

# PLANO OPERATIVO

REPOSITÓRIO DE DADOS DE PESQUISA DA FIOCRUZ  
ARCA DADOS



Rio de Janeiro  
Abril 2026

## **Fundação Oswaldo Cruz**

### **Presidência**

Mario Santos Moreira

### **Chefe de Gabinete**

Rivaldo Venâncio

### **Diretoria Executiva**

Juliano de Carvalho Lima

### **Vice-Presidência de Ambiente, Atenção e Promoção da Saúde (VPAAPS)**

Valcler Rangel Fernandes e Patricia Canto Ribeiro (adjunta)

### **Vice-Presidência de Educação, Informação e Comunicação (VPEIC)**

Marly Marques da Cruz e Eduarda Angela Pessoa Cesse (adjunta)

### **Vice-Presidência de Pesquisa e Coleções Biológicas (VPPCB)**

Alda Maria da Cruz e Marcia de Oliveira Teixeira (adjunta)

### **Vice-Presidência de Produção e Inovação em Saúde (VPPIS)**

Priscila Ferraz Soares e Marco Aurelio de Carvalho Nascimento (adjunto)

### **Vice-Presidência de Saúde Global e Relações Internacionais (VPSGRI)**

Maria de Lourdes Aguiar Oliveira e Marcelo Pelajo (adjunto)

## **Comitê Gestor do Arca Dados**

### **Coordenação**

Vanessa Jorge (VPEIC)

### **Coordenação Adjunta**

Claudete Fernandes de Queiroz (ICICT)

## **Membros**

### **Titulares**

Camila M. de Souza Ribeiro (Gestec/VPPIS)

Karina Veras Praxedes dos Santos (COC)

Paulo Vitor Amaral (COGETIC)

Renato Marins Domingues (VPPCB)

### **Suplentes**

Aline da Silva Alves (ICICT)

Anderson Luiz Cabral dos Santos (VPEIC)

Cristiane Elisa Boar (VPPCB)

Ivone Pereira de Sá (COC)

Jean Gabriel Nguema Ngomo (COGETIC)

## **Elaboração técnica**

Vanessa Jorge, Anderson Luiz Cabral dos Santos, Claudete Queiroz, Aline da Silva Alves, Paloma Helena Fernandes Shimabukuru e Karina Veras Praxedes dos Santos

Esta obra é regida pela Política de Acesso Aberto ao Conhecimento, que busca garantir à sociedade o acesso gratuito, público e aberto ao conteúdo integral de toda obra intelectual produzida pela Fiocruz. Os conteúdos deste documento protegidos por direitos autorais e podem ser utilizados para todos os fins não comerciais, respeitados e reservados os direitos morais dos autores.

## **Fundação Oswaldo Cruz**

Av. Brasil, 4365 – Manguinhos

CEP 21045-900 - Rio de Janeiro, RJ, Brasil

Tel. 55 21 2598-4242



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>3</b>
<b>2</b>	<b>FINALIDADE DO ARCA DADOS</b>	<b>3</b>
<b>3</b>	<b>INSTÂNCIAS DE GOVERNANÇA DO ARCA DADOS</b>	<b>4</b>
3.1	FÓRUM DE CIÊNCIA ABERTA	4
3.2	COMITÊ GESTOR DO ARCA DADOS	4
3.3	NÚCLEOS DE CIÊNCIA ABERTA	5
<b>4</b>	<b>INFRAESTRUTURA TECNOLÓGICA</b>	<b>6</b>
<b>5</b>	<b>ORGANIZAÇÃO DO CONTEÚDO</b>	<b>7</b>
<b>6</b>	<b>DIRETRIZES DE CONTEÚDO</b>	<b>8</b>
6.1	FORMATOS DE ARQUIVOS	8
6.2	TAMANHO DE ARQUIVOS	9
<b>7</b>	<b>POLÍTICA DE CADASTRO E DEPÓSITO</b>	<b>9</b>
7.1	CADASTRO	9
7.2	DEPÓSITO	11
<b>8</b>	<b>LICENÇAS, TERMOS DE USO E CESSÃO</b>	<b>12</b>
8.1	LICENÇAS	12
8.2	TERMOS DE USO	12
8.3	TERMOS DE CESSÃO	12
<b>9</b>	<b>FLUXOS DE OPERAÇÃO</b>	<b>12</b>
<b>10</b>	<b>PRESERVAÇÃO DIGITAL</b>	<b>13</b>
<b>11</b>	<b>REVISÃO DO DOCUMENTO</b>	<b>13</b>

## 1 INTRODUÇÃO

Desde 2017, a Fundação Oswaldo Cruz (Fiocruz) vem ampliando, de forma sistemática, o debate institucional no campo da Ciência Aberta, com ênfase na gestão, compartilhamento e abertura de dados para pesquisa. Esse processo tem orientado a formulação de estratégias, a definição de estruturas organizacionais, o estabelecimento de processos e a consolidação de normas voltadas para a gestão dos dados oriundos das pesquisas científicas desenvolvidas na Instituição.

Nesse contexto, destaca-se a publicação da “Política de gestão, compartilhamento e abertura de dados para pesquisa: princípios e diretrizes”<sup>1</sup>, lançada em 2020, que estabelece orientações para que os dados para pesquisa, reconhecidos como bens públicos, sejam disponibilizados de forma aberta, ética, íntegra, acessível, de maneira oportuna e responsável, em consonância com as políticas estratégicas da pesquisa científica nacional, os interesses institucionais e as normas regulatórias e diretrizes vigentes.

No aspecto prático, tornou-se fundamental consolidar uma infraestrutura tecnológica para a implantação do Repositório de Dados de Pesquisa da FIOCRUZ – Arca Dados. A plataforma foi concebida para a gestão, compartilhamento, e a preservação dos dados de pesquisa produzidos no âmbito da Fundação ou em parceria com outras instituições. O desenvolvimento do Arca Dados teve início em 2019, a partir de um projeto piloto estruturado em duas fases<sup>2</sup>.

Atualmente, o Arca Dados está encontra-se em funcionamento regular como o Repositório Institucional de dados de Pesquisa da Fiocruz, com fluxos operacionais definidos para o depósito, a curadoria, a publicação, o acesso, o compartilhamento e a preservação de dados para pesquisa<sup>3</sup>. Nesse sentido, o presente Plano Operativo tem como objetivo apresentar as instâncias de governança, as diretrizes de organização e gestão de conteúdo, os procedimentos de depósito e publicação, os níveis de acesso, a infraestrutura tecnológica e os fluxos operacionais que estruturam o funcionamento do repositório.

O documento é dirigido, prioritariamente, aos Núcleos de Ciência Aberta das unidades da Fiocruz, constituindo-se como instrumento de orientação e apoio à implementação, à operação e ao alinhamento institucional das ações relacionadas ao Arca Dados.

## 2 FINALIDADE DO ARCA DADOS


O Arca Dados é o repositório oficial da Fiocruz destinado a arquivar, publicar, disseminar, preservar e compartilhar dados digitais para pesquisa produzidos pela comunidade Fiocruz ou em parceria com outros Institutos ou órgãos de pesquisas. Tem

---

<sup>1</sup> Disponível em: <https://www.arca.fiocruz.br/handle/icict/46408>.

<sup>2</sup> Para mais informações sobre o projeto piloto, consulte: <https://doi.org/10.35078/RAWVFZ>.

<sup>3</sup> Arca Dados pode ser acessado em: <https://arcadados.fiocruz.br/>



como objetivo fomentar novas pesquisas, garantir a reprodutibilidade e a replicabilidade de pesquisas existentes e promover a Ciência Aberta e Cidadã.

O repositório busca estimular a ampla circulação do conhecimento científico, fortalecendo o compromisso institucional com a Ciência Aberta e com o livre acesso da informação em saúde. Contribui, também, para a transparência das atividades de pesquisa e para o fortalecimento da colaboração entre pesquisadores, educadores, acadêmicos, gestores, estudantes de pós-graduação, em benefício do avanço do conhecimento e do atendimento às demandas da sociedade.

### **3 INSTÂNCIAS DE GOVERNANÇA DO ARCA DADOS**

O Arca Dados se estrutura a partir de um conjunto de instâncias de governança que atuam de forma articulada, complementar e integrada, abrangendo dimensões estratégicas, normativas, técnicas e operacionais. Essas instâncias são responsáveis por definir diretrizes, coordenar ações, apoiar a implementação, assegurar a conformidade ética, legal e institucional, bem como garantir o funcionamento adequado, a sustentabilidade e a evolução contínua do repositório de dados de pesquisa da Fiocruz:


#### **3.1 FÓRUM DE CIÊNCIA ABERTA**

O Fórum de Ciência Aberta da Fiocruz, instituído por meio da Portaria da Presidência da Fiocruz nº 157, de 12 de maio de 2021, constitui-se em uma instância colegiada de caráter estratégico, executivo e permanente, responsável por pautas da Ciência Aberta na instituição. Sob a coordenação da Vice-presidência de Educação, Informação e Comunicação da Fiocruz (VPEIC), por meio da Coordenação de Informação e Comunicação (CINCO), com coordenação adjunta do Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT), o Fórum conta com a participação de diversas instâncias da Fiocruz, assim como com representantes das unidades.

O Comitê Gestor do Arca Dados é uma das instâncias que compõem o Fórum de Ciência Aberta, levando as questões relativas do Arca Dados para esse espaço de deliberação estratégica, com subsídios técnicos, informacionais e operacionais.

#### **3.2 COMITÊ GESTOR DO ARCA DADOS**

O Comitê Gestor do Arca Dados, definido por meio da Portaria da Presidência da Fiocruz nº 489, de 20 de setembro de 2021, é coordenado pela VPEIC, com coordenação-adjunta do ICICT, contando com representantes da Vice-Presidência de Pesquisa e Coleções Biológicas (VPPCB), da Casa de Osvaldo Cruz (COC), da Coordenação-Geral de Gestão de Tecnologia de Informação (COGETIC) e da Coordenação de Gestão Tecnológica (GESTEC).



O Comitê Gestor tem por objetivo definir as linhas de ação, o plano de trabalho e as diretrizes gerais de funcionamento do Arca Dados, bem como estimular o alinhamento operativo entre as unidades. Compreende funções relacionadas ao apoio aos Núcleos de Ciência Aberta nas ações que envolvem a gestão dos dados, como na curadoria, depósito, preservação e segurança dos dados para pesquisa.

Cabe ao Comitê Gestor monitorar e avaliar a condução e desempenho do repositório, bem como propor a implementação de melhorias e adaptações relacionadas a mudanças tecnológicas, de gestão ou de novas demandas operacionais. O Comitê Gestor do Arca Dados produzirá os relatórios de impacto à proteção de dados pessoais, quando solicitados pela Autoridade Nacional de Proteção de Dados (ANPD), e realizará a comunicação de incidentes de segurança, que acarretar risco ou dano relevante aos titulares de dados pessoais, com apoio técnico das áreas responsáveis pela gestão e operação do repositório.

### 3.3 NÚCLEOS DE CIÊNCIA ABERTA

Os Núcleos de Ciência Aberta (NCAs) são instâncias colegiadas de caráter executivo, atuantes em todas as unidades da Fiocruz, responsáveis pela operação, participação, promoção e monitoramento das ações de implantação da Ciência Aberta na Unidade a que corresponde, incluindo ações à avaliação da viabilidade da disponibilização de dados e informações científicas.

Os NCAs são responsáveis por gerenciar os dataverses de sua Unidade no Arca Dados e produzir relatórios gerenciais pertinentes às suas atividades no repositório de dados. Os membros dos Núcleos de Ciência Aberta serão indicados pelos respectivos diretores das Unidades, e qualquer alteração em sua composição deve ser comunicada ao Fórum de Ciência Aberta.

A composição recomendada dos NCAs é definida pela Portaria nº 157, de 12 de maio de 2021, e para fins de funcionamento do Arca Dados, cada membro terá a seguinte competência:

- a) **Assessoria de Pesquisa:** incentiva e orienta os pesquisadores da unidade quanto ao depósito dos dados de pesquisa produzidos pela Fiocruz no Arca Dados;
- b) **Secretaria Acadêmica:** orienta os alunos dos programas de pós-graduação da Fiocruz quanto ao depósito dos dados de pesquisa oriundos das teses e dissertações no Arca Dados;
- c) **Biblioteca de referência da unidade:** orienta, sensibiliza e capacita usuários para a utilização do Arca Dados, bem como participa do processo de curadoria de dataverses e datasets, em articulação com a coordenação técnica do repositório;
- d) **Serviço ou núcleo de gestão de documentos da unidade:** realiza estudos sobre os prazos de guarda dos dados para elaboração de tabela de temporalidade, envia proposta de temporalidade à Comissão Permanente de Avaliação de

- Documentos e Arquivos da Fiocruz (CPADA), realiza o processo de avaliação e destinação final dos dados com apoio e orientação da CPADA, bem como participa do processo de curadoria de dataverses e datasets;
- e) **Tecnologia de Informação:** oferece suporte ao acesso do usuário no Arca Dados, notifica a COGETIC sobre problemas de instabilidade e segurança no sistema e solicita a resolução de qualquer problema relacionado;
  - f) **Núcleo de Inovação Tecnológica (NIT):** atuar no limite das suas competências legais, avaliando, por exemplo, a necessidade de manutenção de sigilo sobre os dados, seja em razão de contratos e demais acordos firmados ou em razão do potencial inovativo dos dados de pesquisa e a possibilidade de revelação inadvertida de aspectos de criações desenvolvidas institucionalmente, nos termos da Lei de Inovação da Política de Inovação da Fiocruz;
  - g) **Periódicos científicos da Fiocruz:** incentiva e orienta os autores a depositarem os conjuntos de dados associados aos artigos publicados nas revistas das unidades no Arca Dados;
  - h) **Comitê de Ética em Pesquisa (CEP):** orienta quanto a adequação dos Termo de Consentimento Livre e Esclarecido (TCLE) para o reuso dos dados para pesquisa envolvendo seres humanos e questões éticas;
  - i) **Curador:** coordena e gerencia os conteúdos da Unidade e/ou projeto no Arca Dados, responde a consultas de usuários; cria e administra dataverses, edita e qualifica metadados, realiza a curadoria dos conjuntos de dados, dos objetos digitais e da documentação suplementar (Plano de Gestão de Dados - PGD, dicionário de dados, *readme*, fichas prosopográficas entre outros); verifica a conformidade técnica, normativa e informacional para realização do depósito; certifica e aprova a publicação de dataverses e datasets, em conformidade com as diretrizes institucionais do Arca Dados. Para ser elegível ao papel de curador, o profissional deve ter concluído o Programa de Desenvolvimento de Pessoas do Ecossistema da Ciência Aberta.

#### 4 INFRAESTRUTURA TECNOLÓGICA

O software Dataverse<sup>4</sup>, escolhido para a implementação do Arca Dados, é um sistema de código aberto, desenvolvido pelo Instituto de Ciências Sociais Quantitativas (IQSS) da Universidade de Harvard. A plataforma permite armazenar, compartilhar, publicar, citar, explorar e analisar dados de pesquisa, facilitando a disponibilização e a replicação de dados para a comunidade científica.

O Dataverse oferece suporte a identificadores persistentes para conjuntos de dados para pesquisa, como o Digital Object Identifier (DOI), sendo compatível com ferramentas de preservação a longo prazo, como o Archivematica, adotado pela Fiocruz. Também suporta o armazenamento na nuvem e outras formas de armazenamento

---

<sup>4</sup> The Dataverse Project – Disponível em: <<https://dataverse.org>>

distribuído.

O Arca Dados possui duas instalações do Dataverse, sendo uma para homologação e outra de produção. Ambas as instalações operam em servidores institucionais sob responsabilidade da COGETIC.

A capacidade de armazenamento não é definida, utilizando o S3 do Nutanix que é expansível. Os procedimentos de Backup são divididos em incremental, que é realizado em uma vez por dia, e o full, realizado nos fins de semana. As medidas segurança aplicadas para os dados depositados são os mesmos utilizados na sala cofre da Fiocruz. Há também integração com plataformas externas, como a Comunidade Acadêmica Federada (CAFe) e a Datacite, para a geração e o gerenciamento de identificadores DOI.

## 5 ORGANIZAÇÃO DO CONTEÚDO

O Arca Dados é a instalação do software que hospeda vários arquivos virtuais chamados dataverses. Cada dataverse contém conjuntos de dados chamados datasets, e cada dataset contém objetos digitais (arquivos) e metadados, podendo incluir arquivos de dados descritivos, documentação complementar e códigos de programação. Como método de organização, os dataverses também podem conter outros dataverses, conforme as imagens abaixo.

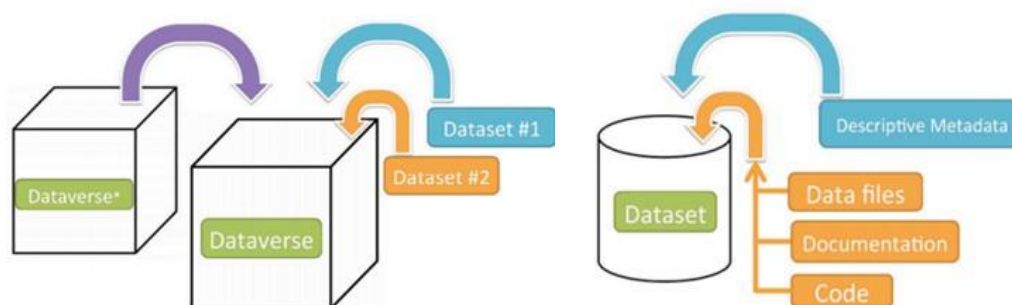


Figura 1 - Organização de conteúdo

O Arca Dados possui um dataverse raiz denominado Fiocruz, que abriga diversas comunidades (dataverses dentro do dataverse raiz Fiocruz), podendo ser organizadas da seguinte forma:

- **Unidade** – Dataverses vinculados a uma Unidade organizacional da Fiocruz, no qual somente os pesquisadores daquela Unidade possuem permissão para criar dataverses e datasets;
- **Projeto** – Dataverses ligados a um projeto de pesquisa, no qual somente os membros daquele projeto podem criar dataverses e datasets;
- **Coleções temáticas** – Dataverses que reúnem conjuntos de dados sobre uma mesma temática de interesse institucional;
- **Publicações** – Conjuntos de dados relacionados a artigos publicados em periódicos da Fiocruz que oferecem a opção de depósito dos dados de pesquisa

- no Arca Dados;
- **Instituição** – Dataverses que contém conjuntos de dados de pesquisas de instituições parceiras.

A criação de comunidades dentro do dataverse raiz Fiocruz e a aprovação de cadastros de usuários depositantes são atribuições exclusivas da Coordenação Geral do Comitê Gestor do Arca Dados. Cada comunidade deverá designar um Curador que trabalhará em conjunto com a equipe da Coordenação para garantir que as políticas e diretrizes do Arca Dados sejam aplicadas dentro da coleção.


## 6 DIRETRIZES DE CONTEÚDO

Os conjuntos de dados e documentos relacionados permitidos para depósito no Arca Dados devem obedecer aos seguintes critérios:

- a) Devem ser produto de pesquisa técnico-científica realizada por pesquisadores, alunos e funcionários da Fiocruz; pesquisas financiadas, apoiadas ou realizadas em parceria com a Fiocruz e com outras instituições; pesquisas em temáticas relevantes para saúde a partir de parcerias firmadas;
- b) Atendem aos requisitos determinados no Termo de Uso do Depositante;
- c) Os dados depositados não devem ter sigilo previsto segundo as diversas legislações vigentes ou em virtude de acordos e contratos previamente firmados;
- d) Os dados com o potencial de gerar novos produtos ou serviços ou protegidos por propriedade intelectual não devem ser depositados neste repositório, até que (i) a proteção intelectual de tais ativos seja efetuada, especialmente, tratando-se de direitos de propriedade industrial (ii) e/ou o Sistema Gestec-NIT tenha opinado quanto sua divulgação;
- e) Os dados pessoais devem ser anonimizados ou pseudonimizados antes do depósito, no caso de pseudonimização os dados complementares deverão ser depositados em ambiente apartados e seguros;
- f) Devem ser de interesse de longo prazo para a instituição;
- g) Devem estar em formato digital, completos e prontos para distribuição;
- h) O autor dos dados deve ceder à Fiocruz o direito de preservar e distribuir os conjuntos de dados e documentação complementar por meio de assinatura do Termo de Cessão de Direitos específico para o depósito no Arca Dados;
- i) Devem ser formalizados todos os contratos, termos e autorizações necessários junto aos eventuais coautores e instituições cotitulares.

### 6.1 FORMATOS DE ARQUIVOS

O Dataverse suporta uma variedade de tipos e formatos de arquivo, incluindo planilhas, bases de dados, tabelas e documentação suplementares, publicações associadas a dados, apresentações, *shapefiles*, imagens e arquivos compactados. No entanto,



somente arquivos de dados tabulares – por exemplo, Stata, SPSS, R, Excel (xlsx) e CSV – podem ser visualizados na própria interface do repositório, por meio do recurso Data Explorer.

Embora qualquer formato de arquivo possa ser depositado no repositório, sempre que possível devem ser escolhidos formatos que ajudem a preservar e garantir o acesso a longo prazo aos arquivos depositados no Arca Dados, com preferência por arquivos em formatos abertos e não-proprietários. Por exemplo, PDF/A e ASCII são formatos preferidos para conteúdo textual; TIFF e JPEG2000 são preferidos para imagens<sup>5</sup>.

## 6.2 TAMANHO DE ARQUIVOS

Em sua configuração padrão, o Dataverse estabelece limites para o armazenamento de arquivos individuais. No Arca Dados, esses parâmetros foram ajustados para atender às necessidades institucionais, permitindo o depósito de arquivos individuais de até **1,5 GB**, respeitados os limites definidos para cada conjunto de dados (dataset).

O limite máximo recomendado para cada conjunto de dados é de até **2,5 GB**, considerando a infraestrutura disponível e os critérios operacionais do repositório. Demandas que envolvam volumes superiores poderão ser avaliadas mediante análise técnica e operacional, em articulação com a COGETIC.

## 7 POLÍTICA DE CADASTRO E DEPÓSITO

### 7.1 CADASTRO

A navegação pelo repositório e o acesso aos metadados dos itens depositados no Arca Dados são abertos a qualquer pessoa interessada.

Os colaboradores da Fiocruz podem realizar o cadastro no repositório utilizando suas credenciais de acesso do login único da Fiocruz, autenticados por meio do protocolo Shibboleth, operado pela rede CAFe<sup>6</sup>. Os usuários externos poderão se cadastrar no repositório criando uma conta local com acesso por e-mail e senha, ou usar a autenticação via Shibboleth, caso sejam vinculados a alguma instituição previamente habilitada pela COGETIC.

Ao registrar-se no repositório, o usuário não possui nenhuma permissão para realizar ações no sistema, sendo necessário que o administrador ou curador atribua um ou mais perfis ao usuário. Os usuários registrados no sistema podem ter atribuídos os seguintes perfis, com suas respectivas permissões:

---

<sup>5</sup> Para informações detalhadas sobre formatos recomendados para preservação, consultar o Programa de Preservação Digital de Acervos da Fiocruz, disponível em:

<https://www.arca.fiocruz.br/handle/iciict/44220>.

<sup>6</sup> A Comunidade Acadêmica Federada (CAFe), provida pela Rede Nacional de Ensino e Pesquisa (RNP), permite que usuários utilizem login e senha institucionais para acessar serviços de sua própria instituição e de outras organizações que participam da federação

a) **Admin:** adicionar dataverse, adicionar dataset, ver dataverse não publicado, ver dataset não publicado, baixar arquivo, editar dataverse, editar dataset, gerenciar permissões do dataverse, gerenciar permissões do dataset, publicar dataverse, publicar dataset, deletar dataverse e deletar rascunho de dataset.

O perfil de administrador tem como papel principal e exclusivo a concessão da permissão de curador(es) de um dataverse. Este perfil será concedido aos coordenadores gerais do Arca Dados, principalmente no processo de transição e aprendizado.

**Contribuidor:** ver dataset não publicado, baixar arquivo, editar dataset e deletar rascunho do dataset. Esse perfil não pode alterar dataset já publicado.

b) **Curador:** adicionar dataverse, adicionar dataset, ver dataverse não publicado, ver dataset não publicado, baixar arquivo, editar dataset, gerenciar permissões do dataset, publicar dataset e deletar rascunho do dataset.

Além das permissões atribuídas ao curador pelo sistema, esse perfil também será responsável por:

- i. Coordenar os pedidos de criação de novos dataverses, datasets e cadastros de usuários depositantes;
- ii. Personalizar as páginas da comunidade com logotipos apropriados, texto descritivo, e informações de licença para uso e acesso aos dados;
- iii. Compreender e observar as políticas de gestão e compartilhamento de dados da Fiocruz e educar os participantes da comunidade sobre essas políticas;
- iv. Garantir o funcionamento do fluxo de operação para depósito de dados;
- v. Supervisionar os processos de aprovação de submissão;
- vi. Zelar pela segurança dos dados e notificar à Tecnologia da Informação qualquer incidente de segurança e realizar as operações necessárias para minimização dos seus impactos;
- vii. Notificar a coordenação geral do Comitê Gestor do Arca Dados sobre mudanças organizacionais que afetem o fluxo de operação.

c) **Criador de dataset:** adicionar dataset.

Os pesquisadores, alunos e colaboradores poderão receber este perfil dentro dos dataverses do qual participam, para inserir nele o conjunto de dados a ser depositado.

d) **Criador de dataverse + dataset:** adicionar dataverse e adicionar dataset.

Estes perfis serão definidos pelo curador do dataverse da unidade


e) **Criador de dataverse:** adicionar dataverse.

Estes perfis serão definidos pelo curador do dataverse da unidade

f) **Download de arquivo:** baixar arquivo<sup>7</sup>.

---

<sup>7</sup> Para arquivos em acesso aberto não é necessário ter este perfil ou autenticar no Arca Dados. Qualquer pessoa pode baixar arquivos quando estão em acesso aberto.



Estes perfis serão definidos pelo curador do dataverse da unidade. Normalmente este perfil será indicado para editores científicos relacionados a processos de submissão de artigo ou em compartilhamentos de arquivos restritos aprovados pelo pesquisador.

g) **Membro**: ver dataverse não publicado, ver dataset não publicado e baixar arquivo.

Estes perfis serão definidos pelo curador do dataverse da unidade

## 7.2 DEPÓSITO

Embora qualquer pessoa que possua um login único da Fiocruz possa criar uma conta no Arca Dados, somente usuários devidamente autorizados, cadastrados com os perfis *Admin*, *Curador*, *Criador de dataverse + dataset* e *Criador de dataset* poderão realizar o depósito de dados de pesquisa. Todos os usuários que desejam depositar dados no Arca Dados devem ler e cumprir os Termos de Uso de Usuário Depositante e o Termo de Cessão de Direitos.

As regras de depósito e os níveis de acesso aos dados poderão ser propostos pelo autor dos dados de pesquisa, considerando aspectos legais, éticos, estratégicos e contratuais. A aplicação, validação e implementação dessas regras caberão ao curador responsável, que deverá avaliar a conformidade das condições de acesso com as políticas institucionais do Arca Dados e da Fiocruz.

No momento da publicação, o nível de acesso a um arquivo ou conjunto de arquivos de um dataset pode ser definido como:

- a) **Público** – quando o pesquisador autoriza o acesso, reuso e redistribuição de dados para pesquisa e não há restrições legais, éticas e estratégicas. Nessa modalidade, o autor e/ou gestor dos dados poderão definir questionários e termos a serem preenchidos e aceitos pelos usuários para o efetivo acesso aos datasets;
- b) **Restrito** – quando os dados para pesquisa podem ser compartilhados sob condições específicas, formalizadas por meio de termos, acordos ou autorizações, conforme definido pelo autor e validado pelo curador;
- c) **Embargado** – quando o pesquisador autoriza o acesso aos dados para pesquisa somente após um período previamente estabelecido.

Em um dataset publicado, os metadados permanecem sempre disponíveis e apenas os arquivos podem ser restritos. É importante destacar que a publicação dos metadados é indicada como padrão, pois esta ação pode resguardar os arquivos e/ou datasets que por motivos legítimos necessitem de proteção, dando visibilidade à existência dos dados para possíveis colaborações e reuso.

Todos os conjuntos de dados depositados no Arca Dados receberão um identificador persistente (Digital Object Identifier – DOI), que poderá ser utilizado para citação em publicações científicas.

## **8 LICENÇAS, TERMOS DE USO E CESSÃO**

### **8.1 LICENÇAS**

Por padrão, o Arca Dados atribui a licença Creative Commons Atribuição–Não Comercial (CC BY-NC) aos itens depositados no repositório. Essa indicação tem caráter orientativo, podendo o depositante definir termos de licença específicos para uso e acesso aos conjuntos de dados, desde que compatíveis com a legislação vigente, as políticas institucionais da Fiocruz e os princípios da Ciência Aberta.

### **8.2 TERMOS DE USO**

O Arca Dados possui Termos de Uso individuais para duas categorias de usuários, a saber: Usuário Comum e Usuário Depositante.

Os Termos de Uso para Usuário Comum, destinado a quem acessa e/ou faz uso dos dados depositados no Arca Dados, estão disponíveis em: <https://doi.org/10.35078/RAWVFZ>, em português, espanhol e inglês.

O Usuário Depositante, ao se cadastrar, acessar e depositar dados de pesquisa no Arca Dados deve aceitar todas as regras e condições constantes nos “Termos de Uso para Usuário Depositante do Arca Dados”, onde estão definidas as obrigações e responsabilidades do usuário. O documento está disponível em: <https://doi.org/10.35078/RAWVFZ>, em português e inglês.

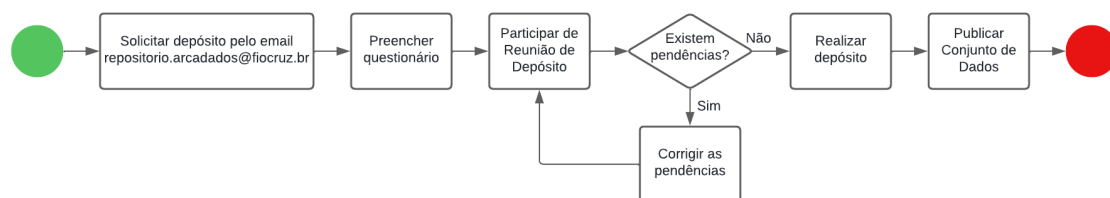
### **8.3 TERMOS DE CESSÃO**

Para que o Arca Dados armazene, distribua e preserve os dados para pesquisa depositados, o Usuário Depositante deve ler e assinar o Termo de Cessão de Direitos sobre Dados de Pesquisa específico para os conjuntos de dados e documentação complementar a serem depositados no Arca Dados, conforme “Procedimento para gerenciamento de documentos jurídicos (TERMOS) no SEI necessários ao depósito no Arca Dados”.

## **9 FLUXOS DE OPERAÇÃO**

Atualmente, se encontram em operação os fluxos dos processos de depósito, de publicação e de acesso aos conjuntos de dados restritos, que serão ajustados e atualizados no curso da implementação do Arca Dados e disponibilizados na seção documentos operativos do repositório.

## Fluxo de Depósito



## 10 PRESERVAÇÃO DIGITAL

Considerando a preservação digital como uma estratégia que envolve diferentes níveis de ações, como o uso de identificadores persistentes e backups automatizados, incremental e full (descritos no item 4), representam algumas estratégias de salvaguarda dos dados já realizadas na operação do repositório.

Outra ação que está em fase final de implementação é a integração do Arca Dados com o Repositório Digital de Preservação (RDP), a partir de interoperabilidade direta entre os softwares. O Archivematica é o software gratuito e de código aberto adotado pelo Preservo – Complexo de Acervos da Fiocruz, em 2018, enquanto solução tecnológica para seu RDP. Em atendimento às diretrizes e requisitos da ISO 16363 e da ISO 14721, a preservação digital dos dados para pesquisa da Fiocruz deve ser realizada mediante uso do RDP.

## 11 REVISÃO DO DOCUMENTO

O Plano Operativo do Arca Dados deve ser revisto e atualizado pelo Comitê Gestor a cada 2 (dois) anos a partir da data de publicação da última versão, desde que não haja outros fatores que demandem a sua revisão antecipada.

A atualização extemporânea do Plano Operativo pode ser necessária nos seguintes casos, entre outros:

- a) Atualização das políticas institucionais de gestão, acesso e preservação de dados para pesquisa;
- b) Mudança estrutural nas instâncias de gestão e operação do Arca Dados;
- c) Mudança ou atualização da infraestrutura tecnológica para a gestão de dados;
- d) Novas diretrizes relativas ao conteúdo do repositório de dados, incluindo mudanças nos termos de uso e cessão;
- e) Alteração nos fluxos de trabalho de depósito, publicação, curadoria e acesso aos dados para pesquisa.
- f) Atualização nos padrões e melhores práticas para gestão de dados para pesquisa;
- g) Alterações legais ou normativas que impactem no tratamento dos dados.