26º ENANCIB
Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
A Amazônia na sociedade global em rede: assimetrias e potencialidades informacionais, tecnológicas e ambientais
Belém, Pará, 09 a 13 de novembro de 2026
Universidade Federal do Pará

ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
ISSN 2177-3688

GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação

DADOS DE PESQUISA DE UNIVERSIDADES BRASILEIRAS EM REPOSITÓRIOS INTERNACIONAIS: UM ESTUDO NO ZENODO E DRYAD

RESEARCH DATA FROM BRAZILIAN INSTITUTIONS IN INTERNATIONAL REPOSITORIES: A STUDY ON ZENODO AND DRYAD

Marilete da Silva Pereira – Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)/Universidade de Brasília (UnB)
Michelli Pereira da Costa – Universidade de Brasília (UnB)

Modalidade: Trabalho Completo

Resumo: A Ciência Aberta tem ampliado a importância do compartilhamento de dados de pesquisa e fortalecido o papel dos repositórios digitais na comunicação científica. Nesse contexto, pesquisadores vinculados a universidades brasileiras têm utilizado repositórios internacionais para disponibilizar conjuntos de dados. O estudo objetiva analisar as características dos dados de pesquisa associados a instituições brasileiras depositados nos repositórios *Zenodo* e *Dryad*. Trata-se de uma pesquisa documental, com abordagem quali-quantitativa, baseada na recuperação e análise de metadados dos registros identificados nessas plataformas. Foram examinadas informações relacionadas à distribuição temporal dos depósitos, instituições de afiliação, áreas do conhecimento e indicadores de impacto. Os resultados identificaram 1.762 conjuntos de dados, com predominância do *Dryad*, crescimento dos depósitos a partir de 2016 e concentração em universidades públicas de maior produção científica, especialmente nas Ciências Biológicas e da Vida. Conclui-se que a presença desses dados em repositórios internacionais evidencia o avanço das práticas de Ciência Aberta e a inserção da produção científica brasileira em ambientes globais de compartilhamento.
Palavras-chave: dados de pesquisa; repositório de dados; ciência aberta.

Abstract: Open Science has heightened the importance of sharing research data and strengthened the role of digital repositories in scholarly communication. In this context, researchers affiliated with Brazilian universities have been using international repositories to

make datasets available. This study aims to analyze the characteristics of research data associated with Brazilian institutions that have been deposited in the Zenodo and Dryad repositories. It is a documentary study employing a mixed-methods (qualitative-quantitative) approach, based on the retrieval and analysis of metadata from records identified on these platforms. Information regarding the temporal distribution of deposits, affiliated institutions, fields of knowledge, and impact indicators was examined. The results identified 1,762 datasets, with a predominance of deposits in Dryad, an increase in submissions starting in 2016, and a concentration among public universities with high scientific output, particularly in the Biological and Life Sciences. The study concludes that the presence of this data in international repositories highlights the advancement of Open Science practices and the integration of Brazilian scientific output into global sharing environments.
Keywords: research data; data repository; open science.

1 INTRODUÇÃO

A intensificação da Ciência Aberta tem transformado os modos de produzir, comunicar e preservar o conhecimento científico. Impulsionado por demandas de transparência, reprodutibilidade, acesso e reutilização dos resultados de pesquisa, esse movimento reposiciona dados, métodos e infraestruturas informacionais como elementos centrais do processo científico. Nesse contexto, "[...] os dados deixam de ser simples subprodutos das atividades de pesquisa e se tornam recursos informacionais de primeira grandeza, caracterizando um novo paradigma" (Sayão; Sales, 2016, p. 90).

No Brasil, o debate sobre dados de pesquisa tem avançado por meio de iniciativas institucionais, políticas de gestão e discussões relacionadas à Ciência Aberta. Entretanto, persistem limitações quanto à disponibilidade de repositórios especializados, às políticas de compartilhamento e às práticas adotadas pelos pesquisadores (Gabriel Junior *et al.*, 2019; Matos; Ribeiro Júnior, 2025; Nicodemo, 2025). Como alternativa, diversas instituições passaram a incorporar conjuntos de dados em repositórios institucionais originalmente destinados às publicações científicas, ampliando sua visibilidade e acesso, embora também enfrentem desafios relacionados à curadoria, interoperabilidade e políticas específicas para esses recursos (Felicíssimo *et al.*, 2024; Lopes *et al.*, 2023; Silveira; Vanz; Pavão, 2024).

Além das iniciativas nacionais, pesquisadores vinculados a instituições brasileiras recorrem a repositórios internacionais para disponibilizar seus dados de pesquisa. Segundo Brumatti, Pinto e Alencar (2026), essa prática amplia a visibilidade da produção científica nacional e favorece sua inserção em redes globais de compartilhamento de dados, ao mesmo tempo em que evidencia possíveis lacunas na infraestrutura nacional destinada à sua gestão.

Embora estudos recentes tenham analisado a infraestrutura brasileira de dados de

pesquisa e as iniciativas relacionadas à Ciência Aberta, ainda são escassas investigações voltadas à caracterização dos conjuntos de dados produzidos por instituições brasileiras e disponibilizados em repositórios internacionais. Questões relacionadas à distribuição institucional dos depósitos, às áreas do conhecimento predominantes e ao alcance desses conjuntos de dados permanecem pouco exploradas. A análise de registros depositados no *Zenodo* e no *Dryad* contribui para compreender o estágio atual da abertura de dados no Brasil e suas relações com a infraestrutura científica nacional. Considerando que a oferta de repositórios especializados ainda é limitada (Gabriel Junior *et al.*, 2019; Vieira, 2022), as práticas de depósito em repositórios internacionais podem revelar demandas institucionais, tendências de compartilhamento e desafios para a gestão e o acesso aos dados de pesquisa.

Diante desse contexto, busca-se responder à seguinte questão: quais são as características dos conjuntos de dados abertos de universidades brasileiras disponibilizados em repositórios internacionais?

2 REFERENCIAL TEÓRICO

A intensificação das práticas associadas à Ciência Aberta tem provocado transformações significativas nos processos de produção, comunicação e circulação do conhecimento científico. Mais do que ampliar o acesso às publicações, a Ciência Aberta representa uma reconfiguração das dinâmicas de pesquisa baseada em princípios de transparência, colaboração, compartilhamento e reutilização dos resultados científicos. Nesse contexto, o conhecimento passa a ser compreendido como um bem coletivo, cuja produção e disseminação devem ocorrer de forma mais aberta, inclusiva e acessível (Albagli, 2017; UNESCO, 2021).

A consolidação desse movimento está associada às iniciativas de acesso aberto surgidas a partir do final da década de 1990, especialmente com a *Open Archives Initiative* e as declarações de Budapeste, Bethesda e Berlim. Posteriormente, a Ciência Aberta ampliou seu escopo para incorporar dados de pesquisa, metodologias, softwares, infraestruturas e processos colaborativos de produção do conhecimento (Rios; Lucas; Amorim, 2019; UNESCO, 2021).

Paralelamente, observa-se a expansão de repositórios especializados em dados de pesquisa. Estudos desenvolvidos por Silva e Ottoni (2023), Bonetti *et al.* (2024), Marques Felipe e Santos (2024) e Danin, Oliveira e Benchimol (2025) demonstram o crescimento desse

ecossistema no Brasil. Iniciativas como o Lattes Data (Groehs *et al.*, 2023) e o Deposita Dados (Bonetti *et al.*, 2025a, 2025b) representam avanços na construção de uma infraestrutura nacional voltada à gestão e ao compartilhamento de dados científicos.

Nesse cenário, os dados de pesquisa assumem papel central e são reconhecidos como produtos científicos relevantes, cuja disponibilização favorece a verificação dos resultados, a reutilização das informações e o avanço cumulativo da ciência (Mellinger; Hanson, 2022). A literatura os compreende como registros produzidos, coletados, observados ou processados durante investigações científicas (Rodrigues *et al.*, 2010), cujos significados dependem dos contextos disciplinares e metodológicos em que são produzidos (Nielsen; Hjørland, 2014).

O reconhecimento dos dados como produtos da atividade científica impulsionou políticas e diretrizes voltadas ao seu compartilhamento. Organismos internacionais, agências de fomento e instituições de pesquisa passaram a incentivar práticas de gestão e abertura dos dados científicos, tornando cada vez mais frequente a exigência de planos de gestão e depósito em repositórios (Mellinger; Hanson, 2022; Rocha; Caregnato; Gabriel Junior, 2018; Sayão; Sales, 2016).

A disponibilização dos dados de pesquisa depende de infraestruturas capazes de assegurar armazenamento, organização, preservação e recuperação. Nesse contexto, os repositórios de dados de pesquisa constituem ambientes destinados ao depósito, gerenciamento e compartilhamento de conjuntos de dados científicos (Rodrigues *et al.*, 2010). A crescente relevância dos repositórios de dados está associada ao desenvolvimento de princípios internacionais destinados a promover a qualidade, a interoperabilidade e a reutilização dos dados científicos. Entre eles destacam-se os princípios *Findable, Accessible, Interoperable and Reusable* (FAIR), propostos por Wilkinson *et al.* (2016), que estabelecem diretrizes para tornar os dados encontráveis, acessíveis, interoperáveis e reutilizáveis.

Embora os repositórios de dados de pesquisa sejam reconhecidos como as infraestruturas mais adequadas para armazenamento, preservação e compartilhamento desses recursos, diferentes arranjos institucionais têm sido adotados para viabilizar tais práticas. Dube (2025), Vieira (2022) e Silveira, Vanz e Pavão (2024) demonstram que repositórios institucionais passaram a assumir funções relacionadas à gestão de dados de pesquisa, coexistindo com repositórios especializados.

Estudos internacionais apontam crescimento dos repositórios de dados de pesquisa, embora persistam assimetrias geográficas e institucionais, com maior concentração dessas

infraestruturas em países do Norte Global e limitações observadas em países do BRICS (Khan; Loan; Parray; Rashid, 2024; Misgar; Bhat; Wani, 2022; Naheem; Mir, 2024).

No Brasil, as discussões sobre dados de pesquisa ganharam força na última década. Gabriel Junior *et al.* (2019) e Vieira (2022) identificaram limitações na oferta de repositórios especializados, enquanto estudos mais recentes apontam avanços nas iniciativas de gestão e compartilhamento de dados, embora a infraestrutura disponível ainda permaneça limitada diante das demandas da comunidade científica (Bonetti *et al.*, 2024; Silva; Carvalho; Freitas, 2023; Rodrigues; Dias, 2025).

Diante desse cenário, pesquisadores brasileiros recorrem tanto a repositórios nacionais quanto a plataformas internacionais para disponibilização de seus conjuntos de dados. Conforme observam Brumatti, Pinto e Alencar (2026), infraestruturas como *Zenodo* e *Dryad* ampliam a visibilidade internacional da produção científica brasileira e favorecem sua inserção em ecossistemas globais de informação científica.

3 PROCEDIMENTOS METODOLÓGICOS

A pesquisa caracteriza-se como documental, de abordagem quali-quantitativa, com caráter descritivo e exploratório, tendo como objetivo mapear e analisar conjuntos de dados de pesquisa associados a universidades brasileiras e disponibilizados nos repositórios internacionais *Zenodo* e *Dryad*. O estudo concentra-se na identificação, extração, tratamento e análise de metadados descritivos e indicadores de uso dos conjuntos de dados, considerando especialmente o tipo de recurso, a área do conhecimento, a data de publicação, a afiliação institucional dos autores e as métricas de impacto disponíveis, como visualizações e downloads.

O *Zenodo* foi selecionado por se tratar de um repositório digital aberto, multidisciplinar e de alcance internacional, desenvolvido no âmbito do OpenAIRE e operado pelo CERN. A plataforma tem como função disponibilizar e preservar diferentes produtos resultantes de atividades de pesquisa, incluindo publicações, conjuntos de dados e softwares. Para isso, atribui identificadores persistentes aos registros e disponibiliza metadados que favorecem sua descoberta, acesso e reutilização. Essas características tornam o *Zenodo* uma infraestrutura relevante para o compartilhamento de diferentes tipos de produção científica em ambiente internacional (CERN; OpenAIRE, 2026?).

O *Dryad*, por sua vez, constitui uma plataforma voltada à publicação, preservação e

compartilhamento de dados de pesquisa. Sua função está centrada na disponibilização de conjuntos de dados associados à produção científica, favorecendo sua descoberta, acesso e reutilização por outros pesquisadores. A plataforma adota procedimentos de curadoria dos dados depositados e disponibiliza metadados para identificação e recuperação dos registros (*Dryad*, c2026). A inclusão do *Dryad* na pesquisa deve-se, portanto, à sua especialização na publicação e preservação de dados de pesquisa e ao seu alcance internacional.

Os dados[1] foram coletados, com auxílio de *scripts* em *Python*[2], por meio da recuperação de metadados disponíveis no *Zenodo* e *Dryada,* as estratégias de busca foram baseadas na sigla, no nome por extenso em português e no nome em inglês das universidades federais e estaduais brasileiras. Os registros foram identificados a partir da ocorrência dessas denominações nos metadados, especialmente nas informações de afiliação, sendo considerados os registros com vínculo identificável às universidades brasileiras. Em casos de múltiplas afiliações, foi considerada a primeira universidade identificada no registro. Foram excluídos registros sem afiliação identificável e aqueles fora do escopo da pesquisa. Os dados foram consolidados em planilha eletrônica e submetidos a procedimentos de tratamento e normalização.

A coleta dos dados foi realizada entre 05 de maio e 09 de junho de 2026, por meio das *Application Programming Interfaces* (APIs) oficiais do *Zenodo* (*InvenioRDM*) e do *Dryad* (API v2). Os metadados recuperados foram tratados e normalizados, incluindo a padronização das instituições e a classificação das áreas do conhecimento com base na Tabela de Áreas do Conhecimento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (CNPq, 2025).

Foram considerados os seguintes metadados: identificador do registro (ID), *Digital Object Identifier* (DOI), título, autores, afiliações, data de publicação, descrição, palavras-chave, área do conhecimento, licença, comunidades, visualizações, downloads e URL. Para a identificação e caracterização das instituições, foram registrados também os campos: instituição detectada, nome da universidade e tipo de universidade. Adicionalmente, foi considerada a variável tipo de dado específico, utilizada na caracterização dos conjuntos de

[1] Os dados tratados, a documentação das variáveis e os scripts utilizados serão disponibilizados em acesso aberto juntamente com a publicação do trabalho, em consonância com os princípios da Ciência Aberta, transparência e reprodutibilidade da pesquisa.

[2] *Python* é uma linguagem de programação de alto nível, de propósito geral, utilizada para desenvolver aplicações, automatizar tarefas, processar e analisar dados (*Python Software Foundation*, 2026).

dados.

A análise foi conduzida de maneira descritiva a partir das seguintes categorias analíticas: distribuição temporal dos depósitos, a participação das universidades brasileiras e a concentração por áreas do conhecimento. Adicionalmente, foram analisados indicadores de visualização e *download*, visando identificar padrões de circulação e uso dos conjuntos de dados nos repositórios estudados.

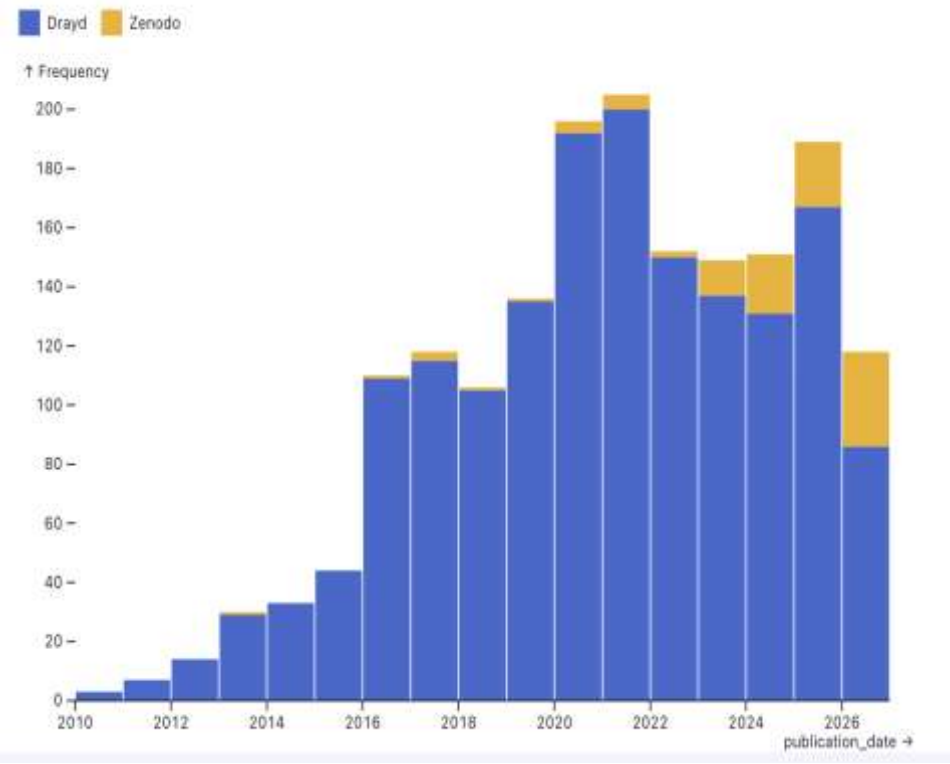
4 RESULTADOS E DISCUSSÃO

Foram identificados 1.762 conjuntos de dados vinculados a universidades brasileiras nos repositórios Zenodo e *Dryad*. Desse total, 1.657 registros (94,0%) encontram-se no *Dryad* e 105 (6,0%) no *Zenodo*, evidenciando a predominância do primeiro repositório como ambiente de compartilhamento de dados por pesquisadores brasileiros. A análise desses registros permitiu identificar características relacionadas à evolução temporal dos depósitos, à distribuição institucional e disciplinar da produção de dados e aos indicadores de impacto associados à sua utilização e visibilidade.

4.1 RASTRO TEMPORAL

A análise da distribuição temporal dos conjuntos de dados, demonstrada no Gráfico 1, permitiu identificar marcos importantes na trajetória de depósito de dados por universidades brasileiras nos repositórios analisados.

Gráfico 1 - Distribuição temporal - *Dryad* e *Zenodo*



Fonte: elaboração própria (2026[3]).

Embora sejam observados registros isolados desde 2010, o crescimento torna-se mais expressivo a partir de 2016, quando o número de depósitos ultrapassa uma centena de conjuntos de dados. Esse marco coincide com a publicação dos princípios FAIR (Wilkinson *et al*., 2016), que estabeleceram diretrizes de gestão e compartilhamento de dados científicos.

Entre 2020 e 2022 observa-se o período de maior expansão dos depósitos, alcançando 196 registros em 2020 (11,1% do total analisado) e 205 em 2021 (11,6%). Em 2022, foram identificados 152 registros (8,6%). Esse crescimento ocorre em um contexto de fortalecimento das práticas de Ciência Aberta em âmbito internacional, marcado pela ampliação das exigências de agências de fomento e periódicos relacionadas ao compartilhamento de dados de pesquisa e pela publicação da Recomendação da UNESCO sobre Ciência Aberta, em 2021.

Após o pico registrado em 2021, com 205 conjuntos de dados (11,6%), os depósitos mantiveram-se em patamares elevados nos anos seguintes, alcançando 152 registros em 2022 (8,6%), 149 em 2023 (8,5%), 151 em 2024 (8,6%) e 189 em 2025 (10,7%). Esse comportamento sugere que o compartilhamento de dados em repositórios internacionais deixou de

[3] Os registros de 2026 correspondem apenas ao período compreendido entre janeiro e 9 de junho, não sendo diretamente comparáveis aos anos completos da série.

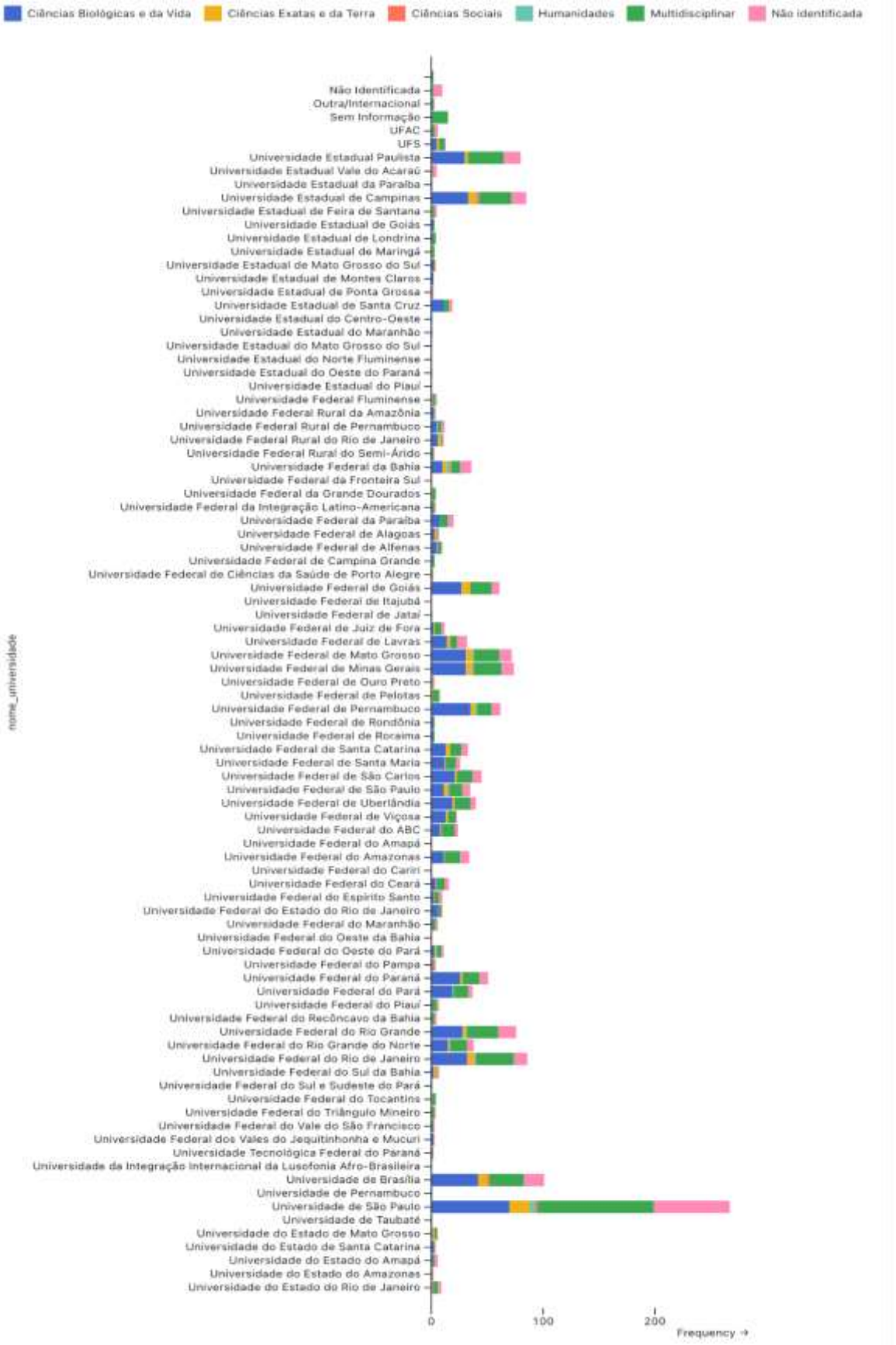
representar uma prática isolada para se tornar uma estratégia cada vez mais incorporada às atividades de pesquisa. Os resultados corroboram estudos que apontam a expansão da infraestrutura global de dados de pesquisa e a crescente participação brasileira nesse ecossistema (Brumatti, Pinto; Alencar, 2026; Khan; Loan; Parray; Rashid, 2024).

Destaca-se ainda o aumento da participação do *Zenodo* entre 2024 e 2025. Embora o *Dryad* permaneça predominante ao longo de toda a série histórica, o crescimento dos depósitos no *Zenodo* sugere uma diversificação das estratégias de compartilhamento adotadas pelos pesquisadores brasileiros.

4.2 ESPAÇOS DE PRODUÇÃO DOS DADOS: INSTITUIÇÕES E ÁREAS DO CONHECIMENTO

Além da evolução temporal dos depósitos, torna-se relevante compreender quem são os principais atores envolvidos na produção e compartilhamento dos dados de pesquisa. Nesse sentido, foram analisadas as universidades de vínculo dos conjuntos de dados e suas respectivas áreas do conhecimento, permitindo identificar padrões de concentração institucional e disciplinar nos repositórios *Zenodo* e *Dryad*. O Gráfico 2 apresenta a distribuição dos conjuntos de dados por universidade e área do conhecimento.

Gráfico 2 - Universidades e áreas do conhecimento



Fonte: elaboração própria (2026).

A distribuição institucional dos conjuntos de dados revela forte concentração em universidades públicas com elevada produção científica e significativa participação na pesquisa nacional. A Universidade de São Paulo (USP) apresenta o maior volume de depósitos, totalizando 267 conjuntos de dados (15,2% do total), seguida pela Universidade de Brasília (UnB), com 101 registros (5,7%), pela Universidade Federal do Rio de Janeiro (UFRJ), com 86 (4,9%), pela Universidade Estadual de Campinas (Unicamp), com 85 (4,8%), e pela Universidade Estadual Paulista (UNESP), com 80 (4,5%). Em conjunto, essas cinco instituições respondem por 619 registros, correspondentes a 35,1% dos depósitos identificados nos repositórios analisados, evidenciando a concentração dos registros em instituições de pesquisa de maior participação no conjunto analisado.

Quanto às áreas do conhecimento, observa-se predominância das Ciências Biológicas e da Vida, com 650 conjuntos de dados (36,9% do total), seguida pela categoria Multidisciplinar, com 618 registros (35,1%). As Ciências Exatas e da Terra totalizam 116 conjuntos de dados (6,6%), enquanto Humanidades correspondem a 21 registros (1,2%) e Ciências Sociais a 17 (1,0%). Em 340 registros (19,3%), não foi possível identificar a área do conhecimento a partir dos metadados analisados. A distribuição evidencia uma participação expressiva das Ciências Biológicas e da Vida e da categoria Multidisciplinar no conjunto de dados analisado, enquanto as demais áreas apresentam menor frequência de registros.

A predominância das Ciências Biológicas e da Vida acompanha tendências observadas em estudos internacionais sobre compartilhamento de dados de pesquisa. Pesquisas realizadas a partir do diretório Re3Data demonstram que áreas como Biologia, Ecologia e Ciências Ambientais figuram entre aquelas com maior tradição na produção, preservação e abertura de conjuntos de dados. Tal comportamento é frequentemente associado à natureza empírica dessas disciplinas, à produção de grandes volumes de dados observacionais e experimentais e às exigências de transparência, validação e reprodutibilidade dos resultados científicos (Khan; Loan; Parray; Rashid, 2024; Mellinger; Hanson, 2022).

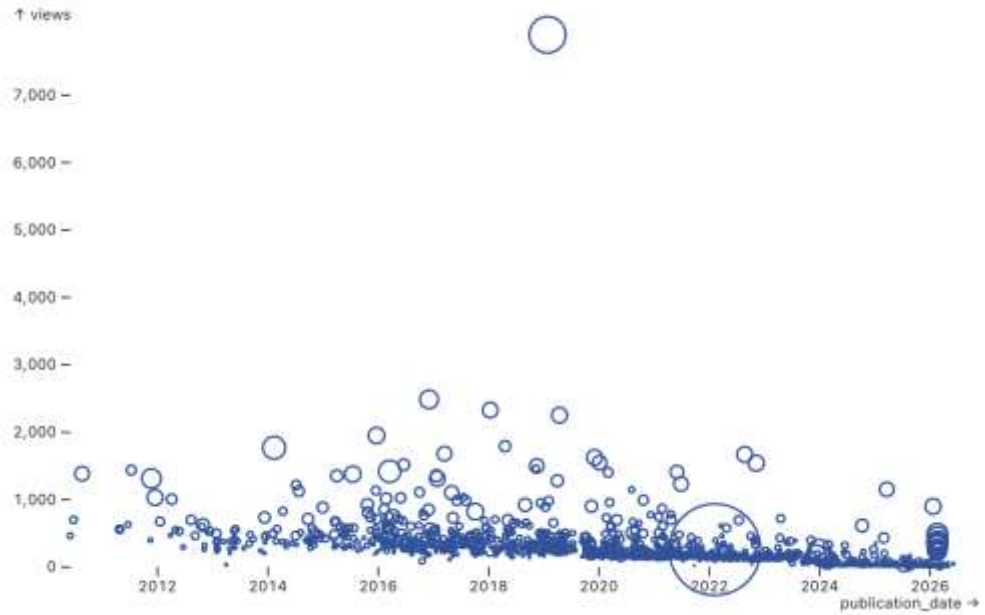
A abertura de dados de pesquisa depende não apenas da produção científica, mas também da existência de recursos, infraestrutura e políticas institucionais capazes de sustentar práticas de gestão, curadoria e compartilhamento ao longo do tempo.

4.3 INDICADORES DE IMPACTO: MÉTRICAS DE VISUALIZAÇÃO E *DOWNLOAD*

Fundamentado nos conjuntos de dados disponibilizados, é importante compreender o

alcance e a visibilidade dessas publicações no ambiente digital. Nesse sentido, as métricas de visualização e *download* constituem indicadores relevantes para avaliar o nível de atenção recebido pelos dados de pesquisa e seu potencial de circulação entre diferentes comunidades científicas. Embora não permitam mensurar diretamente a reutilização efetiva dos dados, essas métricas oferecem evidências sobre o interesse despertado e sua visibilidade. O Gráfico 3 apresenta a distribuição das visualizações ao longo do período analisado, permitindo identificar padrões de atenção e impacto associados aos depósitos realizados por universidades brasileiras nos repositórios analisados.

Gráfico 3 - Indicadores de impacto - visualização e *download*



Fonte: elaboração própria (2026).

A análise das métricas de visualização e download evidencia maior circulação dos registros publicados principalmente entre 2018 e 2022.

Esse comportamento indica maior visibilidade e circulação desses registros no período analisado. Embora o número de depósitos continue crescendo nos anos mais recentes, os conjuntos de dados publicados após 2022 apresentam, de modo geral, níveis inferiores de visualização e download, o que pode estar relacionado ao menor tempo de exposição desses registros e, consequentemente, à menor oportunidade de acumular acessos e downloads.

A predominância de registros com maior impacto entre 2018 e 2022 também sugere um processo de amadurecimento das práticas de compartilhamento de dados associado ao

avanço da Ciência Aberta. Como destacam Wilkinson *et al.* (2016), dados disponibilizados de acordo com os princípios FAIR possuem maior potencial de descoberta, acesso e reutilização. Dessa forma, a visibilidade alcançada por parte dos conjuntos de dados analisados pode estar relacionada tanto às características das plataformas utilizadas quanto à crescente valorização da abertura de dados no contexto científico internacional.

Quando os indicadores de impacto são analisados em conjunto com os resultados apresentados nas seções anteriores, observa-se que os conjuntos de dados mais visíveis tendem a estar associados às instituições que também concentram maior volume de depósitos. Universidades como USP (15,2% dos depósitos), UnB (5,7%), UFRJ (4,9%), Unicamp (4,8%) e UNESP (4,5%) destacaram-se tanto na quantidade de conjuntos de dados disponibilizados quanto na presença de registros com elevados níveis de atenção. Em conjunto, essas instituições concentram 35,1% dos conjuntos de dados identificados, evidenciando a relação entre capacidade institucional, produção científica e visibilidade dos dados compartilhados.

Por fim, a presença de conjuntos de dados brasileiros com elevados índices de visualização e download em repositórios internacionais evidencia a capacidade dessas plataformas de ampliar a exposição da produção científica nacional. Esse resultado converge com as observações de Brumatti, Pinto e Alencar (2026), que destacam o papel dos repositórios internacionais na inserção da ciência brasileira em redes globais de compartilhamento de dados.

Em suma, as métricas analisadas indicam que os conjuntos de dados produzidos por universidades brasileiras não apenas estão sendo disponibilizados em ambientes internacionais, mas também alcançam níveis relevantes de atenção e circulação, ampliando seu potencial de uso e reutilização pela comunidade científica.

5 CONCLUSÃO

A crescente presença de dados produzidos por universidades brasileiras em repositórios internacionais evidencia não apenas a ampliação das práticas de compartilhamento científico, mas também a necessidade de compreender como esses dados estão sendo disponibilizados, quais áreas do conhecimento apresentam maior participação, quais instituições se destacam nesse processo e quais indicadores podem revelar seu alcance e utilização. Os resultados permitem responder à questão de pesquisa ao evidenciar que os

conjuntos de dados abertos de universidades brasileiras disponibilizados nos repositórios *Zenodo* e *Dryad* caracterizam-se pela concentração dos depósitos em determinadas instituições e áreas do conhecimento, pela predominância de registros no *Dryad* e por diferentes níveis de visibilidade, identificados a partir das métricas de visualizações e *downloads*. Nesse sentido, a análise dos depósitos realizados em plataformas internacionais constitui uma oportunidade para compreender aspectos relevantes da inserção da ciência brasileira nos fluxos globais de compartilhamento de dados de pesquisa.

A descrição da prática identificada ainda pode ser um indicativo para as universidades destacadas da necessidade de seus pesquisadores de estruturas adequadas para o armazenamento, gestão e compartilhamento de dados. Nesse sentido, os resultados do estudo podem auxiliar as instituições a compreenderem parte dessa demanda já existente e que vem amadurecendo em suas práticas de ciência aberta.

Os resultados evidenciam que a participação brasileira nos repositórios internacionais de dados de pesquisa está concentrada em universidades mais consolidadas. A forma como a prática se manifesta no Brasil também precisa ser observada a partir do recorte das áreas do conhecimento, da qual se evidencia mais uma vez a tendência de aproximação das Ciências Biológicas e da Vida, como mais engajadas com práticas de compartilhamento de dados e abertura das estruturas da ciência. Esse cenário aproxima-se das tendências observadas internacionalmente e demonstra avanços na inserção da ciência brasileira nos ecossistemas globais de Ciência Aberta. Entretanto, a concentração dos depósitos em um número restrito de instituições também revela assimetrias que ainda caracterizam o desenvolvimento da infraestrutura nacional de gestão e compartilhamento de dados de pesquisa, indicando a necessidade de ampliação de políticas, investimentos e mecanismos de apoio para outras instituições e áreas do conhecimento.

Por fim, o estudo contribui para a compreensão das dinâmicas de compartilhamento de dados de pesquisa por universidades brasileiras em repositórios internacionais, oferecendo evidências empíricas sobre a distribuição temporal, institucional e temática desses depósitos. Além disso, fornece subsídios para o planejamento de políticas institucionais de gestão de dados e para o fortalecimento da infraestrutura nacional de Ciência Aberta.

REFERÊNCIAS

ALBAGLI, Sarita. Ciência aberta como instrumento de democratização do saber. Trabalho, Educação e Saúde, Rio de Janeiro, v. 15, n. 3, p. 659-660, set./dez. 2017. Disponível em: https://doi.org/10.1590/1981-7746-sol00093. Acesso em: 1 nov. 2025.

BONETTI, Letícia Guarany; SILVA, Tatyane Guedes Martins da; GABRIEL JUNIOR, Rene Faustino; SANTOS, Blena Estevam dos *et al.* Características dos repositórios de dados de pesquisa brasileiros: uma análise a partir do diretório re3data. *In*: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 30., 2024, Recife. **Anais** [...]. Recife: FEBAB, 2024. p. 15-28. Disponível em: https://portal.febab.org.br/cbbd2024/article/view/3305. Acesso em: 7 maio 2025.

BONETTI, Letícia Guarany; SILVA, Tatyane Guedes Martins da; PAVÃO, Caterina Marta Groposo; VANZ, Samile Andrea de Souza *et al.* Curadoria dos dados de pesquisa no repositório nacional Deposita Dados. *In*: CONFERÊNCIA INTERNACIONAL BIREDIAL-ISTEC, 14., 2025, Brasília. **Anais** [...]. Brasília: ISTEC, 2025. Disponível em: https://submissions.istec.org/index.php/biredial-istec/article/view/379. Acesso em: 7 set. 2026.

BONETTI, Letícia Guarany; SILVA, Tatyane Guedes Martins da; PAVÃO, Caterina Marta Groposo; VANZ, Samile Andrea de Souza *et al.* Desafios na curadoria de dados de pesquisa no repositório nacional Deposita Dados. **Revista Ciência Aberta Lusófona**, [*s. l.*], v. 1, p. 179-186, 2025. DOI: https://doi.org/10.82524/recal.2025.38. Disponível em: https://recal.rcaap.pt/recal/article/view/38. Acesso em: 2 set. 2026.

BRUMATTI, Josimara Dias; PINTO, Fabiana de Melo Amaral Gonçalves; ALENCAR, Maria Simone de Menezes. Dados de pesquisa no Brasil: iniciativas nacionais e evidências de compartilhamento a partir da OpenAlex. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 25., 2025, Rio de Janeiro. **Anais** [...]. Rio de Janeiro: ENANCIB, 2026. v. 5, n. 1. Disponível em: https://enancib.emnuvens.com.br/enancib/pt_BR/article/view/1821. Acesso em: 7 set. 2026.

CERN; OPENAIRE. **About Zenodo**. [Genebra]: Zenodo, 2026?. Disponível em: https://about.zenodo.org/. Acesso em: 14 jun. 2026.

CONSELHO NACIONAL DE DESENVOLVIMENTO CIENTÍFICO E TECNOLÓGICO (CNPq). **Tabela de áreas do conhecimento**. Brasília, DF: CNPq, 2025. Disponível em: https://lattes.cnpq.br/web/dgp/arvore-do-conhecimento. Acesso em: 7 set. 2026.

DANIN, Gisela Fernanda Monteiro; OLIVEIRA, Hamilton Vieira de; BENCHIMOL, Alegria. A presença brasileira no registro de repositórios de dados de pesquisa. **Revista Ciência Aberta Lusófona**, [*s. l.*], v. 1, p. 187-192, 2025. Disponível em: https://recal.rcaap.pt/recal/article/view/42. Acesso em: 12 set. 2026.

DRYAD. Who we are. [*S. l.*], c2026. Disponível em: https://datadryad.org/about. Acesso em: 14 jun. 2026.

DUBE, Tinyiko Vivian. Research data management in academic libraries: institutional repositories as a reservoir for research data. **Library Management**, [*s. l.*], v. 46, n. 5, p. 319-331, 2025. Disponível em: https://doi.org/10.1108/LM-06-2024-0070. Acesso em: 6 mar. 2026.

FELICÍSSIMO, Carolina Howard; AMARO, Bianca; CUNHA, Dileine Amaral da; CIUFFO, Leandro Neumann *et al.* Núcleo de Dados de Pesquisa: um modelo de governança para a iniciativa de rede brasileira de repositórios institucionais de dados de pesquisa. **Ciência da Informação**, Brasília, v. 53, n. 3, p. 54-59, 2024. Disponível em: https://revista.ibict.br/ciinf/article/view/7200. Acesso em: 2 set. 2026.

GABRIEL JUNIOR, Rene Faustino; ROCHA, Rafael Port da; CAREGNATO, Sônia Elisa; PAVÃO, Caterina Marta Groposo *et al.* Acesso aberto a dados de pesquisa no Brasil: mapeamento de repositórios, práticas e percepções dos pesquisadores e tecnologias. **Ciência da Informação**, Brasília, v. 48, n. 3, suplemento, p. 87-101, set./dez. 2019. Disponível em: https://revista.ibict.br/ciinf/article/view/4958. Acesso em: 15 nov. 2025.

GROES, Adriane; SILVEIRA, Crislaine Zurilda; LIMA, Elaine Maria de; DITTRICH, Eliane. LattesData e a adoção aos Princípios FAIR: uma análise usando a F-UJI Automated FAIR Data Assessment Tool. **Em Questão**, Porto Alegre, v. 29, e-130018, 2023. Disponível em: https://doi.org/10.1590/1808-5245.29.130018. Acesso em: 4 nov. 2026.

KHAN, Aasif Mohammad; LOAN, Fayaz Ahmad; PARRAY, Umer Yousuf; RASHID, Sozia. Global overview of research data repositories: an analysis of re3data registry. **Information Discovery and Delivery**, [*s. l.*], v. 52, n. 1, p. 53-61, 2024. Disponível em: https://doi.org/10.1108/IDD-07-2022-0069. Acesso em: 7 mar. 2026.

LOPES, Ana Julia; FELICISSIMO, Carolina Howard; PAVÃO, Caterina Groposo; CUNHA, Dileine Amaral da *et al.* Um modelo de redes colaborativas representado na iniciativa de rede brasileira de repositórios institucionais de dados de pesquisa. **BiblioCanto**, Natal, v. 9, n. 2, p. 1-15, 2023. DOI: https://doi.org/10.21680/2447-7842.2023v9n2ID33659. Disponível em: https://periodicos.ufrn.br/bibliocanto/article/view/33659. Acesso em: 7 mar. 2026.

FELIPE, Carla Beatriz Marques; SANTOS, Raimunda Fernanda dos. Panorama dos repositórios de dados de pesquisa brasileiros. *In*: CONFERÊNCIA INTERNACIONAL DE BIBLIOTECAS E REPOSITÓRIOS DIGITAIS, 13., 2024, Santiago, Chile. **Anais** [...]. Santiago: Ibero-American Science and Technology Education Consortium, 2024. p. 164-174. Disponível em: https://sedici.unlp.edu.ar/handle/10915/180657. Acesso em: 12 set. 2026.

MATOS, Igor Yure Ramos; RIBEIRO JÚNIOR, Divino Ignácio. A concepção de uma política de gestão de dados de pesquisa nas universidades brasileiras. *In*: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS, 23., 2025, São Paulo. **Anais** [...]. São Paulo: FEBAB, 2025. p. 1-13. Disponível em: https://portal.febab.org.br/snbu2025/article/view/3913. Acesso em: 2 set. 2026.

MELLINGER, Christopher D.; HANSON, Thomas A. Research data. *In*: ZANETTIN, Federico; RUNDLE, Christopher (ed.). **The Routledge handbook of translation and methodology**.

London: Routledge, 2022. p. 307-323. Disponível em: https://www.taylorfrancis.com/chapters/edit/10.4324/9781315158945-23/research-data-christopher-mellinger-thomas-hanson. Acesso em: 12 jan. 2026.

MISGAR, Safat Mushtaq; BHAT, Ajra; WANI, Zahid Ashraf. A study of open access research data repositories developed by BRICS countries. **Digital Library Perspectives**, [*s. l.*], v. 38, n. 1, p. 45-54, 2022. Disponível em: https://www.emerald.com/dlp/article/38/1/45/1273954. Acesso em: 5 mar. 2026.

NAHEEM, K. T.; MIR, Aasif Ahmad. Analyzing research data repositories (RDR) from BRICS nations: a comprehensive study. **Library Management**, [*s. l.*], v. 45, n. 6-7, p. 429-441, 2024. Disponível em: https://www.emerald.com/lm/article/45/6-7/429/1212533. Acesso em: 5 mar. 2026.

NICODEMO, Thiago Lima. Dados abertos de pesquisa no Brasil: diagnóstico e perspectivas futuras. **Ciência e Cultura**, [*s. l.*], v. 77, n. 1, p. 33-41, 2025. DOI: https://doi.org/10.48207/2317-6660.20250005. Disponível em: http://cienciaecultura.bvs.br/pdf/cic/v77n1/a05v77n1.pdf. Acesso em: 4 mar. 2026.

NIELSEN, Hans Jørn; HJØRLAND, Birger. Curating research data: the potential roles of libraries and information professionals. **Journal of Documentation**, [*s. l.*], v. 70, n. 2, p. 221-240, 2014. Disponível em: https://doi.org/10.1108/JD-03-2013-0034. Acesso em: 7 set. 2026.

PYTHON SOFTWARE FOUNDATION. What is Python? **Python.org**, [*s. l.*], 2026. Disponível em: https://www.python.org/about/. Acesso em: 22 jun. 2026.

RIOS, Fahima Pinto; LUCAS, Elaine Rosangela Oliveira; AMORIM, Igor Soares. Manifestos do movimento de acesso aberto: Análise de Domínio a partir de periódicos brasileiros. **Revista Brasileira de Biblioteconomia e Documentação**, [*s. l.*], v. 15, n. 1, p. 148-169, 2019. Disponível em: https://rbbd.febab.org.br/rbbd/article/view/1152. Acesso em: 14 fev. 2026.

RODRIGUES, Adriana Alves; DIAS, Guilherme Ataíde. Repositórios de dados abertos de pesquisa em plataformas de universidades brasileiras. **Ciência da Informação em Revista**, Maceió, v. 12, e17799, 2025. Disponível em: https://periodicos.ufal.br/cir/article/view/17799. Acesso em: 7 set. 2026.

RODRIGUES, Eloy; BAPTISTA, Ana Alice; SOARES, Ana; COSTA, Carla *et al.* **Os repositórios de dados científicos**: estado da arte. Braga: Universidade do Minho. Serviços de Documentação, 2010. 54 p. Disponível em: http://repositorium.sdum.uminho.pt/handle/1822/10830. Acesso em: 5 set. 2026.

ROCHA, Rafael Port da; CAREGNATO, Sônia Elisa; GABRIEL JUNIOR, Rene Faustino. Aspectos de inovação na implantação de um centro de digitalização e gestão de dados da pesquisa. **Encontros Bibli**: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 23, n. esp. 1, p. 1-15, 2018. DOI: https://doi.org/10.5007/1518-2924.2018v23nespp1. Disponível em: https://brapci.inf.br/v/35605. Acesso em: 12 jan. 2026.

SAYÃO, Luis Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa**. Informação & Informação**, Londrina, v. 21, n. 2, p. 90-115, maio/ago., 2016. Disponível em: https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/27939/20122. Acesso em: 18 jan. 2026.

SILVA, Fabiane Cristina Carvalho da; CARVALHO, Kátia de; FREITAS, Angela Maria da Cunha. Diagnóstico dos repositórios de dados no Brasil. **Brazilian Journal of Information Science**: Research Trends, Marília, v. 17, e023031, 2023. DOI: https://doi.org/10.36311/1981-1640.2023.v17.e023031. Disponível em: https://revistas.marilia.unesp.br/index.php/bjis/article/view/14675. Acesso em: 2 set. 2026.

SILVA, Fabrício Silva; OTTONI, Heloisa. Repositórios temáticos de dados de pesquisa em ambientes brasileiros: questões em debate. **BiblioCanto**, Natal, v. 9, n. 2, p. 139-144, 2023. DOI: https://doi.org/10.21680/2447-7842.2023v9n2ID33755. Disponível em: https://periodicos.ufrn.br/bibliocanto/article/view/33755. Acesso em: 2 set. 2026.

SILVEIRA, J. I.; VANZ, Samile Andrea de Souza; PAVÃO, C. M. G. Disponibilização de dados de pesquisa em repositórios universitários da América Latina. *In*: CONFERÊNCIA INTERNACIONAL BIREDIAL-ISTEC, 12., 2023, Montevideo. **Anais** [...]. La Plata: Universidad Nacional de La Plata; SEDICI, 2024. Disponível em: http://sedici.unlp.edu.ar/handle/10915/167232. Acesso em: 2 set. 2026.

UNESCO. Recomendação da Unesco sobre Ciência Aberta. Paris: Unesco, 2021. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000379949_por. Acesso em: 15 abr. 2026.

VIEIRA, Bruna Melissa. **Repositórios institucionais de dados de pesquisa como estratégia do movimento de acesso aberto à informação científica**. 2022. Dissertação (Mestrado em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2022. Disponível em: https://lume.ufrgs.br/handle/10183/253084. Acesso em: 12 mar. 2026.

WILKINSON, Mark D.; DUMONTIER, Michel; AALBERSBERG, Ijsbrand; APPLETON, Gabrielle *et al.* The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, [*s. l.*], v. 3, 160018, 2016. Disponível em: https://www.nature.com/articles/sdata201618. Acesso em: 12 set. 2026.