

Pecha kucha

Dados de Pesquisa Abertos: Panorama da Enfermagem a partir da Base de Dados OpenAlex

Open Research Data: an Overview of Nursing Based on the OpenAlex Database

Datos de Investigación Abiertos: Panorama de la enfermería a partir de la Base de Datos OpenAlex

Josimara Dias Brumatti

Doutoranda em Ciência da Informação

Instituto Brasileiro de Informação em Ciência e Tecnologia

ORCID: <https://orcid.org/0000-0001-7199-2404>

E-mail: josimaradias@id.uff.br

Jessica Silva Brunoni

Doutoranda em Enfermagem e Biociências

Universidade Federal do Estado do Rio de Janeiro

ORCID: <https://orcid.org/0000-0002-6581-2179>

E-mail: jessica.brunoni@edu.unirio.br

Marcia Rodrigues dos Santos

Doutoranda em Enfermagem e Biociências

Universidade Federal do Estado do Rio de Janeiro

ORCID: <https://orcid.org/0000-0002-1562-9026>

E-mail: marcia.santos@edu.unirio.br

Maria Simone de Menezes Alencar

Professora adjunta da UNIRIO, doutora na área de Gestão e Inovação Tecnológica

Universidade Federal do Estado do Rio de Janeiro

ORCID: <https://orcid.org/0000-0002-2992-2215>

E-mail: simone.alencar@unirio.br

Fabiana de Melo Amaral Pinto

Doutora em Ciência da Informação

Universidade Federal Fluminense

ORCID: <https://orcid.org/0000-0002-8118-3165>

E-mail: fabianamelo@id.uff.br

Resumo

Este estudo analisa as práticas de ciência aberta na publicação de dados de pesquisa em enfermagem, com ênfase no uso de repositórios e no acesso aberto. A pesquisa, de caráter descritivo e exploratório, utilizou dados da base OpenAlex, com recorte dos últimos cinco anos, identificando 5.017 conjuntos de dados relacionados à área de enfermagem. Os resultados revelaram que 95% dos dados estão depositados em repositórios, com predominância do acesso aberto (73%); enquanto os 5% restantes estão depositados em periódicos, sendo o acesso restrito majoritário (84%). Figshare, Harvard Dataverse e Zenodo foram os repositórios mais utilizados. A maioria das publicações era de autoria de pesquisadores dos Estados Unidos, China e Reino Unido; o Brasil contribuiu com apenas 2% dos depósitos, apesar do expressivo financiamento público na área. A análise aponta para avanços no compartilhamento de dados, mas também sinaliza desafios como: a baixa adesão da América Latina, a concentração em repositórios generalistas e a necessidade de maior adesão aos princípios FAIR. Conclui-se que é essencial fortalecer políticas institucionais de gestão de dados e fomentar uma cultura de transparência e colaboração na enfermagem, ampliando a representatividade e a adoção efetiva da ciência aberta.

Palavras-chave: Gestão de Dados de Pesquisa; Repositório de Dados de Pesquisa; Enfermagem

Abstract

This study analyzes open science practices in the publication of nursing research data, with a focus on the use of repositories and open access. The research was descriptive and exploratory, drawing on data from the OpenAlex database covering the last five years, and identified 5,017 nursing-related datasets. The results revealed that 95% of the data were deposited in repositories, with a predominance of open access (73%); while the 5% were deposited in journals, with restricted access being the majority (84%). Figshare, Harvard Dataverse, and Zenodo were the most widely used repositories. Most publications were authored by researchers from the United States, China, and the United Kingdom; Brazil contributed only 2% of the deposits, despite significant public funding in the field. The analysis highlights progress in data sharing but also points to challenges such as low adoption in Latin America, the concentration in generalist repositories, and the need for greater adherence to the FAIR principles. The study concludes that it is essential to strengthen institutional data management policies and foster a culture of transparency and collaboration in nursing, thereby expanding the representation and effective adoption of open science.

Keywords: Research Data Management; Research Data Repository; Nursing

Resumen

Este estudio analiza las prácticas de ciencia abierta en la publicación de datos de investigación en enfermería, con énfasis en el uso de repositorios y acceso abierto. La investigación fue descriptiva y exploratoria, basada en datos de la base OpenAlex, correspondientes a los últimos cinco años, e identificó 5017 conjuntos de datos relacionados con enfermería. Los resultados revelaron que el 95% de los datos se depositaron en repositorios, con predominio del acceso abierto (73%), mientras que el 5% restante se depositaron en revistas, en el que el acceso restringido fue mayoritario (84%). Figshare, Harvard Dataverse

y Zenodo fueron los repositorios más utilizados. La mayoría de las publicaciones eran de investigadores de Estados Unidos, China y el Reino Unido; Brasil contribuyó solo con el 2% de los depósitos, a pesar de haber un importante financiamiento público en el área. El análisis destaca avances en el intercambio de datos, pero también señala desafíos como: la baja adopción en Latinoamérica, la concentración en repositorios generalistas y la necesidad de una mayor adhesión a los principios FAIR. Se concluye que es fundamental fortalecer las políticas institucionales de gestión de datos y fomentar una cultura de transparencia y colaboración en enfermería, ampliando la representación y la adopción efectiva de la ciencia abierta.

Palabras clave: Gestión de datos de investigación; Repositorio de datos de investigación; Enfermería

Introdução

Os dados de pesquisa são essenciais para a validação científica e a reprodutibilidade dos estudos. Com o avanço da ciência aberta, seu compartilhamento é cada vez mais incentivado, promovendo transparência e colaboração e permitindo uma utilização mais eficiente e econômica dos dados, especialmente em pesquisas financiadas com recursos públicos (Prosser et al., 2023).

Na América Latina e especialmente no Brasil, iniciativas como o SciELO Data¹ e o Arca Dados² destacam-se como repositórios de dados de pesquisa abertos. No entanto, assim como nos artigos de periódicos, a apropriação comercial da ciência aberta para os dados de pesquisa pode gerar futuras tensões no que se refere à custódia dos dados. A Declaração de Barcelona (2024) reforça a importância da segurança e da preservação dos dados científicos. Na enfermagem, o compartilhamento de dados ainda é limitado, afetando a transparência e a colaboração (Pinto, 2022). Apesar dos avanços, a falta de sistematização e a adesão a padrões internacionais seguem como desafios (Anglada, L. et al, 2018; Borgman et al, 2015; Tenopir et al., 2020).

O OpenAlex³ se caracteriza como uma base de dados bibliográfica totalmente aberta, cuja coleção é obtida a partir do harvesting em outros repositórios. Ela é considerada uma fonte com alta cobertura de publicações em acesso aberto e desempenha um papel fundamental na facilitação da análise de dados científicos abertos (Priem, Piwowar & Orr, 2022).

No entanto, a adoção de repositórios ainda é limitada, e ferramentas como a OpenAlex podem ajudar a identificar padrões e lacunas. O objetivo deste estudo é analisar as práticas de ciência aberta na publicação de dados de pesquisa em enfermagem, incentivando maior transparência, reprodutibilidade e colaboração.

¹ O SciELO Data é um repositório multidisciplinar para depósito, preservação e disseminação de dados de pesquisa de artigos submetidos e aprovados para publicação, já publicados em periódicos da Rede SciELO ou depositados no SciELO Preprints. Fonte: <https://data.scielo.org/>.

² O Arca Dados é o repositório oficial de dados para pesquisa da Fiocruz. Tem a missão de ser a plataforma para arquivar, publicar, disseminar, preservar e compartilhar os dados digitais para pesquisa produzidos pela comunidade Fiocruz e/ou em parceria com outros institutos ou órgãos de pesquisa, seguindo os princípios FAIR (Findable, Accessible, Interoperable, and Reusable). Fonte: <https://arcadados.fiocruz.br/>.

³ <https://openalex.org/>.

Metodologia

Trata-se de um estudo descritivo e exploratório, com abordagem qualitativa e quantitativa, tendo como fontes documentais os dados de pesquisas compartilhados e indexados na base de dados OpenAlex. Foram extraídos dados numéricos para subsidiar a análise quantitativa.

Como refinamento da pesquisa, utilizamos os seguintes metadados disponíveis no sistema de busca da plataforma:

primary_topic.field.display_name (área do conhecimento): Nursing;

Year (período analisado): 2020-2024, últimos 5 anos; e

Type (tipo de documento): dataset.

Foram utilizadas informações sobre os conjuntos de dados depositados pela área de enfermagem e recuperados por meio da OpenAlex, com o objetivo de analisar a frequência de depósitos, considerando: o local de depósito (repositório ou periódico), as cinco principais subáreas da enfermagem e a nacionalidade do autor (país de origem). A busca foi realizada em março de 2025, tendo sido localizados 5.017 conjuntos de dados, dos quais cerca de 5% não puderam ser analisados pois estavam com metadados em branco, o que comprometeria a análise.

Resultados e Discussão

Por meio da busca na OpenAlex, identificou-se que 95% dos 5.017 conjuntos de dados de pesquisa em enfermagem estão depositados em repositórios, enquanto apenas 5% estão em periódicos (Gráfico 1). Esses achados corroboram Pinto (2022), que destacou a preferência dos pesquisadores da área pelo depósito de dados em repositórios. No entanto, essa constatação, embora positiva em termos quantitativos, revela contradições estruturais no avanço da ciência aberta.

Gráfico 1- Tipos de Depósito de Datasets em Enfermagem

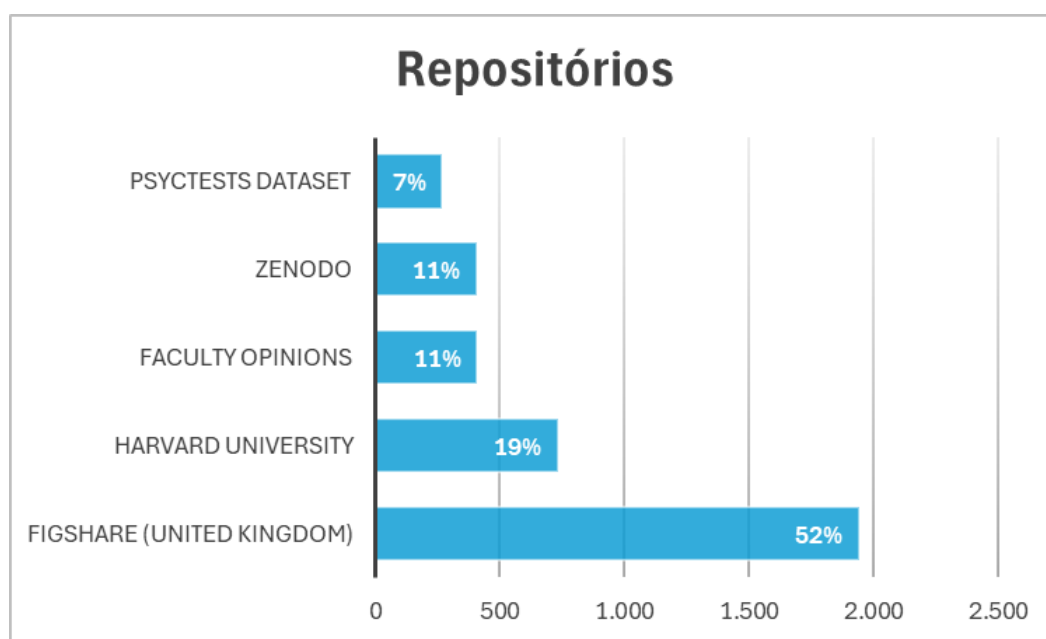


Fonte: elaboração própria, 2025.

Entre os datasets em repositórios, 73% estavam em acesso aberto e 27% em acesso restrito. Nos periódicos a situação foi oposta: a maioria dos datasets estava em acesso fechado (84%). Essa discrepância sugere uma tensão entre os princípios da ciência aberta e a lógica editorial dominante, que ainda privilegia modelos de restrição. A permanência de dados sob paywalls compromete a transparência e a reprodutibilidade, levantando questões éticas e políticas sobre o papel das editoras comerciais no ecossistema da ciência aberta.

A análise mostrou que os principais repositórios utilizados são: Figshare (52%), Harvard Dataverse (19%), Faculty Opinions (11%), Zenodo (11%) e, por fim, o PsycTests (7%) (Gráfico 2). O resultado evidencia a forte dependência de plataformas generalistas e externas, enquanto iniciativas institucionais e temáticas da área da saúde, especialmente na América Latina, permanecem pouco expressivas. Essa dependência acarreta riscos quanto à padronização disciplinar e à sustentabilidade das práticas de preservação de dados, além de reforçar a ausência de infraestruturas próprias para a enfermagem.

Gráfico 2- Tipos de repositório utilizados para depósito



Fonte: elaboração própria, 2025.

No recorte geográfico, os Estados Unidos lideram os depósitos (30%), seguidos por China (7%) e Reino Unido (7%). O Brasil aparece com apenas 39 conjuntos de dados (2%), apesar de ter elevado volume de financiamento público em pesquisas de enfermagem. Esse paradoxo revela uma fragilidade institucional: embora financiada por recursos públicos, a produção de dados ainda não oferece retorno à sociedade de forma ampla e acessível. A falta de políticas de gestão de dados de pesquisa que exijam, por exemplo, planos de gestão por agências de fomento pode explicar parte dessa lacuna (Abadal & Anglada, 2020).

Outro aspecto crítico refere-se à qualidade dos metadados: cerca de 5% dos conjuntos analisados estavam sem metadados preenchidos. Essa limitação compromete diretamente a aplicação dos princípios FAIR (Findable, Accessible, Interoperable, Reusable), que sustentam a efetividade da ciência aberta. Disponibilizar dados sem a documentação adequada significa perder grande parte de seu potencial científico, reduzindo a possibilidade de reuso, interoperabilidade e impacto social (Macedo, E., Henning, A., & Souza, S. 2021).

Além disso, a abertura de dados na enfermagem traz desafios específicos relacionados à ética e à proteção de informações sensíveis em saúde (Sayão, L. F., & Sales, L. F., 2020). A tensão entre garantir privacidade e promover transparência científica deve ser permanentemente debatida, sob pena de a abertura indiscriminada produzir riscos para sujeitos e comunidades envolvidos nas pesquisas.

Assim, embora os resultados revelem avanços importantes, eles também evidenciam que a prática da ciência aberta na enfermagem ainda opera sob contradições e desigualdades globais, apontando para a necessidade de políticas públicas e institucionais mais consistentes.

Considerações Finais

A análise dos dados extraídos da OpenAlex evidenciou avanços importantes na adesão ao depósito de dados em repositórios, especialmente em acesso aberto. Contudo, esses avanços quantitativos não garantem, por si só, a consolidação da ciência aberta na enfermagem.

A predominância de repositórios generalistas, a baixa representatividade da América Latina e a persistência do acesso fechado em periódicos revelam contradições que limitam a efetividade desse movimento. A situação brasileira é emblemática: mesmo com forte investimento público em pesquisa, a prática de abertura de dados permanece modesta, o que reforça a necessidade de políticas institucionais mais robustas de gestão e compartilhamento.

A pesquisa revelou que as práticas de ciência aberta na publicação de dados de pesquisa na área de enfermagem são bastante positivas, com 95% dos dados depositados em repositórios, predominantemente em acesso aberto. Nos periódicos, entretanto, prevalece o acesso fechado, indicando que os periódicos da área ainda não aderiram de forma significativa à abertura dos dados. Repositórios como Figshare, Harvard Dataverse e Zenodo se destacaram, e a produção de dados está concentrada em países como Estados Unidos, China e Reino Unido.

Do ponto de vista crítico, a desigualdade geopolítica no compartilhamento de dados expõe um risco de dependência e invisibilidade para a produção científica da enfermagem na América Latina. Caso não haja esforços coordenados de fortalecimento de repositórios regionais, formação de profissionais em gestão de dados e integração

às redes internacionais, a ciência aberta corre o risco de reproduzir assimetrias globais em vez de superá-las.

A adoção consistente dos princípios FAIR, associada a iniciativas como a Rede GO FAIR Brasil Saúde Enfermagem, pode contribuir para superar tais barreiras, desde que acompanhada de políticas de fomento que exijam planos de gestão de dados e mecanismos claros de monitoramento.

Conclui-se que os dados abertos em enfermagem representam não apenas uma oportunidade de ampliar a transparência e a colaboração científica, mas também um campo de disputa política, econômica e ética. O fortalecimento da ciência aberta dependerá não só do volume de dados depositados, mas, sobretudo, da qualidade, da acessibilidade e da equidade com que esses dados circulam, de modo a assegurar que o investimento público se converta em benefícios coletivos e em impacto social efetivo.

Conflito de Interesses

Os autores declaram não haver conflitos de interesses.

Disponibilização dos Dados de Investigação

Disponível em: <https://doi.org/10.48472/deposita/J9RWYL>

CRedit – Contribuições dos Autores

Josimara Dias | Concetualização, Escrita – redação original, Supervisão, Recolha de dados
Jessica Silva | Concetualização, Curadoria de dados, Escrita – revisão e edição
Marcia Santos | Concetualização, Escrita – revisão e edição, Supervisão
Simone Alencar | Escrita – revisão e edição, Supervisão
Fabiana Pinto | Escrita – revisão e edição, Supervisão

Referências

- Abadal, E., & Anglada, L. (2020). *Ciencia abierta: Cómo han evolucionado la denominación y el concepto*. *Anales de Documentación*, 23(1). <https://doi.org/10.6018/analesdoc.378171>
- Anglada, L., & Abadal, E. (2018). ¿Qué es la ciencia abierta? *Anuario ThinkEPI*, 12, 292–298.
- Borgman, C. L., Van De Sompel, H., Scharnhorst, A., Van Den Berg, H., & Treloar, A.. (2015). Who uses the digital data archive? An exploratory study of DANS. *Proceedings of the Association for Information Science and Technology*, 52(1), 1–4. <https://doi.org/10.1002/pa2.2015.145052010096>
- Macedo, E., Henning, A., & Souza, S. (2021). Rede GO FAIR Brasil Saúde Enfermagem: Onde estamos e aonde queremos chegar? In L. F. Sales, V. S. Veiga, P. Henning, & L. F. Sayão (Orgs.), *Princípios FAIR aplicados à gestão de dados de pesquisa* (pp. 243–252). IBICT. <https://doi.org/10.22477/9786589167242.cap17>

OpenAlex. (2024). API Entities: Works - Work Object. https://docs.openalex.org/api-entities/works/work-object#oa_status

Pinto, F. M. A. G. (2022). *Publicação de dados de pesquisa em Enfermagem: Novas práticas em Ciência Aberta* [Tese de doutorado, Universidade Federal do Rio de Janeiro]. Instituto Brasileiro de Informação em Ciência e Tecnologia. <http://app.uff.br/riuff/handle/1/26650>

Priem, J., Piwowar, H., & Orr, R. (2022). OpenAlex: Um índice totalmente aberto de trabalhos acadêmicos, autores, locais, instituições e conceitos. *arXiv*. <https://arxiv.org/abs/2205.01833>

Prosser, A. M., et al. (2023). When open data closes the door: A critical examination of the past, present and the potential future for open data guidelines in journals. *British Journal of Social Psychology*. <https://doi.org/10.1111/bjso.12576>

Sayão, L. F., & Sales, L. F. (2020). Afinal, o que é dado de pesquisa? BIBLOS - Revista do Instituto de Ciências Humanas e da Informação, 34(2). <https://doi.org/10.14295/biblos.v34i2.11875>

SemOpenAlex. (2024). About. <https://semopenalex.org/resource/semopenalex:UniversalSearch>

Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., ... & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PloS one*, 15(3), e0229003. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7065823>