD)</b>	Um padrão para codificar auxiliares de busca de arquivos usando XML em repositórios de arquivos e manuscritos, implementando as recomendações do Conselho Internacional de Arquivos
<b>MIDAS-Heritage</b>	Padrão Britânico para registrar informações sobre prédios, sítios arqueológicos, naufrágios, parques e jardins, campos de batalha, áreas de interesse e artefatos
<b>OAI-ORE - Open Archives Initiative Object Reuse and Exchange</b>	Define padrões para a descrição e troca de agregações de recursos da Web.
<b>QuDEx - Qualitative Data Exchange Format</b>	Um modelo qualitativo de troca de dados para arquivamento e intercâmbio de dados
<b>SDMX - Statistical Data and Metadata Exchange</b>	Um conjunto de padrões e diretrizes técnicas e estatísticas comuns a serem usados para o intercâmbio e o compartilhamento eficientes de dados e metadados estatísticos.

# Identificar padrões de metadados do campo da saúde

## Biology - Metadata Standards

<b>ABCD - Access to Biological Collection Data</b>	Uma norma para o acesso e o intercâmbio de dados primários de biodiversidade, incluindo amostras e observações.
<b>Darwin Core</b>	Um corpo de padrões, incluindo um glossário de termos (em outros contextos, esses podem ser chamados de propriedades, elementos, campos, colunas, atributos ou conceitos) destinados a facilitar o compartilhamento de informações sobre diversidade biológica, fornecendo definições de referência, exemplos e comentários
<b>EML - Ecological Metadata Language Ecological Metadata Language (EML)</b>	A Linguagem de Metadados Ecológicos (EML) é uma especificação de metadados especialmente desenvolvida para a disciplina de ecologia.
<b>Genome Metadata</b>	Dados descritivos sobre genomas únicos no Pathosystems Resource Integration Center.
<b>ISA-Tab</b>	Uma estrutura de uso geral com a qual capturar e comunicar metadados para arquivos de dados de experimentos 'baseados em ics' que empregam combinações de tecnologias
<b>MIBBI - Minimum Information for Biological and Biomedical Investigations</b>	Um portal comum para um grupo de listas de verificação de informações mínimas em quase 40 disciplinas biológicas.

## Biology - Metadata Standards

<b>Observ-OM</b>	Usado para integrar e comparar dados de observação em projetos experimentais, bancos de dados de doenças e biobancos clínicos.
<b>OME-XML - Open Microscopy Environment XML</b>	Um padrão de metadados e formato de arquivo de dados para dados de microscopia de luz biológica.
<b>PDBx/mmCIF – Protein Data Bank Exchange Dictionary and the Macromolecular Crystallographic Information Framework</b>	PDBx / mmCIF é o formato de arquivo padrão usado pelo Protein Data Bank (PDB). Fornece metadados e dados de acordo com as propriedades definidas no PDB Exchange Dictionary e no Macromolecular Crystallographic Information Framework (mmCIF).
<b>Protocol Data Element Definitions</b>	Usado pelos Institutos Nacionais de Saúde (EUA) no site ClinicalTrials.gov. O Sistema de Registro de Protocolo (PRS) é usado para registrar estudos clínicos em seres humanos
<b>Repository-Developed Metadata Schemas</b>	Alguns repositórios decidiram que os padrões atuais não atendem às suas necessidades de metadados e, portanto, criaram seus próprios requisitos

# Identificar padrões de metadados do campo da saúde

General Research Data - Metadata Standards	
<b>CERIF - Common European Research Information Format</b>	O CERIF é o padrão que a União Européia recomenda a seus estados membros para registrar informações sobre atividades de pesquisa.
<b>Data Package</b>	A especificação do pacote de dados é um formato genérico de wrapper para troca de dados, consistindo em uma pasta contendo arquivos de dados e um arquivo descritor.
<b>DataCite Metadata Schema</b>	Uma lista independente de domínio das principais propriedades de metadados escolhidas para a identificação precisa e consistente de dados para fins de citação e recuperação.
<b>DCAT - Data Catalog Vocabulary</b>	DCAT é um vocabulário RDF projetado para facilitar a interoperabilidade entre catálogos de dados publicados na Web.
<b>Dublin Core</b>	Um padrão básico, independente de domínio, que pode ser facilmente entendido e implementado e, como tal, é um dos padrões de metadados mais conhecidos e mais amplamente utilizados.
<b>OAI-ORE - Open Archives Initiative</b>	Define padrões para a descrição e troca de agregações de recursos da Web.
<b>Observations and Measurements</b>	Este é um padrão que especifica a implementação XML para o modelo conceitual OGC e Observações e Medidas ISO (O&M), incluindo um esquema para Recursos de Amostragem.
<b>PREMIS - The PREMIS (Preservation Metadata: Implementation Strategies)</b>	O Dicionário de dados define um conjunto de metadados que a maioria dos repositórios de objetos digitais precisaria gravar e usar para preservar esses objetos a longo prazo.
<b>PROV</b>	Uma especificação que fornece um vocabulário para trocar informações de proveniência.
<b>RDF Data Cube Vocabulary</b>	O vocabulário do Cubo de Dados é uma base essencial que suporta vocabulários de extensão para permitir a publicação de outros aspectos de fluxos de dados estatísticos ou outros conjuntos de dados multidimensionais.
<b>Repository-Developed Metadata Schemas</b>	Alguns repositórios decidiram que os padrões atuais não atendem às suas necessidades de metadados e, portanto, criaram seus próprios requisitos.



# Aplicação de questionário

---

- ✓ Elaboração de um questionário visando identificar os padrões de metadados e os metadados específicos das áreas de Life Science do Re3Data;
- ✓ Questionário enviado em Julho de 2020 aos gestores dos Repositórios selecionados;
- ✓ 89 respostas;
- ✓ Tabulação e análise.



## GO FAIR Brazil Health - Metadata

Dear participants,

The attached quiz is part of research that is being developed within the scope of the GO FAIR Brazil Health network <<https://portal.fiocruz.br/go-fair-brasil-saude>> coordinated by the Oswaldo Cruz Foundation (Fiocruz) <<https://portal.fiocruz.br/>> in Brazil. GO FAIR Brazil Health is a member of the GO FAIR Brazil Network <<https://www.go-fair.org/go-fair-initiative/go-fair-offices/go-fair-brazil-office/>> linked to the initiative GO FAIR International <<https://www.go-fair.org/>>.

Our goal is identified the specific metadata and metadata schemes or standards that are being adopted by the Data Repositories registered in Re3Data, according to their respective areas of knowledge.

The result of this survey will provide us support recommendations for standards and metadata in the health field, which will be used in the Brazilian Data Health Repositories.

Your participation is extremely important for us! Please collaborate by answering a just seven of a short questionnaire, which will take no more than 10 minutes. The result will be shared with the respondents of this research and published to the scientific community. We thank you in advance for your participation and if you have any questions, please contact the Coordinator of GOFAIR Brazil Health.

Kind regards,

Professor. Viviane Veiga – [viviane.veiga@icict.fiocruz.br](mailto:viviane.veiga@icict.fiocruz.br)  
Coordination of GOFAIR Brazil Health

Próxima

## GO FAIR Brazil Health - Metadata

\*Obrigatório

### Identification

Institution Name: \*

Sua resposta

Country Name: \*

Escolher

Voltar

Próxima



## About the Data Repository

What software (Platform) does your Institution use? \*

- Dataverse
- Dspace
- Figshare
- Zenodo
- Own development
- I don't know
- Outro: \_\_\_\_\_

Indicate the name of the repository \*

Sua resposta \_\_\_\_\_

Are there any data management policy in your Institution Data Repository? \*

- Yes
- No
- I don't know

# GO FAIR Brazil Health - Metadata

\*Obrigatório

## About the LIFE SCIENCES areas presented in RE3DATA

About the LIFE SCIENCES areas presented in RE3DATA, in wich one is your Institution Data Repository registred? \*

Escolher

Biology

Medicine

Agriculture, Forestry, Horticulture and Veterinary Medicine

[Privacidade](#)

[Termos de Serviço](#) - [Política de](#)

# GO FAIR Brazil Health - Metadata

\*Obrigatório

## Biology

About the LIFE SCIENCES subareas presented in RE3DATA, in wich one is your Institution Data Repository registred? \*

Check all that apply

Basic Biology and Medical Research

Plant Sciences

Zoology

Outro: \_\_\_\_\_

Voltar

Próxima



## GO FAIR Brazil Health - Metadata

\*Obrigatório

### About standards or metadata schema

Does your Institution Data Repository use any standard or metadata scheme specific to your area of knowledge? \*

- Yes
- No
- I don't know

Voltar

Próxima

## GO FAIR Brazil Health - Metadata

\*Obrigatório

### Metadata standards

What are the standards metadata used by your Institution Data Repository? \*

Check all that apply

- Dublin Core Standard
- MODS Standard
- Distance Learning Standard
- ANSI / NISO Z39.87 standard
- METS Standard
- PREMIS Standard
- Darwin Core Standard
- Outro: \_\_\_\_\_

Voltar

Próxima

## Generic metadata

Could you select the generic metadata used in your Institution Data Repository? \*

Check all that apply

- Title
- Author
- Affiliation
- License
- Creation Date
- Publication Date
- Financier
- Versioning
- Data Type
- Unique identifier of the data set
- Unique identifier for metadata
- Unique identifier of the author
- Related document
- Type of document
- Uploading the Document
- Language
- Description
- Keywords

Is there any specific metadata used in your area of knowledge? \*

- Yes
- No
- I don't know

Voltar

Próxima

## GO FAIR Brazil Health - Metadata

\*Obrigatório

### Specific metadata

Describe these specific metadata. \*

Separated by commas.

Sua resposta

Voltar

Próxima



## GO FAIR Brazil Health - Metadata

### General comments:

Please, make any comments you think are necessary.

Sua resposta

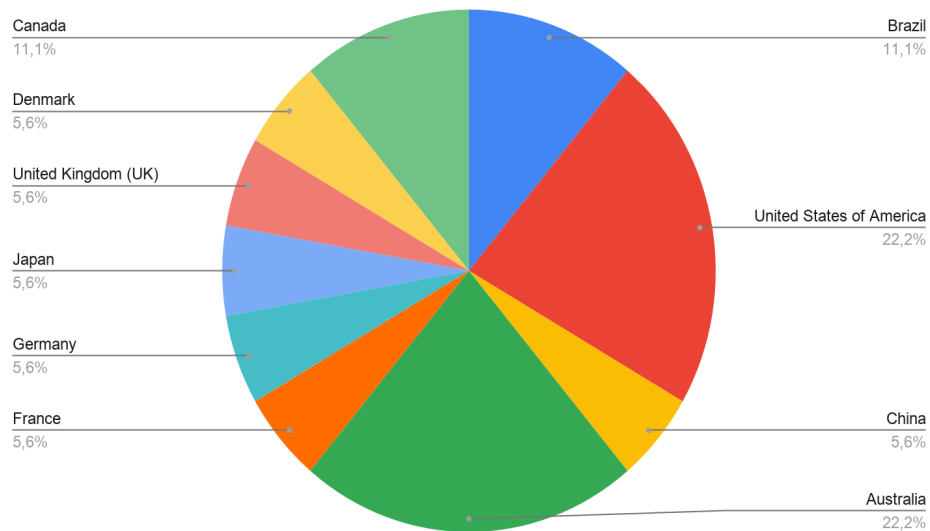
---

Voltar

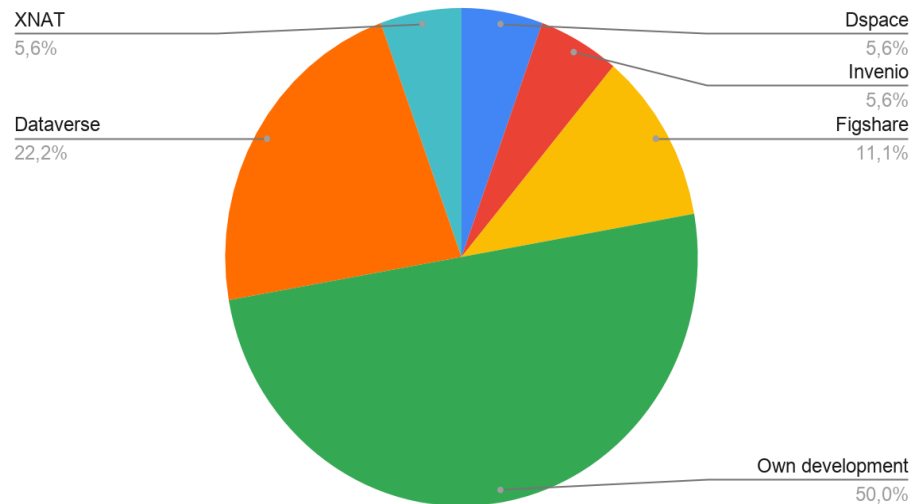
Enviar

# GO FAIR Brazil Health - Metadata

## Country Name:

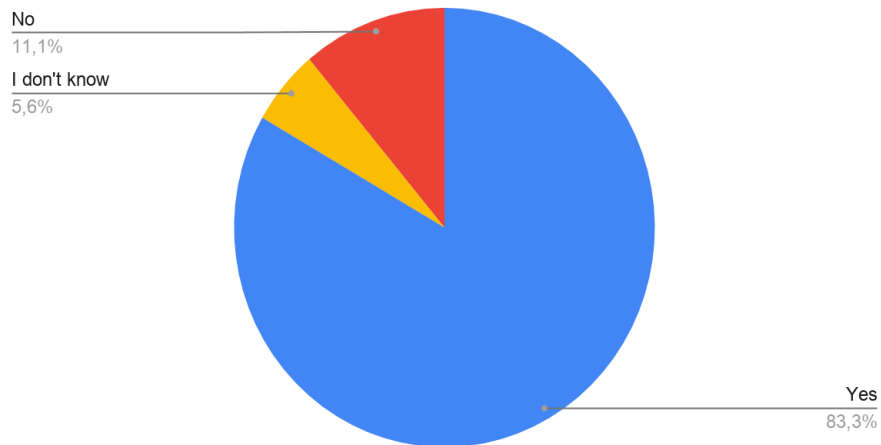


## What software (Platform) does your Institution use?

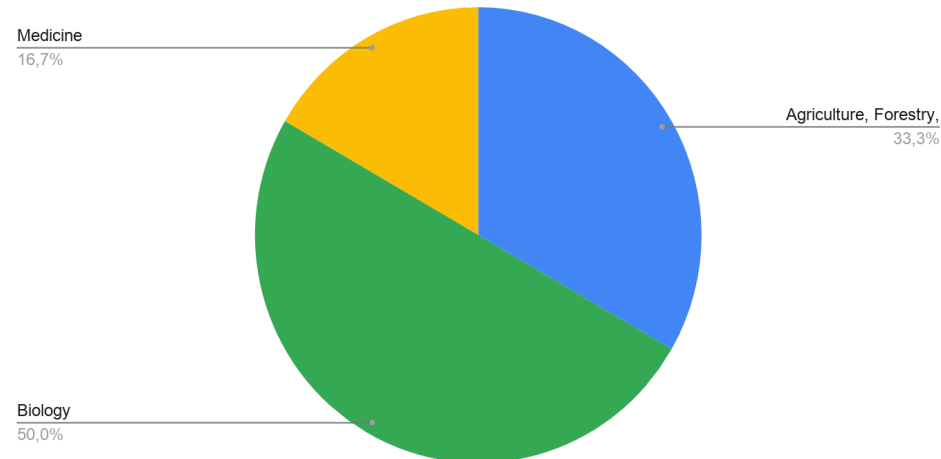


# GO FAIR Brazil Health - Metadata

Are there any data management policy in your Institution Data Repository?



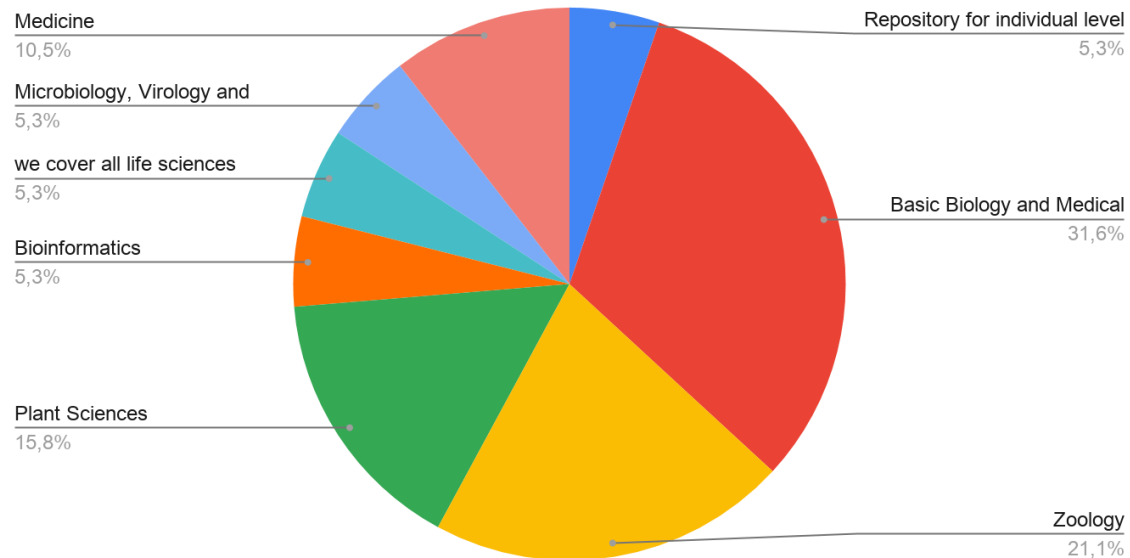
About the LIFE SCIENCES areas presented in RE3DATA, in wich one is your Institution Data Repository registered?





# GO FAIR Brazil Health - Metadata

About the LIFE SCIENCES subareas presented in RE3DATA, in wich one is your Institution Data Repository registred?

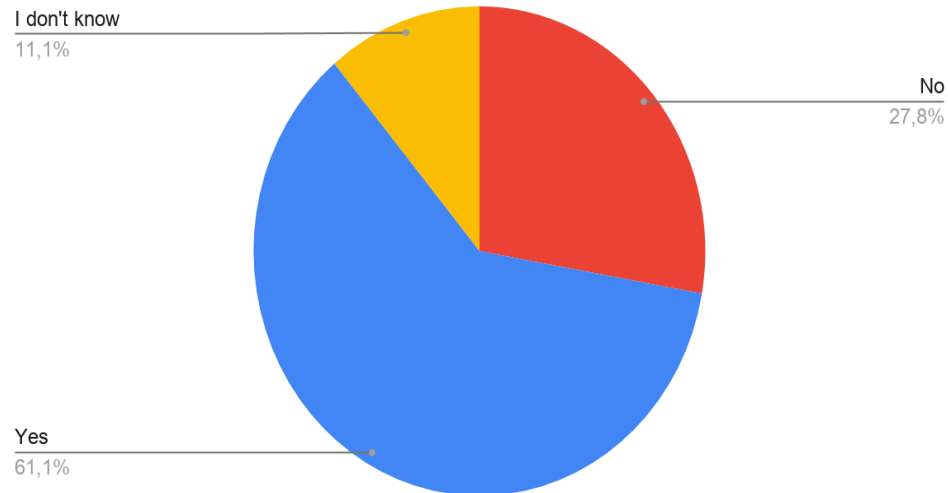




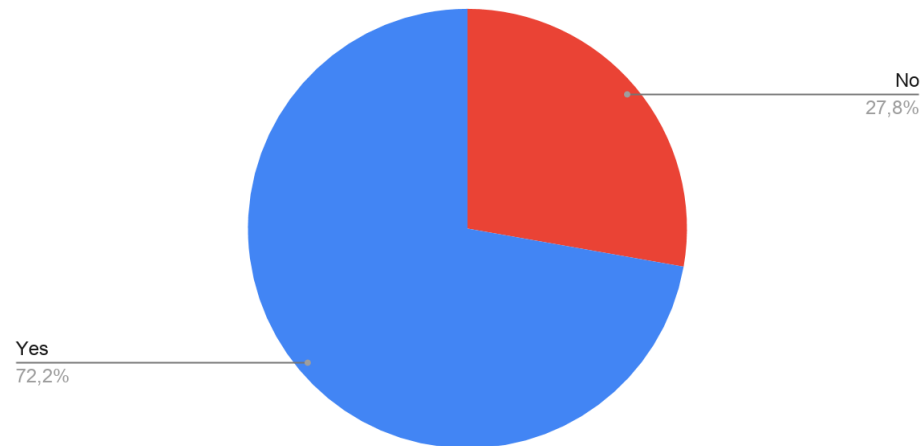
# GO FAIR Brazil Health - Metadata

---

Is there any specific metadata used in your area of knowledge?

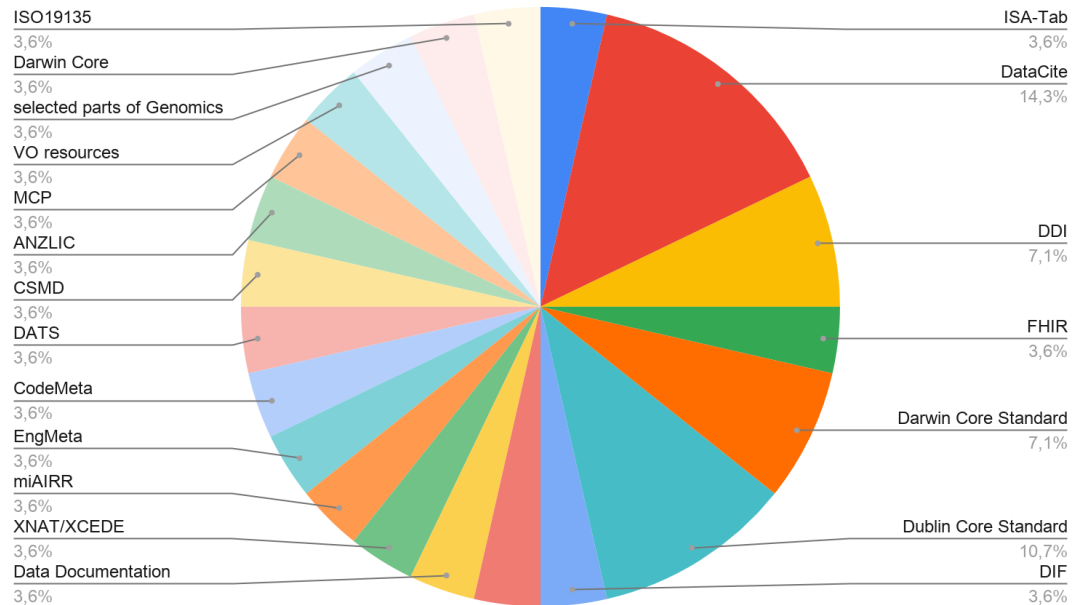


Does your Institution Data Repository use any standard or metadata scheme specific to your area of knowledge?



# GO FAIR Brazil Health - Metadata

What are the standards metadata used by your Institution Data Repository?





# NAMES OF REPOSITORIES

Indicate the name of the repository

Base de Dados Científicos (BDC) da UFPR

CaltechDATA

DataShare

Database of Genotypes and Phenotypes (dbGaP)

GigaDB

Atlas of Living Australia

Data INRAE

DaRUS

CSIRO Data Access Portal

Arctic and Antarctic Data archive System (ADS)

xzfhdf

Ibict Cariniana Dataverse Network

European Variation Archive

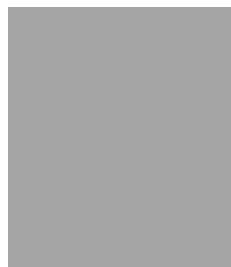
DNBC

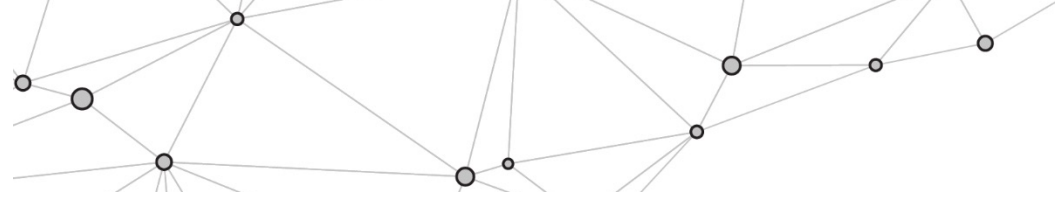
University of Guelph Research Data Repositories

NITRC-IR

iReceptor

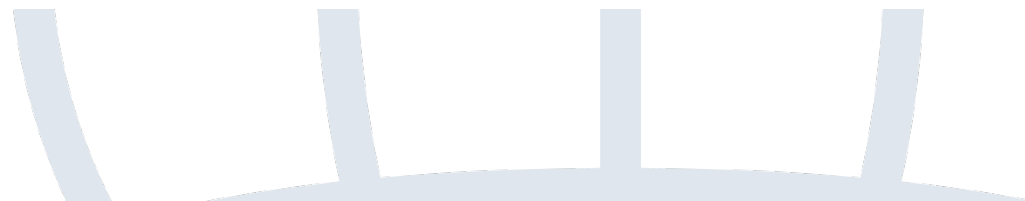
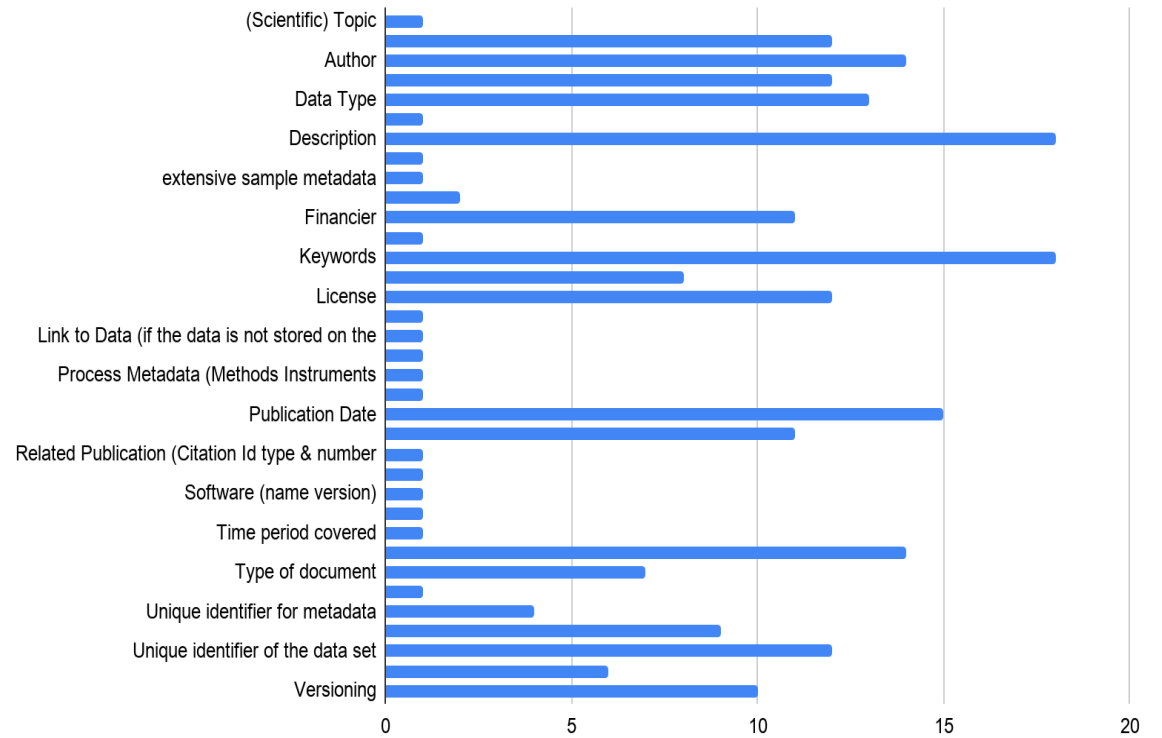
Bridges





# GENERIC METADATA

Could you select the generic metadata used in your Institution Data Repository?





# Resultado esperado

---

Disponibilizar uma lista de de padrões de metadados e plataformas do campo da saúde alinhados aos princípios FAIR.



# Mais informações

[rededebibliotecas@icict.fiocruz.br](mailto:rededebibliotecas@icict.fiocruz.br)

[gofairbrasilsaude@icict.fiocruz.br](mailto:gofairbrasilsaude@icict.fiocruz.br)