

Pecha Kucha

Infraestruturas de Apoio para Abertura de Dados de Pesquisa em Repositórios Digitais

Supporting Infrastructures for Opening Research Data in Digital Repositories

Infraestructuras de apoyo para apertura de datos de investigación en repositorios digitales

Carolina Howard Felicíssimo*

Rede Nacional de Ensino e Pesquisa (RNP), Brasil

ORCID: <https://orcid.org/0000-0001-5065-2689>

Lattes: <http://lattes.cnpq.br/9285740086584040>

carolina.felicissimo@rnp.br

Leandro Neumann Ciuffo

Rede Nacional de Ensino e Pesquisa (RNP), Brasil

ORCID: <https://orcid.org/0000-0002-9747-2279>

Lattes: <http://lattes.cnpq.br/0951542162499636>

leandro.ciuffo@rnp.br

Resumo

Justificada pela recomendação da UNESCO sobre Ciência Aberta (2021), onde é destacada a responsabilidade das Redes Nacionais de Ensino e Pesquisa (do termo, em inglês, *National Research and Education Networks* – NRENs), a Rede Nacional de Ensino e Pesquisa brasileira (RNP) vem atuando em parceria com o Núcleo de Dados de Pesquisa (NDP), um grupo de trabalho da Rede Brasileira de Repositórios Digitais (RBRD), para promoção do acesso aberto a dados de pesquisa em repositórios digitais. Este artigo apresenta infraestruturas de apoio para repositórios de dados de pesquisa, considerando qualidade com economia de escala.

Palavras-chave:

Infraestruturas de apoio; Repositórios de dados de pesquisa; Acesso Aberto

Abstract

Justified by the UNESCO Recommendation on Open Science (2021), which highlights the responsibility of National Research and Education Networks (NRENs), the Brazilian National Research and Education Network (RNP) has been collaborating with the Research Data Nucleus (NDP), a working group of the Brazilian Network of Digital Repositories (RBRD), to promote open access to research data in digital repositories. This article presents supporting infrastructures for research data repositories, considering quality with economy of scale.

Keywords:

Supporting infrastructure; Research Data Repositories; Open Access

Resumen

Justificada por la Recomendación de la UNESCO sobre la Ciencia Abierta (2021), que destaca la responsabilidad de las Redes Nacionales de Investigación y Educación (NREN), la Red Nacional de Investigación y Educación de Brasil (RNP) ha colaborado con el Nucleous de Datos de Investigación (NDP), un grupo de trabajo de la Red Brasileña de Repositorios Digitales (RBRD), para promover el acceso abierto a los datos de investigación en repositorios digitales. Este artículo propone infraestructuras de apoyo para repositorios de datos de investigación, considerando calidad con economía de escala.

Palabras clave:

Infraestructuras de apoyo; Repositorios de datos de investigación; Acceso abierto.

Introdução

Alinhado à recomendação da UNESCO para Ciência Aberta (UNESCO, 2021), o Núcleo de Dados de Pesquisa (NDP) se estabelece no Brasil, como um grupo de trabalho da Rede Brasileira de Repositórios Digitais (RBRD - <https://rbrd.ibict.br/>), para atuar na promoção do acesso aberto a repositórios de dados de pesquisa (Felicissimo et al., 2025). No entanto, não está claro qual suporte pesquisadores e suas respectivas instituições precisam ter com infraestruturas de apoio, nacionais/regionais, além de como benefícios coletivos podem ser maximizados no NDP, reduzindo-se custos. Este artigo visa apresentar iniciativas atuais do NDP para abertura de dados

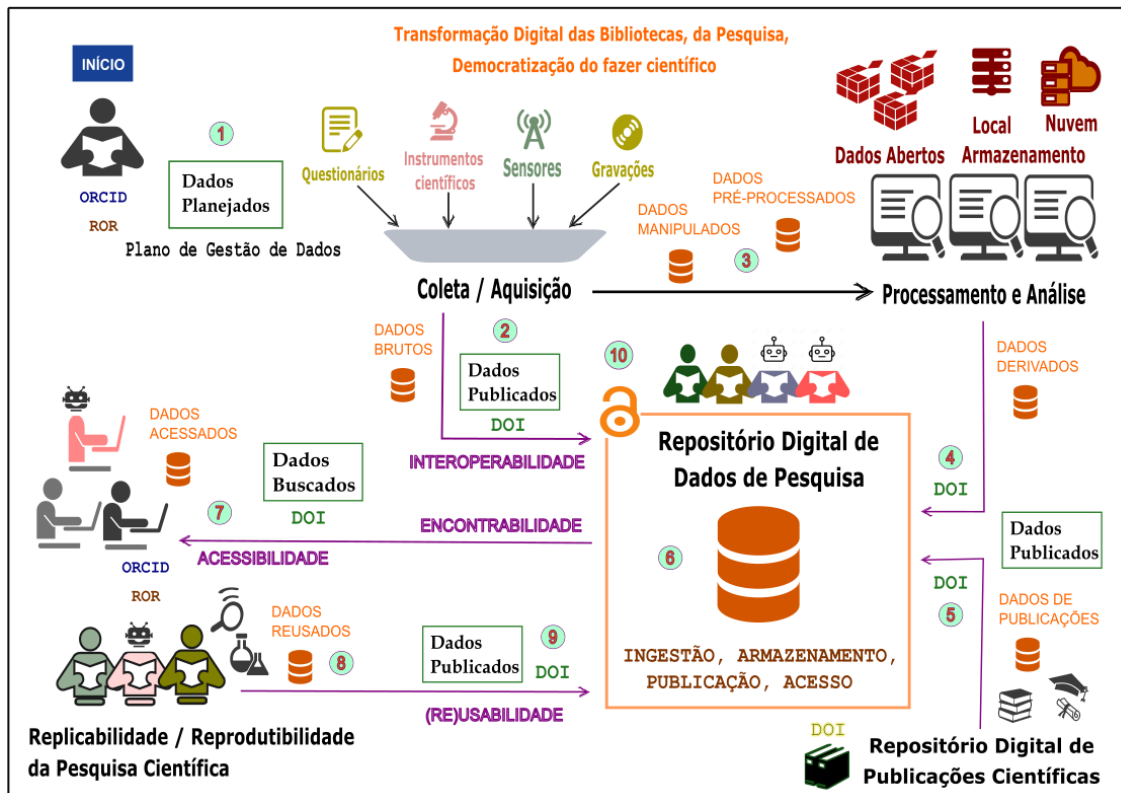
de pesquisa em repositórios digitais com alguns exemplos de projetos correntes no Brasil na área. Finalizamos o artigo com nossas conclusões e planos de trabalhos futuros.

Abertura de Dados de Pesquisa em Repositórios Digitais

Ao longo de uma pesquisa científica dados são manipulados por toda cadeia produtiva. Acentuados pelo volume, variedade e velocidade que as tecnologias trazem, cada vez mais frequente, faz-se necessário infraestruturas de apoio que suportem toda a complexidade para abertura de dados de pesquisa em repositórios digitais. A figura 1 ilustra 10 etapas do ciclo de vida de dados de pesquisa suportada por uma instância de repositório digital específica para dados de pesquisa.

Figura 1

Ciclo de Vida de Dados de Pesquisa em Repositório Digital



Nota. (Felicissimo, 2026).

A etapa 1 inicia-se com dados planejados, oriundos de um plano de gestão de dados do pesquisador. Na etapa 2, os dados brutos são coletados e/ou adquiridos por fontes de informações diversas e podem/devem ser publicados, mesmo que por um conjunto parcial, para registro inicial de onde a pesquisa se iniciou. Quando publicados, identificadores persistentes são

atribuídos de forma que os dados podem ser recuperados por mecanismos de buscas, os quais recorrem a metadados qualificados para (re)uso de forma a se replicar e/ou reproduzir a pesquisa, além de visibilidade da propriedade intelectual.

Na etapa 3, os dados são manipulados e pré-processados para permitirem processamento e análise em conjunto com outros dados armazenados em locais diversos, criando dados derivados. Na etapa 4, os dados derivados devem ser publicados de forma a permitirem rastreabilidade dos resultados obtidos. Na etapa 5, os dados apresentados em publicações científicas devem ser publicados em repositórios de dados de pesquisa, os quais são apropriados para melhor descrição por metadados específicos para dados. Ambos repositórios consideram conteúdos e descritores (metadados) como entidades intelectuais distintas.

Na etapa 6, com a implementação de requisitos mínimos de representação e utilização de dados, como os princípios FAIR (Wilkinson, 2016), dados podem ser manipulados por humanos e/ou máquinas. Na etapa 7, os dados são encontrados e acessados para que, na etapa 8, pares científicos (humanos e/ou máquinas) possam realizar novas pesquisas, seja por replicabilidade e/ou reprodutibilidade. Na etapa 9, os dados reutilizáveis são publicados para permitir novos acessos por humanos e/ou máquinas (etapa 10), democratizando, assim, o acesso aberto a dados oriundos de uma pesquisa científica.

Apesar de tratar-se de um poderoso recurso digital, pondera-se que os pesquisadores nas instituições precisam de apoio no processo de abertura de dados em repositórios digitais para sua correta representação e utilização por outros. Iniciativas internacionais como o GO FAIR (<https://www.go-fair.org/fair-principles/>) apostam em três pilares para implementação dos princípios FAIR: Cultura; Treinamento e Infraestrutura, contando com alinhamentos gerais internacionais do próprio GO FAIR, além de execução de ações apoiadas por meio de escritórios nacionais, como o GO FAIR Brasil. No Brasil, a RNP tem atuado junto ao NDP/RBRD, apoiando pesquisadores e instituições em uma melhor governança e gestão de dados de pesquisa em repositórios digitais. Na próxima seção, apresentaremos exemplos de iniciativas para abertura de dados em repositórios digitais, considerando algumas perspectivas, como tecnológicas, informacionais, disciplinares e organizacionais.

Iniciativas para Abertura de Dados de Pesquisa em Repositórios Digitais

Por meio de execução de projetos de pesquisa, desenvolvimento e inovação (PD&I), a RNP tem conseguido avançar em um melhor entendimento sobre quais serviços e infraestruturas de pesquisa possuem maior potencial de beneficiar pesquisadores e instituições na jornada de abertura de dados em repositórios digitais. Tal entendimento é constantemente validado junto à rede de colaboração do NDP/RBRD, além dos parceiros dos projetos de PD&I. Percebe-se maior celeridade e assertividade quando pesquisadores são apoiados por equipes multidisciplinares, por

exemplo, das áreas de Ciência da Informação (CI), para uma melhor descrição informacional, e de Tecnologia da Informação (TI), para suporte tecnológico no uso de repositórios digitais.

Em 2024, a RNP atuou em conjunto com as Universidades Federais de Goiás (UFG) e do Rio Grande do Sul (UFRGS) em um estudo para atribuição de identificadores persistentes na modalidade de DOIs (Digital Object Identifiers) para o Sistema de Informação para a Biodiversidade Brasileira (SiBBR), nó brasileiro do Global Biodiversity Information System (GBIF) (Rezende, 2024). Três frentes de atuação foram executadas: (i) prospecção do estado da arte para emissão de DOIs via próprio GBIF; (ii) possibilidade de adesão ao consórcio nacional do DataCite no Brasil; e, (iii) apresentação das possibilidades prospectadas com prós e contras para a escolha definitiva. O estudo possibilitou a adesão do SiBBR à solução de emissão de DOIs pelo próprio GBIF por garantia de prefixo exclusivo para acesso direto aos dados do SiBBR, além de não haver custos extras por essa solução já estar contemplada ao pacote de serviços do GBIF para nós nacionais.

Um outro exemplo, desde 2021, a RNP e as Universidades Federais do Rio Grande do Sul (UFRGS) e do Rio de Janeiro (UFRJ) têm atuado para divulgação e disseminação de pesquisas estatísticas censitárias educacionais em repositórios digitais (Caregnato, 2025). Experimentações tecnológicas vêm sendo conduzidas utilizando-se os softwares Dataverse (<https://dataverse.org/>) para abertura de dados de pesquisa e Dspace (<https://dspace.org/>) para publicações correlatas na divulgação de produtos institucionais, considerando modelos de referência internacionais para modelagem de pesquisas estatísticas em contexto organizacional. Identificou-se que o software Dataverse oferece suporte abrangente para descrição de metadados disciplinares, sendo robusto suficiente para a representação de dados oriundos de estudos estatísticos. A estruturação de comunidades destinadas para dados de pesquisa e de coleções para publicações correlatas consideram o uso dos repositórios Dataverse e Dspace, operando em conjunto. Desafios englobam gestão de mudanças para variáveis e descritores, mantendo-se qualidade, em escala crescente, ao longo de ciclos anuais das pesquisas estatísticas censitárias educacionais. Rosa et al. (2025) apresenta uma proposta de construção de terminologia, contextualizada por conteúdo informacional para o domínio educacional, baseada na modelagem e harmonização de termos e definições para pesquisas estatísticas anuais. O método proposto considera a expansão por um sistema conceitual representativo, baseado em uma árvore de domínio em substituição a modelos lineares, como dicionários de dados.

Dentre estes exemplos e outros, os quais surgem com utilização crescente de repositórios digitais de dados de pesquisa no Brasil, percebe-se especificidades de conhecimentos distintos de profissionais qualificados para manipulação de dados de pesquisa (do termo em inglês, data stewards). Profissionais de TI destacam-se na configuração de soluções tecnológicas customizadas com aspectos informacionais e disciplinares para repositórios de dados de pesquisa. Profissionais de CI são essenciais na criação de templates com pré-cadastro de registros padronizados, além de sinalização de uso de vocabulários controlados por ontologias de domínios. Assim, percebe-se

benefícios a pesquisadores e instituições quando são apoiados por equipes multidisciplinares na jornada de abertura de dados (FAIR) em repositórios digitais, permitindo que os dados sejam melhor encontráveis, interoperáveis, acessíveis e (re)usáveis.

Conclusões

Neste artigo, apresentamos como a RNP vem atuando no Núcleo de Dados de Pesquisa (NDP) da Rede Brasileira de Repositórios Digitais (RBRD) na oferta de infraestruturas de apoio para abertura de dados de pesquisa em repositórios digitais. Assemelha-se à iniciativa Re.Data (<https://redata.pt/>), uma rede para a gestão de dados de investigação de Portugal, em seu objetivo de ofertar infraestruturas e serviços para suporte à gestão de dados de investigação FAIR.

Ao longo do ano de 2026, a RNP avança na modelagem de seu serviço de repositório de dados de pesquisa, utilizando-se de sua infraestrutura de tecnologia da informação federada. Espera-se prover economicidade em escala nacional/regional para instituições/pesquisadores interessados na abertura de dados de pesquisa em repositórios digitais. Entende-se que não se trata apenas de uma instalação tecnológica, mas, sim, de uma infraestrutura onde especialistas apoiam as diferentes etapas do ciclo de vida de dados de pesquisa de forma a torná-los FAIR para acesso aberto em repositório digital.

Conflito de Interesses

Os autores declaram não haver conflitos de interesses.

CRedit – Contribuições dos Autores

Carolina Felicíssimo | Concetualização, Escrita – redação original
Leandro Ciuffo | Revisão

Referências

- Caregnato, S., Tavares, M., Campos, M., Rocha, R., Ferreira, M., Vagner, P., Prata, E., Morais, G., Sousa, J. & Felicíssimo, C. (2025). Implantação e integração de repositórios institucionais para censos educacionais: desafios e oportunidades na disseminação e consumo de dados e publicações. *Revista Ciência Aberta Lusófona*, 1, 34-42. <https://doi.org/10.82524/recal.2025.37>.
- Felicissimo, C., Ciuffo, L., Amaro, B. & Segundo, W. (2025). O estabelecimento do Núcleo de Dados de Pesquisa, um grupo de trabalho da Rede Brasileira de Repositórios Digitais. *Revista Ciência Aberta Lusófona*, 1, 24-33. <https://doi.org/10.82524/recal.2025.30>.

Felicissimo, C. H. (2026). Dados FAIR em Repositório Digital de Dados de Pesquisa. Zenodo. <https://doi.org/10.5281/zenodo.15374871>.

Rezende, L., Ribeiro, G., Castro, M., Oliveira, C., Silva, F.; Felicissimo, C., Juarez, K., Fonseca, C., & Ditzel, Y. (2024). Understanding DOI attribution in biodiversity repositories: A Brazilian case study. Zenodo. Open Repositories, Gotenburgo. <https://doi.org/10.5281/zenodo.12579382>.

Rosa, S.S, Ferreira Laipelt, R., Caregnato, S., Felicíssimo, C., Rocha, R. da, Tavares, M., & Machado Campos, M. (2025). Aplicabilidade das normas ISO 704:2022 e NBR 13790:1997 nos processos de modelagem e harmonização de termos e definições dos Censos Educacionais Brasileiros para aplicação em Sistemas de Organização do Conhecimento. *ISKO Brasil*, 8. <https://isko.org.br/ojs/index.php/iskobrasil/article/view/81>.

UNESCO. (2021). *Recommendation on Open Science*. <https://doi.org/10.54677/MNMH8546>.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. <https://doi.org/10.1038/sdata.2016.18>