



Seminário FAU-USP
GESTÃO DE DADOS CIENTÍFICOS: POR
ONDE COMEÇAR?

**Gestão,
compartilhamento e
abertura de dados para
pesquisa: as estratégias
da Fiocruz**

26 de outubro de 2020
Vanessa de Arruda Jorge
vanessa.jorge@fiocruz.br



PATRIMÔNIO
DA SOCIEDADE
BRASILEIRA

Conhecimento e informação científica

1
9
0
1

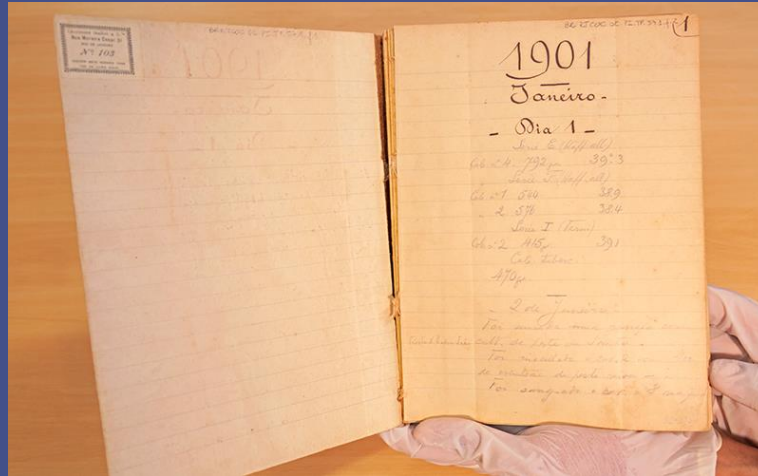


Foto: Acervo COC

1
9
0
9

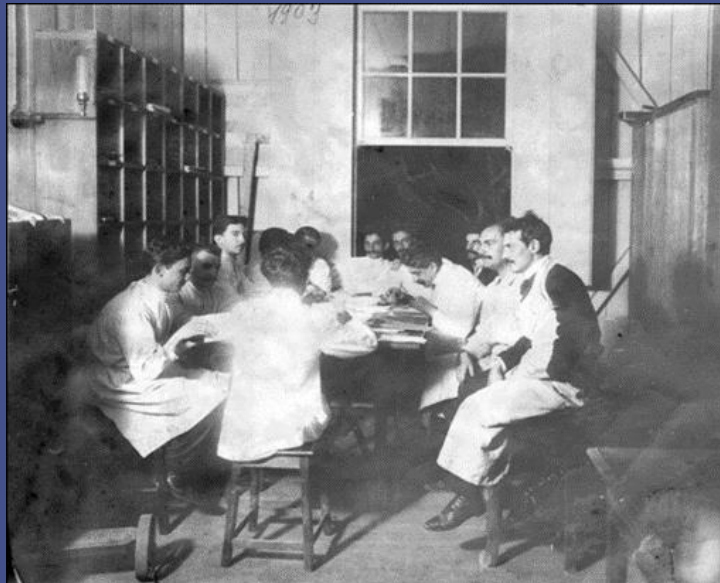


Foto: Acervo COC

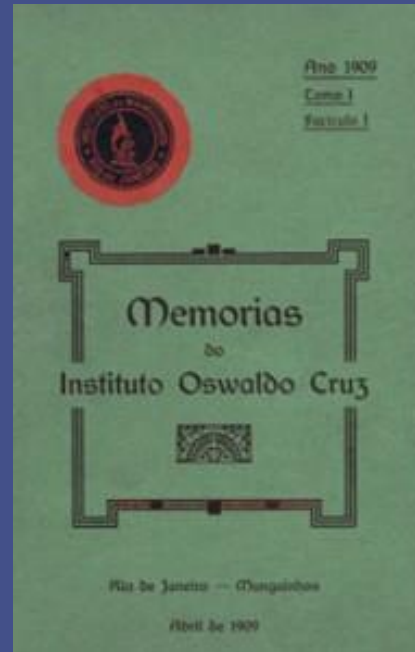


Foto: Gutemberg Brito / IOC

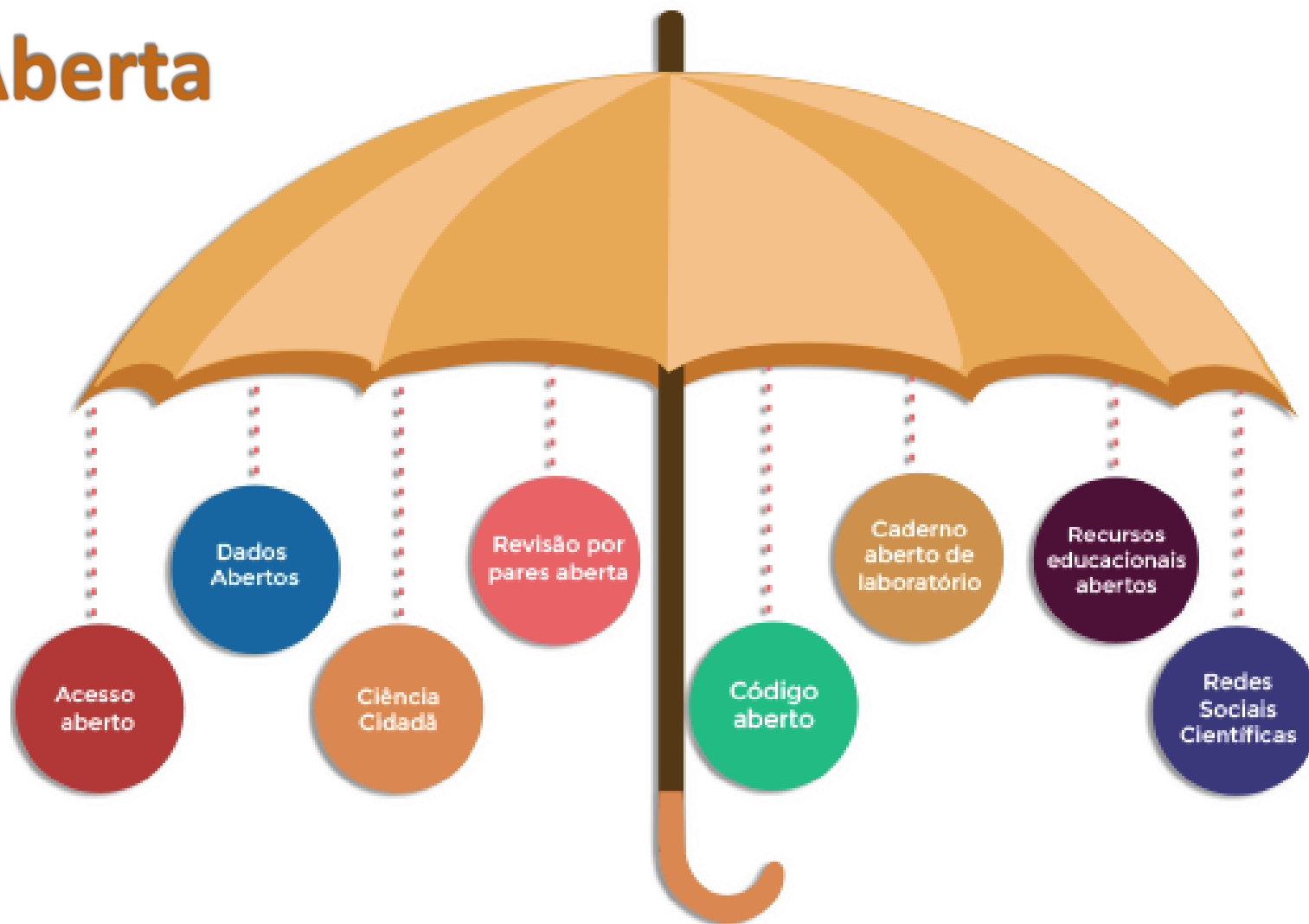
Rede de Bibliotecas da Fiocruz



Instituto de Comunicação e Informação Científica e Tecnológica em Saúde



Ciência Aberta



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

Construção coletiva da Política de Acesso Aberto ao Conhecimento na Fiocruz



2010

VI Congresso Interno

6) Priorizar a política de “acesso livre” na gestão da informação e do conhecimento produzido na Fiocruz

2011

2012

GT Multidisciplinar
Portaria 795/2012-PR

Consulta interna

2013

Debate nas Câmaras Técnicas

Aprovação dos princípios norteadores no Conselho Deliberativo (CD) Fiocruz

2014

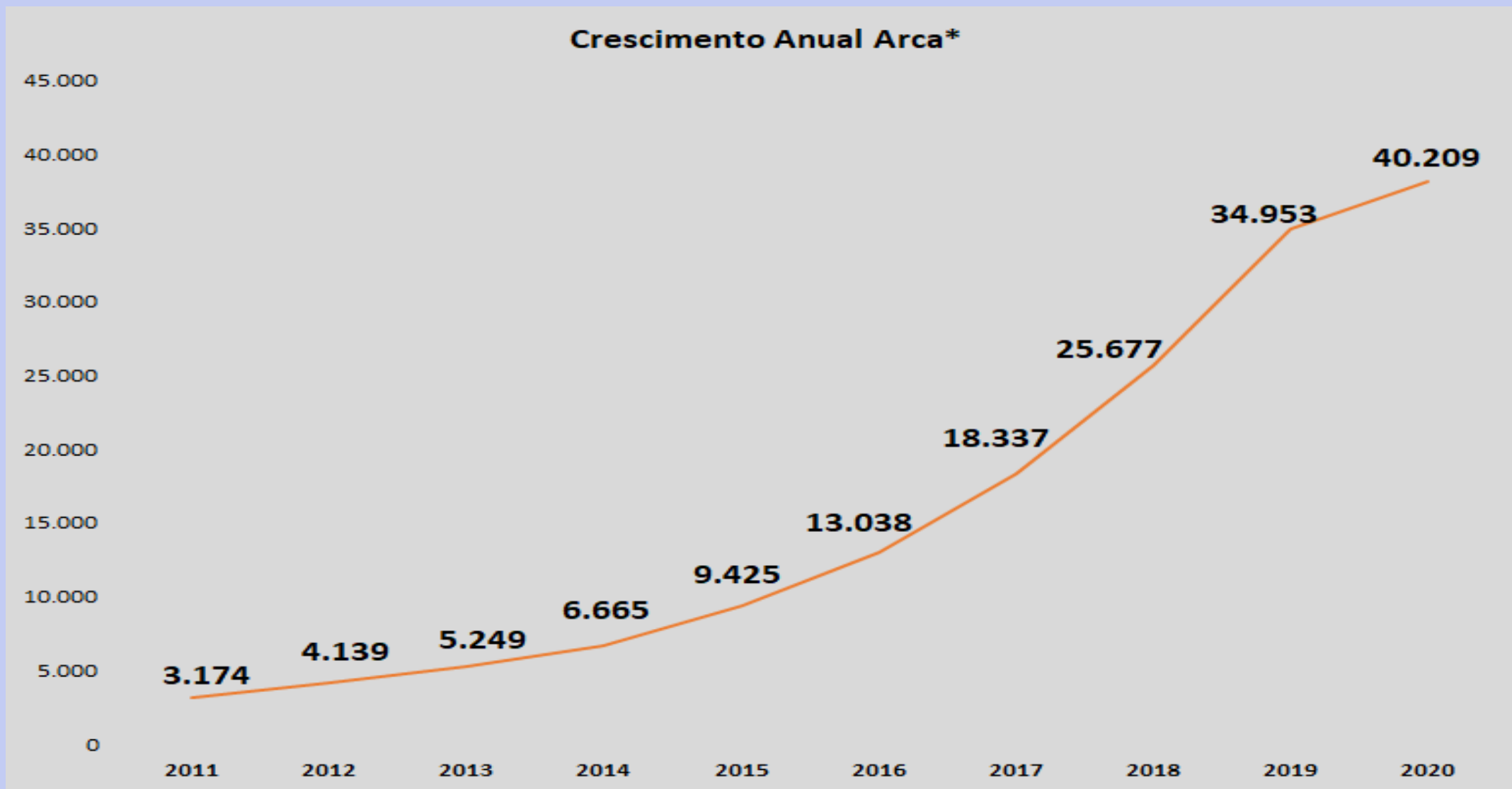
Aprovação da Política CD Fiocruz

Publicação de Portarias
329/2014-PR retificada por
382/2014-PR (Política)
501/2014-PR (Governança)

Plano Operativo do Repositório Institucional Arca



Implementação da política de **ACESSO ABERTO** da Fiocruz





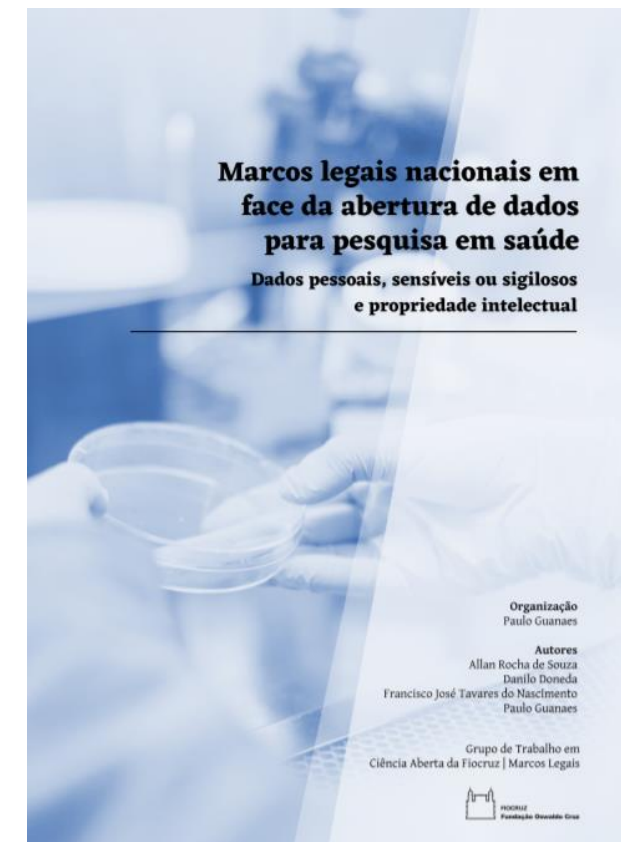
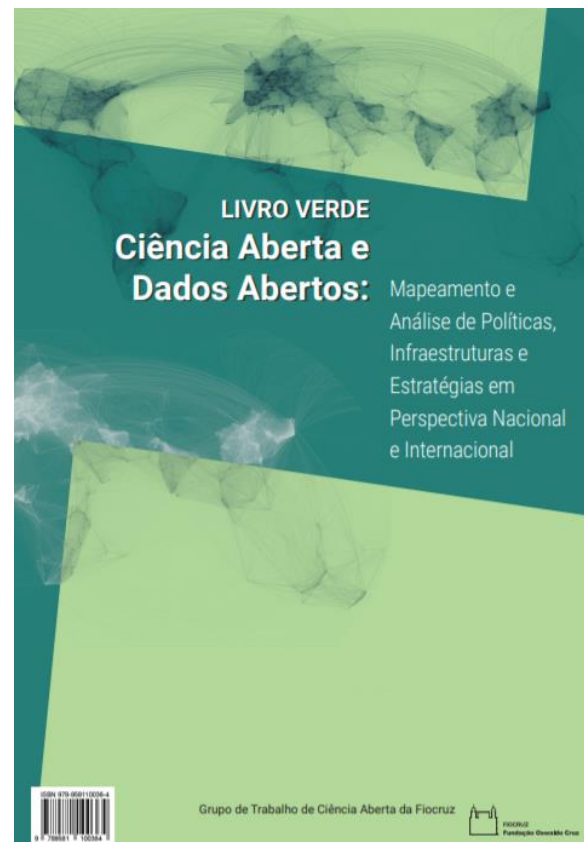
*Gestão,
Compartilhamento e Abertura de
dados na Fiocruz*

Por onde começamos?

Criação de Grupo de Trabalho em Ciência Aberta

- ◊ **Coordenado pela Vice-presidência de Educação, Informação e Comunicação**
- ◊ **Financiado inicialmente pelo CIDACS/Fiocruz**
- ◊ **Multidisciplinar (bibliotecários, arquivistas, TI, advogados, representantes da área da Qualidade, gestores e pesquisadores de diversas áreas)**
- ◊ **Articula-se com pontos focais indicados em todas as unidades técnico-científicas da Fiocruz**
- ◊ **A partir de 2019, pesquisadores de diferentes unidades passaram a participar do GTCA Ampliado**

Pesquisa inicial



Estratégias iniciais

Política institucional

- TR
- Consulta interna

Capacitação

- Curso EAD
- Oficinas
- Disciplina transversal

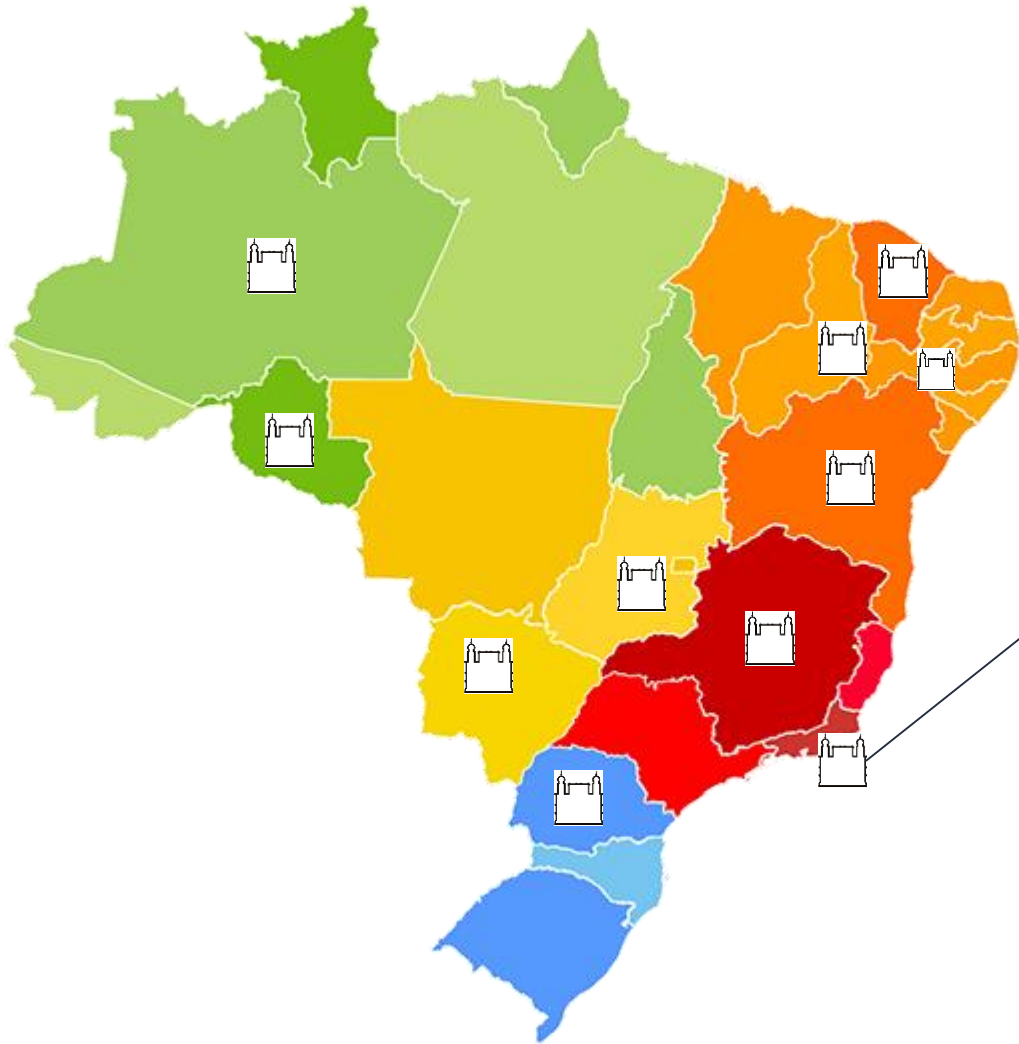
Ferramentas e infraestrutura

- PGD
- 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

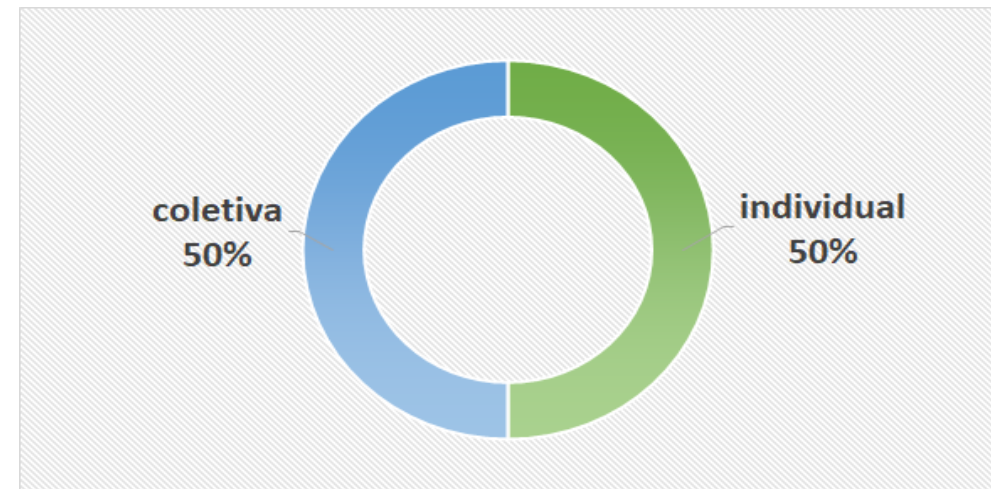
EPSJV

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



Capacitação

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

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

*(cerca de 16.200 mil alunos inscritos até setembro/2020)

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

Oferecida a partir de 2020

- **Oficinas para os diversos públicos**

Formação de formadores – Universidade do Minho

FIOCRUZ CAMPUS VIRTUAL

Séries do Programa

Série 1 (20h)	Série 2 (20h)	Série 3 (20h)	Série 4 (20h)
Fundamentos da Ciência Aberta	Marcos Legais	Pesquisa Aberta	Educação Aberta
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

Ministério da Saúde
FIOCRUZ
Fundação Oswaldo Cruz
Coordenação-Geral de Gestão de Pessoas

FIOCRUZ 120 ANOS
PATRIMÔNIO DA SOCIEDADE BRASILEIRA

ESCOLA CORPORATIVA FIOCRUZ

Universidade do Minho

Plano de Dados da Fiocruz



Sistema desenvolvido por:



Piloto RNP e IBICT - Repositório de dados

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



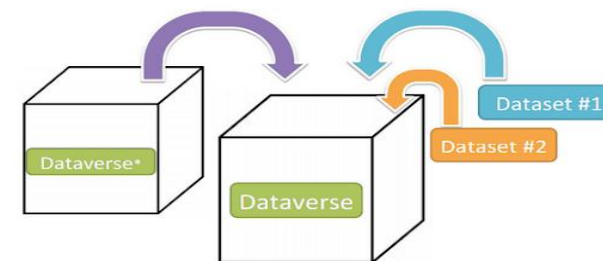
MINISTÉRIO DA
CIÊNCIA, TECNOLOGIA,
INOVAÇÕES E COMUNICAÇÕES



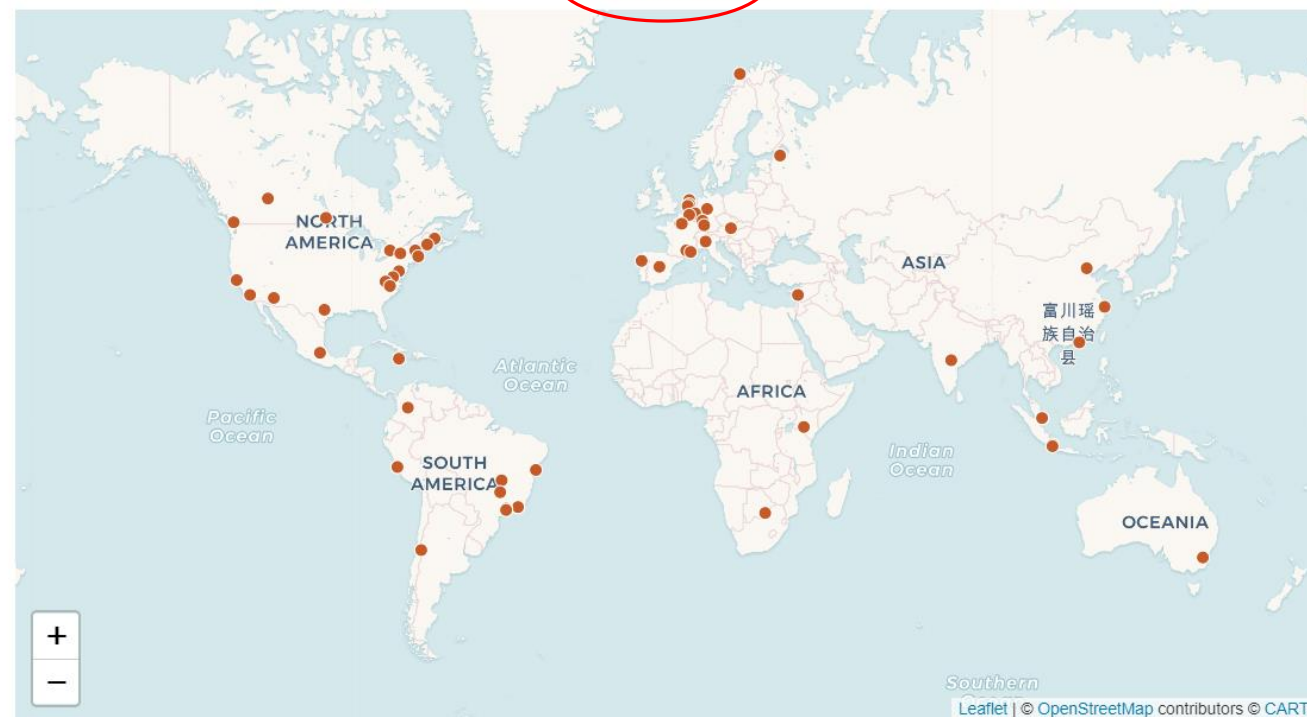
Algumas vantagens do software

- ❖ Código aberto
- ❖ Permite o versionamento, 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.

Dataverse



63 Installations





8.000 arquivos disponibilizados no Dataverse se constituem em **documentos** relacionados aos **98.000 “conjuntos de dados”** de pesquisa

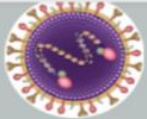
Documentação

Dados de Pesquisa







Códigos

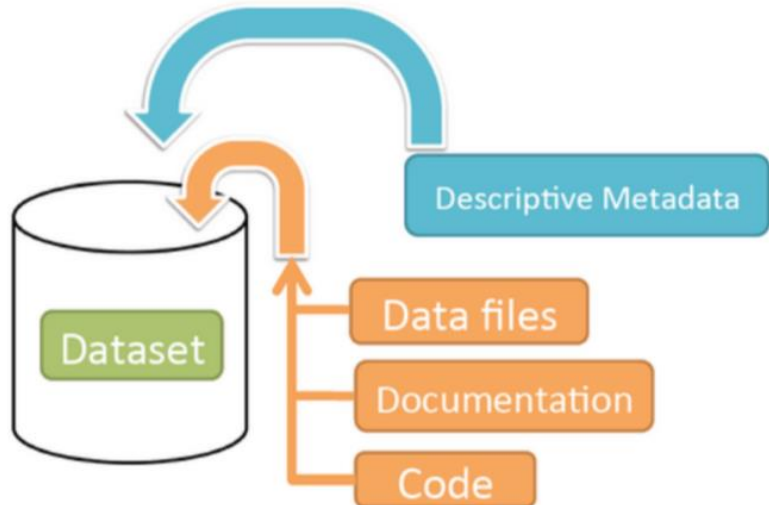
HARVARD
Dataverse

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



17

Linha do tempo



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



239

Datasets



Testes realizados



**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**



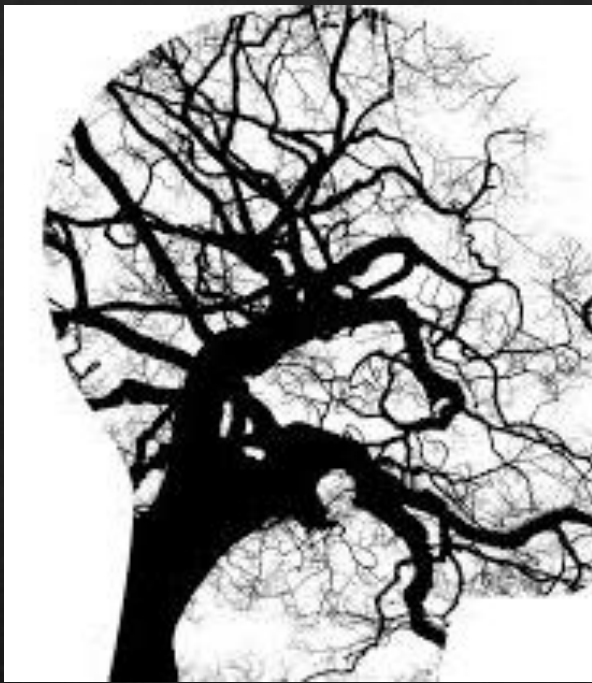
**Interoperabilidade
com Archivematica**



**Coleta de dados
pelo OASISBR**

Fiocruz - Interoperabilidade global





Pilotos de gestão, Compartilhamento e Abertura de dados para pesquisa da Fiocruz

- ◆ 1 - Memória da Saúde Mental e Programa de Volta Para Casa (PVC)
- ◆ 2 - Programa de Controle da Doença de Chagas (PCDCh) - Minas Gerais
- ◆ 3 - Rede de Pesquisa Clínica e Aplicada em Chikungunya (Replick)
- ◆ 4 - Pesquisa em modelagem matemática e estatística de sistemas biológicos – PROCC - Fiocruz

Etapas do Piloto



Ciclo de vida dos dados na Fiocruz

Planejamento



Digitalbevaring.dk

Produção, coleta, reutilização e análises



Depósito, publicação e acesso



Information	Access	Licenses	Persistent Identifier	Certificates and Standards	Policy	Reviewed
<input type="checkbox"/>		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	
<input type="checkbox"/>		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Preservação (dados com guarda permanente)





Obrigada!



Fiocruz:
health for all