

**CIÊNCIA ABERTA E O REÚSO DE DADOS DE PESQUISA***OPEN SCIENCE AND THE REUSE OF RESEARCH DATA*Renato Varella Bueno ¹

Universidade Federal de Minas Gerais

Benildes Coura Moreira dos Santos Maculan ²

Universidade Federal de Minas Gerais

RESUMO

A ciência aberta considera que dados de pesquisa com acesso aberto podem contribuir e acelerar o desenvolvimento científico com o reuso desses dados. Entretanto, isso ainda é um desafio, pois são escassos os relatos do efetivo reuso de dados de pesquisa. A proposta de estudo é mapear a literatura e descrever as metodologias utilizadas no reuso de dados de pesquisa, buscando evidências na área da Ciência da Informação. A metodologia empregada para mapear a literatura será a Revisão de Literatura Sistemática (RSL) e a análise dos dados empregará o método da Análise de Conteúdo. A análise dos resultados poderá proporcionar conhecimento das práticas de reuso de dados de pesquisa que podem orientar novas pesquisas.

Palavras-chave: Ciência Aberta; Dados de Pesquisa; Reuso de dados.

Keywords: Open Science; Research Data; Reuse of Research Data.

1 INTRODUÇÃO

A Ciência Aberta é um movimento universal que busca eliminar barreiras que possam impedir o acesso ao conhecimento científico. Uma das suas principais vertentes é o compartilhamento de dados de pesquisa que, tradicionalmente, são mantidos em sigilo, sendo divulgados apenas em artigos científicos. Uma das opções de compartilhamento de dados de pesquisa é em repositórios, que,

¹ Doutorando pelo Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Orcid: <https://orcid.org/0000-0002-4855-880X>. E-mail: renatovarellab@gmail.com.

² Docente do Programa de Pós-Graduação em Gestão e Organização do Conhecimento da Escola de Ciência da Informação da Universidade Federal de Minas Gerais. Orcid: <https://orcid.org/0000-0003-4303-9071>. E-mail: benildes@gmail.com.

geralmente, fornecem identificadores únicos para cada conjunto de dados, permitindo que sejam citados e referenciados adequadamente. Alguns exemplos populares são o Zenodo, Dryad e Figshare. No Brasil, há, por exemplo, o lançamento, em 2023, do repositório de pesquisa LattesData³ pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Reusar dados de pesquisa pode oferecer várias vantagens, como economia de tempo e recursos, validação de resultados anteriores, novas descobertas e contribuições para a literatura científica, conforme relatam Faniel e Zimmerman (2011). Na literatura se encontra com abundância reflexões acerca da importância do compartilhamento de dados para o avanço da ciência e também descrições de estratégias para a publicação dos dados brutos, tais como a consideração da natureza dos dados, o contexto da pesquisa e as questões éticas (CURTY, 2019). Uma vez que reusar dados de pesquisa envolve aproveitar conjuntos de dados existentes para realizar novas análises, estudos ou investigações, na prática os pesquisadores devem se atentar para questões tais como a qualidade e confiabilidade dos conjuntos de dados, conhecer a estrutura e o formato dos dados disponibilizados. O problema é que o reúso de dados ainda é um desafio para os pesquisadores, seja por barreiras técnicas, falta de qualidade nos dados ou por documentação insuficiente. Agravando esse problema, na literatura são escassos os relatos de reúso efetivo de dados de pesquisa em acesso aberto, que poderiam orientar aqueles que buscam realizar esse tipo de investigação.

Tendo em vista esse problema, neste estudo busca-se mapear a literatura e descrever as metodologias utilizadas no reúso de dados de pesquisa, buscando evidências na área da Ciência da Informação. Pretende-se entender em que medida a reutilização dos dados de pesquisa publicados em dados abertos estão produzindo novos conhecimentos na Ciência da Informação. Com isso, acredita-se permitir que outros pesquisadores possam verificar, reproduzir e aplicar as mesmas estratégias metodológicas de reúso de dados de pesquisa, possibilitando que pesquisas mais

³ Disponível em: <https://www.gov.br/cnpq/pt-br/cnpq-e-ibict-realizam-lancamento-do-lattes-data-repositorio-de-dados-de-pesquisa> Acesso em: 26 jun. 2023.

avançadas sejam conduzidas, pois outros cientistas podem explorar conjuntos de dados maiores e mais diversos.

2 FUNDAMENTAÇÃO TEÓRICA

A Ciência Aberta promove a transparência, acessibilidade e compartilhamento livre de conhecimento científico, tendo em vista tornar a pesquisa científica mais colaborativa, inclusiva e impactante. Ela tem em vista “tornar a pesquisa e os dados científicos acessíveis a todos” (CIÊNCIA, 2023, n. p.). Sayão e Sales (2020, p. 32) afirmam, porém, que tornar os dados abertos não é tarefa fácil, pois exige “uma forte contextualização de muitas faces, especialmente nas dimensões disciplinares e comunitárias”. Para os autores, “dado de pesquisa é todo e qualquer tipo de registro coletado, observado, gerado ou utilizado no âmbito da pesquisa científica”, sendo que seja possível que seja “interpretado, tratado e aceito como evidência pela comunidade científica e necessário para analisar, validar e produzir resultados de pesquisa” (SAYÃO; SALES, 2020, p. 32).

Tornar os dados de pesquisa acessíveis e compartilháveis não é tarefa corriqueira, pois apresenta algumas dificuldades e desafios, tais como: questões éticas e legais, anonimização e desidentificação dos dados, uso de padrões e metadados interoperáveis, sustentabilidade e preservação. Superar essas dificuldades requer esforços colaborativos entre pesquisadores, instituições acadêmicas, financiadores e organizações que mantêm repositórios de dados, exigindo diretrizes claras e boas práticas de gerenciamento de dados (SANTOS *et al.*, 2014).

3 METODOLOGIA

Esta pesquisa se caracteriza como básica, exploratória e descritiva, com abordagem qualitativa de análise dos dados. Para cumprir o objetivo geral do estudo será executado o método da Revisão Sistemática da Literatura (RSL). Os procedimentos metodológicos se baseiam em Biolchini *et al.* (2005) e Machi (2012) e se compõem em três etapas básicas: Planejamento, com a elaboração do protocolo

da RSL; Execução, para identificar e selecionar as bases de dados; e Seleção, com a compilação, sumarização e análise dos estudos recuperados.

A análise dos dados será realizada pelo método da Análise de Conteúdo, a partir da proposta de Bardin (2011, p. 48), que é um “conjunto de técnicas de análise [...] visando a obter [...] descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos”, organizada em três fases: pré-análise, exploração do material e tratamento dos resultados, inferência e interpretação.

4 CONSIDERAÇÕES FINAIS

A pesquisa está ainda em fase embrionária, e tem em vista apresentar as metodologias aplicadas na reutilização de dados abertos de pesquisa, e ainda não houve coleta de dados. A lacuna encontrada até o momento é de que o reaproveitamento de dados de pesquisa não é explorado na Ciência da Informação. Além disso, parece que existe empecilhos ou desconhecimento das regras para compartilhar dados. Estima-se que as evidências encontradas poderão contribuir para promover o reúso de dados nas comunidades científicas.

REFERÊNCIAS

BARDIN, L. **Analyse de contenu**. São Paulo: Éditions 70, 2011.

BIOLCHINI, J. *et al.* Systematic review in software engineering. **System Engineering and Computer Science Department**, COPPE/UFRJ, Technical Report ES, v. 679, n. 5, p. 45, 2005.

CIÊNCIA aberta no Brasil. Organização das Nações Unidas para a Educação, a Ciência e a Cultura, Unesco, 25 de janeiro de 2023. Disponível em: <https://www.unesco.org/pt/fieldoffice/brasil/>. Acesso em: 11 jun. 2023.

CURTY, R. G. Abordagens de reúso e a questão da reusabilidade dos dados científicos| Approaches for data reuse and the issue of scientific data reusability. **Liinc em Revista**, Rio de Janeiro, v. 15, n. 2, 2019. Dados de Pesquisa| Research Data| Datos de Investigación, v. 24, n. 2.

FANIEL, I. M.; ZIMMERMAN, A. Beyond the data deluge: a research agenda for large-scale data sharing and reuse. **International Journal of Digital Curation**, [S.l.], v. 6, n. 1, p. 58-69, 2011.

MACHI, L.A; MCEVOY, B. T. **The literature review**: six steps to success . 2. ed. Thousand Oaks, Calif.: Corwin Press, 2012.

SANTOS, P. X. et al. Política de Acesso Aberto ao Conhecimento: análise da experiência da Fundação Oswaldo Cruz/Fiocruz. **RECIIS**: Revista Eletrônica de Comunicação, Informação & Inovação em Saúde, Rio de Janeiro, v. 8, n. 2, p. 210-226, 2014.

SAYÃO, L. F.; SALES, L. F. Afinal, o que é dado de pesquisa?. **Biblos**, Rio Grande, v. 34, n. 2, p. 32-51, 2020.