

Repositórios abertos digitais: avaliação da confiabilidade com base nos critérios da ISO 16363

Digital open repositories: reliability evaluation based on ISO 16363 criteria

Keli Rodrigues do Amaral Benin

Universidade Estadual de Londrina, Brasil.

E-mail: keliamaral@hotmail.com

ORCID: <https://orcid.org/0000-0002-2564-2531>

Raíssa Yuri Hamanaka

Universidade Estadual de Londrina, Brasil.

E-mail: raissa.hamanaka@uel.br

ORCID: <https://orcid.org/0000-0001-9516-5825>

Paula Regina Ventura Amorim Gonçalves

Universidade Federal do Espírito Santo, Brasil.

E-mail: paulaventuramorim@gmail.com

ORCID: <https://orcid.org/0000-0002-5480-4106>

RESUMO

A Ciência Aberta se constitui a partir do acesso aberto; dados abertos; pesquisa reprodutível aberta; avaliação da Ciência Aberta; políticas de Ciência Aberta; ferramentas de Ciência Aberta; educação aberta; licenciamento aberto; Ciência Cidadã; preservação digital e inovação aberta. Nesta pesquisa, pretende-se abordar o acesso aberto, constituído pela publicação aberta em repositórios digitais (RDs). RDs são *locus* de preservação de objetos digitais, mas como garantir que um documento armazenado em um RD estará disponível para acesso a longo prazo? A confiabilidade de RDs se relaciona com a preservação a longo prazo de seus documentos digitais. A ISO 16363:2012 *Space data and information transfer systems – Audit and certification*

Como citar: Benin, K. R. A.; Hamanaka, R. Y.; & Gonçalves, P. R. V. A. (2022). Repositórios abertos digitais: avaliação da confiabilidade com base nos critérios da ISO 16363. En T. M. R. Dias (Ed.), *Informação, Dados e Tecnologia. Advanced Notes in Information Science, volume 2* (pp. 121-130). Tallinn, Estonia: ColNes Publishing. DOI: 10.47909/anis.978-9916-9760-3-6.90.

Copyright: © 2022, The author(s). This is an open access work distributed under the terms of the CC BY-NC 4.0 license which permits copying and redistributing the material in any medium or format, adapting, transforming and building upon the material as long as the license terms are followed.

of trustworthy digital repositories estabelece um conjunto de atributos, responsabilidades e requisitos que permitem a auditoria e certificação de repositórios digitais. O objetivo da pesquisa é avaliar o Repositório Institucional da Universidade Federal do Espírito Santo, Repositório Institucional do Instituto Federal do Amazonas e Repositório Digital da Universidade Federal do Rio Grande do Sul conforme os requisitos exigidos pela ISO 16363:2012. Pesquisa de abordagem qualitativa, com objetivo exploratório e descritivo. Foi enviado um survey para os colaboradores do repositório para que realizassem sua autoavaliação segundo os critérios de confiabilidade da norma. A preservação a longo prazo e a autenticidade de informações em meio digital se tornam necessárias para que os RDs sejam efetivamente confiáveis.

Palavras-chave: repositórios digitais confiáveis, preservação digital, ISO 16363

ABSTRACT

Open science is constituted of open access, open data, open reproducible search, open science assessment; open science policies; open science tools; open education; open licensing, citizen science, digital preservation, and open innovation. In this research, we intend to approach open access, constituted by open publication in digital repositories (DRs). DRs are the locus of preservation of digital objects, but how do we ensure that a document stored in a DR will be available for long-term access? The reliability of DRs relates to the long-term preservation of their digital records. The ISO 16363:2012 Space data and information transfer systems – Audit and certification of trustworthy digital repositories establish a set of attributes, responsibilities, and requirements that allow the audit and certification of repositories. The objective of the research is to evaluate the institutional repository of the Federal University of Espírito Santo, the Institutional Repository of the Federal Institute of Amazonas, and the Digital Repository of the Federal University of Rio Grande do Sul according to the requirements demanded by ISO 16363:2012. This article is qualitative research with an exploratory and descriptive objective. We sent a survey to the repository collaborators to conduct the self-assessment according to the ISO's reliability criteria. Long-term preservation and authenticity of the information in digital media become necessary for DRs to be effectively trusted.

Keywords: trustworthy digital repository, digital preservation, ISO 16363

INTRODUÇÃO

A CIÊNCIA ABERTA pode ser definida como investigações científicas colaborativas que preveem a disponibilização gratuita dos dados científicos, permitindo seu uso, reuso e reprodutibilidade (Fortaleza & Bertin, 2019; Oliveira & Silva, 2016). Vicente-Saez e Martinez-Fuentes (2018, p. 428) a definem como o “[...] conhecimento transparente e acessível que é compartilhado e desenvolvido por meio de redes colaborativas”.

Sayão e Sales (2014, p. 77) apontam o principal objetivo da Ciência Aberta, no qual “[...] o conhecimento científico é um patrimônio da humanidade e, que, portanto, deve estar disponível livremente para que as pessoas – cientistas ou não – possam usá-lo, reusá-lo e distribuí-lo sem constrangimentos tecnológicos, econômicos, sociais ou legais”.

É possível pensar essa ciência colaborativa por meio de uma taxonomia, na qual o termo mais geral seria Ciência Aberta, e os termos mais específicos suas facetas e subcategorias. Silveira *et al.* (2021) propõem a atualização da taxonomia desenvolvida pelo grupo Facilitate Open Science Training for European Research (Foster), disponível em Pontika *et al.* (2015), na perspectiva de pesquisadores brasileiros.

De acordo com Silveira *et al.* (2021), a Ciência Aberta se constitui a partir do acesso aberto; dados abertos; pesquisa reproduzível aberta; avaliação da Ciência Aberta; políticas de Ciência Aberta; ferramentas de Ciência Aberta; educação aberta; licenciamento aberto; Ciência Cidadã; preservação digital e inovação aberta. Nesta pesquisa, pretende-se abordar o acesso aberto, constituído pela publicação aberta que pode ocorrer pela via dourada (portais de periódicos, revistas de dados e *megajournals*) ou pela via verde (repositórios institucionais ou temáticos).

O compartilhamento de dados científicos permite a reprodutibilidade de pesquisas, a divulgação de resultados financiados com verba pública, o reuso de dados científicos e o avanço do estado da arte da ciência (Borgman, 2012). Os repositórios digitais (RDs) facilitam o compartilhamento de dados contribuindo para o movimento de acesso aberto e da Ciência Aberta (Sayão & Sales, 2016; Silva Júnior & Santos, 2019).

RDs são *locus* de preservação de objetos digitais (Shintaku *et al.*, 2016), mas como garantir que um documento armazenado em um RD estará disponível para acesso a longo prazo? Como preservar objetos digitais enquanto forem necessários? Essas questões norteiam a necessidade de existência de RDs confiáveis. Um RD confiável “[...] é aquele cuja missão é fornecer acesso confiável e de longo prazo a recursos digitais gerenciados para sua comunidade designada, agora e no futuro” (Research Libraries Group, 2002, p. 5). Dessa forma, RDs confiáveis devem garantir a integralidade, a proveniência e a preservação de seus objetos digitais de modo persistente (Gonzalez, 2017a).

A ISO 16363:2012 *Space data and information transfer systems – Audit and certification of trustworthy digital repositories*, estabelece um conjunto de atributos, responsabilidades e requisitos que permitem a auditoria e certificação de RDs.

A presente pesquisa pretende avaliar a confiabilidade de RDs após sua implementação a partir da ISO 16363:2012.

OBJETIVO

Avaliar o Repositório Institucional da Universidade Federal do Espírito Santo (RiUfes - <https://repositorio.ufes.br/>), o Repositório Institucional do Instituto Federal do Amazonas (Ifam - http://repositorio.ifam.edu.br/jspui/?subject_page=4) e o Repositório Digital da Universidade Federal do Rio Grande do Sul (Lume - <https://lume.ufrgs.br/>) conforme os requisitos exigidos pela ISO 16363:2012.

PROCEDIMENTOS METODOLÓGICOS

Foram realizados três estudos de caso: um no RiUfes, um no Repositório Institucional do Ifam e um no Lume. Utilizou-se como técnica a aplicação de *survey* para averiguar se os repositórios estavam em conformidade com os critérios de confiabilidade estabelecidos pela ISO 16363:2012.

O *survey* foi estruturado conforme a norma em três seções: infraestrutura organizacional (16 questões), gerenciamento de objetos digitais (24 questões) e infraestrutura e gestão de riscos de segurança (11 questões). Essas questões foram fechadas, e avaliadas segundo cinco níveis de maturidade de confiabilidade: nível um, inexistente (não implementa nenhum processo que atenda ao critério da norma); nível dois, incipiente (os processos para atender à norma não se encontram devidamente formalizados ou são realizados de forma *ad hoc*); nível três, em formação (possui processos definidos que satisfazem o requisito normativo); nível quatro, operacional (há políticas, procedimentos e processos implementados que satisfazem o critério da norma) e nível cinco, proativo (políticas, procedimentos e processos implementados que satisfazem o critério da norma e são constantemente melhorados com base em objetivos, metas e métricas) (Carvalho *et al.*, 2014), também havia a opção “não sei responder”. Além das 51 perguntas fechadas, houve uma pergunta aberta para que as dificuldades na criação e manutenção dos repositórios fossem descritas. No total foram elaboradas 52 questões.

Foi realizado o pré-teste do questionário com uma professora especialista em RDs confiáveis. O questionário foi enviado aos três repositórios no mês de julho de 2021. Foram recebidas três respostas, uma de cada repositório. O questionário não passou pelo Comitê de Ética em Pesquisa das universidades das proponentes, pois não houve identificação dos participantes desta pesquisa (Resolução nº 510, 2016).

Dentre os estudos correlatos sobre RDs confiáveis, podem-se elencar: Carvalho *et al.* (2014); Gonçalves (2017a); Gonçalves (2017b) e Lanzellote (2019). O questionário desta pesquisa foi baseado no criado por

Lanzellote (2019), na medição dos cinco níveis de maturidade em repositórios de Carvalho *et al.* (2014) e nos critérios estabelecidos pela ISO 16363:2012.

RESULTADOS

O objetivo da pesquisa foi propor uma forma de avaliação da confiabilidade de RDs. Dentre as 52 perguntas feitas ao RiUfes, 9 receberam “não sei responder” e, portanto, não foram computadas no cálculo da maturidade alcançada pelo repositório em relação aos critérios da norma.

Os requisitos do critério “infraestrutura organizacional” tiveram uma média de 3,25. A missão e documentação das ações do repositório, as licenças de propriedade intelectual e restrições de uso e a notificação ao depositante, são requisitos com maturidade proativa. O desenvolvimento de procedimentos de transparência, a atualização da política do repositório, as ações de auditoria e a autoavaliação, são requisitos em formação. E a estrutura organizacional e de pessoal possuem maturidade incipiente.

Os requisitos do critério “gestão de objetos digitais” tiveram uma autoavaliação com média de 2,7. O uso de identificadores únicos e de metadados descritivos, a recuperação de documentos pela comunidade alvo e as políticas de uso e acesso estão em estado proativo. O registro de falhas no acesso e anomalias no sistema e a transparência nos processos de auditoria estão em estado operacional. A elaboração da missão e da política do repositório, o uso de padrão de metadados, a autenticação dos depositantes, os relatórios de fluxo de trabalho, a verificação da integridade dos dados e a documentação do sistema e arquitetura técnica estão em formação. A elaboração de estratégias de armazenamento e preservação dos objetos digitais estão incipientes.

Os requisitos do critério “infraestrutura e gestão da segurança” tiveram uma média de 3,12. O monitoramento dos riscos é incipiente; o *backup*, o plano de preservação e a manutenção de infraestrutura tecnológica estão em formação e a infraestrutura do sistema e os processos de *software* e *hardware* estão operacionais.

O RiUfes pode ser avaliado com um nível de maturidade aproximadamente igual a três (em formação), com processos definidos, mas não totalmente implementados. Conforme extraído da pergunta aberta, a principal dificuldade encontrada pelo repositório é sua integração com a comunidade alvo.

No Lume, 13 perguntas receberam “não sei responder” e não foram contabilizadas no cálculo da média da maturidade dos critérios da norma. Os requisitos do critério “infraestrutura organizacional” tiveram uma média de 2,88. A quantidade de colaboradores e a sustentabilidade

financeira estão no estado de maturidade incipiente; atividades relativas à transparência e treinamentos estão em formação e o desenvolvimento da política de propriedade intelectual e restrições de uso estão operacionais.

Os requisitos do critério “gestão de objetos digitais” tiveram uma média de 2,83. A política de preservação, a missão, os procedimentos de aquisição de objetos digitais e os processos de auditoria estão em estado incipiente. A documentação do que o usuário pode ou não fazer está em formação. O uso de padrão de metadados, de pacotes de informação, a especificação de formatos de ficheiros, a autenticação de depositantes, o uso de identificadores persistentes, a verificação da integridade dos dados, a recuperação dos documentos do repositório pela comunidade alvo, a criação de diferentes políticas de acesso para diferentes objetos digitais, a documentação do sistema e sua arquitetura técnica e o monitoramento de falhas no acesso e segurança do sistema, todos esses requisitos se encontram operacionais. O uso de padrão de metadados, o uso de identificadores persistentes de acordo com padrões reconhecidos e a descoberta pela comunidade dos materiais de interesse estão em estado proativo.

Os requisitos do critério “infraestrutura e gestão da segurança” tiveram uma média de 3,65. O monitoramento e a manutenção da infraestrutura tecnológica, o *backup* e o plano de preservação estão operacionais. A especificação de recursos financeiros para a aquisição de *hardware* e *software* está em estado incipiente. O Lume pode ser avaliado com um nível de maturidade igual a três, com a confiabilidade em formação. A principal dificuldade encontrada pelo repositório é em relação às restrições de direito autoral, que demandam tempo na busca por autorizações dependendo do tipo de documento.

No Repositório Institucional do Ifam, 16 perguntas receberam “não sei responder” e não foram contabilizadas no cálculo da maturidade dos critérios da norma. Os requisitos do critério “infraestrutura organizacional” tiveram uma média de 2,46. A política e missão do repositório estão em estado operacional, a notificação aos depositantes, a realização de autoavaliações e o uso de uma matriz de responsabilidades estão incipientes.

Os requisitos do critério “gestão de objetos digitais” tiveram uma média de 2,79. As verificações de integridade dos dados, a documentação dos objetos digitais, o uso de metadados descritivos, o uso pela comunidade alvo do repositório, a documentação do sistema e sua arquitetura técnica, a política de propriedade intelectual e restrições de uso e os procedimentos de monitoramento de falha do sistema são requisitos em maturidade operacional. Enquanto a política de preservação, o uso de

metadados de preservação e a transparência nos processos de auditoria são requisitos em estado incipiente.

Os requisitos do critério “infraestrutura e gestão da segurança” tiveram uma média de autoavaliação três. O emprego de boas práticas relativas à gestão da segurança da informação está em estado operacional e a identificação do número de cópias dos objetos digitais está incipiente.

O Repositório Institucional do Ifam pode ser avaliado com um nível de maturidade aproximadamente igual a três, com a confiabilidade em formação. A principal dificuldade em sua implementação é o diálogo com a equipe de tecnologia da informação, responsável pelas mudanças infra-estruturais e tecnológicas do repositório.

CONSIDERAÇÕES FINAIS

A avaliação do nível de maturidade da confiabilidade do RiUfes, Repositório Institucional do Ifam e Lume, com base nos critérios da norma, foi a mesma com confiabilidade em formação. Ressalta-se que apesar de maturidade semelhante, os repositórios apresentaram diferentes dificuldades. Dificuldades técnicas (no trabalho integrado com a equipe de tecnologia da informação), legais (no gasto de tempo para a obtenção de autorizações relacionadas a direitos autorais) e no relacionamento integrado com a comunidade alvo.

Na era do *Big Data*, na qual um grande volume de dados é produzido e precisa ser processado e analisado rapidamente, surgem novos desafios e possibilidades de tratamento da informação. A preservação a longo prazo e a autenticidade de informações em meio digital se tornam necessárias para que os RDs sejam efetivamente confiáveis.

Algumas limitações da pesquisa podem ser elencadas: o escopo reduzido de repositórios avaliados e a quantidade de respondentes do questionário, o que inviabiliza a generalização dos resultados. Por outro lado, o instrumento de medição da confiabilidade de RDs pôde ser validado e foi a solução proposta para a avaliação da confiabilidade destes. Além disso, a pesquisa não levou em consideração o modelo de referência para sistemas de preservação de objetos digitais, o documento *Trusted Digital Repositories: Attributes and Responsibilities*, os princípios FAIR (*Findable, Accessible, Interoperable, Reusable*), os critérios para certificação de repositórios de dados de pesquisa (princípios TRUST e certificação *CoreTrust-Seal*). Como pesquisa futura sugere-se a contemplação destes aspectos, a ampliação do diagnóstico de medição de maturidade da confiabilidade de RDs, a proposição de planos de ação para o aumento da confiabilidade e a criação de um sistema *web* que permita aos gestores dos repositórios realizarem uma autoavaliação, com a exibição gráfica de relatórios.

CONFLITOS DE INTERESSE

As autoras declaram que não há conflitos de interesse.

DECLARAÇÃO DE CONTRIBUIÇÃO

Conceptualização, curadoria de dados, análise formal, pesquisa, metodologia, administração do projeto, validação, visualização, escrita do rascunho inicial e escrita da revisão e edição: Keli Rodrigues do Amaral Benin, Raíssa Yuri Hamanaka e Paula Regina Ventura Amorim Gonçalves.

DECLARAÇÃO DE CONSENTIMENTO DE DADOS

Os dados gerados durante este artigo foram incluídos no manuscrito.

AGRADECIMENTOS

As autoras gostariam de agradecer a Diego Leonardo de Souza Fonseca, Janise Silva Borges da Costa, Márcia Auzier e Morgana Carneiro de Andrade. 

REFERÊNCIAS

- BORGMAN, C. L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science and Technology*, 63(6), 1059-1078. <https://doi.org/10.1002/asi.22634>
- RESOLUÇÃO Nº 510, DE 07 DE ABRIL DE 2016. (2016, 24 maio). *Dispõe sobre as normas aplicáveis a pesquisas em Ciências Humanas e Sociais*. Ministério da Saúde. Recuperado de https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/22917581
- CARVALHO, J. et al. (2014). Auditoria ISO 16363 a repositórios institucionais. *Cadernos BAD*, (2), 29-39. Recuperado de <https://publicacoes.bad.pt/revistas/index.php/cadernos/article/view/1175/pdf>
- FORTALEZA, J. M., & BERTIN, P. R. B. (2019). A parceria para Governo Aberto e o compromisso pela Ciência Aberta. In Shintaku, M., & Sales, L. F. (Org.), *Ciência aberta para editores científicos* (pp. 21-28). Botucatu: ABEC.
- GONÇALEZ, P. R. V. A. (2017a). Recomendações para certificação ou medição de confiabilidade para repositórios arquivísticos digitais confiáveis com ênfase no acesso. *Informação & Informação*, 22(1), 215-241. <https://doi.org/10.5433/1981-8920.2017v22n1p215>

- GONÇALEZ, P. R. V. A. (2017b). *Repositórios arquivísticos digitais confiáveis: identificação de requisitos com ênfase no acesso à informação*. (Tese de Doutorado, Universidade Estadual Paulista, Marília, Brasil). Recuperado de https://repositorio.unesp.br/bitstream/handle/11449/150028/goncalez_prva_dr_mar.pdf?sequence=3
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. (2012). *ISO 16363: Space data information transfer systems: audit and certification of trustworthy digital repositories*. Genebra: The British Standards Institution.
- LANZELLOTE, M. F. (2019). *Avaliação de confiabilidade do repositório Arca da Fiocruz*. (Dissertação de Mestrado, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, Brasil). Recuperado de <http://www.repositorio-bc.unirio.br:8080/xmlui/bitstream/handle/unirio/12883/PPGB DISSERTACAO MICHELLE FL 2019.pdf?sequence=1>
- OLIVEIRA, A. C. S., & SILVA, E. M. (2016). Ciência Aberta: dimensões para um novo fazer científico. *Informação & Informação*, 21(2), 5-39. <https://doi.org/10.5433/1981-8920.2016v21n2p5>
- PONTIKA, N. et al. (2015). *Fostering Open Science to Research using a Taxonomy and an eLearning Portal*. Artigo apresentado na 15th International Conference on Knowledge Technologies and Data Driven Business, Graz, Áustria. Recuperado de http://oro.open.ac.uk/44719/2/kmi_foster_iknow.pdf
- RESEARCH LIBRARIES GROUP. (2002). *Trustworthy digital repositories: attributes and responsibilities*. Mountain View: RLG. Recuperado de <https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf>
- SAYÃO, L. F., & SALES, L. F. (2016). Algumas considerações sobre os repositórios digitais de dados de pesquisa. *Informação & Informação*, 21(2), 90-115. <https://doi.org/10.5433/1981-8920.2016v21n2p90>
- SAYÃO, L. F., & SALES, L. F. (2014). Dados abertos de pesquisa: ampliando o conceito de acesso livre. *RECIIS: Revista Eletrônica de Comunicação em Informação, Inovação e Saúde*, 8(2), 76-92. <https://doi.org/10.3395/reciis.v8i2.934.pt>
- SHINTAKU, M., DUQUE, C., & SUAIDEN, E. J. (2016). Análise da adesão às tendências da Ciência pelos repositórios institucionais brasileiros. *IN-CID: Revista de Ciência da Informação e Documentação*, 6(2), 148-169. <https://doi.org/10.11606/issn.2178-2075.v6i2p148-169>
- SILVA JÚNIOR, L. P., & SANTOS, T. H. N. (2019). Repositórios de dados científicos: um panorama teórico-prático. In Dias, G. A., & Oliveira, B. M. J. F. (Org.), *Dados científicos: perspectivas e desafios* (pp. 89-112). João Pessoa: Ed. UFPB.
- DA SILVEIRA, L., RIBEIRO, N. C., DE OLIVEIRA SANTOS, S. R., DE ALMEIDA SILVA, F. M., DA SILVA, F. C. C., CAREGNATO, S. E., ... & ARAÚJO, R. F. (2021). Ciência aberta na perspectiva de especialistas brasileiros:

proposta de taxonomia: propuesta de taxonomía. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 26, 1-27. <https://doi.org/10.5007/1518-2924.2021.e79646>

VICENTE-SAEZ, R., & MARTINEZ-FUENTES, C. (2018). Open Science now: a systematic literature review for an integrated definition. *Journal of business research*, 88, 428-436. <https://doi.org/10.1016/j.jbusres.2017.12.043>

