

QUALIFICAÇÃO DE REPOSITÓRIOS DE DADOS E DE PUBLICAÇÕES: uma proposta de critérios alinhada à Ciência Aberta

QUALIFICATION OF DATA REPOSITERS AND PUBLICATIONS: a proposal of guidelines in support of Open Science

Priscila Machado Borges Sena¹, Tatyane Guedes Martins da Silva², Juliana Araujo Gomes de Sousa³, Washington Luís Ribeiro de Carvalho Segundo⁴, Bianca Amaro de Melo⁵

(1) Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), SAUS Quadra 5 - Lote 6, Bloco H, Brasília - DF, 70070-912, priscilasena@ibict.br.

(2) Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), SAUS Quadra 5 - Lote 6, Bloco H, Brasília - DF, 70070-912, tatyanesilva@ibict.br

(3) Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), SAUS Quadra 5 - Lote 6, Bloco H, Brasília - DF, 70070-912, julianasousa@ibict.br

(4) Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), SAUS Quadra 5 - Lote 6, Bloco H, Brasília - DF, 70070-912, washingtonsegundo@ibict.br

(5) Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), SAUS Quadra 5 - Lote 6, Bloco H, Brasília - DF, 70070-912, bianca@ibict.br

Resumo:

A Ciência Aberta abarca práticas voltadas para a transparência da pesquisa científica, promovendo o detalhamento dos processos de elaboração de metodologias e gestão de dados científicos. Posto isso, o objetivo de relatar o processo de elaboração de uma proposta de critérios de qualificação de repositórios de dados e de publicações, torna-se relevante e pertinente para a comunidade acadêmica e industrial que trabalham com dados no Brasil. Justifica-se por se tratar de um dos 11 Marcos do Compromisso 8 - Construir uma proposta de modelo de avaliação que fomente a Ciência Aberta (Marco 2 - Proposição de critérios de Qualificação de Repositórios de Dados, de Repositórios de Publicações), no 5º Plano de Ação Nacional, na Parceria para Governo Aberto (OGP). Utilizou-se de um relato de experiência, pautado no registro documental-analítico, o que permite caracterizar este trabalho como uma pesquisa documental e descritiva, de abordagem qualitativa. Como resultado se descreveram as três atividades estabelecidas para o desenvolvimento da proposta. O grupo de trabalho envolvido espera produzir um conjunto de recomendações normativas para auxiliar, principalmente, gestores de repositórios a fazerem uma avaliação da qualidade e confiabilidade dos repositórios de dados e/ou de publicações.

Palavras-chave: Ciência Aberta; Qualificação de repositórios; Repositório de Dados; Repositórios de publicações.

Abstract:

Open Science embraces practices aimed at the transparency of scientific research, promoting the detailing of the processes of elaboration of methodologies and management of scientific data. Therefore, the objective of reporting on the process of developing a proposal of qualification criteria for data and publication repositories becomes relevant and pertinent to the academic and industrial community that works with data in Brazil. It is justified because it is one of the 11 Milestones of Commitment 8 - Build a proposal for an evaluation model that fosters Open Science (Milestone 2 - Propose Qualification Criteria for Data and Publication Repositories), in the 5th National Action Plan, in the Open Government Partnership (OGP). An experience report was used, based on a documental-analytical register, which allows characterizing this work as a documental and descriptive research, with a qualitative approach. As a result, the three activities established for the development of the proposal were described. The working group involved hopes to produce a set of normative recommendations to help, mainly, repository managers to make an evaluation of the quality and reliability of data and/or publication repositories.

Keywords: Open Science; Repository Qualification; Data Repository; Publication Repositories.

1 Introdução

Conforme o Digital Repositories JISC Briefing Paper (2005), um repositório digital é onde conteúdos digitais e recursos estão armazenados, e podem ser pesquisados e

recuperados para uso futuro. Desse modo, suporta mecanismos de importação, exportação, identificação, armazenamento e recuperação de recursos digitais.

Os repositórios digitais constam definidos no Tesouro Brasileiro de Ciência da

Informação, como dispositivos para “administrar, armazenar e preservar conteúdos informacionais em formato eletrônico”, e que podem ser norteados por um assunto (repositórios temáticos) ou pela produção científica de uma instituição (repositórios institucionais) (PINHEIRO; FERREZ, 2014, p. 195).

Compreende-se os repositórios como ferramentas que registram o desenvolvimento da pesquisa por meio dos dados coletados, e ainda podem auxiliar a metodologia científica, em casos de experimentos sem êxito, cujos dados também podem ser registrados (SAYÃO; SALES, 2016).

Evidencia-se nas definições de repositórios digitais, alinhamento com as abordagens de Ciência Aberta, uma vez que, como um “movimento dos movimentos” (ALBAGLI, 2019), ela abarca práticas voltadas para a transparência da pesquisa científica, promovendo o detalhamento dos processos de elaboração de metodologias e gestão de dados científicos, para que estes sejam distribuídos, reutilizados e estejam acessíveis sem custos a todos os níveis da sociedade (SILVA; SILVEIRA, 2019).

Os repositórios de dados possibilitam o armazenamento e a gestão de dados vislumbrando o aprimoramento da recuperação, o que aumenta as potencialidades de reuso dos dados entre os pesquisadores. Deste modo, colabora para otimizar os processos da investigação científica, resultando no avanço da ciência (MONTEIRO, 2021).

Os repositórios de publicações embora sem conceito específico, alicerçam-se nas definições de repositórios institucionais. De acordo com Crow (2002), os repositórios institucionais são uma coleção digital que captura e preserva a produção intelectual da comunidade acadêmica. De forma que devem estar focados no armazenamento e promoção do acesso aberto para as publicações revisadas pelos pares (HANARD, 2001). A partir do posicionamento dos autores citados e para compreensão do trabalho apresentado, considera-se que nos repositórios de publicações os conjuntos de dados de pesquisa não fazem parte de seu escopo.

Posto isso, relatar o processo de elaboração de uma proposta de critérios de qualificação de repositórios de dados e de publicações é o objetivo delimitador deste trabalho.

Este objetivo é justificado por se tratar também de um dos 11 Marcos do Compromisso 8 - Construir uma proposta de modelo de avaliação que fomente a Ciência Aberta (Marco 2 - Proposição de critérios de Qualificação de Repositórios de Dados, de Repositórios de Publicações), no 5º Plano de Ação Nacional, na Parceria para Governo Aberto (do inglês *Open Government Partnership* - OGP) (TRANSPARÊNCIA EM CIÊNCIA..., 2022). Logo, torna-se relevante e pertinente para a comunidade acadêmica e industrial que trabalham com dados no Brasil, conforme visa o Workshop de Informação, Dados e Tecnologia (WIDaT), especificamente no subtópico “Ciência Aberta” do tópico “Dados”.

2 Procedimentos Metodológicos

Relatar um processo constitui-se em uma ferramenta valiosa para a compreensão da diversidade intrínseca à ciência contemporânea. O Relato de Experiência, de acordo com Daltro e Faria (2019), pode ser empregado como um instrumento político/social, devido ao registro documental-analítico realizado. O que permite caracterizar este trabalho como uma pesquisa documental e descritiva, de abordagem qualitativa.

Para tanto, o processo aqui detalhado refere-se às perspectivas de seis instituições diferentes: Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), Associação Brasileira de Editores Científicos (Abec), *Scientific Electronic Library Online* (SciELO), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação Oswaldo Cruz (Fiocruz) e Empresa Brasileira de Pesquisa Agropecuária (Embrapa).

O Ibict como órgão/entidade responsável por coordenar a execução do Marco 2 - Proposição de critérios de Qualificação de Repositórios de Dados, de Repositórios de Publicações, optou por utilizar a Plataforma

Google Meet¹ para realizar as reuniões online com as demais instituições.

3 Resultados Preliminares

Postas as definições utilizadas para a compreensão do grupo de trabalho em relação à repositório de dados e a repositório de publicações, foram estabelecidas as seguintes atividades para alcançar a concretização esperada para o Marco 2, a saber:

Atividade 1: Verificação de um conjunto inicial de critérios (exemplo: existência de identificadores persistentes, política de preservação). Essa atividade teve como ponto de partida as diretrizes OpenAIRE para repositórios de dados² e o quadro para boas práticas em repositórios da comunidade COAR³ além do conhecimento e das experiências práticas dos integrantes do grupo. O Quadro 1 contém os indicadores levantados pelos integrantes do grupo e pela literatura.

Quadro 1 - Indicadores preliminares do Marco 2 do Compromisso 8 do 5º Plano de Ação brasileiro na Parceria para Governo Aberto

Nº	Indicadores
1	Número de acessos ao repositório
2	Número de downloads no repositório
3	Percentual de identificadores perenes existentes no repositório
4	Qualidade dos metadados
5	Percentual de registros que permitem acesso ao texto completo ou aos conjuntos de dados
6	Existência de instrumentos de gestão; política

¹ Mais informações em: <https://meet.google.com/>.

² <https://livroaberto.ibict.br/bitstream/123456789/1086/2/Diretrizes%20OpenAIRE%20para%20repositórios%20de%20dados.pdf>

³ <https://livroaberto.ibict.br/bitstream/123456789/1089/2/Quadro%20para%20Boas%20Práticas%20em%20Repositórios%20da%20Comunidade%20COAR.docx.pdf>

Nº	Indicadores
7	Serviços ofertados pelo repositório, como ferramentas de análise de dados, modelagem, estatísticas de uso dos dados, consultoria para elaboração de PGD, serviços de referência para encontrar outros dados, mecanismos que auxiliam a citação
8	Interoperabilidade
9	Visibilidade aos dados
10	Recompensa do pesquisador por meio de estatística de citação
11	Acompanha os avanços tecnológicos
12	Atende as demandas do pesquisador

Fonte: Dados da Pesquisa (2022)

Atividade 2: Pesquisa sobre a literatura existente (levantamento bibliográfico e de grupos existentes no mesmo tema); considerando também as tecnologias envolvidas (Exemplo: *DSpace 7* e *Dataverse*) e pesquisa de critérios já levantados e aplicados sobre o ARCA Fiocruz (ISO-16363 e critérios *Trustworthy Repositories Audit and Certification: Criteria and Checklist - TRAC / TRUST / FAIR* para Repositórios de Dados).

Atividade 3: Utilização da matriz do *Resource Description and Access (RDA)*⁴ para criação de critérios destinados a repositórios de dados e para repositórios em geral. A primeira fase compreendeu a tradução das sugestões da matriz do inglês para o português. Seguindo para a separação das sugestões em colunas: “Quem”, “O que”, “Para quê”, “Tipo de critério”, “Aplicável somente a repositórios de dados?”. A partir dessa separação, os integrantes apresentaram sugestões de critérios que atendessem a demanda de cada perfil e sinalizaram se aquele critério poderia ser aplicado para repositórios em geral. Ainda com base na matriz de critérios da RDA, o grupo identificou e definiu os papéis que cada perfil profissional tem dentro do cenário que compõe o universo dos repositórios de dados e de publicação.

⁴ Ver mais em: <https://miro.com/app/board/uXjVOE-p7AI/>.

4 Considerações Finais

Considerando as definições de repositório digital adotadas no decorrer do texto e os objetivos da Ciência Aberta, que tem como princípio fundamental o acesso livre aos conjuntos de dados de pesquisa e aos resultados das pesquisas financiadas com recursos públicos, evidencia-se que a implementação e a gestão de serviços que garantem a exibição, recuperação, exportação, identificação, preservação, uso e reuso de recursos digitais vão exigir iniciativas que apoiem a disseminação de boas práticas para auxiliar os repositórios a alcançarem um nível eficaz no que tange a interoperabilidade.

Nessa direção, o grupo espera produzir um conjunto de recomendações normativas para auxiliar, principalmente, gestores de repositórios a fazerem uma avaliação da qualidade e confiabilidade dos repositórios de dados e/ou de publicações. De igual modo, almeja-se contribuir para a criação de novos repositórios estruturados em consonância com a Ciência Aberta.

Dessa maneira, vislumbra-se que a publicação, disseminação e adoção das recomendações possam proporcionar maior adequação dos repositórios brasileiros às práticas de compartilhamento e interoperabilidade disseminadas por iniciativas internacionais.

Por fim, conclui-se que o objetivo determinado para este trabalho se concretizou, uma vez que se obteve o relato almejado do processo de elaboração de uma proposta de critérios de qualificação de repositórios de dados e de publicações. Proposta essa basilar para a efetivação do Compromisso 8 - Construir uma proposta de modelo de avaliação que fomente a Ciência Aberta. Isso porque, as publicações e dados científicos, bem como seus processos de gestão são imprescindíveis para o impulsionamento e sustentabilidade da Ciência Aberta.

Agradecimentos

À Controladoria-Geral da União por toda a gestão e suporte para a concretização do

Compromisso 8 do 5º Plano de Ação Brasileiro.

À todas as pessoas e instituições participantes do grupo de trabalho referente ao Marco 2 do Compromisso.

À Financiadora de Estudos e Projetos (Finep) pelo subsídio financeiro concedido.

Referências

ALBAGLI, Sarita. O que é Ciência Aberta e qual o papel das agências de fomento diante deste fenômeno? In: Encontro Capes de Ciência Aberta. Tema: direitos de propriedade intelectual e políticas institucionais. Brasília, [s. n.], 2019.

Disponível em:

<http://capes.gov.br/conteudo/2-encontro-capes-de-ciencia-aberta/>.

CROW, Raym et al. The case for institutional repositories: A SPARC position paper. ARL Bimonthly Report, n. 223, 2002. Disponível em:

https://ils.unc.edu/courses/2014_fall/inls690_109/Readings/Crow2002-CaseforInstitutionalRepositoriesSPARCPaper.pdf

DALTRO, Mônica Ramos; DE FARIA, Anna Amélia. Relato de experiência: Uma narrativa científica na pós-modernidade. **Estudos e pesquisas em psicologia**, v. 19, n. 1, p. 223-237, 2019. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/revispsi/article/view/43015>.

DIGITAL REPOSITORIES JISC. **Digital repositories**: helping universities and colleges. 2005. Disponível em: <https://www.yumpu.com/en/document/view/15608201/digital-repositories-jisc>.

HARNAD, Stevan. Research access, impact and assessment. Times higher education supplement, n. 18. maio 2001. Disponível em:

<https://eprints.soton.ac.uk/255950/1/thes1.html>

MONTEIRO, Elizabete Cristina de Souza de Aguiar. **Operacionalização de repositórios**

de dados: uma análise sobre as perspectivas e atitudes dos pesquisadores nas questões de autoria e licença. Tese (Doutorado em Ciência da Informação). Marília/SP: Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, 2021. 268 f. Disponível em: <https://repositorio.unesp.br/handle/11449/214671>.

PINHEIRO, Lena Vania Ribeiro; FERREZ, Helena Dodd. **Tesouro Brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia, 2014. Disponível em: <http://sitehistorico.ibict.br/publicacoes-e-institucionais/tesouro-brasileiro-de-ciencia-da-informacao-1>.

SAYÃO, Luis Fernando; SALES, Luana Farias. Algumas considerações sobre os repositórios digitais de dados de pesquisa. **Informação & Informação**, v. 21, n. 2, p. 90-115, 2016. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/27939/20122>.

SILVA, Fabiano Couto Corrêa da; SILVEIRA, Lúcia da. O ecossistema da Ciência Aberta. **Transinformação**, v. 31, 2019. Disponível em: <https://www.scielo.br/j/tinf/a/dJ89vRg94Qxtf6Y7M49Hztr/?lang=pt>.

TRANSPARÊNCIA EM CIÊNCIA PARA O AVANÇO DA CIÊNCIA ABERTA - 5º PLANO DE AÇÃO OGP Brasil. 2022. Disponível em: <https://wiki.rnp.br/x/So-QCQ>.