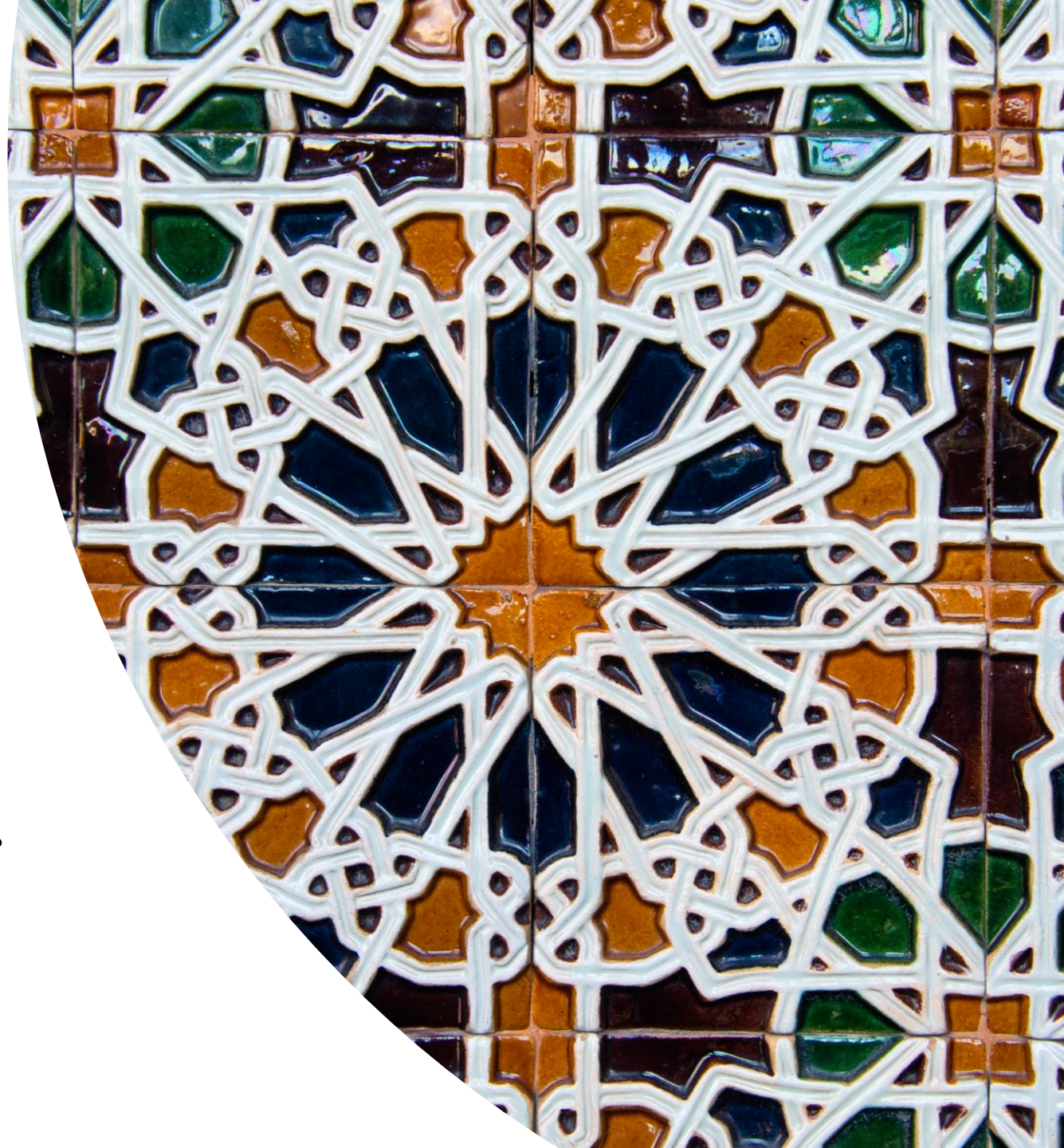


# Ciência Aberta na Fiocruz: novos desafios com implementação da Política de gestão, compartilhamento e abertura de dados

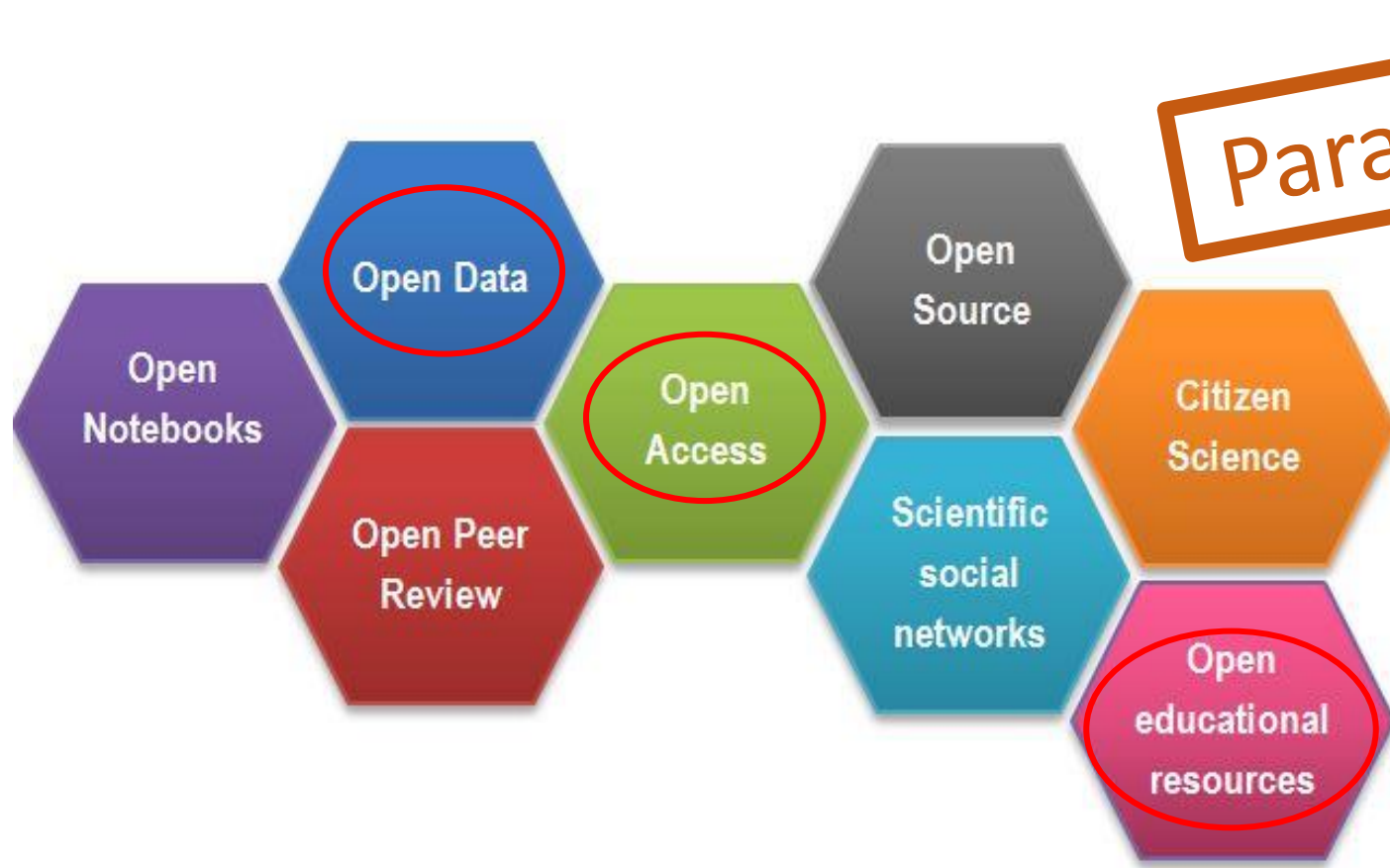
---

*Vice-Presidência de Educação, Informação e  
Comunicação/Coordenação de Informação e  
Comunicação*

Instituto René Rachou - Fiocruz Minas  
13 de Abril de 2021



# Contexto - Movimento da Ciência Aberta



Para que?

- ✓ Maior transparência;
- ✓ Ciência mais eficiente e com um maior fator de impacto;
- ✓ Mais colaborativa
- ✓ Velocidade nas respostas científicas
- ✓ ...maior visibilidade!!!

“**Ciência aberta** é, por sua vez, um termo **guarda-chuva**, que envolve **múltiplos níveis** e **escopos** de **abertura**”  
(ABLAGI; APPEL; MACIEL, 2014)

# ACESSO ABERTO

#

# CIÊNCIA ABERTA

O **acesso aberto (open access)** é o acesso imediato, gratuito, sem requisitos de registro, a materiais digitais educacionais, acadêmicos, científicos etc., principalmente os artigos de revistas científicas revisadas por pares.



<https://www.opensocietyfoundations.org/explainers/what-open-access>



“É uma mudança de paradigma na forma de fazer ciência”



“Aberta, colaborativa e feita com e para a sociedade”  
(ANLADA & ABADAL, 2018)

A **ciência aberta** vai além do **acesso aberto** à literatura científica e dados de pesquisa, incluindo a **abertura de todo o processo de pesquisa...**

# Por que compartilhar dados?

## Pesquisadores /instituições

- Maior visibilidade e impacto => citações dos dados
- Qualidade / melhores práticas
- Promove a colaboração

OPEN

## Revistas Científicas

- Transparência
- Ética e integridade científica
- Facilita a reprodutibilidade

OPEN

OPEN

## Agências de Fomento

- Aumenta a visibilidade e a reuso das pesquisas financiadas
- Maior impacto / retorno financeiro

OPEN

PGD

## Políticas Públicas

- Promove políticas públicas baseadas em evidências /dados
- Promoção dos direitos humanos e da democracia

OPEN

## Sociedade

- Transferência mais rápida de conhecimento
- Promove o avanço da ciência
- Participação social

OPEN

OPEN

Exigência de compartilhar/disponibilizar

OPEN

Incentivo por compartilhar/disponibilizar

PGD

Exigência de Plano de gestão de dados

# O QUE SÃO DADOS DE/PARA PESQUISA?

Os dados podem ser:

- **numéricos, qualitativos, descritivos ou visuais;**
- **em estado bruto ou analisado;**
- **experimentais ou observacionais.**



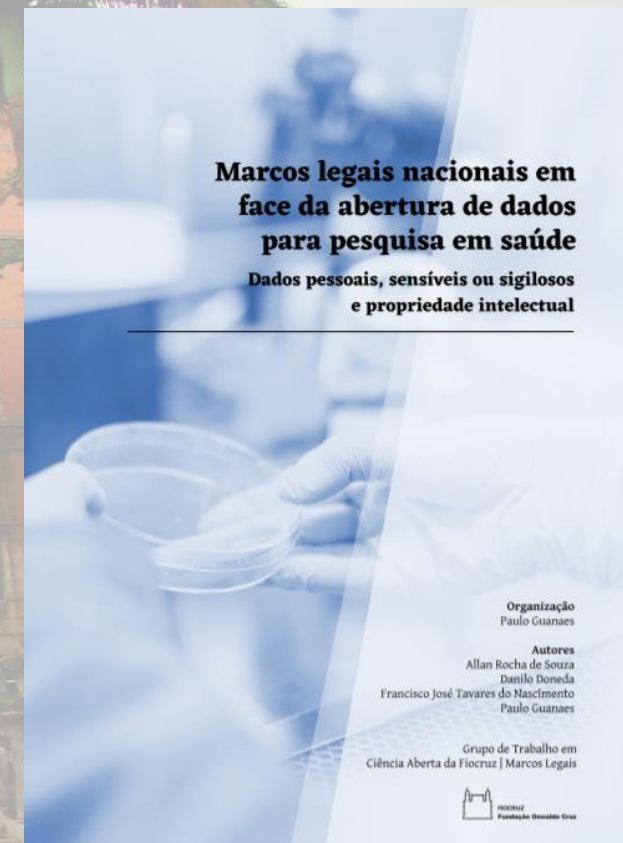
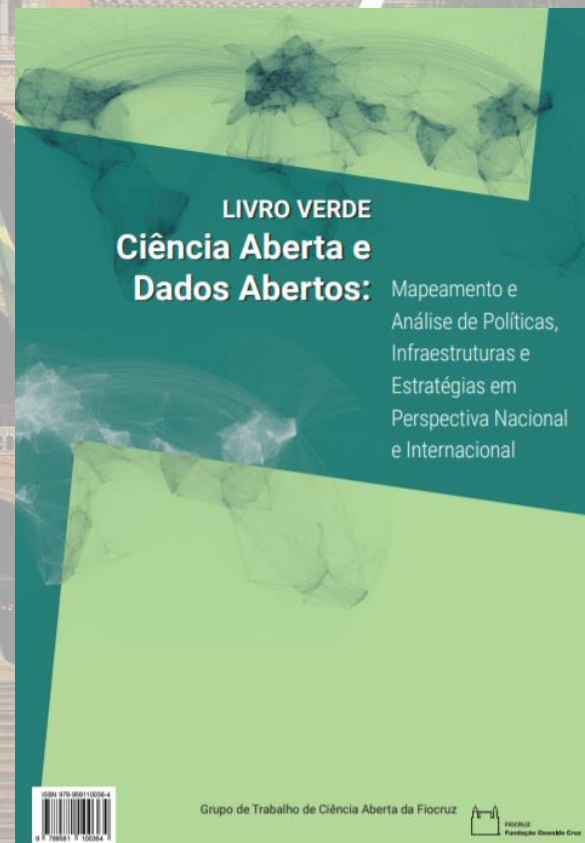
Boulton G (2012). *Open your minds and share your results*. Nature 486: 441. doi:10.1038/486441a

As informações sobre a **origem dos dados** também podem incluir: **como, quando, onde foi coletado e com que (por exemplo, instrumentos)**. O **código de software** usado para **gerar, comentar ou analisar** os dados também podem ser considerados **dados**.

Grupo de Trabajo de “Depósito y Gestión de datos en Acceso Abierto” del proyecto RECOLECTA. La conservación y reutilización de los datos científicos en España. Informe del grupo de trabajo de buenas prácticas [en línea] Madrid: Fundación Española para la Ciencia y la Tecnología, FECYT (2012) [Consulta 14/01/2013]. Disponible en WWW.FECYT.ES

*Como  
começamos?*

*Pesquisa inicial*



# ESTRATÉGIAS INICIAIS DA FIOCRUZ

## *Dados para pesquisa*


### Política institucional

- Termo de Referência
- Consulta interna
- Publicação
- Governança

### Capacitação

- Curso EAD
- Oficinas
- Disciplina transversal

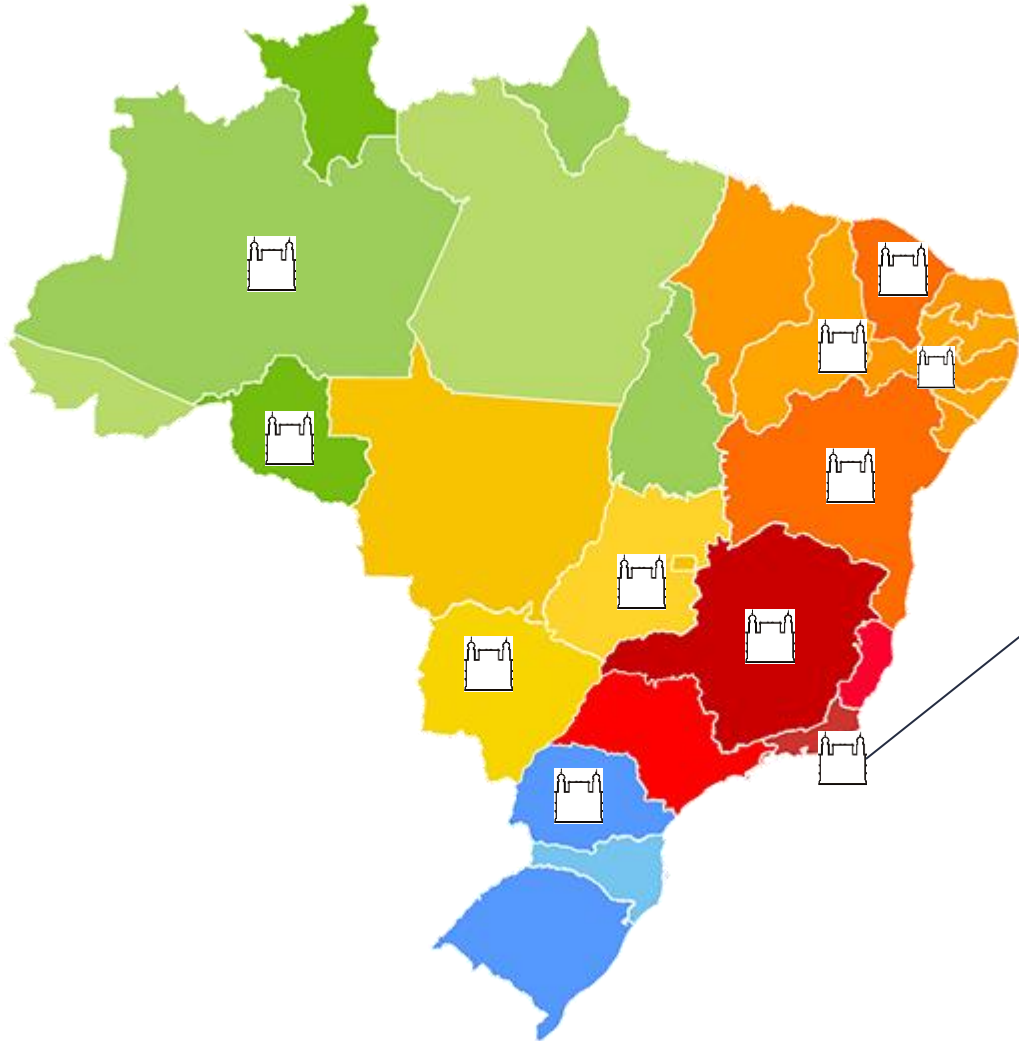
### Ferramentas e infraestrutura

- Plano de Gestão de Dados
  - Repositório de dados
- 

### Projetos pilotos

- Identificação de tipos de dados
- Aspectos éticos e jurídicos
- Fluxos

# TR e consulta interna pública



## Rio de Janeiro

**Bio-Manguinhos**

**ICTB**

**ICICT**

**IFF**

**INI**

**INCQS**

**IOC**

**COC**

**ENSP**

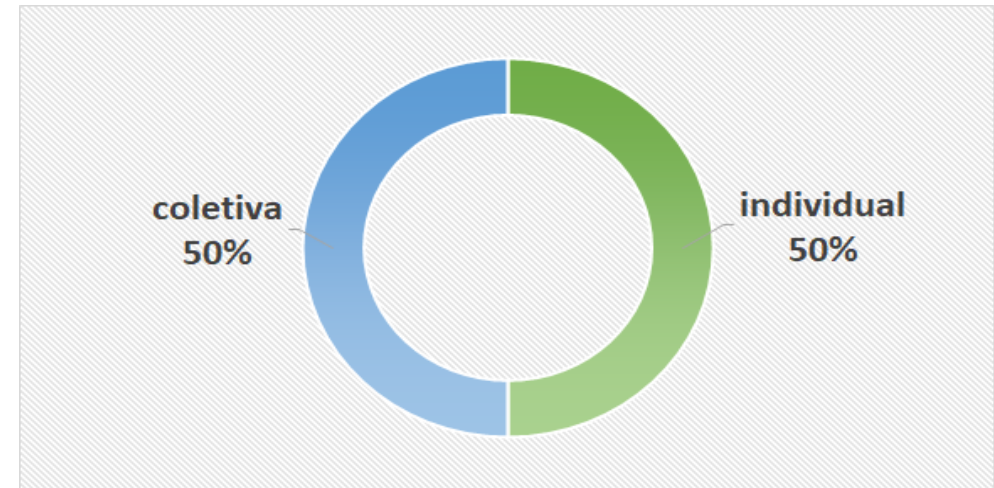
**EPSJV**

**Farmanguinhos**

**10 Fóruns internos  
(inclusive Câmaras  
Técnicas)**



286 contribuições





# Política de Gestão, compartilhamento e abertura de dados para pesquisa na Fiocruz

## Princípios e Diretrizes

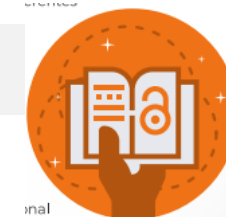
Aprovada no CD da  
Fiocruz em 24/09/2020

Princípio 1 – Interesse Público

Princípio 2 - Gestão, Compartilhamento e Abertura de Dados para Pesquisa

Princípio 3 - Marcos regulatórios

Princípio 4 – Desenvolvimento de Capacidades e Sustentabilidade



07/12/2020

SEI/FIOCRUZ - 0437852 - Portaria da Presidência



Ministério da Saúde

FIOCRUZ  
Fundação Oswaldo Cruz

Boletim de Serviço Eletrônico em 07/12/2020

#### Portaria da Presidência

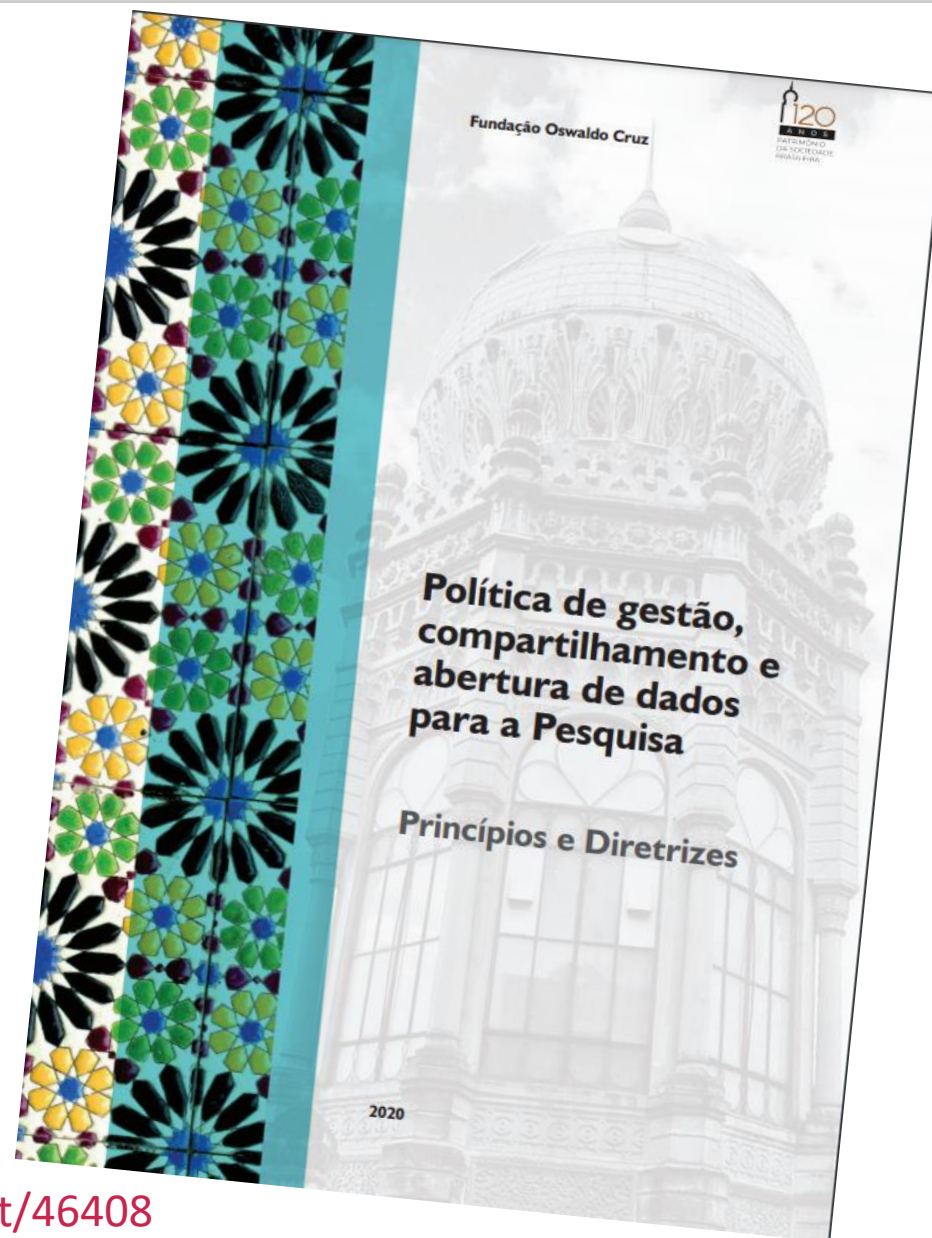
PORTARIA Nº 5760, de 07 de dezembro de 2020

A Presidente da Fundação Oswaldo Cruz, no uso das atribuições que lhe são conferidas pelo Decreto de 03 de janeiro de 2017, da Presidência da República e pelo Decreto nº 8.932, de 14 de dezembro de 2016 - Estatuto da Fiocruz,

#### RESOLVE:

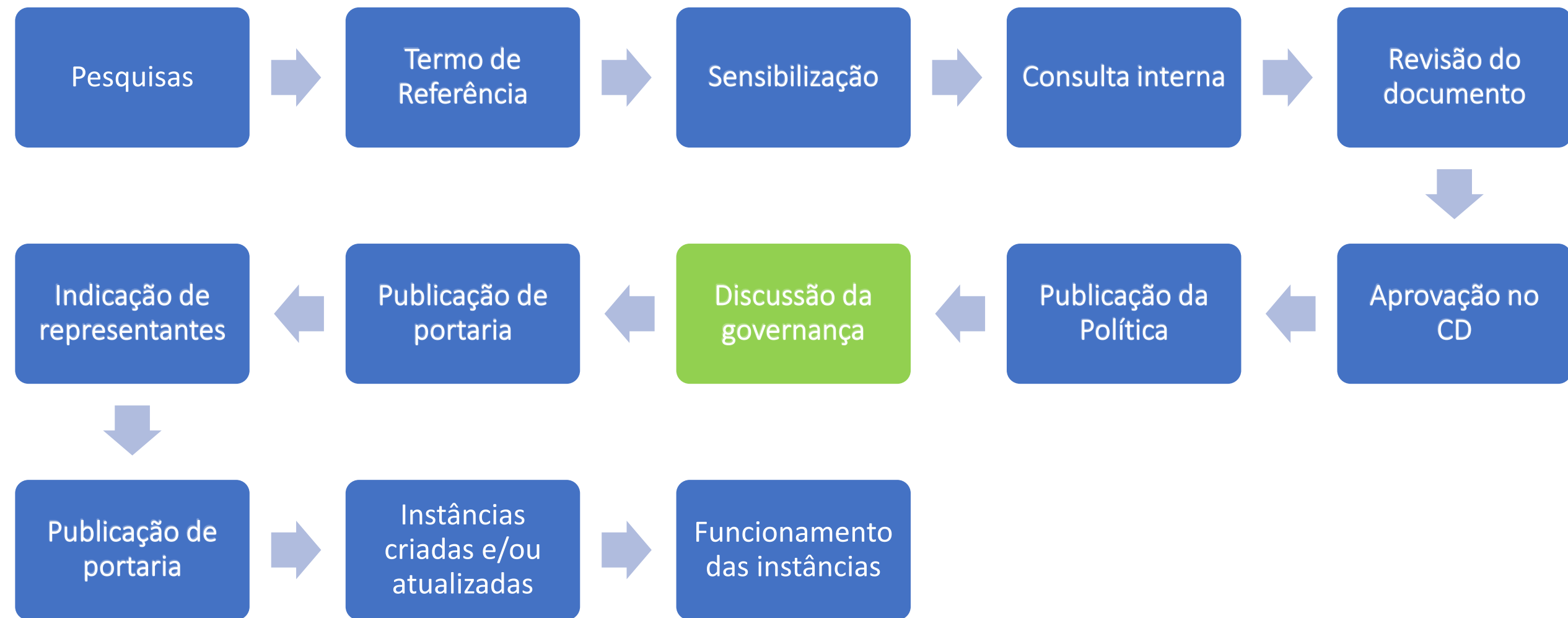
##### 1.0 - PROPÓSITO

Instituir a "Política de gestão, compartilhamento e abertura de dados para pesquisa", aprovada pelo Conselho Deliberativo da Fiocruz em reunião de 24 de setembro de 2020.



<https://www.arca.fiocruz.br/handle/icict/46408>

# Política de gestão, compartilhamento e abertura de dados para pesquisa da Fiocruz



# Capacitação

<https://campusvirtual.fiocruz.br/portal/>

- **Programa de Formação Modular sobre Ciência Aberta (curso EAD)**

Promoção: VPEIC, Escola Corporativa, Campus Virtual e Universidade do Minho.

\*(cerca de 18.690 alunos inscritos até fevereiro/2021)

- **Disciplina transversal aos Programas de Pós-Graduação**

Oferecida a partir de 2021

- **Oficinas**

A partir do lançamento das ferramentas

**FIOCRUZ CAMPUS VIRTUAL**

*Séries do Programa*

Série 1 (20h)	Série 2 (20h)	Série 3 (20h)	Série 4 (20h)
<b>Fundamentos da Ciência Aberta</b>	<b>Marcos Legais</b>	<b>Pesquisa Aberta</b>	<b>Educação Aberta</b>
Curso 1 (10h)   7 aulas O que é Ciência Aberta?	Curso 1 (10h)   4 aulas Propriedade Intelectual aplicada à Ciência Aberta	Curso 1 (10h)   4 aulas Acesso Aberto	Curso 1 (10h)   4 aulas Panorama da Educação Aberta
Curso 2 (10h)   7 aulas Panorama Histórico da Ciência Aberta	Curso 2 (10h)   5 aulas Dados Abertos Direito de Acesso à Informação e Proteção de Dados Pessoais	Curso 2 (10h)   5 aulas Dados Abertos	Curso 2 (10h)   4 aulas Recursos Educacionais Abertos
<a href="#">Saiba mais</a> <a href="#">Iniciar curso</a>	<a href="#">Saiba mais</a> <a href="#">Iniciar curso</a>	<a href="#">Saiba mais</a> <a href="#">Iniciar curso</a>	<a href="#">Saiba mais</a> <a href="#">Iniciar curso</a>

Logos at the bottom: Ministério da Saúde, FIOCRUZ Fundação Oswaldo Cruz, 120 ANOS PATRIMÔNIO DA SOCIEDADE BRASILEIRA, ESCOLA CORPORATIVA FIOCRUZ, Universidade do Minho.

# Plano de Dados da Fiocruz



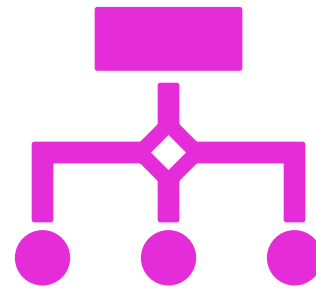
Sistema desenvolvido por:



# Repositório institucional de dados para pesquisa da Fiocruz



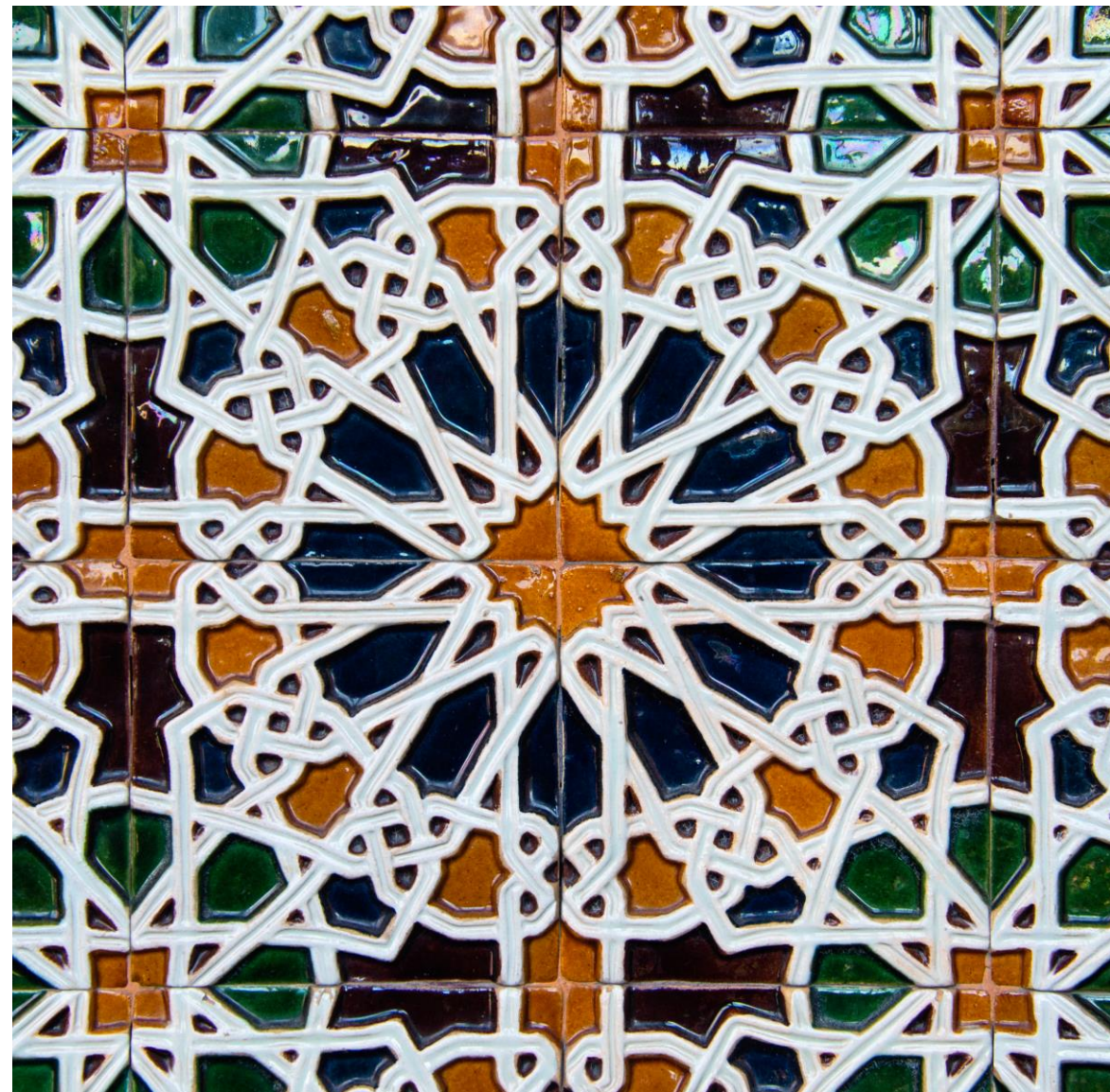
Fase 1 – Funcionalidades e aspectos tecnológicos da ferramenta Dataverse



Fase2 – Depósito e publicação, documentos e fluxos para gestão, compartilhamento e abertura dos dados para pesquisa da Fiocruz



## Fase 1 – DATAVERSE



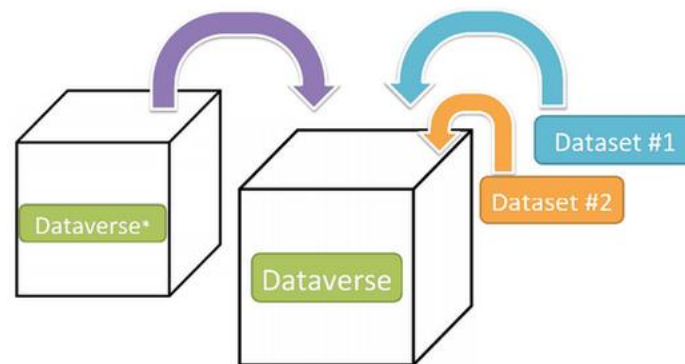


## Ferramenta Escolhida

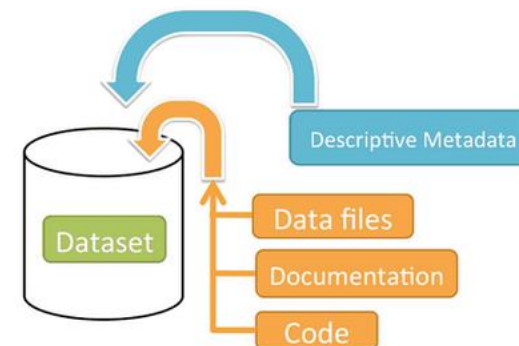
Lançado em **2006** pelo Harvard's Institute for Quantitative Social Science: **15 anos de Dataverse**

O **Dataverse** é um **Repositório de Dados** de código aberto para **compartilhar, preservar, citar, explorar e analisar dados de pesquisa**. **Pesquisadores, autores de dados, editores, distribuidores de dados e instituições afiliadas** recebem o **crédito apropriado** por meio de uma **citação de dados** com um **identificador persistente** (p. ex.: DOI ou Handle). Cada **conjunto de dados** contém **metadados descritivos** e **arquivos de dados** (incluindo documentação e código que acompanham os dados) (DATAVERSE PROJECT, 2018, tradução nossa).





Container for your **Datasets** and/or **Dataverses\***



Container for your data, documentation, and code.

\* Dataverses can now contain other Dataverses (this replaces Collections & Subnetworks)

## Harvard Dataverse é o **maior** site de repositórios de dados com:

- Código aberto;
- Permite o versionamento, armazenamento, restrição, compartilhamento e abertura do conjunto ou de arquivos específicos;
- Permite criar conexões em diferentes coleções;
- Permite citar, explorar e analisar dados de pesquisa;
- Permite replicar o trabalho de outras pessoas;
- Permite interoperabilidade com diversos softwares já utilizados na pesquisa.



Fonte: <https://dataverse.harvard.edu/>



**14.016 arquivos** disponibilizados no Dataverse se constituem em **documentos** relacionados aos **107.704 “conjuntos de dados”** de pesquisa

**Documentação**

**Dados de Pesquisa**

**Códigos**

**HARVARD Dataverse** Search - About User Guide Support Sign Up Log In

Virus Epidemiology and Control (VEC) Dataverse (Kemri Wellcome Trust Research Programme, Kilifi, Kenya) Population dynamics of viral pathogens informing intervention strategies

Harvard Dataverse > KWTRP Research Data Repository > Virus Epidemiology and Control (VEC) Dataverse > **Replication Data for: Whole genome sequencing and phylogenetic analysis of Human metapneumovirus strains from Kenya and Zambia**

<input type="checkbox"/>	<b>EKamau_HMPV_WGS_Readme.txt</b> Plain Text - 4.5 KB - Aug 5, 2019 - 0 Downloads MD5: 94e1f85ded6a0a8b4e99f460ba7de65f Dataset readme file <b>Documentation</b>	<input type="button" value="Download"/>
<input type="checkbox"/>	<b>Identity_graph_HMPVA_Ggene.csv</b> Comma Separated Values - 3.2 KB - Aug 5, 2019 - 0 Downloads MD5: 85b9d82a093f56f425a618da56dbba64 <b>Data</b>	
<input type="checkbox"/>	<b>Identity_graph_HMPVA_SHgene.csv</b> Comma Separated Values - 2.4 KB - Aug 5, 2019 - 0 Downloads MD5: 5eec8e812e0c9cdd1a81e7d31a7cf551 <b>Data</b>	
<input type="checkbox"/>	<b>Identity_graph_HMPVB_Ggene.csv</b> Comma Separated Values - 3.6 KB - Aug 5, 2019 - 0 Downloads MD5: 991131141a43d62276cd3083fd78a7d9 <b>Data</b>	
<input type="checkbox"/>	<b>Identity_graph_HMPVB_SHgene.csv</b> Comma Separated Values - 2.6 KB - Aug 5, 2019 - 0 Downloads MD5: c8a3d807c5e88443678bcf3b68291802 <b>Data</b>	<input type="button" value="Download"/>
<input type="checkbox"/>	<b>script_2Jul2019.R</b> R Syntax - 3.0 KB - Aug 5, 2019 - 0 Downloads MD5: 64531365d4f6caaeaf95549d170fdccd Replication code in R <b>Code</b> @mercecrosas	<input type="button" value="Download"/>

**Dataset**

**Descriptive Metadata**

**Data files**

**Documentation**

**Code**

**Access**  
Public (897,112)  
Restricted (43,370)

17

# Fase 1 - Repositório de dados

Compromisso 3 que tem como propósito “estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil”



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA,  
INOVAÇÕES E COMUNICAÇÕES



OPEN KNOWLEDGE



# Linha do tempo – Fase 1



**5.451 *datasets* da Fiocruz disponíveis em repositórios internacionais**

Pesquisa em 10/02/2020 no Portal OpenAIRE

**Convite em julho de 2019**

**Início em agosto 2019**

**Atualização do planejamento em outubro de 2019**

**Entrega em Abril de 2020**

*Decisão de usar dados disponíveis em repositórios de dados multidisciplinares*



 **arca** | Repositório Institucional da Fiocruz



**zenodo**



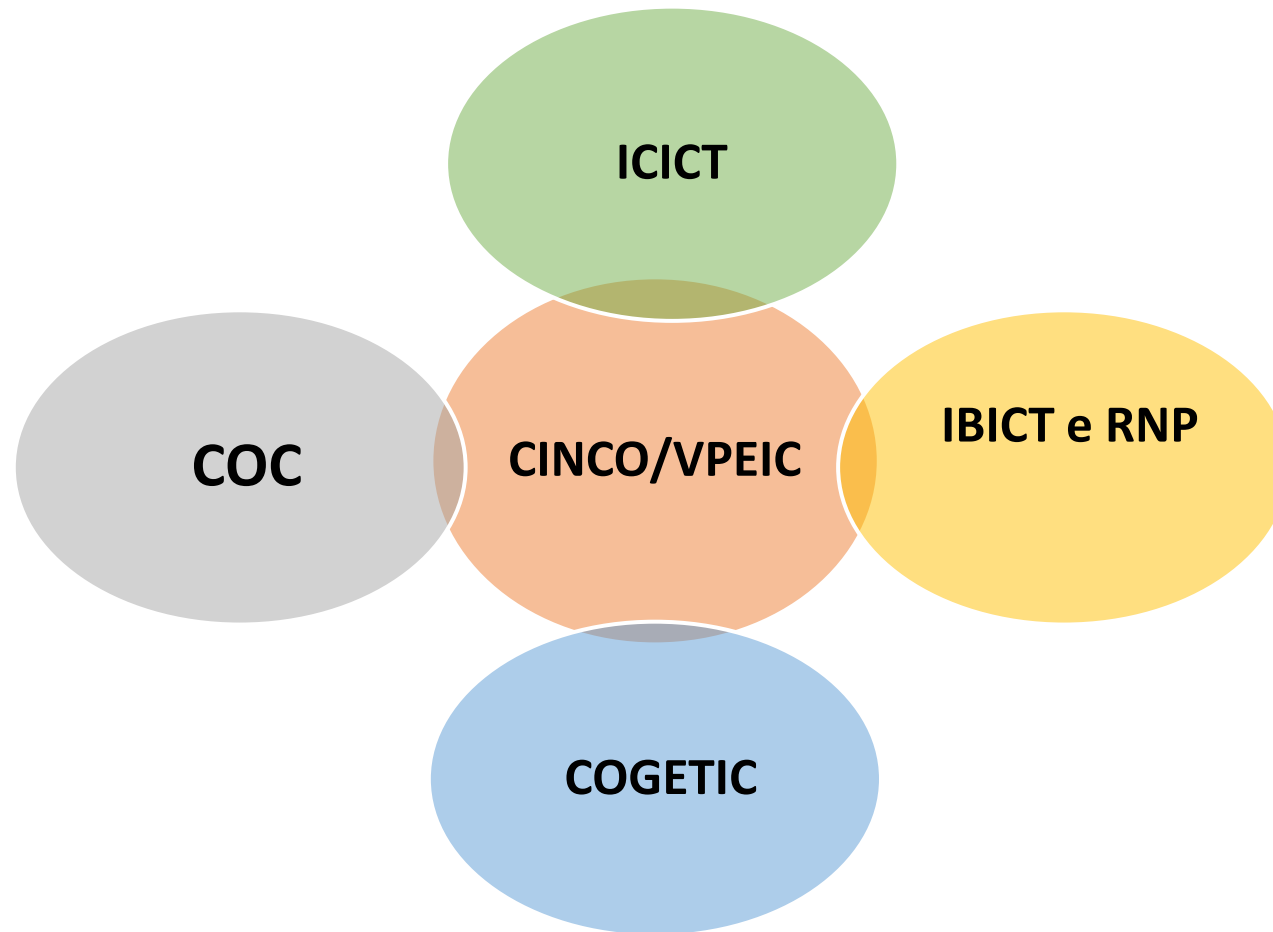
 **figshare**



**239**

Datasets

# Composição de Grupo de Trabalho - Fase 1





# Testes realizados - Fase 1



**Cadastro e  
Importação  
dos dados**



**Links de  
datasets**



**Metadados  
padrão**



**Tamanho  
dos  
arquivos**



**Perfis**



**Publicação,  
embargo e  
acesso**



**Remoção e  
restrição de  
arquivos**



**Segurança**



**Autenticação institucional  
via Shibboleth**



**Interoperabilidade  
com Archivematica**



**Coleta de dados  
pelo OASISBR**

# Principais destaques

Dataverse de Teste para Validação (FUNDAÇÃO OSWALDO CRUZ)

FIOCRUZ > Dataverse de Teste para Validação >

## (TESTE DO DATAVERSE) Modelos matemáticos aplicados a Sars-CoV-2: Plataforma interativa para avaliação de cenários e modelos matemáticos aplicados a Sars-CoV-2

Version 1.1

Citação



Oliveira, Juliane Fonseca de, 2021, "(TESTE DO DATAVERSE) Modelos matemáticos aplicados a Sars-CoV-2: Plataforma interativa para avaliação de cenários e modelos matemáticos aplicados a Sars-CoV-2", <https://doi.org/10.80458/OWDUC9>, FIOCRUZ, V1

Cite Dataset ▾

Learn about Data Citation Standards.

Identificador persistente

Versionamento

Access Dataset ▾

Contact Owner

Share

Dataset Metrics ?

Files

Metadata

Terms

Versions

1 Download ?

Search this dataset...

Find

Filter by

File Type: All ▾

Access: All ▾

File Tag: All ▾

1 to 2 of 2 Files



N1.2.pdf

Adobe PDF - 287.2 KB - Jan 28, 2021 - 1 Download  
MD5: e8afb310a86ce1c3dc2745677d154a94



N1.pdf

Adobe PDF - 287.2 KB - Jan 28, 2021 - 0 Downloads  
MD5: e8afb310a86ce1c3dc2745677d154a94

Tag PDF Documentação

Controle do tipo de acesso

Files Metadata Terms Versions

Todas as informações do dataset

Dataset	Summary	Contributors	Published
1.1	<a href="#">View Details</a>	Hataanderson Luiz Cabral dos Santos	Jan 28, 2021
1.0	This is the first published version.	Hataanderson Luiz Cabral dos Santos	Jan 28, 2021

# Tipos de acesso no Arca Dados

---







# Entregas realizadas - Fase 1

---

*Fluxos de depósito e publicação*

---

*Estratégias segurança do repositório*

---

*Estratégias para preservação dos dados*

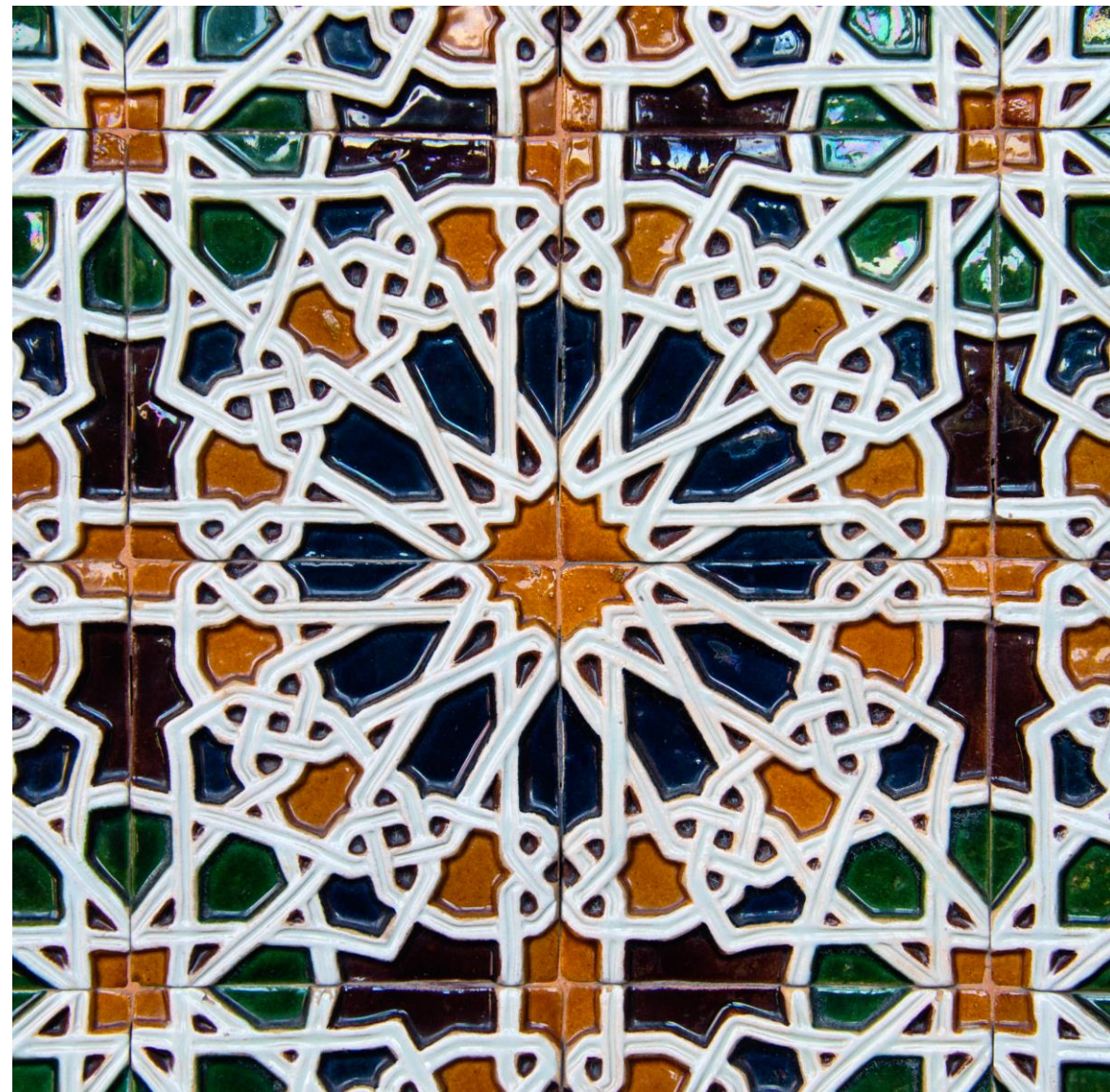
---

*Aquisição do DOI via CoNciência (CNPq)*

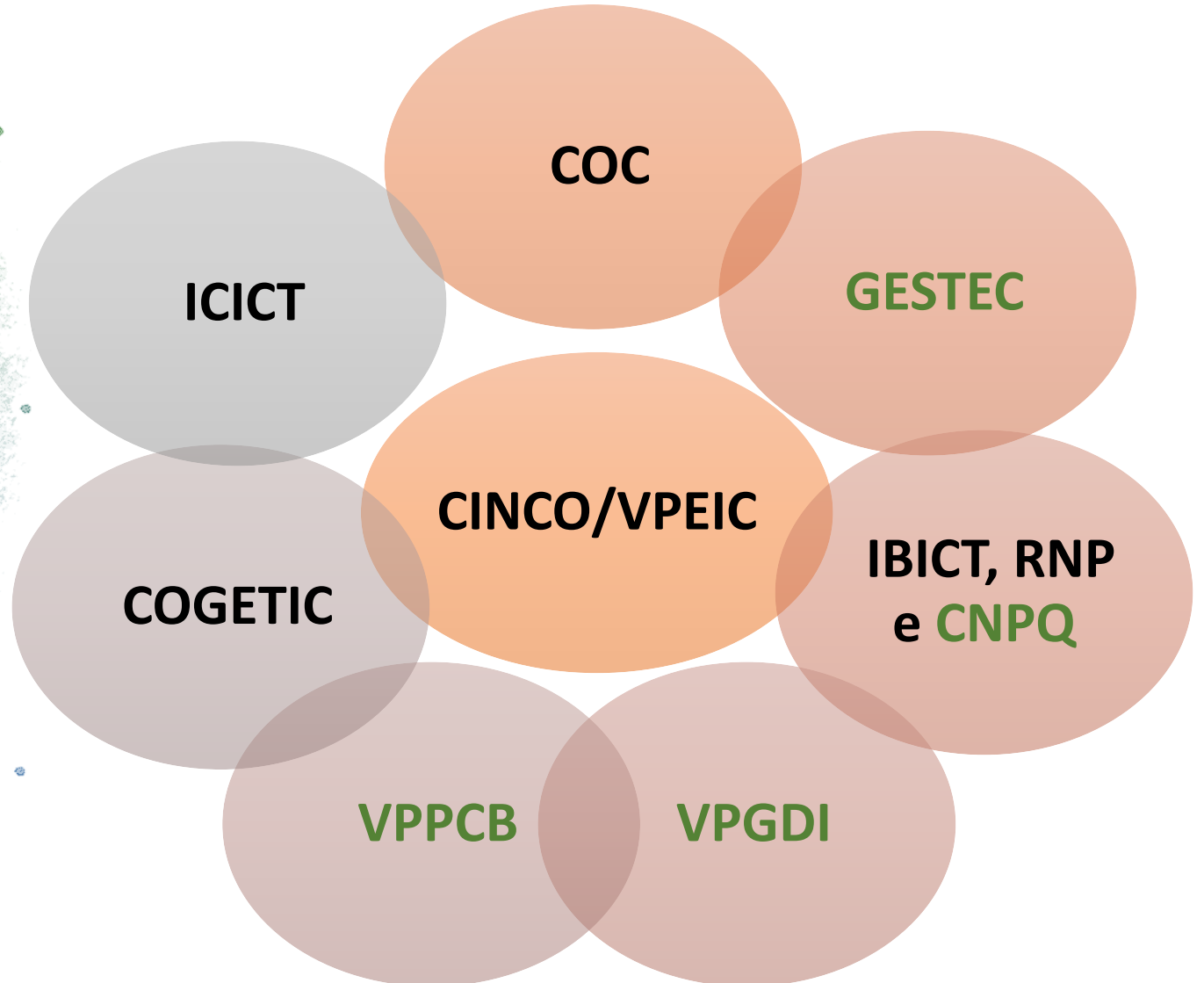
---

Relatório Final

Fase 2 – ARCA DADOS  
e processos  
necessários para  
depósito e publicação  
de dados



# Fase 2 – Composição do Grupos de Trabalho



Preenchimento do Plano de Gestão de Dados

Identificação e análise dos tipos de dados de pesquisa

Levantamento da conformidade em aspectos éticos e jurídicos

Avaliação se dados tem potencial de patentes

Infraestrutura tecnológica e segurança da informação

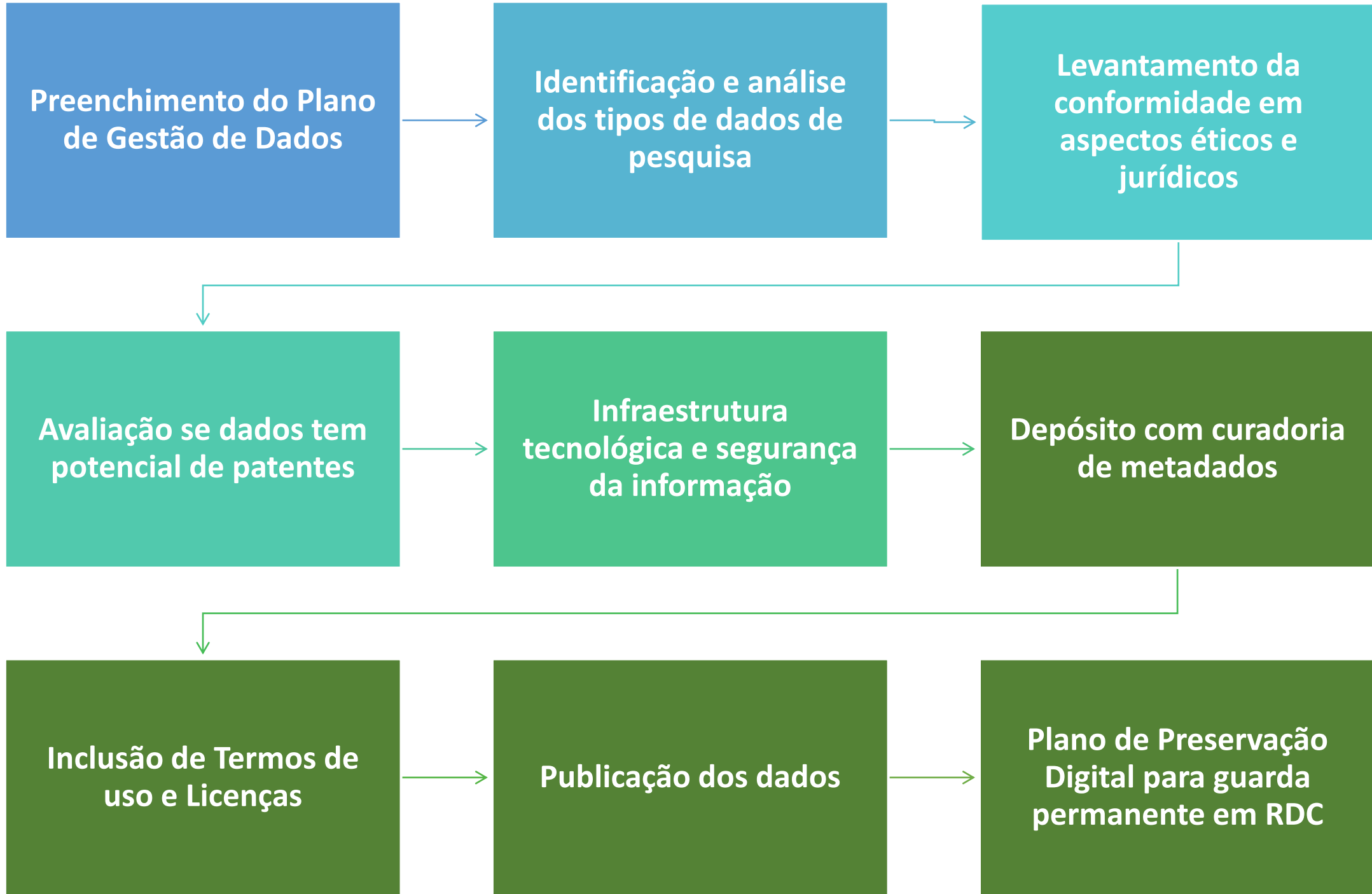
Depósito com curadoria de metadados

Inclusão de Termos de uso e Licenças

Publicação dos dados

Plano de Preservação Digital para guarda permanente em RDC

## Etapas do Piloto – Fase 2





# Temáticas dos Pilotos



**Saúde Mental – “Morar em Liberdade”- Luta Antimanicomial**

**Vigilância da Doença de Chagas em Minas Gerais – Programa de controle da Doença de Chagas**

**Rede de Pesquisa Clínica e Aplicada em Chikungunya (Replick) - Consórcio de estudos clínicos aplicados**

**Monitoramento de casos de síndrome respiratória aguda grave (SRAG) notificados no SIVEP-Gripe - Infogripe**

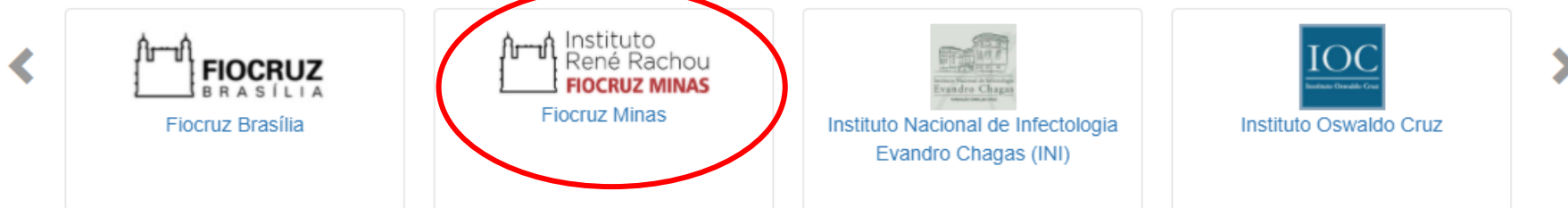
Fundação Oswaldo Cruz (FIOCRUZ) Fiocruz

# ARCA DADOS

Metrics 91 Downloads

Contact Share Edit

Criada em 25 de maio de 1900 - com o nome de Instituto Soroterápico Federal -, a Fundação Oswaldo Cruz (Fiocruz) nasceu com a missão de combater os grandes problemas da saúde pública brasileira. Hoje, a instituição, vinculada ao Ministério da Saúde, tem como missão "promover a saúde e o desenvolvimento social, gerar e difundir conhecimento científico e tecnológico, ser um agente da cidadania".



Search this dataverse...

Find Advanced Search

+ Add Data

Dataverses (22)

Datasets (668)

Files (192,835)

**Dataverse Category**

Organization or Institution (11)

Research Project (4)

Laboratory (3)

Research Group (2)

Department (1)

**Publication Year**

2021 (12)

2020 (19)

1 to 10 of 690 Results

Sort

Fiocruz Bahia (FUNDAÇÃO OSWALDO CRUZ)

Apr 9, 2021



A Fiocruz Bahia tem como missão promover a melhoria da qualidade de vida da população através da geração e difusão de conhecimento científico e tecnológico no estado da Bahia e no Brasil.

Fiocruz Amazônia (FUNDAÇÃO OSWALDO CRUZ)

Apr 9, 2021



O Instituto Leônidas & Maria Deane (ILMD/Fiocruz Amazônia) é a unidade técnico-científica da Fundação Oswaldo Cruz no Amazonas. Sediado em Manaus, sua missão é contribuir para a melhoria das condições de vida e saúde das populações amazônicas e para o desenvolvimento científico e...

Domingos Sávio. Extrato de Entrevista - Memórias da Saúde Mental

Draft Unpublished

Apr 8, 2021 - Sujeitos Históricos da Reforma Psiquiátrica no Brasil

https://dadosdepesquisa.fiocruz.br/



# Desafios para o Repositório de Dados da Fiocruz

## Depósito

Tipos de dados ✓

Padrões de metadados ✓

Vocabulários temáticos ✓

Interoperabilidade ✓

Integração com plataformas ✓

Licenças, cessão e termos de uso ✓

## Publicação

Aplicação do DOI ✓

Aberto ✓

Compartilhado ✓

Fechado ✓

## Preservação digital

Prazos de guarda para os dados

Integração com Archivemática para os dados com longos prazos de guarda ✓

## Certificação

Core Trust Seal

## Acesso

Citação ✓

Indicadores

Métricas ✓

Onde estamos  
na  
implementação  
do Arca Dados?

Finalização das pendências da Fase 2 com produção de documentos de apoio

Discussão da governança e fluxos do Arca Dados para lançamento na Fiocruz

D2.2 - Apoiar, de forma prioritária, os projetos de pesquisa sobre os quais recaem exigências dos órgãos financiadores sobre gestão de dados para pesquisa, e a solicitação de dados por revistas científicas.

Retorno as unidades para apresentar os avanços do trabalho



# Como solicitar este apoio?



Envie um e-mail para  
[dadosabertos@fiocruz.br](mailto:dadosabertos@fiocruz.br)



*Obrigada!*

[vanessa.jorge@fiocruz.br](mailto:vanessa.jorge@fiocruz.br)  
[cinco@fiocruz.br](mailto:cinco@fiocruz.br)