



A GESTÃO DA PESQUISA CIENTÍFICA NO INSTITUTO NACIONAL DA MATA ATLÂNTICA: diretrizes basilares

Juliana Lazzarotto Freitas

 <https://orcid.org/0000-0003-2572-9407>.

✉ julianalf@ufpr.br.

🏢 Instituto Nacional da Mata Atlântica (INMA) |

ROR: <https://ror.org/0395f2d85> | Santa Teresa, Brasil.

Fábio Mascarenhas e Silva

 <https://orcid.org/0000-0001-5566-5120>.

✉ fabio.mascarenhas@ufpe.br.

🏢 Universidade Federal de Pernambuco (UFPE) |

ROR: <https://ror.org/047908t24> | Recife, Brasil.

Alejandro Caballero Rivero

 <https://orcid.org/0000-0003-1061-0534>.

✉ caballero.alecaba@gmail.com.

🏢 Instituto Nacional da Mata Atlântica (INMA) |

ROR: <https://ror.org/0395f2d85> | Santa Teresa, Brasil.

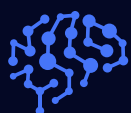
Eixo temático: Acesso Aberto, Ciência Aberta e Dados Abertos

Modalidade: Resumo expandido

DOI: 10.22477/ix.ebbc.300

Resumo: Políticas de gestão de dados de pesquisa têm sido focais para o desenvolvimento de boas práticas científicas nas instituições. Definem-se, aqui, a partir de resultados preliminares de pesquisa, diretrizes para a política de gestão de dados de pesquisa do Instituto Nacional da Mata Atlântica. De forma específica, identificam-se modelos de gestão de dados de pesquisa para gerir e compartilhar conjuntos de publicações relacionadas à Mata Atlântica Brasileira. É um estudo exploratório e propositivo que adota a análise documental para analisar as melhores práticas de gestão de dados de pesquisa no âmbito nacional. Recomendam-se diretrizes para a implementação de uma política de gestão de dados de pesquisa científica, priorizando princípios da ciência aberta para gerir e compartilhar conjuntos de dados de uma Instituição de Ciência e Tecnologia.

Palavras-Chave: Política de gestão de pesquisa. Mata Atlântica. Plano de gestão de dados. Produção científica. Instituto Nacional da Mata Atlântica.



1 INTRODUÇÃO

A gestão de dados de pesquisa consiste em um domínio estratégico para a democratização do acesso ao conhecimento e para o desenvolvimento científico e tecnológico. Esse domínio é reconhecido pela comunidade científica como parte importante das boas práticas de pesquisa, visto que os dados gerados no processo de pesquisa, que muitas vezes são considerados subprodutos desta atividade, quando disponíveis para acesso e reuso, tornam-se recursos informacionais valiosos para fomento de novas pesquisas (Piccolo *et al.*, 2021; Sales; Sayão, 2019).

Um Instituto de pesquisa é o lócus, por excelência, de produção de conhecimento, ciência e tecnologia e de transferência desse conhecimento para a sociedade, com forte potencial de contribuição para o desenvolvimento econômico, social e cultural de um país. Os dados e publicações produzidos e organizados pelo Instituto Nacional da Mata Atlântica (INMA) e/ou com sua colaboração; que tratam da Mata Atlântica, são recursos significativos que, quando devidamente organizados, descritos, representados, geridos, protegidos e compartilhados podem fomentar o desenvolvimento da ciência sobre o bioma.

Uma política de gestão de dados de pesquisa para o INMA tem como intenção principal garantir o acesso universal aos conjuntos de dados de pesquisas produzidos e coletados pela instituição, beneficiando os pesquisadores que atuam no domínio de pesquisa sobre a Mata Atlântica, permitindo-lhes localizar e acessar as publicações científicas e os seus dados brutos, assegurando-lhes também, dados bem representados, armazenamento seguro e a possibilidade de compartilhá-los, usá-los e reutilizá-los. Estimula-se, assim, a colaboração científica entre pesquisadores; a troca e o compartilhamento de conhecimento de forma transparente e ética no âmbito institucional.

Portanto, este estudo contempla a gestão de dados de pesquisa como atividade estratégica que busca favorecer a integridade, a qualidade e o uso da pesquisa, favorecendo também o registro da memória institucional. Questionam-se, portanto, quais as melhores práticas de gestão de dados para subsidiar a implementação de uma política de gestão de pesquisa no INMA?

Considerando que uma das metas do INMA, de acordo com seu plano estratégico (2022-2030) é consolidar-se como um centro de excelência na produção de sínteses de informação sobre a biodiversidade da Mata Atlântica (INMA, 2023), é essencial que a Instituição constitua uma política de informação para os próprios dados científicos que gera a partir das pesquisas realizadas por seus colaboradores, de modo a regular e gerir um dos seus mais importantes produtos de informação e possibilitar que sejam acessados e compartilhados.

O objetivo deste estudo é propor diretrizes norteadoras preliminares de uma política de gestão de dados da pesquisa científica para o INMA. De forma específica, identificam-se metodologias de gestão de dados de pesquisa que possam nortear a gestão de conjuntos de dados referentes à produção científica sobre a Mata Atlântica brasileira, coletado ao longo do Programa de Capacitação Institucional (PCI) do INMA, entre os anos de 2019 e 2023. Ainda, explora-se a literatura científica sobre planos e políticas de gestão de dados, bem como os modelos adotados por outras instituições de pesquisa nacionais e internacionais.

2 REFERENCIAL TEÓRICO

As transformações das práticas científicas mostram que a ciência passa por um momento de valorização do trabalho colaborativo e do compartilhamento de dados de pesquisa, com a finalidade de acelerar o desenvolvimento científico e tecnológico e aumentar a eficiência nos investimentos. Logo, há uma intensa produção e consumo de dados motivados pela geração de novos conhecimentos, o que gera a demanda por dispositivos para controle e organização desses dados produzidos.

O gerenciamento e a preservação dos dados de pesquisas são questões que têm sido centrais para as instituições, agências de fomento e pesquisadores, a fim de desenvolver boas práticas que permitam a valorização e a reutilização dos conjuntos de dados gerados e, conseqüentemente, o desenvolvimento de novas formas de pesquisar a partir desses conjuntos de dados (Aventurier; Alencar, 2016).

No cenário da ciência aberta, com a produção massiva de dados pelos pesquisadores, a importância da gestão de dados de pesquisa evoluiu (Bonetti; Moreno, 2021) e, pode-se dizer que o tema tem sido tratado pela academia de forma mais enfática nos últimos 10 anos. Em estudo bibliométrico feito nas bases de dados *Web of Science* e *Scopus*, sobre a qualidade de gestão de dados de pesquisa entre 1984 e 2020, constatou-se que o primeiro artigo sobre o tema foi publicado em 2013 (Piccolo *et al.*, 2021).

Instituições de pesquisa e agências de fomento, nacionais e internacionais, já iniciaram a demanda para que projetos financiados requeiram um plano de gestão de dados de pesquisa para o compartilhamento em acesso aberto. As instituições de pesquisa, em curto ou médio prazo precisarão estar preparadas para atender às exigências dos órgãos financiadores (Monteiro, 2019). Portanto, estas instituições vêm exigindo novas condutas nesse contexto imposto pela ciência aberta, com a elaboração de planos de gestão de dados que permitam ao pesquisador gerenciar e disponibilizar os dados brutos de suas pesquisas, valorizando o compartilhamento e a abertura dos dados, bem como seu reuso em benefício dos próprios pesquisadores e da sociedade (Guandalini; Furnival; Arakaki, 2019).

3 METODOLOGIA

Esse estudo caracteriza-se como exploratório e propositivo, e adota a análise documental como meio para explorar a literatura e os modelos de planos e políticas de gestão de dados de pesquisa nacionais e internacionais (Fapesp, 2023; Fiocruz (Veiga *et al.*, 2023); Ibict, 2023; *SciELO*, 2023; *The Dataverse Project*, 2023; Unirio, 2023; PUC-RS, 2023; *Ohio State University*, 2023).¹

A análise permite recomendar a adoção de sistemas de repositórios abertos preferidos com padrões de metadados reconhecidos internacionalmente para organizar, representar e compartilhar as pesquisas produzidas pelos pesquisadores do INMA e facilitar o acesso, a recuperação e o uso dessa informação.

O PCI do INMA é composto por equipes de pesquisadores de distintas áreas. A título experimental,

¹ Documentos indicados nas referências.



foram reunidos, organizados e elegeram-se os dados produzidos por três pesquisadores da Ciência da Informação (CI). Procedeu-se a descrição desses dados com respectivos metadados, bem como a padronização, revisão e publicação desses dados no repositório *Deposita Dados* do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)² com apoio técnico da mesma Instituição. Os conjuntos de dados do INMA foram depositados no *DataVerse* do IBICT referente à região Sudeste, e podem ser consultados em <https://depositadados.ibict.br/dataverse/sudeste>. Os conjuntos estão alinhados à política de dados do Deposita Dados, disponíveis em <https://depositadados.ibict.br/about.xhtml>.

4 RESULTADOS PRELIMINARES

A política de gestão de dados de pesquisa do INMA tem como papel principal assegurar a qualidade e integridade dos dados de pesquisa institucional, e permitir o compartilhamento e armazenamento seguro desses dados para geração de novos conhecimentos, possibilitando o uso e reuso desses dados em benefício do desenvolvimento das pesquisas sobre a Mata Atlântica Brasileira, além do registro da memória de pesquisa do INMA. Considerando os resultados obtidos, recomendam-se diretrizes basilares para a implementação de uma política de gestão de dados de pesquisa.

Dentre as melhores práticas de gestão de pesquisa de livre acesso, sugere-se o uso de softwares de repositórios de dados de pesquisa de código aberto que atendam aos princípios *FAIR*, acrônimo para *findable, accessible, interoperable* e *reusable*, postulados pela Iniciativa Global *Open Fair*³, como por exemplo, o repositório *DataVerse* (*Harvard University*), repositório de dados de pesquisa de código aberto para armazenamento, preservação, compartilhamento, exploração e análise de dados)⁴ ou Repositório Invenior⁵.

A organização, descrição, publicação e compartilhamento dos conjuntos de dados relativos à produção científica e patentes coletadas pelos três pesquisadores de CI do INMA foi viabilizada por meio do Deposita Dados⁶, repositório construído e mantido pelo IBICT, baseado no *software DataVerse*, que garante o auto arquivamento dos conjuntos de dados de instituições desprovidas de recursos e infraestrutura para constituir seus próprios repositórios, a fim de que possam compartilhar seus conjuntos de dados publicamente e manter seus dados seguros e protegidos a longo prazo, sob licenças adequadas, de forma alinhada à legislação nacional e às boas práticas da ciência aberta.

O Deposita Dados também permitiu a criação do *Digital Object Identifier* (DOI), identificador persistente, que facilita a identificação e localização dos conjuntos de dados disponibilizados.

O compartilhamento dos conjuntos de dados foi realizado sob a licença descrita na Política do Deposita Dados, Licença *Creative Commons* CC BY 4.0⁷, que permite compartilhar, copiar e redistribuir o

² Disponível em: <https://depositadados.ibict.br>. Acesso em: 02 mar. 2024.

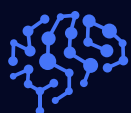
³ Disponível em: <http://go-fair-brasil.ibict.br>. Acesso em: 04 mar. 2024.

⁴ Disponível em: <https://dataverse.org>. Acesso em: 04 mar. 2024.

⁵ Disponível em: <https://inveniordm.web.cern.ch>. Acesso em: 04 mar. 2024.

⁶ Disponível em: <https://depositadados.ibict.br>. Acesso em: 04 mar. 2024.

⁷ Disponível em: <https://creativecommons.org/licenses/by/4.0/deed.pt-br#ref-technological-measures>. Acesso em: 04 mar. 2024.



material em qualquer suporte ou formato e adaptar, remixar, transformar, e criar a partir do material para qualquer fim.

Foi realizada a curadoria de dados nos conjuntos de dados publicados do INMA, com a definição e seleção de metadados específicos a serem preenchidos, como por exemplo, campos de financiamentos de pesquisa com números de processos de bolsas; com a padronização do uso de termos e siglas nos metadados de afiliações institucionais e também com a padronização de palavras-chave representativas dos conjuntos de dados, conforme tabelas de termos e tesouros adotados pelo sistema. Ainda, especificaram-se informações sobre autores e colaboradores, identificando códigos de *Orcids* e afiliações institucionais.

Além disso, relacionaram-se os conjuntos de dados às publicações e subprodutos de divulgação científica correspondentes a eles.

A validação das diretrizes aqui estabelecidas foi realizada com o compartilhamento dos conjuntos de dados relativos a publicações científicas e patentes sobre a Mata Atlântica, coletadas nas bases de dados *Web of Science*, *Scopus*, *SciELO*, Biblioteca Digital de Teses e Dissertações e *Derwent*. Os referidos conjuntos foram coletados e analisados no decorrer do PCI-INMA pelos pesquisadores da área de CI e compartilhados conforme as diretrizes aqui explicitadas.

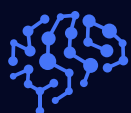
5 CONSIDERAÇÕES FINAIS

A gestão da pesquisa científica em instituições de pesquisa tornou-se uma tendência mundial e o INMA, a partir deste estudo, conta com material que subsidia a gestão de seus dados de pesquisa. Além disso, é a primeira instituição de pesquisa brasileira a depositar dados no Repositório do Deposita Dados, recentemente lançado no XXII Seminário Nacional de Bibliotecas Universitárias, em dezembro de 2023.

Os objetivos propostos foram cumpridos de modo a constituir proposições iniciais relativas a diretrizes norteadoras para uma política de gestão de dados da pesquisa científica institucional. Os conjuntos de dados de pesquisa relacionados à Mata Atlântica, que foram objeto de estudo dos pesquisadores da CI, colaboradores no subprojeto “Arquivos históricos e produção científica sobre a Mata Atlântica e a Conservação da Natureza no Brasil: organização, análise e difusão da informação” do PCI-INMA, vem sendo reunidos, documentados, descritos e disponibilizados para acesso universal.

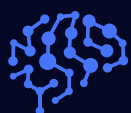
AGRADECIMENTOS

Ao CNPq e ao INMA pelas Bolsas do Programa de Capacitação Institucional, Processos nº: 301435/2024-3 e 301377/2024-3. Também ao professor Rene Gabriel Faustino Junior (UFRGS) e à Tatya-ne Guedes Martins da Silva (IBICT) pela colaboração técnica e apoio na publicação dos dados.



REFERÊNCIAS

- AVENTURIER, Pascal; ALENCAR, Maria Cleofas de. Os desafios de dados de pesquisa abertos. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, v. 10, n. 3, 2016. Disponível em: <https://brapci.inf.br/index.php/res/v/131000>. Acesso em: 03 mar. 2024.
- BONETTI, L.; MORENO, F. P. . Gestão de dados de pesquisa em bibliotecas universitárias brasileiras. **Revista Brasileira de Educação em Ciência da Informação**, São Paulo, v. 8, 2021. DOI: <https://doi.org/10.24208/rebecin.v8i.244>. Disponível em: <https://portal.abecin.org.br/rebecin/article/view/244>. Acesso em: 03 mar. 2024.
- FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO (Fapesp). **Diretrizes para Planos de Gestão de Dados (PGD) para propostas de centros**. *Online*. Disponível em: <https://fapesp.br/14974/diretrizes-para-planos-de-gestao-de-dados-pgd-para-propostas-de-centros>. Acesso em: 02 mar. 2023.
- GUANDALINI, C. A.; FURNIVAL, A. C. M.; ARAKAKI, A. C. S. Boas práticas científicas na elaboração de planos de gestão de dados. **RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação**, v. 17, 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/125479>. Acesso em: 06 mar. 2024.
- INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA (Ibict). **Deposita dados**: repositório comum de dados de pesquisa. *Online*. Disponível em: <https://depositadados.ibict.br>. Acesso em: 03 mar. 2024.
- INSTITUTO NACIONAL DA MATA ATLÂNTICA (INMA). **Plano Estratégico (2022-2030)**. *Online*. Disponível em: <https://www.gov.br/inma/pt-br/acesso-a-informacao/institucional/planejamento-estrategico>. Acesso em: 03 mar. 2024.
- MONTEIRO, Gabriela. Dados científicos abertos: políticas de gestão e agências de fomento. *In*: CONGRESSO DE GESTÃO ESTRATÉGICA DA INFORMAÇÃO, EMPREENDEDORISMO E INOVAÇÃO, v. 1, v. 2, 2019, **Anais [...]**. Disponível em: <https://brapci.inf.br/index.php/res/v/147436>. Acesso em: 6 mar. 2024.
- OHIO STATE UNIVERSITY. **University policies**. *Online*. Disponível em: <https://www.ohio.edu/policy>. Acesso em: 03 mar. 2023.
- PICCOLO, Daiane Marcela *et al.* Qualidade de dados em gestão de dados de pesquisa: um estudo bibliométrico. **Em Questão**, n. online, 2021. DOI: <https://doi.org/10.19132/1808-5245281.159-184>. Disponível em: <https://brapci.inf.br/index.php/res/v/165261>. Acesso em: 6 mar. 2024.
- PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO GRANDE DO SUL (PUC-RS). **Política de pesquisa**. *Online*. Disponível em: <https://www.pucrs.br/pesquisa/politicas-de-pesquisa>. Acesso em: 03 mar. 2024.
- SCIENTIFIC ELECTRONIC LIBRARY ONLINE (SciELO). **Gestão de dados científicos**: da coleta à preservação. *Online*. Disponível em: https://blog.scielo.org/blog/2018/06/22/gestao-de-dados-cientificos-da-coleta-a-preservacao/#.Y_9d4HbMKUk. Acesso em: 02 mar. 2024.
- SALES, Luana Farias; SAYÃO, Luís Fernando. Uma proposta de taxonomia para dados de pesquisa. **Revista Conhecimento em Ação**, n. 1, v. 4, p. 31-48, 2019. DOI: <https://doi.org/10.47681/rca.v4i1.26337>. Disponível em: <https://brapci.inf.br/index.php/res/v/118944>. Acesso em 06 mar. 2024.



THE DATAVERSE PROJECT. **Institutions.** *Online.* Disponível em: <https://dataverse.org>. Acesso em: 04 mar. 2024.

UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO (Unirio). **Política de acesso aberto à informação técnico científica e aos dados de pesquisa.** *Online.* Disponível em: <http://www.unirio.br/biblioteca-central/arquivos/200bpolitica-de-acesso-aberto-a-informacao-tecnico-cientifica-e-aos-dados-de-pesquisa>. Acesso em: 04 mar. 2024.

VEIGA, Viviane Santos de Oliveira *et al.* **Plano de gestão de dados de pesquisa - PGD:** guia de elaboração. Rio de Janeiro: Fiocruz/Icict, ago. 2022. 32 p.