



10° EBBC

Encontro Brasileiro de
Bibliometria e Cientometria



MÉTRICAS DE DEPÓSITO NO REPOSITÓRIO DE DADOS DE PESQUISA DEPOSITA DADOS

Letícia Guarany Bonetti

<https://orcid.org/0000-0002-3012-8465>
leticiaBonetti@ibict.br
Instituto Brasileiro de Informação em
Ciência e Tecnologia (Ibict)

Tatyane Guedes Martins da Silva

<https://orcid.org/0000-0002-1743-0467>
tatyaneSilva@ibict.br
Instituto Brasileiro de Informação em
Ciência e Tecnologia (Ibict)

Caterina Groposo Pavão

<https://orcid.org/0000-0003-3712-7200>
caterina@ufrgs.br
Universidade Federal do Rio Grande
do Sul (UFRGS)

Samile Andrea de Souza Vanz

<https://orcid.org/0000-0003-0549-4567>
samile.vanz@ufrgs.br
Universidade Federal do Rio Grande
do Sul (UFRGS)

Rene Faustino Gabriel Junior

<https://orcid.org/0000-0003-1021-3360>
rene.gabriel@ufrgs.br
Universidade Federal do Rio Grande
do Sul (UFRGS)

Marcel Garcia de Souza

<https://orcid.org/0000-0003-2255-199X>
marcelsouza@ibict.br
Instituto Brasileiro de Informação em
Ciência e Tecnologia (Ibict)

Resumo:

Este estudo tem como objetivo analisar as métricas de depósito do Deposita Dados, no período de novembro de 2023 a dezembro de 2025, a fim de identificar suas principais características e padrões. A pesquisa descritiva e quantitativa analisou: número de depósitos por ano, afiliação institucional dos depositantes, áreas temáticas, licenças e formatos de arquivos. Resultados indicam crescimento progressivo no número de depósitos, predominância de formatos abertos e da atribuição da licença CC BY. Há uma concentração de depósitos de pesquisadores das instituições UFAM e UDESC, e a área temática predominante é Computação e Ciência da Informação. Conclui-se que o repositório contribui para o fortalecimento da gestão de dados de pesquisa no Brasil, apoiando a soberania científica nacional.

Palavras-Chave: Dados de pesquisa - Gestão. Repositório de dados. Ciência Aberta.

EIXO TEMÁTICO:

Bases e Fontes de Dados

MODALIDADE:

Resumo expandido

DOI: 10.22477/x.ebbc.1003

COMO CITAR ESTE ARTIGO:

BONETTI, Letícia Guarany; SILVA, Tatyane Guedes Martins da; PAVÃO, Caterina Groposo; VANZ, Samile Andrea de Souza; GABRIEL JUNIOR, Rene Faustino; SOUZA, Marcel Garcia de. Métricas de depósito no repositório de dados de pesquisa Deposita Dados. In: EBBC - Encontro Brasileiro de Bibliometria e Cientometria. 10. Anais [...]. Curitiba, 2026. DOI: 10.22477/x.ebbc.1003. Disponível em: <https://doi.org/10.22477/x.ebbc.1003>.



1 INTRODUÇÃO

Os dados de pesquisa ganharam um papel de destaque na comunicação científica contemporânea, especialmente diante das exigências de compartilhamento das agências de fomento, universidades e editoras científicas. O compartilhamento desses dados otimiza o retorno dos investimentos para as instituições e comunidades de pesquisadores e acelera a descoberta científica (Antonio *et al.*, 2020; Mosha; Ngulube, 2023; Bishoff; Johnston, 2015).

Somado a isso, Sayão e Sales (2022) apontam que a gestão adequada dos dados de pesquisa resulta no cumprimento das metas de transparência relativas aos métodos e aos fluxos de trabalho, bem como na obtenção de níveis satisfatórios de reprodutibilidade. Logo, percebe-se um ganho não só para a sociedade e para os pesquisadores, mas também para as próprias instituições que financiam a pesquisa.

No Brasil, a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) já “[...] considera necessário que os dados resultantes de projetos financiados pela Fundação sejam gerenciados e compartilhados de forma a garantir o maior benefício possível para o avanço científico e tecnológico” (FAPESP, 2020). No cenário internacional, agências como *National Science Foundation* (NSF), *National Institutes of Health* (NIH) e *Wellcome Trust* também vêm fortalecendo políticas de gestão de dados de pesquisa (Silva; Prado; Araújo, 2024; Bishoff; Johnston, 2015). Nesse contexto, os repositórios de dados assumem um papel estratégico no apoio ao cumprimento das exigências de agências de fomento e instituições, além de promoverem práticas de Ciência Aberta (Lin *et al.*, 2024; Mosha; Ngulube, 2023).

Para auxiliar os pesquisadores frente a essa demanda, o Instituto Brasileiro de In-

formação em Ciência e Tecnologia (Ibict) desenvolveu o Deposita Dados, um repositório de dados de pesquisa em acesso aberto. Sua criação fortalece a soberania científica nacional ao oferecer uma infraestrutura confiável para o armazenamento, a preservação e a disseminação de dados de pesquisa em território nacional. O Deposita Dados foi lançado em novembro de 2023 durante o XXII Seminário Nacional de Bibliotecas Universitárias (SNBU) e, desde então, apresenta um crescimento significativo no número de depósitos.

Tendo em vista que o Deposita Dados foi lançado há pouco mais de dois anos, evidencia-se a importância de realizar uma análise das métricas de depósito, a fim de identificar as principais características e padrões. Embora a literatura avance nas discussões sobre a importância dos repositórios para o compartilhamento de dados de pesquisa (Antonio *et al.*, 2020; Silva; Prado; Araújo, 2024), ainda é limitada a compreensão sobre o comportamento de pesquisadores brasileiros em relação aos depósitos (Silveira; Neubert; Dias, 2025). Diante disso, este estudo busca responder à seguinte questão de pesquisa: quais são as características e os padrões dos depósitos realizados no Deposita Dados?

2 PROCEDIMENTOS METODOLÓGICOS

A pesquisa caracteriza-se como descritiva, de abordagem quantitativa, e tem como objetivo analisar as métricas de depósito do Deposita Dados no período de novembro de 2023 a dezembro de 2025, a fim de identificar as principais características e os padrões dos depósitos realizados. Os dados foram extraídos a partir das opções de facetas configuradas no Dataverse, *software* adotado pelo repositório.

O estudo foi realizado em duas etapas principais: 1) levantamento dos dados a partir do sistema do Dataverse, e 2) validação individual e sistematização dos da-

dos em uma planilha para fins de limpeza, tratamento e análise estatística. Foram selecionadas cinco variáveis para análise: 1) depósitos por ano; 2) afiliação institucional dos depositantes; 3) áreas temáticas; 4) licenças e 5) formatos dos arquivos. A planilha contendo os dados tratados e consolidados encontra-se disponível em acesso aberto no repositório Aleia¹.

3 RESULTADOS

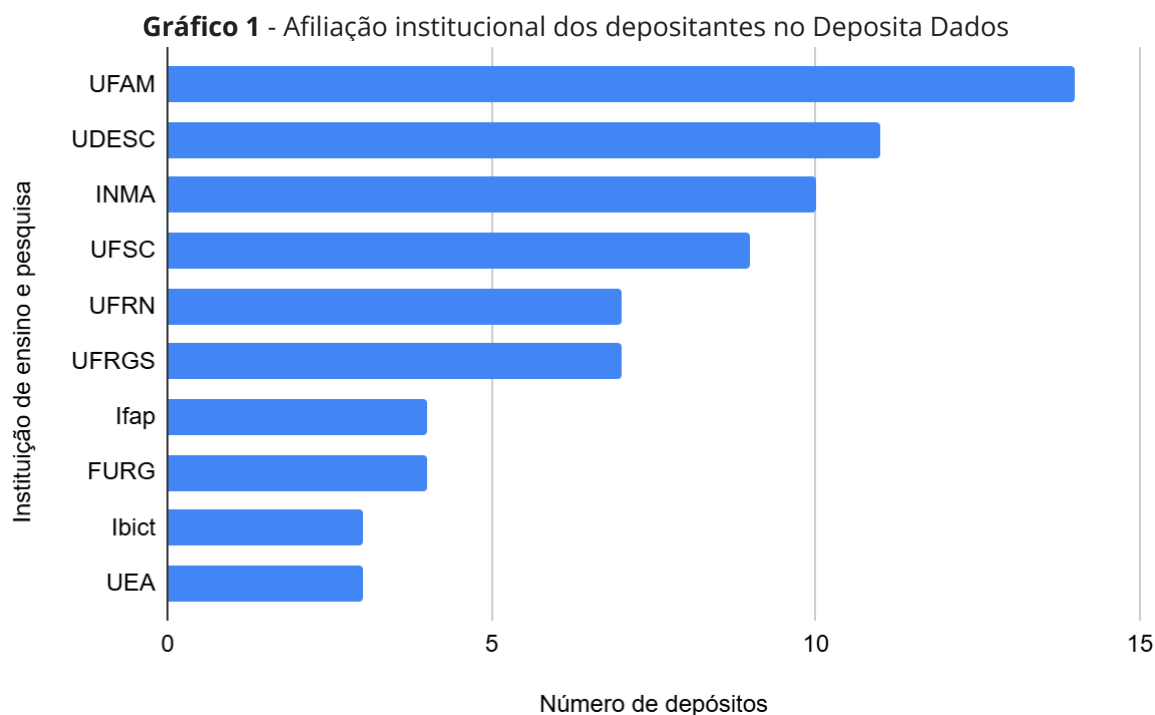
Foram analisados 85 conjuntos de dados de pesquisa depositados entre novembro de 2023 e dezembro de 2025. Em 2023, ano de lançamento, o Deposita Dados recebeu sete depósitos. Já em 2024, primeiro ano completo de funcionamento do serviço, o repositório recebeu 36 conjuntos de dados, sendo dezembro o período com o maior número de depósitos: 25. Essa concentração pode estar associada à iniciativa do Núcleo de Dados de Pesquisa (NDP), que apoiou pesquisadores de instituições brasileiras na realização de depósitos como parte de uma fase de experimentação (Oliveira *et al.*, 2025). O NDP é uma iniciativa colaborativa promovida pela Rede Brasileira de Repositórios Digitais (RBRD), pelo Ibict e pela Rede Nacional de Ensino e Pesquisa (RNP).

Em 2025, o repositório registrou 42 depósitos, superando o ano anterior. Diferentemente de 2024, as submissões ocorreram de forma mais distribuída ao longo dos meses, com pelo menos um depósito mensal. O mês de maior destaque foi abril, com 14 depósitos. A regularidade observada em 2025 pode indicar uma transição do uso experimental para uma incorporação mais estável do repositório às rotinas de pesquisa da comunidade científica no Brasil.

No Gráfico 1, é possível visualizar a afiliação dos pesquisadores depositantes. O gráfico exibe as instituições que depositaram pelo menos três conjuntos de da-

¹ Disponível em: <https://doi.org/10.48472/aleia/MTRLUG>

dos. Ressalta-se que um mesmo dado pode ter autores de diferentes instituições quando há coautoria.

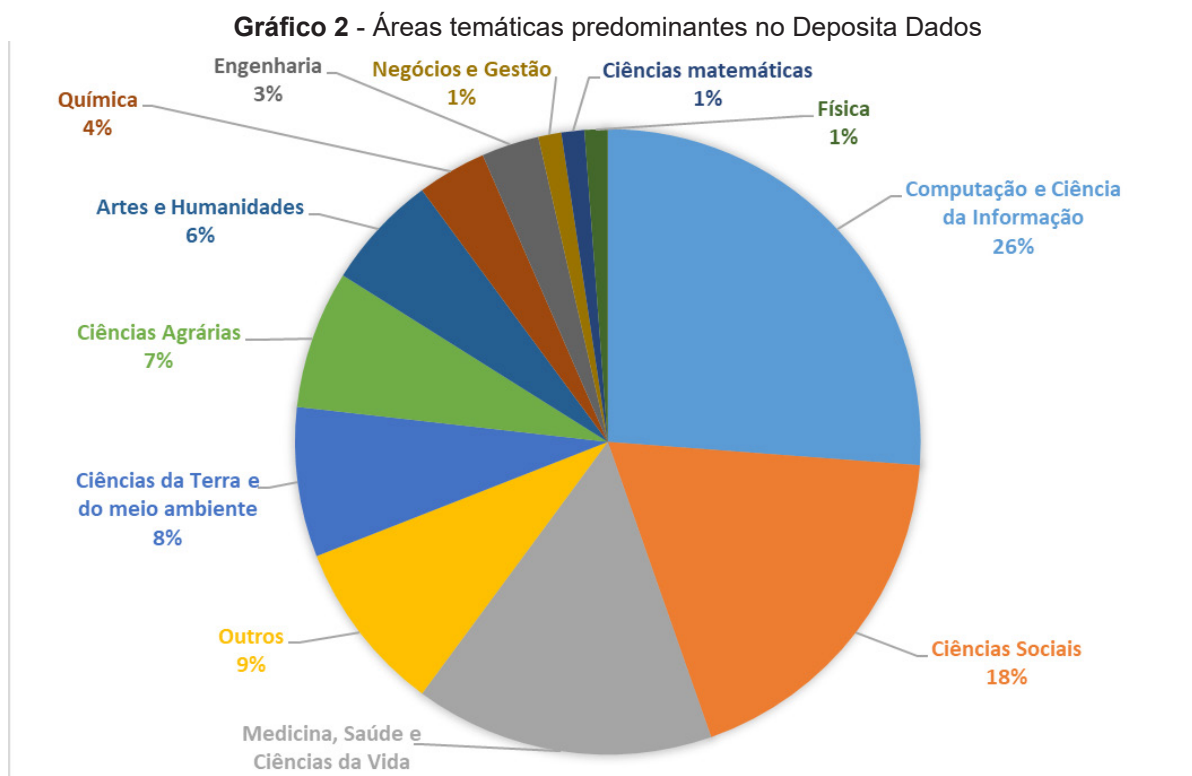


Fonte: elaborado pelos próprios autores.

O Deposita Dados é destinado a pesquisadores vinculados a instituições que ainda não possuem um repositório de dados de pesquisa. A métrica referente ao número de depósitos por instituição no Deposita Dados pode fornecer subsídios para que bibliotecas ou outros setores responsáveis justifiquem a criação do repositório institucional com base em uma demanda já existente.

A instituição com o maior número de depósitos (14) é a Universidade Federal do Amazonas (UFAM), seguida da Universidade do Estado de Santa Catarina (UDESC), com 11 depósitos. O Instituto Nacional da Mata Atlântica (INMA) possui 10 conjuntos de dados e a Universidade Federal de Santa Catarina (UFSC), oito. A UFSC e a UFAM têm um repositório de publicações (teses, artigos, anais, dentre outros), mas não implementaram ainda um repositório dedicado aos dados de pesquisa. Já a UDESC lançou, em 2025, um repositório institucional de dados de pesquisa,

o que a torna apta para a repatriação dos conjuntos de dados armazenados no Deposita Dados.



Fonte: elaborado pelos próprios autores.

No momento do depósito, o pesquisador pode selecionar uma ou mais áreas temáticas para atribuir ao conjunto de dados. No gráfico 2, é possível identificar alguns assuntos predominantes: Computação e Ciência da Informação (26%), Ciências Sociais (18%) e Medicina, Saúde e Ciências da Vida (16%). Essa predominância pode estar associada a diversos fatores como: 1) políticas mandatórias de compartilhamento de dados para a publicação em revistas ou financiamento; e 2) práticas mais consolidadas de Ciência Aberta em determinadas áreas do conhecimento.

Em contrapartida, dois assuntos ainda não foram atribuídos a nenhum conjunto de dados: 1) Astronomia e Astrofísica e 2) Direito. Um ponto a ser considerado é que, em algumas áreas, pode existir uma preferência pelo depósito em repositório.

rios disciplinares ao invés de multidisciplinares como o Deposita Dados.

Já com relação à licença, o pesquisador pode selecionar uma das opções *Creative Commons* (CC) no repositório: 1) CC BY; 2) CC BY-NC-SA; 3) CC BY-NC ou 4) CC BY-SA. Dos 85 conjuntos de dados, 84 têm a licença CC BY atribuída, que permite que qualquer pessoa copie, distribua, exiba, adapte e crie obras derivadas do material original, mesmo para fins comerciais, desde que atribua o crédito ao autor. Essa predominância do uso da CC BY pode se dar por dois motivos principais: 1) trata-se da licença padrão atribuída pelo sistema quando o usuário não realiza a seleção manual; 2) a comunidade que utiliza os repositórios tende a ser mais adepta ao movimento do Acesso Aberto e, portanto, busca compartilhar sua produção e seus dados da forma mais aberta possível sem comprometer seu direito de ser citado como autor.

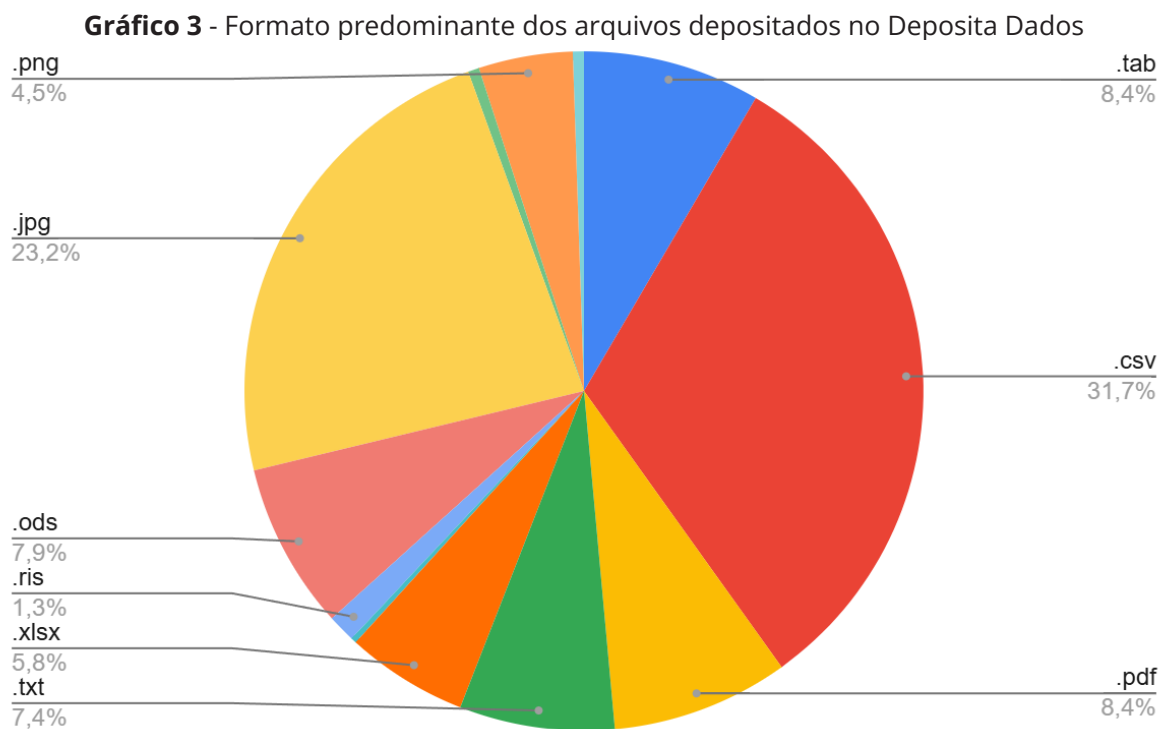
Apenas um pesquisador optou por utilizar a licença CC BY-SA, que permite que terceiros distribuam, remixem, adaptem e criem obras derivadas do material, mesmo para fins comerciais, desde que atribuam o crédito devido ao criador original e licenciem as novas criações sob os mesmos termos.

No Guia de Usuário², é apontado que o Deposita Dados suporta uma ampla variedade de tipos e formatos de arquivos. No gráfico 3, é possível observar essa diversidade, um reflexo da heterogeneidade dos dados de pesquisa em um repositório multidisciplinar. Ressalta-se que um conjunto de dados pode ter mais de um arquivo e eles podem estar em formatos distintos.

O formato mais utilizado é o *Comma-Separated Values* (csv), um arquivo de texto que armazena dados em formato de tabela com vírgulas separando cada célula na linha. O csv é um formato não proprietário, alinhado às práticas da Ciência

² Disponível em: <https://depositadados.ibict.br/guide.xhtml>

Aberta e à proposta de Dados Abertos 5 estrelas do Tim Berners-Lee³. Além do csv, outros formatos abertos também são encontrados no Deposita Dados: tab (8,4%), ods (7,9%) e txt (7,4%).



Fonte: elaborado pelos próprios autores.

Essa predominância pode estar ligada a dois fatores: 1) o Guia de Usuário apresenta uma tabela com os formatos preferidos para depósito: os não proprietários; 2) a equipe de curadoria orienta os depositantes para a submissão de formatos abertos. Assim, é possível reduzir o número de arquivos proprietários no repositório, que podem dificultar o acesso e a preservação a longo prazo e, portanto, a reutilização em novos contextos. Outro formato muito presente são as imagens em jpg (23,2%), utilizado principalmente nas áreas de Ciências Agrárias e Ciências da Terra e do Meio Ambiente. Um único conjunto de dados chegou a conter mais de 60 imagens nesse formato.

³ Disponível em: <https://5stardata.info/en/>

4 CONSIDERAÇÕES FINAIS

A análise das métricas de depósito do Deposita Dados no período de novembro de 2023 a dezembro de 2025 permitiu identificar padrões relacionados ao crescimento, ao perfil institucional dos depositantes e às características dos conjuntos de dados publicados. Observou-se um aumento progressivo no número de depósitos, com destaque para 2025, com 42 conjuntos de dados. Como visto, as análises apontam para uma incorporação mais estável do repositório à rotina da comunidade científica.

A concentração de depósitos de pesquisadores das instituições UFAM, UDESC e INMA evidencia o papel estratégico do Deposita Dados como solução transitória até que a instituição implemente seu próprio repositório de dados, como foi o caso da UDESC. Nesse sentido, o Deposita Dados fortalece a soberania científica nacional ao mesmo tempo em que fornece subsídios para fundamentar a implementação institucional de repositórios de dados de pesquisa em apoio à comunidade científica.

Quanto às áreas temáticas, verifica-se a predominância de Computação e Ciência da Informação, Ciências Sociais e áreas da Saúde, o que pode refletir políticas editoriais e exigências para financiamento. Percebe-se também, pela distribuição das áreas, que existe um perfil diverso de pesquisadores que utilizam o Deposita Dados.

A ampla adoção da licença CC BY e de formatos não proprietários como csv, tab e ods apontam para um alinhamento dos pesquisadores depositantes com os princípios fundamentais da Ciência Aberta. Entretanto, parte dessas escolhas também pode estar associada à configuração padrão do sistema e à atuação da equipe de

curadoria, o que ressalta a importância da construção de guias e materiais de consulta. Tais instrumentos são essenciais para assegurar que as decisões técnicas sejam compreendidas como escolhas conscientes de governança informacional, e não apenas como decorrências operacionais do ambiente tecnológico. Somado a isso, a variedade de formatos dos arquivos destaca a heterogeneidade dos dados de pesquisa. Em estudos futuros, pretende-se ampliar as variáveis analisadas para se obter uma visão mais ampla dos padrões de depósito no Deposita Dados.

REFERÊNCIAS

ANTONIO, M. G.; SCHICK-MAKAROFF, K.; DOIRON, J. M.; SHIELDS, L.; WHITE, L.; MOLZAHN, A. Qualitative data management and analysis within a data repository. **Western Journal of Nursing Research**, v. 42, n. 8, p. 640-648, 2020. DOI: <https://doi.org/10.1177/0193945919881706>. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31665999/>. Acesso em: 17 jan. 2026.

BISHOFF, C.; JOHNSTON, L. Approaches to data sharing: an analysis of NSF data management plans from a large research university. **Journal of Librarianship and Scholarly Communication**, [S. l.], v. 3, n. 2, eP1231, 2015. DOI: <https://doi.org/10.7710/2162-3309.1231>. Disponível em: <https://www.iastatedigitalpress.com/jlsc/article/id/12752/>. Acesso em: 20 fev. 2026.

FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO (FAPESP). Gestão de dados. 2020. Disponível em: <https://fapesp.br/gestaodedados>. Acesso em: 12 fev. 2026.

LIN, D.; MCAULIFFE, M.; PRUITT, K. D.; GURURAJ, A.; MELCHIOR, C.; SCHMITT, C.; WRIGHT, S. N. Biomedical Data Repository Concepts and Management Principles. **Scientific data**, v. 11, n. 1, 13 jun. 2024. DOI: <https://doi.org/10.1038/s41597-024-03449-z>. Disponível em: <https://www.nature.com/articles/s41597-024-03449-z#citeas>. Acesso em: 20 fev. 2026.

MOSHA, N. F.; NGULUBE, P. The utilisation of open research data repositories for storing and sharing research data in higher learning institutions in Tanzania. **Library Management**, v. 44, n. 8-9, p. 566-580, 2023. DOI: <https://doi.org/10.1108/LM-05-2023-0042>. Disponível em: <https://www.emerald.com/lm/article/44/8-9/566/261270/The-utilisation-of-open-research-data-repositories>. Acesso em: 27 fev. 2026.

OLIVEIRA, A.; ROSA, I. F. da; ALMEIDA, A. C. L. de; GALVES, J. M.; PAVÃO, C. M. G.; BARBALHO, C. R. S.; ALVES, A. S. Núcleo de Dados de Pesquisa (NDP): desafios e oportunidades para a implantação de repositórios no Brasil. **Revista Ciência Aberta Lusófona**, [S. l.], v. 1, 2025. DOI: <https://doi.org/10.82524/recal.2025.20>. Disponível em: <https://recal.rcaap.pt/recal/article/view/20>. Acesso em: 10 fev. 2026.



SAYÃO, L. F.; SALES, L. F. Plataformas de gestão de dados de pesquisa: expandindo o conceito de repositórios de dados. **Palavra clave**, v. 12, n. 1, e171, 2022. DOI: <https://doi.org/10.24215/18539912e171>. Disponível em: <https://www.palavraclave.fahce.unlp.edu.ar/article/view/pce171>. Acesso em: 14 fev. 2026.

SILVA, L. C.; PRADO, C. B.; ARAÚJO, E. A. Fatores que influenciam o processo de compartilhamento e reuso de dados de pesquisa: revisão sistemática de literatura a partir da base de dados BRAPCI. **AtoZ: Novas Práticas em Informação e Conhecimento**, Curitiba, v. 13, p. 1-15, 2024. DOI: <https://doi.org/10.5380/atoz.v13i0.91352>. Disponível em: <https://revistas.ufpr.br/atoz/article/view/91352>. Acesso em: 12 fev. 2026.

SILVEIRA, C. Z.; NEUBERT, P. da S.; DIAS, T. M. R. Explorando a cobertura temática dos repositórios associados aos dados de pesquisa descritos em data papers. In: XIV Conferência Internacional BIREDIAL-ISTEC, 2025, Brasília. **Anais** [...] Brasília: BIREDIAL-ISTEC, 2025. Disponível em: <https://submissions.istec.org/index.php/biredial-istec/article/view/390>. Acesso em: 24 fev. 2026.