

Modelo de avaliação diagnóstica FAIR para fontes de informação científicas na internet endereçadas a bibliotecas digitais abertas e agregadas

FAIR diagnostic assessment model for scientific information sources on the internet addressed to open and aggregated digital libraries

Heloene Gonçalves Santos Passos¹



<https://lattes.cnpq.br/2362442830790668>

 <https://orcid.org/0009-0009-5603-3323>

Daniela Lucas da Silva Lemos²



<http://lattes.cnpq.br/9280443047358807>

 <https://orcid.org/0000-0003-1565-7366>

Dirceu Flávio Macedo³



<http://lattes.cnpq.br/2765166629600107>

 <https://orcid.org/0000-0003-4837-3618>

Resumo

O artigo propõe um modelo de avaliação diagnóstica baseado nos princípios FAIR (*Findable, Accessible, Interoperable, Reusable*) para avaliar fontes de informação na internet, com foco na qualificação de acervos em bibliotecas digitais abertas e agregadas. A pesquisa, de natureza aplicada, adota uma abordagem quali-quantitativa, exploratória e descritiva, analisando periódicos científicos da área da Museologia. O modelo se fundamenta em seis dimensões analíticas: (i) Padrão de Dados, (ii) Metadados e Indexação, (iii) Licença de Uso, (iv) Infraestrutura e Protocolos de Acesso, (v) Documentação e Proveniência, e (vi) Integração e Conectividade. A aderência FAIR foi mensurada por uma fórmula que avalia os critérios dentro dessas dimensões, utilizando uma escala de pontuação de 0 a 3. Os resultados revelam que 20% dos periódicos analisados atingiram a aderência máxima aos princípios FAIR. Destacou-se que a adoção de procedimentos simples, como o uso de identificadores persistentes, poderia aumentar significativamente a conformidade dos demais periódicos. Além disso, periódicos que utilizam padrões como *Dublin Core* e protocolos interoperáveis apresentaram melhor indexação e acessibilidade. O estudo conclui que o modelo proposto é uma ferramenta eficiente para avaliar a qualidade de fontes de informação e apoiar a seleção, a extração e a organização de coleções em bibliotecas digitais. Os achados oferecem subsídios para gestores e pesquisadores das áreas de Ciência Aberta, Humanidades Digitais e Ciência da Informação, auxiliando na tomada de decisões e na implementação de boas práticas para o aprimoramento da interoperabilidade e reutilização de dados científicos na Web.

¹ Mestranda em Ciência da Informação, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil. Assistente Administrativo do Sistema Integrado de Biblioteca, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil. heloene.passos@ufes.br.

² Doutora em Ciência da Informação, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, MG, Brasil. Professora e Pesquisadora, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil. daniela.l.silva@ufes.br.

³ Mestre em Ciência da Informação, Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil. Pesquisador no Laboratório de Ciência da Informação, Dados e Tecnologia (Cidat Lab), Universidade Federal do Espírito Santo (UFES), Vitória, ES, Brasil. dfmacedo@gmail.com.

Palavras-chave: Princípios FAIR; análise de conteúdo; Modelo de Avaliação Diagnóstica FAIR; fontes de informação; biblioteca digital.

Abstract

The article proposes a diagnostic evaluation model based on the FAIR principles (Findable, Accessible, Interoperable, Reusable) to assess information sources on the internet, focusing on the qualification of collections in open and aggregated digital libraries. The research, applied in nature, adopts a qualitative, exploratory, and descriptive approach, analyzing scientific journals in the field of Museology. The model is based on six analytical dimensions: (i) Data Standards, (ii) Metadata and Indexing, (iii) Usage License, (iv) Infrastructure and Access Protocols, (v) Documentation and Provenance, and (vi) Integration and Connectivity. FAIR adherence was measured using a formula that evaluates criteria within these dimensions, employing a scoring scale from 0 to 3. The results reveal that 20% of the analyzed journals achieved the highest level of adherence to the FAIR principles. It was highlighted that adopting simple procedures, such as the use of persistent identifiers, could significantly improve compliance among other journals. Furthermore, journals that apply standards such as Dublin Core and interoperable protocols demonstrated better indexing and accessibility. The study concludes that the proposed model is an effective tool for evaluating the quality of information sources and supporting the selection, extraction, and organization of collections in digital libraries. The findings provide valuable insights for managers and researchers in the fields of Open Science, Digital Humanities, and Information Science, aiding decision-making and the implementation of best practices to enhance interoperability and the reuse of scientific data on the Web.

Keywords: FAIR Principles; content analysis; FAIR Diagnostic Evaluation Model; information sources; digital library.

1 INTRODUÇÃO

O grande volume de dados gerados pela explosão da informação tem sido um desafio no aspecto da organização e recuperação da informação com qualidade para instituições, governos, pesquisadores e sociedade em geral, que demonstram preocupação com o fluxo informacional presente nos meios digitais devido à percepção de que, muitas vezes, não se verifica a adoção de padrões e regras (Gilliland, 2016) na disponibilização dos conteúdos nas bases de dados.

Lemos e Souza (2024) consideram que encaminhamentos de pesquisas têm sido feitos progressivamente pela Ciência da Informação e Ciência da Computação, cujos esforços são endereçados aos sistemas de recuperação da informação (SRIs) em ambientes digitais a fim de melhorar o escopo de interoperabilidade entre diferentes esquemas de metadados e aplicações.

Nesse contexto, os princípios FAIR (Wilkinson *et al.*, 2016), um acrônimo para (*Findable/Localizável*, *Accessible/Acessível*, *Interoperable/Interoperável* e *Reusable/Reutilizável*), se adentram e podem ser considerados como um norteador de ações a serem desenvolvidas nos SRIs (Martins *et al.*, 2022), desde a seleção, a extração,

o tratamento e a organização de coleções de documentos até o depósito desses e seus registros de metadados em bases de dados, repositórios ou bibliotecas digitais, com a finalidade de garantir localização, acesso, interoperabilidade e reuso dos dados dispostos nas bases ou fontes de informação disponíveis na rede. Para Araújo e Fachin (2015), fonte de informação, vista como um sistema de informação, pode ser um documento, base de dados ou um repositório, tendo a característica de informar algo para alguém.

A disponibilização de fontes de informação na rede para uso e reúso tem forte relação com o movimento Ciência Aberta, que vai ao encontro dos objetivos da Unesco (Unesco, 2022) em valorizar a colaboração e o compartilhamento de conhecimento à humanidade, baseando-se nos sistemas de propriedade intelectual para ampliar o acesso de todos ao conhecimento para o benefício da ciência e da sociedade. A preocupação da Ciência Aberta é “tornar a atividade de pesquisa mais transparente, mais colaborativa e mais eficiente” (Sayão; Sales, 2021, p. 175). Assim sendo, os princípios FAIR facilitam a governança de dados, a cooperação internacional e o acesso universal ao conhecimento científico, pois objetiva garantir a localização, o acesso, a interoperabilidade e o reúso de artefatos digitais disponibilizados pela comunidade científica na rede ou em rede, assumindo, logo, um importante reconhecimento internacional nas boas práticas de gestão de dados em diferentes contextos (Sales *et al.*, 2021).

Adicionalmente, os avanços tecnológicos e a Web, a exemplo do FAIR, possibilitaram um maior alcance da informação de forma rápida, dinâmica e acessível. As Bibliotecas Digitais (BD) são exemplos desse alcance, disponibilizando informação além do espaço físico. Essa característica das Bibliotecas Digitais expande suas possibilidades de alcance fazendo com que se tornem um meio eficaz para distribuição, cooperação e acesso ao conhecimento científico, além de atuarem como núcleo agregador para comunidades segmentadas e distribuídas geograficamente (Sayão, 2021).

Nessa perspectiva, iniciativas de criação e disponibilização de bibliotecas digitais com dados abertos, agregados e científicos têm crescido nos últimos anos, especialmente no âmbito do patrimônio cultural digital (Martins *et al.*, 2022), proporcionando inovação na forma de produção e tratamento de fontes de informação.

Entretanto, um dos principais desafios na construção de BD com dados abertos e agregados de fontes heterogêneas encontra-se na localização e na seleção de fontes de

informações pertinentes ao propósito desse tipo de SRI, além da extração de dados contextualizados e da organização desses dados em coleções temáticas a compor a base de dados agregada (Martins *et al.*, 2022).

Tal cenário dificulta a prestação de serviço de qualidade aos usuários, o que implica em garantir que tenham acesso a informações precisas, atualizadas, acessíveis e confiáveis, contribuindo, portanto, para uma experiência positiva em relação ao SRI. Além disso, dificulta a reutilização de fontes de informação na Web, considerando que os objetos digitais são elementos preponderantes na geração de valor em suas circulações e reuso social. Isso, inclusive, prejudica o avanço do crescimento de campos emergentes, como as Humanidades Digitais (Poole, 2017), que buscam a extração de valor em objetos digitais por meio de ações de reuso para o desenvolvimento econômico, social e humano.

Nesse ensejo, formula-se a questão desta pesquisa em responder: como avaliar a qualidade de fontes de informações científicas para fins de seleção, extração e organização de coleções temáticas em bibliotecas digitais abertas e agregadas à luz dos princípios FAIR? As fontes de informações científicas escolhidas como amostragem para o percurso da presente pesquisa foram as da área da Museologia pelo fato de que há em andamento um projeto de pesquisa mais abrangente envolvendo a construção de uma biblioteca digital aberta e agregada de conhecimento científico para este domínio.

Assim sendo, o objetivo da presente pesquisa é apresentar um modelo de avaliação diagnóstica fundamentado nos princípios FAIR para a avaliação de fontes de informação científica de interesse à criação de Bibliotecas Digitais abertas e agregadas.

Os resultados desta pesquisa podem ajudar o gestor de um projeto envolvendo BD desta natureza a visualizar, de forma sistematizada, a relação entre fontes de informações candidatas a re(uso), possíveis dimensões analíticas sobre áreas importantes que refletem a qualidade de fontes de informação científica na internet e índices de aderência aos princípios FAIR. Essa visualização pode ajudar o gestor a identificar quais áreas (dimensões) apresentam maior conformidade e desafios, segundo os princípios FAIR, facilitando a tomada de decisões e a análise comparativa entre fontes de informação candidatas a compor uma BD.

Para além, essas dimensões podem ajudar na investigação se as fontes de informação seguem os princípios FAIR de maneira eficaz. Uma abordagem sistemática

com métricas de informação pode revelar quais fontes estão mais alinhadas às boas práticas e quais precisam de melhorias para se tornarem mais encontráveis, acessíveis, interoperáveis e reutilizáveis.

2 PROCEDIMENTOS METODOLÓGICOS

A presente pesquisa é classificada como de natureza aplicada, de abordagem quali-quantitativa, de objetivo exploratório e descritivo em sites de periódicos nacionais da área da Museologia, adotando pesquisa bibliográfica, documental e mapeamento de fontes de informação.

O primeiro ponto a considerar nas etapas metodológicas foi a concepção das dimensões analíticas, para as quais se usou o método de análise de conteúdo (Bardin, 2016) em artigos científicos ligados à Ciência da Informação que retratam áreas importantes sobre qualidade de fontes de informação científicas na internet para uso e reúso, com garantia literária, conforme revisão bibliográfica. As dimensões são enumeradas e elucidadas abaixo:

- a) **Padrão de dados:** recomendações para a padronização de estrutura (ex.: padrão de metadados *Dublin Core*), valor (ex.: vocabulário controlado *Art & Architecture Thesaurus - AAT*), conteúdo (ex.: regra de catalogação *Anglo American Cataloguing Rules - AACR2*) e comunicação de dados, numa linguagem legível para a máquina (ex.: *Dublin Core Extensible Markup Language - XML*), para a produção qualificada de metadados.
- b) **Indexação a partir de metadados:** adoção de padrões de metadados; descritores preenchidos e padronizados; compatibilidade com mecanismos de busca e *harvesters*.
- c) **Licença de uso:** especificação do conjunto de licenças explicitadas para publicação dos dados (ex.: *Copyright; Creative Commons*); indicação clara das permissões e restrições de uso.
- d) **Infraestrutura tecnológica e de comunicação:** especificação do tipo de sistema de recuperação da informação e *softwares* usados (ex.: repositório *Open Journal Systems - OJS*; protocolos de busca e recuperação, tais como o

Hypertext Transfer Protocol – HTTP e o Open Archives Initiative Protocol for Metadata Harvesting - OAI-PMH; Application Programming Interface – APIs para acesso à base de dados) para acessar e manipular dados e metadados.

- e) **Documentação e proveniência:** disponibilidade de documentação clara sobre os dados e a fonte de informação (ex.: catálogo bibliográfico; repositório de dados de pesquisa); e registro de proveniência dos dados (ex: quem criou; processou; modificou; versões anteriores).
- f) **Integração e Conectividade:** uso de identificadores padronizados para autores, instituições e publicações (ex: *Open Researcher and Contributor ID - ORCID; Research Organization Registry -ROR; International Standard Serial Number- ISSN; Digital Object Identifier - DOI*); e capacidade de interligar diferentes conjuntos de dados e metadados abertos (ex.: princípios para *linked open data*).

A segunda etapa foi determinar o critério de seleção das fontes de informação para compor a lista de avaliação (*corpus* central). Assim, dentre vários tipos de fontes passíveis de serem identificadas e selecionadas na rede, o critério adotado foi a seleção inicial de 10 repositórios de periódicos científicos da área da Museologia, visto que a presente pesquisa faz parte de um projeto maior em parceria com o Instituto Brasileiro de Museus (Ibram) em criar uma biblioteca digital aberta e agregada de conhecimento científico dos museus brasileiros. Deste modo, seria possível obter resultados plausíveis para validação e avanços nos estudos.

A última etapa metodológica focou na construção do modelo de avaliação diagnóstica das fontes de informação (periódicos científicos) para o qual contou com uma planilha de dados para pontuação, sendo formatada, em suas linhas, com a relação das fontes de informação, e suas colunas com as 6 dimensões analíticas, cada qual com subníveis alinhados aos critérios FAIR, totalizando 15 variáveis para cada dimensão. É válido ressaltar que os critérios FAIR não se relacionam com apenas uma dimensão. Frequentemente, eles estão presentes em outras dimensões, tornando necessária uma análise rigorosa e criteriosa da aderência desses princípios em áreas importantes da Ciência da Informação. Essas áreas contribuem para a qualidade das fontes de informações científicas disponíveis na Internet para diversos fins, com destaque para a

extração e agregação de dados provenientes de fontes heterogêneas em bases de dados. O Quadro 1 a seguir exhibe a relação dos critérios FAIR juntamente com suas respectivas descrições.

Quadro 1 – Critérios FAIR e suas descrições

Critério	Descrição
F1	Dados devem ter um identificador único e persistente (PID – <i>persistent identifier</i> , em inglês).
F2	Dados devem ser descritos com metadados ricos e precisos.
F3	Metadados devem incluir o identificador do dado que descrevem.
F4	Metadados devem ser registrados ou indexados em um recurso pesquisável.
A1	Dados devem ser recuperáveis por meio de protocolos padronizados.
A1.1	O protocolo deve ser aberto, livre e universalmente implementável.
A1.2	O protocolo deve permitir autenticação e autorização, quando necessário.
A2	Metadados devem permanecer acessíveis, mesmo que os dados não estejam mais disponíveis.
I1	Dados devem usar linguagens, formatos e vocabulários padronizados e amplamente aceitos.
I2	Dados devem usar vocabulários que sigam os princípios FAIR.
I3	Dados devem incluir referências a outros dados para facilitar a integração.
R1	Dados devem ser descritos com atributos precisos e relevantes
R1.1	Dados devem ter uma licença clara e acessível que permita reutilização
R1.2	Dados devem incluir informações sobre sua proveniência.
R1.3	Dados devem seguir padrões de domínio relevantes para permitir reutilização.

Fonte: Elaboração própria.

A escala de pontuação adotada variou de 0 a 3, sendo: **0**, quando o critério FAIR não se aplicava à dimensão avaliada, segundo o julgamento do avaliador; **1**, para casos de baixa aderência; **2**, quando a aderência era considerada média; e **3**, nos casos de alta aderência. A análise foi conduzida de forma manual e criteriosa em fevereiro de 2025, com posterior processamento automático dos dados para cálculo da pontuação atribuída a cada dimensão da fonte de informação.

O cálculo seguiu a seguinte fórmula: Pontuação $i,k = ((\sum \text{Valor } i,j,k) / n)$

Onde:

- “i” é uma **fonte de informação** particular.
- “j” é um **critério** aplicável particular incluído na **dimensão**.
- “k” é uma **dimensão** na **fonte de informação** “i”.

- Pontuação_k é a pontuação para a fonte de informação “i” para o conjunto de critérios pertencente à dimensão “k”.
- Valor_{ij,k} é o valor para o critério “j” de uma dimensão “k” na fonte de informação “i”.
- n é o número de **critérios aplicáveis** a uma **dada dimensão**.

Após a aplicação da pontuação final (dada pela média) para os 10 periódicos científicos, foi possível obter a informação de *ranking*, o que permitiu avaliá-los frente às dimensões analíticas, apresentadas e discutidas na próxima seção.

3 RESULTADOS E DISCUSSÃO

Os resultados foram obtidos a partir da análise do avaliador em relação aos critérios FAIR aplicáveis e não aplicáveis, o que resultou no arranjo que se segue. Em Padrão de Dados, foram identificados 12 critérios aplicáveis e 3 não aplicáveis (F3, A2, R1). Em Metadados e Indexação, foram 13 critérios aplicáveis e 2 não aplicáveis (A1.2, R1.3). Na Licença de Uso, houve 9 critérios aplicáveis e 6 não aplicáveis (F2, F3, F4, A1.2, I2, R1.3). Em Infraestrutura e Protocolos de Acesso, foram 14 critérios aplicáveis e 1 não aplicável (R1.3). Em Documentação e Proveniência, registraram-se 10 critérios aplicáveis e 5 não aplicáveis (F3, F4, A1.2, R1, R1.3). Por fim, em Integração e Conectividade, houve 13 critérios aplicáveis e 2 não aplicáveis (F3, R1.3).

Assim, para os 10 periódicos analisados, dentro das 6 categorias analíticas propostas, constataram-se diferentes níveis de conformidade com os princípios FAIR, variando entre 1 e 3, conforme mostrado no Quadro 2.

Quadro 2 – Ranking dos periódicos científicos frente às dimensões baseadas em FAIR

Periódico	Pontuação	Análise dos periódicos com base nas seis dimensões analíticas e nos princípios FAIR
Anais do Museu Paulista: História e Cultura Material	3	Emprega padrão de metadados <i>Dublin Core</i> . Metadados descritivos e administrativos preenchidos de forma completa e normalizada, o que sugere uso de vocabulário controlado. Apresenta metadados indexados em mecanismos de busca. Licenças de uso claras. Utiliza o <i>software</i> OJS com suporte ao protocolo OAI-PMH. Apresenta documentação de dados e a proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos. Faz uso de PID.

Boletim do Museu Paraense Emílio Goeldi Ciências Naturais	3	Emprega padrão de metadados <i>Dublin Core</i> . Metadados descritivos e administrativos preenchidos de forma completa e normalizada, o que sugere uso de vocabulário controlado. Apresenta metadados indexados em mecanismos de busca. Licenças de uso claras. Utiliza o <i>software</i> OJS com suporte ao protocolo OAI-PMH. Apresenta documentação de dados e a proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos. Faz uso de PID.
Boletim do Museu Paraense Emílio Goeldi Ciências Humanas	2,8	Metadados descritivos e administrativos preenchidos de forma completa e normalizada, o que sugere uso de vocabulário controlado. Apresenta metadados indexados em mecanismos de busca. Licenças de uso claras. Apresenta documentação de dados e a proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos. Faz uso de PID.
Anais do Museu Histórico Nacional	2,7	Emprega padrão de metadados <i>Dublin Core</i> . Metadados descritivos e administrativos preenchidos. Apresenta metadados indexados em mecanismos de busca. Licenças de uso claras. Apresenta documentação de dados. Usa ORCID e ISSN.
Humanidades em Revista	2,16	Emprega padrão de metadados <i>Dublin Core</i> . Metadados descritivos e administrativos preenchidos. Apresenta metadados indexados em mecanismos de busca. Licenças de uso claras. A proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos. Usa apenas o ISSN.
Arquivos do Museu de História Natural e Jardim Botânico da UFMG	2,07	Emprega padrão de metadados <i>Dublin Core</i> . Metadados descritivos e administrativos preenchidos. Apresenta metadados indexados em mecanismos de busca. Utiliza o <i>software</i> OJS com suporte ao protocolo OAI-PMH. A proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos. Usa apenas o ORCID.
Boletim do Museu de Biologia Mello Leitão	2,03	Usa apenas ISSN. Apresenta metadados indexados em mecanismos de busca. Licenças de uso claras. Apresenta documentação de dados. A proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos.
Africanidades: a Revista do Mafro	1,99	Usa apenas ISSN. Emprega padrão de metadados <i>Dublin Core</i> . Metadados descritivos e administrativos preenchidos. Apresenta metadados indexados em mecanismos de busca. Licenças de uso claras. Utiliza o protocolo OAI-PMH. A proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos.
Anuários do Museu Imperial (1940 – 2022)	1,97	Usa apenas ISSN. Apresenta metadados indexados em mecanismos de busca. A proveniência é verificada apenas pelo uso de referências bibliográficas nos artigos.
Boletim Informativo do Museu de Arqueologia e Etnologia da Universidade Federal da Bahia	1,69	Usa apenas o ISSN.

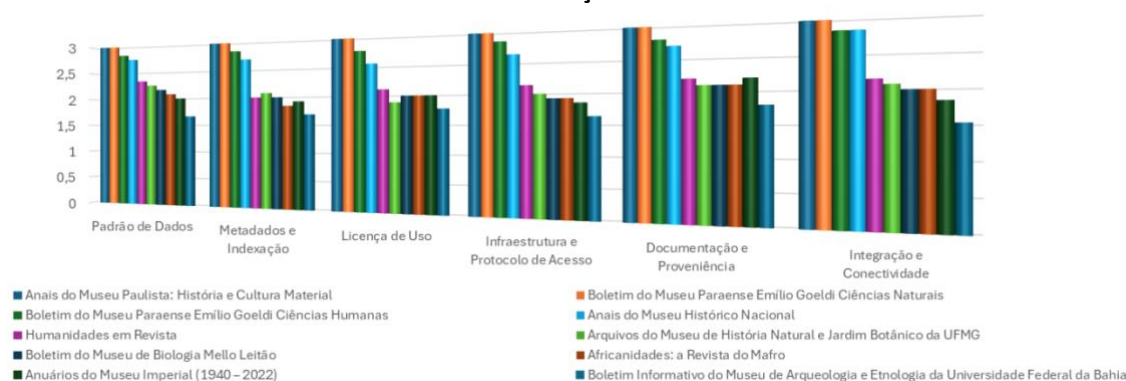
Fonte: Elaboração própria.

Todos os periódicos analisados utilizam protocolo de comunicação padronizada, aberto, gratuito e universalmente implementável (HTTP) e linguagem formal, acessível, compartilhada e amplamente aplicável (XML), viabilizando aspectos relacionados à comunicação de seus dados na rede. Destacam-se os periódicos Anais do Museu Paulista: História e Cultura Material e Boletim do Museu Paraense Emílio Goeldi Ciências Naturais,

que apresentam máxima adesão (3,0) aos princípios nas 6 categorias avaliadas. Já os periódicos Boletim do Museu Paraense Emílio Goeldi Ciências Humanas e Anais do Museu Histórico Nacional têm uma média de boa aderência também, com pontuações de 2,8 e 2,7, respectivamente. Os periódicos Humanidades em Revista, Arquivos do Museu de História Natural e Jardim Botânico da UFMG, Boletim do Museu de Biologia Mello Leitão, Africanidades: a Revista do Mafro, e Anuários do Museu Imperial (1940 – 2022) obtiveram pontuações moderadas de 2,16; 2,07; 2,03; 1,99; e 1,97, respectivamente. Finalmente, o periódico Boletim Informativo do Museu de Arqueologia e Etnologia da Universidade Federal da Bahia recebeu a avaliação mais baixa, com a pontuação de 1,69.

O Gráfico 1 sintetiza os resultados das análises, apresentando a conformidade dos periódicos científicos com as dimensões analíticas baseadas em FAIR. A análise permite identificar os periódicos que atendem melhor aos critérios de Encontrabilidade, Acessibilidade, Interoperabilidade e Reutilização, além de destacar áreas (dimensões analíticas) que necessitam de melhorias.

Gráfico 1 – Análises de cada fonte de informação frente às dimensões baseadas em FAIR



Fonte: Elaboração própria.

Em primeiro lugar, a aplicação dos princípios FAIR como critérios de avaliação permitiu distinguir os repositórios que estão mais alinhados às boas práticas dos que necessitam de melhorias. Os repositórios com pontuação máxima, como os Anais do Museu Paulista: História e Cultura Material e Boletim do Museu Paraense Emílio Goeldi Ciências Naturais, demonstraram uma forte aderência aos princípios FAIR em todas as categorias. Isso se reflete principalmente na implementação eficaz de identificadores persistentes como ORCID e DOI, bem como na utilização de licenças adequadas para garantir a acessibilidade e a reutilização dos dados (Marín-Arraiza; Heredia, 2021).

Por outro lado, repositórios com pontuação inferior, como Boletim Informativo do Museu de Arqueologia e Etnologia da Universidade Federal da Bahia e Anuários do Museu Imperial (1940 - 2022), mostraram lacunas significativas, particularmente nas categorias licença de uso, padrão de dados, e metadados e indexação. A ausência de licenças claras e a falta de metadados completos prejudicam a recuperação e a reutilização da informação, evidenciando a necessidade de atenção nesses aspectos com o objetivo de melhorar a aderência aos princípios FAIR.

A análise das métricas também destacou a importância da padronização dos dados e da infraestrutura tecnológica para a integração e conectividade dos repositórios. Os periódicos que adotam o padrão *Dublin Core*, protocolos como o OAI-PMH, e repositório digital OJS apresentam uma melhor indexação e interoperabilidade, facilitando a comunicação e a troca de dados (Gilliland, 2016) entre diferentes sistemas e repositórios. Isso é fundamental para o desenvolvimento de coleções digitais que sejam consistentes, bem-organizadas e facilmente acessíveis.

Em suma, as métricas aplicadas não apenas identificaram os repositórios que seguem as boas práticas, mas também forneceram um diagnóstico claro das áreas que precisam de melhorias. Esse diagnóstico é essencial para orientar a tomada de decisões e implementar ações corretivas que melhorem a qualidade e a eficácia dos repositórios, contribuindo para o desenvolvimento de coleções digitais mais robustas e funcionais.

Finalmente, essa abordagem permite que gestores ou profissionais da informação identifiquem quais áreas apresentam maior conformidade e desafios em relação aos princípios FAIR, facilitando a tomada de decisões e a análise comparativa entre diferentes fontes de informação endereçadas, por exemplo, a uma BD aberta e agregada.

4 CONSIDERAÇÕES FINAIS

Buscou-se, nesta pesquisa, a proposição de um modelo de avaliação diagnóstica genérico e escalável por meio do qual poder-se-ia reproduzi-lo para outros cenários que lidam com fontes de informação na internet e que demandam a apuração de suas qualidades. Considera-se, portanto, o modelo genérico uma vez que a apuração dos resultados permite avaliar quaisquer dimensões e critérios que se fizerem necessários a

um determinado contexto de pesquisa envolvendo qualidade de fontes de informação, além de levar em consideração somente os critérios FAIR que são aplicáveis a uma dimensão em particular, segundo o ponto de vista de quem está executando a avaliação; e escalável porque permite o seu avanço para outros tipos de fontes de informação, bases de dados, sistemas de informação, enfim, para além dos periódicos científicos.

É importante destacar que este estudo se restringiu a um *corpus* composto por 10 periódicos científicos. Apesar dessa limitação, o conjunto analisado revelou-se adequado e coerente com os resultados obtidos, configurando-se como um pré-teste relevante para a continuidade da pesquisa, especialmente quando se considerar a ampliação do *corpus* e a inclusão de outros tipos de fontes de informação. No que se refere à subjetividade e à dependência do julgamento do avaliador durante o processo de avaliação, recomenda-se, como perspectiva para estudos futuros, o desenvolvimento de regras explícitas para aplicação das 15 variáveis FAIR em cada uma das dimensões envolvidas no modelo de avaliação diagnóstica. Dessa forma, os procedimentos metodológicos poderão ser ainda mais alinhados aos resultados obtidos, favorecendo maior consistência e reprodutibilidade da análise.

Frente ao exposto, tanto a questão quanto o objetivo de pesquisa foram atendidos na apresentação de um modelo de avaliação diagnóstica fundamentado nos princípios FAIR para a avaliação de fontes de informação científicas para fins de seleção, extração e organização de coleções temáticas em Bibliotecas Digitais abertas e agregadas.

Espera-se que os resultados aqui apresentados possam ser úteis às comunidades de Ciência Aberta, Humanidades Digitais, Ciência da Informação e outras interessadas em usar e reusar fontes de informação com qualidade para propósitos de pesquisas diversos. Para além, os resultados serão úteis à continuidade de um estudo mais abrangente envolvendo a construção de uma Biblioteca Digital Temática ao Ibram, incluindo ações de mapeamento de fontes de informação estratégicas, especialmente documentos associados à área da Museologia, incluindo periódicos científicos, teses e dissertações, revistas especializadas, entre outras; coleta dos dados mapeados nessas fontes previamente avaliadas pelo modelo diagnóstico; e agregação de dados coletados com outros dados oriundos de outras fontes documentais.

REFERÊNCIAS

- ARAÚJO, Nelma Camêlo; FACHIN, Juliana. Evolução das fontes de informação. **Biblos Revista do Instituto de Ciências Humanas e da Informação**, [S. l.], v. 29, n. 1, 2015. Disponível em: <https://periodicos.furg.br/biblos/article/view/5463/3570>. Acesso em: 27 fev.2025.
- BARDIN, Laurence. **Análise de conteúdo**. Lisboa: Edições 70, 2016.
- GILLILAND, Anne J. Setting the Stage. *In*: BACA, Murta. (ed.). **Introduction to metadata**. Los Angeles: Getty Research Institute, 2016. Disponível em: <https://www.getty.edu/publications/intrometadata/setting-the-stage/>. Acesso em: 27 fev. 2025.
- LEMOS, Daniela Lucas da Silva; SOUZA, Renato Rocha. Ontologies for semantic annotation: proposal for an ontological multimedia reference model. **Knowledge Organization**, [S. l.], v. 51, n. 8, p. 561-581, 2024. DOI: <https://doi.org/10.5771/0943-7444-2024-8-561>. Disponível em: <https://www.imrpess.com/journal/KO/51/8/10.5771/0943-7444-2024-8-561>. Acesso em: 27 fev. 2025.
- MARÍN-ARRAIZA, Paloma; HEREDIA, Ana. FAIR PIDs: o papel da ORCID no fortalecimento dos Princípios FAIR. *In*: SALES, Luana Farias *et al.* (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p. 15-26. Disponível em: <https://www.arca.fiocruz.br/handle/icict/51281>. Acesso em: 15 fev. 2025.
- MARTINS, Dalton Lopes *et al.* Information organization and representation in digital cultural heritage in Brazil: systematic mapping of information infrastructure in digital collections for data science applications. **Journal of the Association for Information Science and Technology**, [S. l.], 2022. DOI: <https://doi.org/10.1002/asi.24650>. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/asi.24650>. Acesso em: 18 maio 2025.
- POOLE, Alex. H. The conceptual ecology of digital humanities. **Journal of Documentation**, [S. l.], v. 73, n. 1, 2017. Disponível em: https://www.researchgate.net/publication/311921946_The_Conceptual_Ecology_of_Digital_Humanities. Acesso em: 27 fev. 2025.
- UNESCO. **Recomendação da UNESCO sobre ciência aberta**. 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000379949_por?locale=en. Acesso em: 20 fev. 2025.
- SALES, Luana Sales *et al.* Um panorama histórico da iniciativa GO FAIR: da Europa para o Brasil. *In*: SALES, Luana Farias *et al.* (org.). **Princípios FAIR aplicados à gestão de dados de pesquisa**. Rio de Janeiro: Ibict, 2021. p.15-26. DOI: 10.22477/9786589167242. Disponível em: <https://www.arca.fiocruz.br/handle/icict/51281>. Acesso em: 15 fev. 2025.

SAYÃO, Luis Fernando. Afinal, o que é biblioteca digital. *In*: SALES, Luana Farias; VIOLA, Carla Maria Martellote. **Informação digital e suas diversas abordagens pela ótica de um cientista da informação**. Rio de Janeiro: IBICT, 2021. Disponível em:

<https://www.revistas.usp.br/revusp/article/view/13709/15527>. Acesso em: 27 fev. 2025.

SAYÃO, Luis Fernando; SALES, Luana Farias. Dados abertos de pesquisa: ampliando o conceito de acesso livre. *In*: SALES, Luana Farias; VIOLA, Carla Maria Martellote.

Informação digital e suas diversas abordagens pela ótica de um cientista da informação. Rio de Janeiro: IBICT, 2021. Disponível em:

<https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/611/1252>. Acesso em: 27 fev. 2025.

Wilkinson, M. D. *et al*. The FAIR guiding principles for scientific data management and stewardship. **Scientific Data**, [S. l.], v. 3, n. 1, p. 160018, 2016. Disponível em:

<https://www.nature.com/articles/sdata201618>. Acesso em: 25 fev. 2025.



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

Como citar este trabalho:

PASSOS, Heloene Gonçalves Santos; LEMOS, Daniela Lucas da Silva; MACEDO, Dirceu Flávio. Modelo de avaliação diagnóstica FAIR para fontes de informação científicas na internet endereçadas a bibliotecas digitais abertas e agregadas. *In*: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 8., Marília, SP. **Anais [...]**. Marília, SP: Universidade de Marília, 2025. DOI: <https://doi.org/10.22477/viii.widat.234>.