



REDE GO FAIR BRASIL SAÚDE

Grupo de Trabalho Metadados e Plataformas

Viviane Veiga-Fiocruz/ICICT

Simone Dib-Fiocruz/ICICT

Claudete Queiroz-Fiocruz/ICICT

REDE BRASIL SAÚDE

Membros:

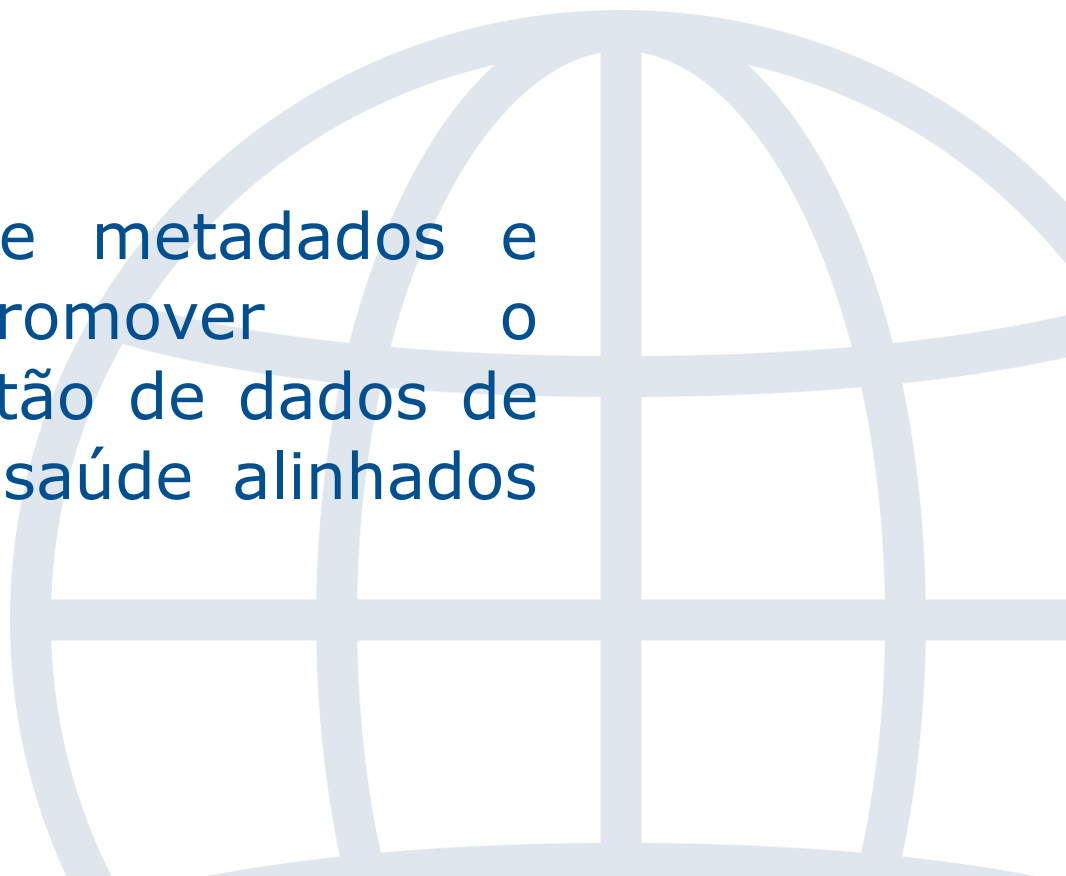
Adrienne Silva - COC/Fiocruz
Adriano Silva - ENSP/Fiocruz
Andrea Suman - Hospital Albert Einsten
Claudio Ribeiro - UNIRIO
Claudete Queiroz - ICICT/Fiocruz
Erick Penedo - ICICT/Fiocruz
Francijane Oliveira - INC
Gizele Rocha - ICICT/Fiocruz

Jefferson Lima - ICICT/Fiocruz
Marcus Silva - COC/Fiocruz
Patrícia Henning - UNIRIO
Patrícia Mendes - ICICT/Fiocruz
Rejane Machado - ICICT/Fiocruz
Simone Dib - ICICT/Fiocruz
Solange Santos - Scielo
Viviane Veiga - ICICT/Fiocruz



Objetivo do GT Metadados e Plataformas

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



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 softwares e plataformas para dados de pesquisa **genéricos**

Característica	Dspace	CKAN	Figshare	Zenodo	ePrints	EUDAT	Dendro
Arquitetura de Implantação	Pacote de Instalação ou Serviço	Pacote de Instalação ou Serviço	Serviço	Serviço	Pacote de Instalação ou Serviço	Serviço	Pacote de Instalação ou Serviço
Localização do Armazenamento	Local ou Remoto	Local ou Remoto	Remoto	Remoto	Local ou Remoto	Remoto	Local ou Remoto
Custos de Manutenção	Gerenciamento de Infraestrutura	Gerenciamento de Infraestrutura	Taxa Mensal	Taxa Mensal	Gerenciamento de Infraestrutura	Taxa Mensal	Gerenciamento de Infraestrutura
Código Aberto	Sim	Sim	Não	Não	Sim	Não	Sim
Customização	Sim	Sim	Não	Sim	Sim	Não	Sim
Suporte a internacionalização	Sim	Sim	Não	Não	Sim	Não	Não
Embargo	Sim	Armazenamento Privado	Armazenamento Privado	Sim	Sim	Sim	Sim
Versionamento de Conteúdo	Não	Sim	Não	Não	Sim	Sim	Sim
Pré-reserva do DOI	Sim	Não	Sim	Sim	Sim	Sim	Não
Metadados e disseminação de esquemas de exportação	Qualquer esquema pré-carregado	Não	DC	DC, MARCXML	DC, METS, MODS, DIDL	DC, MARC, MARCXML	MARCXML, JSON, RDF
Flexibilidade de esquema	Flexível	Flexível	Fixo	Fixo	Fixo	Flexível	Fixo
Validação	Sim	Não	Não	Sim	Sim	Sim	Sim
Versionamento	Não	Sim	Não	Não	Sim	Sim	Sim
OAI-PMH	Sim	Não	Sim	Sim	Sim	Sim	Sim
Especificação de licença	Sim	Sim	Sim	Sim	Sim	Sim	Sim

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

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

Health Science

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

RESULTADOS
PARCIAIS

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

Levantamento

re3data.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

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

Social Science & Humanities - Metadata Standards	
DDI - Data Documentation Initiative	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
Encoded Archival Description (EAD)	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
MIDAS-Heritage	Padrão Britânico para registrar informações sobre prédios, sítios arqueológicos, naufrágios, parques e jardins, campos de batalha, áreas de interesse e artefatos
OAI-ORE - Open Archives Initiative Object Reuse and Exchange	Define padrões para a descrição e troca de agregações de recursos da Web.
QuDEx - Qualitative Data Exchange Format	Um modelo qualitativo de troca de dados para arquivamento e intercâmbio de dados
SDMX - Statistical Data and Metadata Exchange	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	
ABCD - Access to Biological Collection Data	Uma norma para o acesso e o intercâmbio de dados primários de biodiversidade, incluindo amostras e observações.
Darwin Core	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
EML - Ecological Metadata Language Ecological Metadata Language (EML)	A Linguagem de Metadados Ecológicos (EML) é uma especificação de metadados especialmente desenvolvida para a disciplina de ecologia.
Genome Metadata	Dados descritivos sobre genomas únicos no Pathosystems Resource Integration Center.
ISA-Tab	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
MIBBI - Minimum Information for Biological and Biomedical Investigations	Um portal comum para um grupo de listas de verificação de informações mínimas em quase 40 disciplinas biológicas.

Biology - Metadata Standards	
Observ-OM	Usado para integrar e comparar dados de observação em projetos experimentais, bancos de dados de doenças e biobancos clínicos.
OME-XML - Open Microscopy Environment XML	Um padrão de metadados e formato de arquivo de dados para dados de microscopia de luz biológica.
PDBx/mmCIF – Protein Data Bank Exchange Dictionary and the Macromolecular Crystallographic Information Framework	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).
Protocol Data Element Definitions	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
Repository-Developed Metadata Schemas	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

CERIF - Common European Research Information Format	O CERIF é o padrão que a União Européia recomenda a seus estados membros para registrar informações sobre atividades de pesquisa.
Data Package	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.
DataCite Metadata Schema	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.
DCAT - Data Catalog Vocabulary	DCAT é um vocabulário RDF projetado para facilitar a interoperabilidade entre catálogos de dados publicados na Web.
Dublin Core	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.
OAI-ORE - Open Archives Initiative	Define padrões para a descrição e troca de agregações de recursos da Web.
Observations and Measurements	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.
PREMIS - The PREMIS (Preservation Metadata: Implementation Strategies)	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.
PROV	Uma especificação que fornece um vocabulário para trocar informações de proveniência.
RDF Data Cube Vocabulary	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.
Repository-Developed Metadata Schemas	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;
- ✓ Foram enviados 928 questionários no dia 22 de Julho de 2020 aos gestores dos Repositórios selecionados;
- ✓ Até o dia 29 de Julho já tínhamos 89 respostas;
- ✓ Reenvio de questionário;
- ✓ 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
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 registered? *

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 registered? *

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 schema 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.

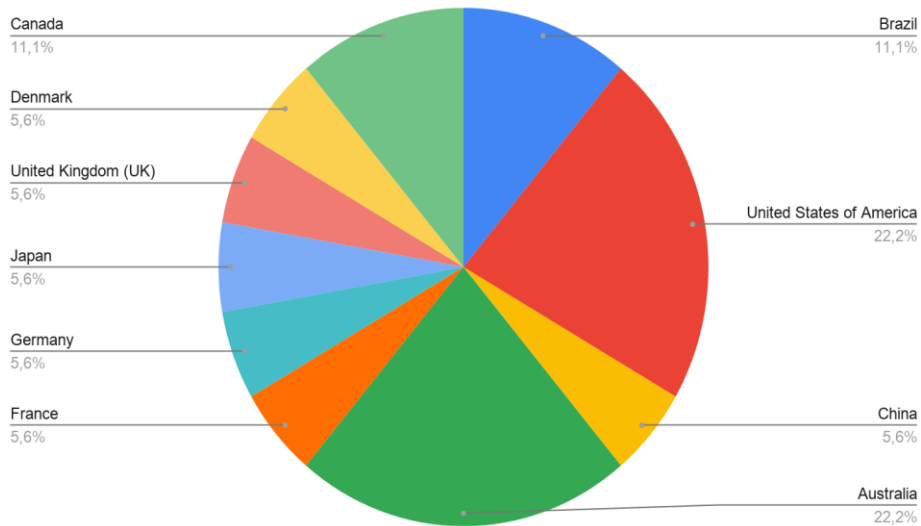
Sua resposta

Voltar

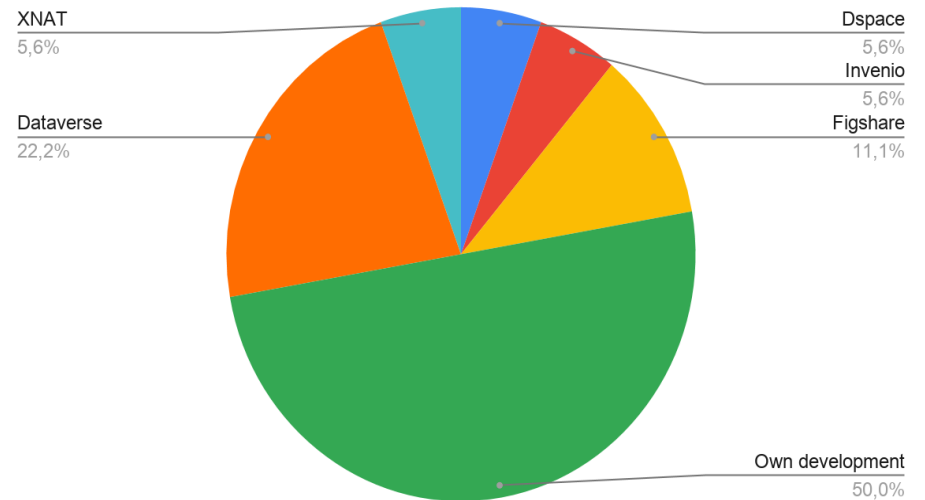
Enviar

GO FAIR Brazil Health - Metadata

Country Name:

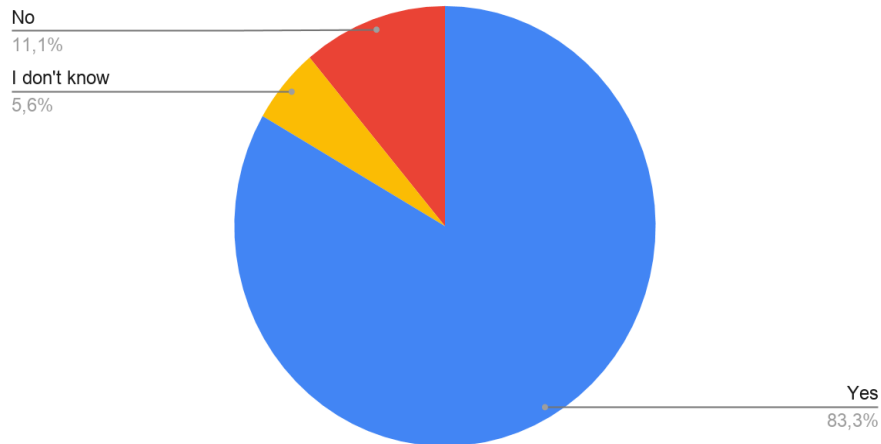


What software (Platform) does your Institution use?

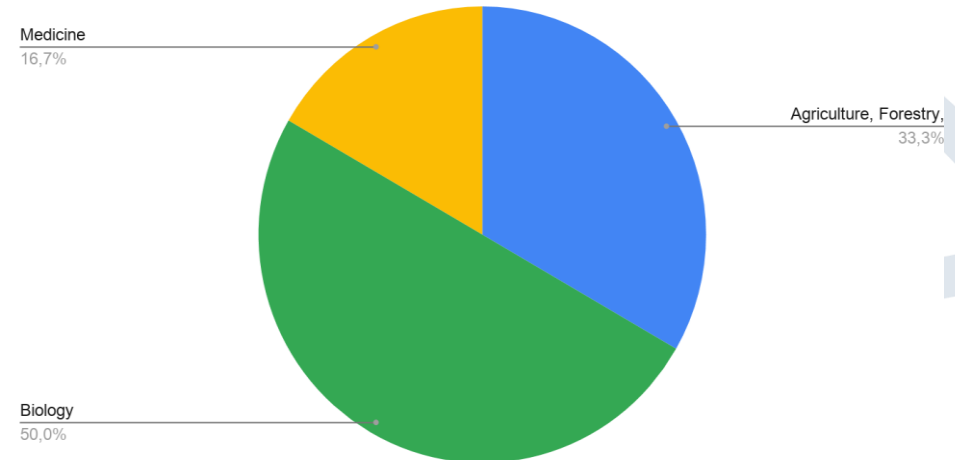


GO FAIR Brazil Health - Metadata

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

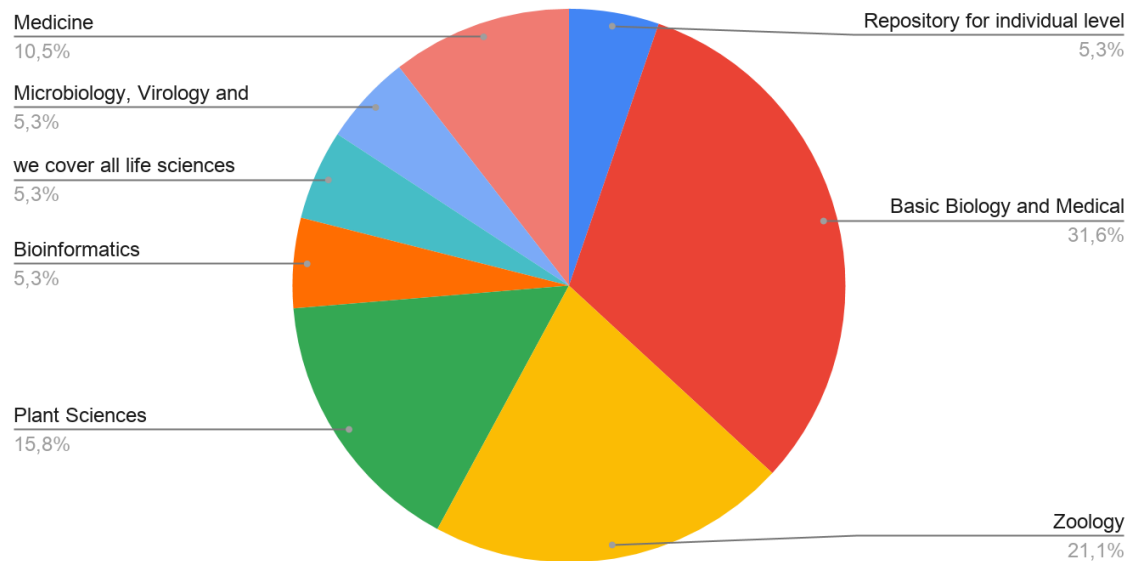


About the LIFE SCIENCES areas presented in RE3DATA, in wich one is your Institution Data Repository registred?



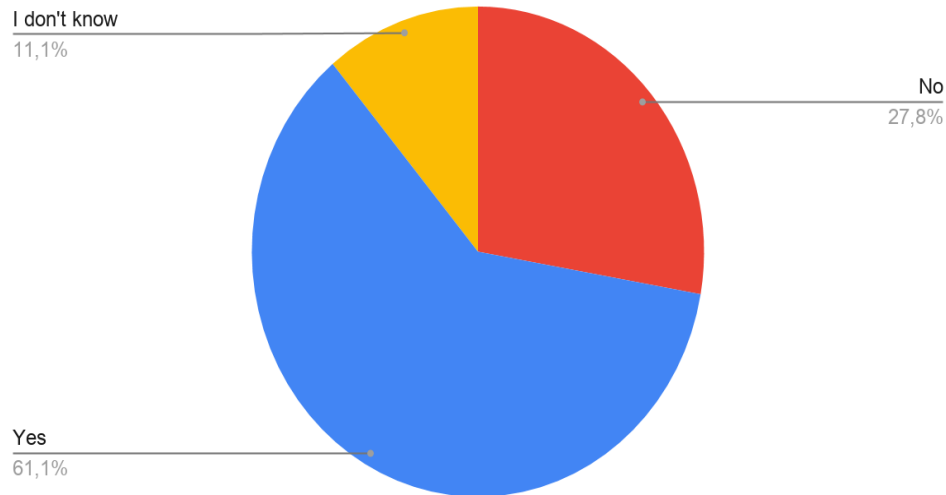
GO FAIR Brazil Health - Metadata

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

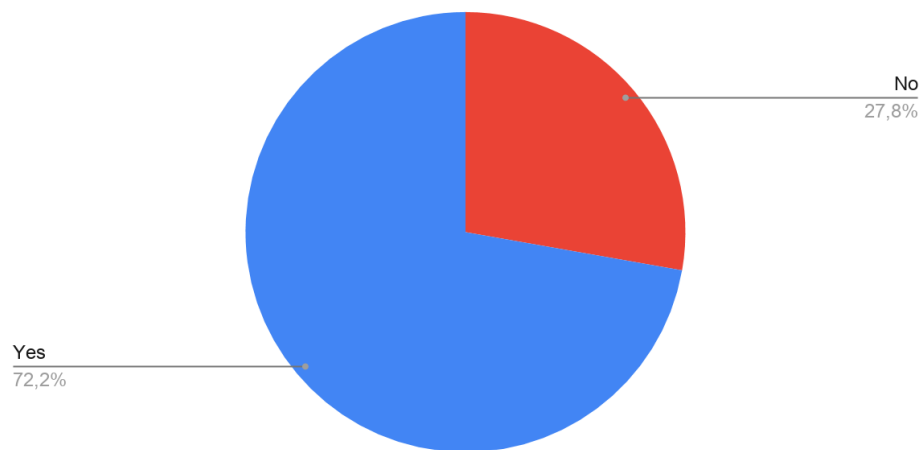


GO FAIR Brazil Health - Metadata

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

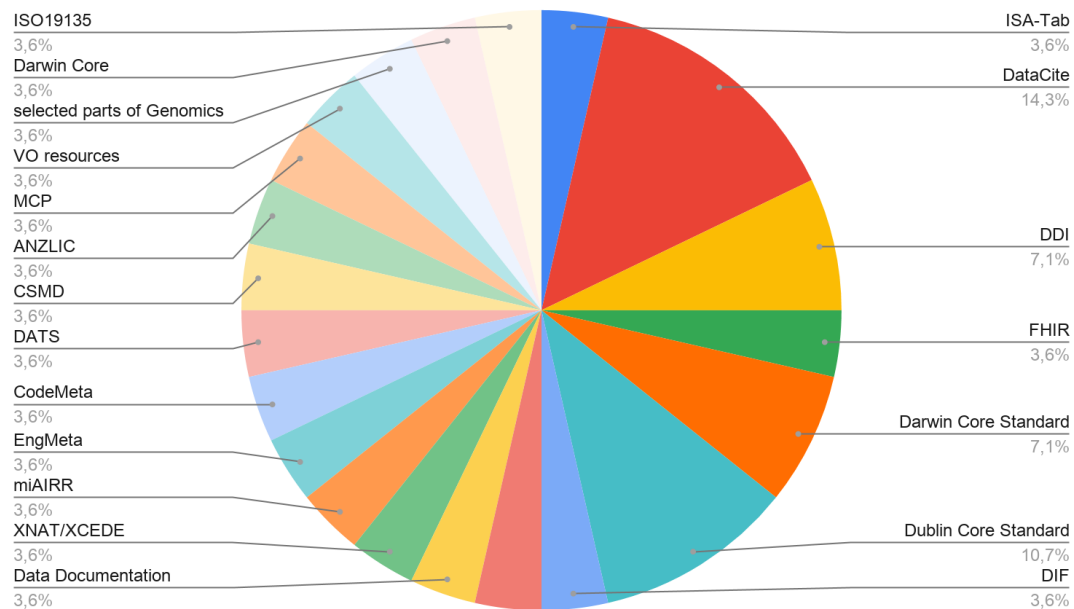


Does your Institution Data Repository use any standard or metadata scheme specific to your area of knowledge?



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

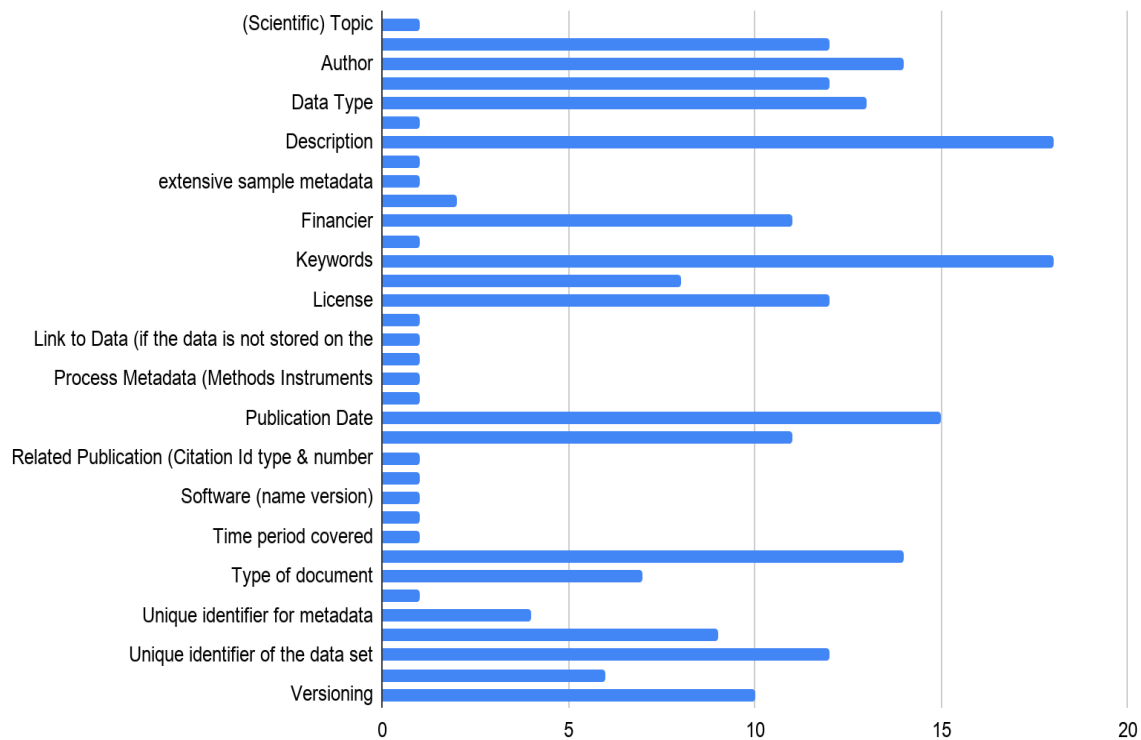
NITRC-IR

iReceptor

Bridges

GENERIC METADATA

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





Próximos passos...

- ✓ Dar continuidade ao levantamento dos softwares e plataformas para dados de pesquisa genéricos e do campo da saúde;
- ✓ Finalizar o mapeamento dos padrões de metadados do campo da saúde;
- ✓ Verificar se os padrões de metadados e os metadados específicos da Life Science e as plataformas estão alinhadas aos princípios FAIR.



Resultado esperado

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



Obrigada!

viviane.veiga@icict.fiocruz.br

simone.dib@icict.fiocruz.br

claudete.queiroz@icict.fiocruz.br

