AKR Revista Advances on Knowledge Representation

Avaliação Diagnóstica da FAIR*ness* de fontes de informação científicas na *Web*: proposição de um modelo baseado em dimensões analíticas

Daniela Lucas da Silva Lemos[1], Heloene Gonçalves Santos Passos[2]

Resumo. *Discute-se a heterogeneidade semântica e estrutural das fontes de informação científicas e seus reflexos na representação formal do conhecimento, afetando a interoperabilidade e a proveniência de objetos digitais em ambientes abertos. A ausência de estruturas conceituais que articulem padrões de metadados, consistência descritiva e controle terminológico limita a qualificação e o compartilhamento de dados. Nesse cenário, os princípios de localização (Findable), acesso (Accessible), interoperabilidade (Interoperable) e reúso (Reusable) (acrônimo FAIR), configuram requisitos essenciais ao tratamento informacional na Web. Contudo, a literatura aponta que modelos de verificação da FAIRness operam com abordagens genéricas, pouco sensíveis às particularidades do ecossistema de publicações científicas. Este artigo propõe e aplica um Modelo de Avaliação Diagnóstica para fontes científicas heterogêneas, fundamentado em dimensões analíticas consideradas, logo, boas práticas de tratamento da informação. A metodologia aplica a análise documental e de Bardin (2016), integrando diretrizes institucionais nacionais e internacionais. Aplicado a um corpus de periódicos e repositórios, o modelo evidenciou a consolidação do Dublin Core associado ao OJS e DSpace como pressuposto técnico de interoperabilidade. Todavia, identificou lacunas como a ausência de identificadores persistentes, licenças e vocabulários controlados, comprometendo o reúso. O modelo oferece uma ferramenta útil e alinhada à qualidade da informação para humanos e máquinas.*

[1] Doutora em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG); Professora Associada do Departamento de Biblioteconomia da Universidade Federal do Espírito Santo (UFES)
Docente permanente do Programa de Pós-Graduação em Ciência da Informação (UFES);
Email: daniela.l.silva@ufes.br.
Lattes: http://lattes.cnpq.br/9280443047358807; Orcid: https://orcid.org/0000-0003-1565 7366.
[2] Mestre em Ciência da Informação pela Universidade Federal do Espírito Santo (UFES);
Técnica Administrativa do Sistema Integrado de Bibliotecas da Universidade Federal do Espírito Santo SIB/UFES)
Email: heloene.passos@ufes.br.
Lattes: http://lattes.cnpq.br/2362442830790668; Orcid: https://orcid.org/0009-0009-5603-3323.

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Palavras-chave: *Ciência Aberta, Fontes de Informação, FAIRness, Interoperabilidade, Modelo de Avaliação, Princípios FAIR.*

1. Introdução

As recomendações das Nações Unidas para a Educação, a Ciência e a Cultura, UNESCO (2021)[3], consolidaram a relevância do livre acesso aos dados de pesquisa, mapeando as vantagens e complexidades dessa transição. Esse cenário é impulsionado pela transformação digital, que acelerou a geração, o fluxo e a custódia de vastos ecossistemas de objetos digitais, demandando o estabelecimento de diretrizes técnicas para garantir a qualidade, a interoperabilidade e o reúso desses registros (Gilliland, 2016; Lemos & Coelho Junior, 2023). No entanto, o ecossistema da *Web* enfrenta a heterogeneidade semântica e estrutural das fontes de informação científicas, cujos registros incompletos ou inconsistentes dificultam a recuperação, a identificação da proveniência e o reúso e qualificação da informação (Tomaél & Alcará, 2016). Essa fragmentação desafia a representação formal do conhecimento, afetando de forma direta a interoperabilidade e a proveniência dos objetos digitais em ambientes abertos e agregados.

A ausência de estruturas conceituais robustas que articulem padrões de metadados, consistência descritiva e controle terminológico limita significativamente a qualificação dos dados e o seu efetivo reúso entre diferentes domínios do conhecimento (Zeng, 2019; Gilliland, 2016). Juty *et al.* (2020) afirmam que, no contexto da *e-Science,* quando os metadados são frágeis ou inconsistentes, a informação torna-se "invisível" para os sistemas automatizados de busca.

Como resposta global a esse cenário de dispersão, surgiram os Princípios FAIR (*Findable, Accessible, Interoperable, Reusable*), propostos por Wilkinson *et al.* (2016). Eles estabelecem diretrizes estruturais e semânticas essenciais para que os dados sejam compreensíveis tanto por humanos quanto por máquinas. Diante da alta disponibilização de dados sem o devido tratamento informacional, torna-se imperativa a avaliação desses ativos por meio de ferramentas destinadas a medir o seu grau de adequação, ou medir a FAIR*ness* (Henning *et al*., 2018). Contudo, a literatura recente aponta limitações nas práticas vigentes de verificação, uma vez que a maioria dos modelos e validadores automáticos opera por meio de abordagens genéricas (Candela *et al.*, 2024; Bahim *et al.,* 2020; Devaraju & Huber, 2020; Devaraju *et al.*, 2021; Emonet *et al.,* 2022; Gonzalez *et al.*, 2022). Esses sistemas falham por não demonstrar sensibilidade às especificidades e dinâmicas próprias de ecossistemas complexos, como o de publicações científicas, representado, neste estudo, por periódicos e repositórios institucionais (RI).

Diante dessa lacuna, *o objetivo deste artigo é propor e apresentar um modelo de avaliação diagnóstica FAIR fundamentado em dimensões analíticas para fontes de*

[3] https://unesdoc.unesco.org/ark:/48223/pf0000379949_por

informação científicas na Web. A investigação justifica-se por fornecer um arcabouço teórico-metodológico e conceitual aplicável à curadoria de objetos digitais, servindo como suporte a gestores e profissionais da informação na auditoria dos acervos sob sua custódia.

Para tanto, o artigo está estruturado em sete seções. A Seção 2 contextualiza os desafios da avaliação de fontes de informação científicas na *Web* a partir de ferramentas FAIR, enquanto a Seção 3 fundamenta as dimensões analíticas que fundamentam a proposta do modelo de avaliação diagnóstica. O percurso metodológico é detalhado na Seção 4, servindo de base para a apresentação do Modelo de Avaliação Diagnóstica na Seção 5, resultado do presente artigo. Por fim, a Seção 6 traz a discussão sobre as contribuições teórico-metodológicas e de aplicação do modelo, seguida pelas considerações finais na Seção 7.

2. Fontes de informação científicas na *Web* e desafios para avaliação FAIR

Os periódicos científicos e os repositórios institucionais constituem os pilares da difusão do conhecimento acadêmico (Pereira *et al.*, 2020; Moraes & Sales, 2019). Embora compartilhem o propósito de dar visibilidade à ciência, configuram-se como fontes heterogêneas em suas naturezas, fluxos, arquiteturas, a exemplo dos *softwares Open Journal Systems* (OJS), e DSpace, e práticas descritivas. Enquanto o periódico atua como veículo de comunicação formal focado no prestígio e indexação internacional (Pereira *et al.*, 2020), o repositório dedica-se à salvaguarda da memória e preservação de longo prazo (Moraes & Sales, 2019). Essa diferença de ecossistemas gera variações em formatos, esquemas de metadados e critérios de preenchimento que desafiam a recuperação da informação e a interoperabilidade global (Kamm *et al*., 2023).

Essa assimetria estrutural estende-se aos mecanismos de identificação persistente. No ecossistema de periódicos, consolida-se o uso do DOI (*Digital Object Identifier*), o qual permite rastrear o uso e o impacto da pesquisa ao longo do tempo. Em contrapartida, os repositórios baseados no *software* DSpace apoiam-se prevalentemente no sistema Handle. Embora ambos realizem a resolução de objetos digitais, o DOI diferencia-se por uma governança global que exige metadados estruturados e ricos, ampliando a rastreabilidade e a integração automatizada com os sistemas formais de comunicação (Sayão, 2007; Silva & Rodrigues, 2021). No contexto do *Open Access*, garantir a qualidade dessa estrutura computacional é vital para a real apropriação social do conhecimento (Araújo *et al.*, 2021).

As infraestruturas distintas desses dois ecossistemas revelam que a avaliação do nível de adequação das fontes de informação aos Princípios FAIR não pode ser homogênea, genérica, ou baseada em verificações binárias de "presença ou ausência" de dados. No entanto, ao transitar para o campo prático da auditoria de qualidade, a literatura aponta que os instrumentos mais frequentes aplicam questionários, listas de verificação e

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

validadores semiautomáticos (González *et al*., 2022; Candela *et al*., 2024), ignorando tais especificidades. Em função disso, diferentes ferramentas aplicadas a um mesmo objeto digital geram pontuações divergentes, uma vez que os validadores exploram metadados que não seguem esquemas universais e são sensíveis a domínios específicos (Emonet *et al*., 2022). Outro obstáculo reside na disparidade das fórmulas de cálculo da conformidade, visto que as análises FAIR apresentam pesos e graus de relevância variáveis de acordo com a ferramenta aplicada.

Embora haja convergência no uso de indicadores baseados no FAIR, persistem assimetrias na granularidade dos critérios e baixa reprodutibilidade dos resultados devido à opacidade metodológica na definição das regras de avaliação (González *et al.*, 2022). Ademais, nota-se que esses instrumentos são majoritariamente orientados a dados de pesquisa estruturados, falhando em capturar a heterogeneidade e a complexidade semântica de fontes de informação científicas e curatoriais mais complexas. Diante dessas limitações, evidencia-se a necessidade de modelos diagnósticos que, além de alinhados às diretrizes internacionais, sejam sensíveis aos ecossistemas avaliados.

3. Dimensões analíticas para avaliação diagnóstica FAIR

A segmentação da avaliação em dimensões realiza a mediação conceitual entre os requisitos técnicos do FAIR e os fundamentos da Ciência da Informação. Essa escolha responde a uma lacuna crítica na literatura: a inadequação de abordagens genéricas ou puramente automatizadas frente a ecossistemas informacionais heterogêneos (González *et al.*, 2022; Emonet *et al*., 2022). Como preconizam os estudos sobre arquitetura e governança de objetos digitais (Vidotti *et al*., 2021), os Princípios FAIR funcionam como diretrizes macro, exigindo um arcabouço intermediário para sua operacionalização. Desse modo, essas dimensões direcionam a auditoria diagnóstica para demandas de tratamento da informação, consistência descritiva e controle terminológico (Chalhub, 2012; Guandalini *et al*., 2018). Essa estrutura garante que a aferição da FAIRness ocorra de forma contextualizada, sensível às particularidades curatoriais e de infraestrutura de periódicos e repositórios institucionais.

3.1 Padrão de dados

A dimensão Padrão de Dados avalia a conformidade das fontes de informação científica em relação às estruturas de representação do conhecimento. Seu escopo teórico e prático abrange a análise integrada de quatro pilares fundamentais da arquitetura da informação digital: os padrões de metadados, a estrutura de sintaxe, as regras de conteúdo e a comunicação dos dados (Gilliland, 2016). O foco reside em garantir que a informação possua uma estrutura legível por máquina e inteligível por humanos, assegurando

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

consistência semântica e representação precisa do conhecimento para além de meros depósitos textuais.

No nível estrutural e de comunicação, a dimensão examina a adoção de padrões consolidados de metadados, como o *Dublin Core*, e sua expressão em sintaxes de intercâmbio voltadas à *Web* Semântica, a exemplo do *Resource Description Framework* (RDF) / *eXtensible Markup Language* (XML). Essa engrenagem técnica, contudo, depende diretamente da qualidade das plataformas escolhidas para a gestão editorial e arquivística. A constatação, por exemplo, do uso de plataformas genéricas como o *WordPress* em detrimento do (OJS) para a publicação de periódicos evidencia um gargalo crítico. A escolha do *WordPress* compromete severamente a interoperabilidade automatizada, visto que o OJS provê uma infraestrutura nativa e específica para a comunicação científica, com foco estrutural em preservação de longo prazo, indexação padronizada e enriquecimento de metadados (Marques *et al*., 2025).

No que tange às regras de conteúdo e aos valores dos dados, o diagnóstico avalia o emprego de códigos de catalogação e o controle terminológico por meio de vocabulários controlados. Como ferramenta semântica aplicada à representação e à recuperação documentária, o vocabulário controlado atua como um elemento mediador essencial entre as linguagens do indexador, do documento e do usuário. É por meio dessa mediação que se garante a consistência descritiva e a precisão na recuperação da informação, conferindo ao usuário maior autonomia e eficiência no processo de busca (Cruz *et al*., 2022).

Ao analisar essas estruturas sob a perspectiva dos Sistemas de Organização do Conhecimento (SOCs), Zeng (2008) as categoriza em quatro grandes grupos funcionais: as listas de termos (como glossários e anéis de sinônimos); os modelos de metadados (listas de autoridade e dicionários geográficos); as estruturas de classificação e categorização (cabeçalhos de assunto, taxonomias e esquemas de classificação); e os modelos de relacionamentos complexos (tesauros, redes semânticas e ontologias).

Independentemente da nomenclatura adotada ou de sua complexidade morfológica, tais instrumentos diferenciam-se em suas arquiteturas internas, mas se complementam no objetivo comum de mitigar a ambiguidade da linguagem natural, favorecendo o tratamento, a governança e a representação da informação em ambientes digitais.

Embora regras tradicionais de conteúdo, como *o Anglo American Cataloging Rules, 2nd* (AACR2), ainda sejam aplicadas com consistência nas instituições de salvaguarda, o cenário atual da *e*-Ciência demanda ações de tratamento da informação em conformidade com as novas exigências das fontes de informação científicas na *Web*. A literatura aponta para uma transição para regras mais dinâmicas, relacionais e semanticamente orientadas, como as propostas pelo modelo RDA (*Resource Description and Access*), configurando um passo indispensável em direção à integração dos metadados na *Web* de Dados (Holanda & Lourenço, 2021; Santos & Arakaki, 2022).

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Assim, ao cruzar estrutura, valor, conteúdo e comunicação, esta dimensão permite aferir se os ecossistemas avaliados realizam um tratamento da informação plenamente legível por máquina, superando o preenchimento meramente formal e alcançando a efetividade semântica exigida pelos Princípios FAIR.

3.2 Indexação a partir de metadados

A indexação determina diretamente o potencial de recuperação e a descoberta de um objeto digital no ecossistema da ciência aberta. Esta dimensão investiga a qualidade, a abrangência e a padronização dos metadados preenchidos na fonte, partindo do princípio de que a existência isolada do dado não garante sua utilidade (Tomaél & Alcará, 2016). A consistência nesta dimensão impacta de forma imediata a localização, a precisão na recuperação da informação e a visibilidade dos trabalhos científicos na *Web*, estabelecendo o elo entre a produção local e as redes globais de conhecimento.

No plano da infraestrutura de comunicação, a indexação plena apoia-se na convergência técnica entre a sintaxe XML, o padrão *Dublin Core* Qualificado e o protocolo *Open Archives Initiative Protocol for Metadata Harvesting (*OAI-PMH) (Marques *et al*., 2025; Santos *et al*., 2023). Essa tríade tecnológica assegura que os metadados gerados pelas plataformas de periódicos e repositórios não fiquem confinados em silos institucionais, fornecendo os requisitos operacionais indispensáveis para que provedores de serviços e grandes indexadores globais (*harvesters*) realizem a coleta automatizada e sistemática das informações.

Contudo, a viabilidade técnica do protocolo de comunicação exige qualidade rigorosa nos níveis de conteúdo e valor dos metadados preenchidos. Nesse aspecto, considera-se que, por um lado, a aplicação das regras clássicas de catalogação AACR2 favorece a padronização formal e a consequente recuperação da informação pelos usuários finais (Machado & Pereira, 2017). Por outro lado, a ausência recorrente de vocabulários controlados para a representação temática representa um gargalo qualitativo. Essa lacuna pode acarretar um déficit semântico que compromete a indexação conceitual e fragiliza a precisão na recuperação da informação por parte dos agregadores automatizados (Berners-Lee, 2006).

Em suma, metadados ricos, padronizados e compatíveis com os mecanismos de busca e *harvesters* se apresentam como determinantes para promover a circulação da literatura científica. A análise desta dimensão implica que a visibilidade e o impacto de longo prazo dos objetos digitais dependem da superação desse descompasso entre a consistência formal da catalogação e o refino semântico das linguagens de indexação.

3.3 Licença de uso

O reúso efetivo da informação científica pressupõe, fundamentalmente, segurança jurídica. Esta dimensão aborda o aspecto regulatório e legal do compartilhamento de dados e metadados, partindo da premissa de que a dimensão jurídica é um pilar indissociável da própria viabilidade do reúso. Demonstra-se que a ausência de uma licença de uso clara e explícita não significa que o dado ou o documento esteja livre para ser livremente apropriado pela sociedade; pelo contrário, a omissão regulatória gera severa incerteza jurídica e bloqueia preventivamente o reúso por parte de pesquisadores e instituições (Jacobsen *et al*., 2020).

Para sanar esse gargalo, o modelo avalia a presença de licenças padronizadas internacionalmente e reconhecidas pela área (como as atribuições *Creative Commons*) declaradas não apenas nas interfaces textuais voltadas aos usuários humanos. O diferencial analítico e metodológico desta abordagem reside em avaliar se tais licenças estejam formalmente codificadas e estruturadas diretamente nos metadados da fonte, tornando-as legíveis por máquina e prontas para processamento automatizado.

Falhas de descrição ignoram a necessidade de uma legibilidade jurídica operada por agentes computacionais, os quais dependem de campos explícitos (como o *dc.rights.uri* no padrão Dublin Core) para indexar, agregar e autorizar o reaproveitamento ético e legal da informação em larga escala (Wilkinson *et al*., 2016).

Dessa forma, ao vincular a governança legal à capacidade de processamento das máquinas, a dimensão evidencia que a abertura da ciência não se consolida apenas com a gratuidade do acesso, mas com a transparência e a precisão dos direitos de uso expressos na arquitetura dos metadados.

3.4 Infraestrutura tecnológica e de comunicação

Esta dimensão analisa minuciosamente a configuração e a maturidade das arquiteturas de *software* (a exemplo do OJS e do DSpace) que sustentam o ecossistema científico, sob a premissa de que a infraestrutura não deve ser reduzida a um mero suporte técnico passivo ou a um depósito de arquivos digitais. Pelo contrário, as plataformas computacionais, como o OJS no âmbito dos periódicos e o DSpace no contexto dos repositórios institucionais, constituem condições ativas, estruturantes e determinantes para a garantia da acessibilidade, da interoperabilidade e da viabilidade da ciência orientada por dados na *Web* Semântica (Marques *et al.*, 2025; Oliveira, 2023; Dimitrova *et al.*, 2020).

Para além do armazenamento estático de documentos em formato PDF, exige-se que essas arquiteturas atuem como nós dinâmicos e comunicantes em redes globais de informação. Diante disso, a avaliação desta dimensão visa contemplar a compatibilidade técnica e a conformidade das plataformas com o protocolo OAI-PMH, estabelecido como

o mecanismo padrão para a interoperabilidade e a coleta automatizada de metadados na comunidade científica internacional. Complementarmente, audita-se a adoção do protocolo HTTPS, o qual atesta a conformidade com os requisitos contemporâneos de segurança e integridade digital (Tariq *et al.*, 2024).

A relevância de analisar o ecossistema sob esse prisma reside no fato de que o isolamento computacional inviabiliza a aplicação prática dos Princípios FAIR. Desse modo, a convergência entre protocolos de comunicação abertos e camadas seguras de transferência de dados reflete uma infraestrutura robusta, alinhada às boas práticas internacionais de comunicação científica e curadoria digital (Santos *et al*., 2023; Marques *et al*., 2025)."

3.5 Documentação e proveniência

A confiança na ciência aberta e a reprodutibilidade dos resultados dependem intrinsecamente da rastreabilidade dos registros. Esta dimensão explora a documentação da fonte e o histórico de dados e metadados, mapeando elementos como autoria, criação, modificações no ciclo de vida, responsabilidade institucional e políticas de preservação digital (Marín-Arraiza & Heredia, 2021). Sob essa ótica, a proveniência robusta não é um mero formalismo, mas a base conceitual e técnica para validar a autenticidade e promover o reúso qualificado do conhecimento na Web.

No ambiente da e-Ciência, a consolidação dessa rastreabilidade exige que as plataformas utilizem identificadores persistentes de forma interconectada. A integração entre o DOI (para o objeto), o ORCID (para pesquisadores) e o ROR (para a filiação institucional) é vital para o ecossistema informacional: ela garante a integridade da autoria, estabelece a precedência temporal do objeto e subsidia agências de fomento no monitoramento e na auditoria dos resultados de pesquisas financiadas (Jacobsen *et al*., 2020).

Esta dimensão confronta as práticas de descrição tradicionais e as demandas da *Web* de Dados. O preenchimento dos metadados seguindo padrões consolidados de catalogação, como o AACR2, no contexto atual, é insuficiente para qualificar o dado para o reúso pleno e automatizado por agentes computacionais (Marín-Arraiza & Heredia, 2021). A negligência com metadados mais profundos de proveniência gera severos prejuízos à comunicação científica. A inconsistência ou a omissão na descrição detalhada da linhagem dos recursos transforma o dado científico em um objeto fora de contexto técnico e jurídico, desprovido de suas amarras de validação (Henning *et al*., 2018). Como consequência direta desse esvaziamento contextual, cria-se um cenário de apagamento que dificulta atribuir o devido e merecido crédito a quem produziu, coletou ou editou originalmente o recurso informacional.

Assim, ao analisar de forma integrada a documentação e a proveniência, esta dimensão demonstra que a sustentação da confiança na Ciência Aberta exige a transição

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

de uma catalogação puramente descritiva para uma governança de metadados focada na rastreabilidade histórica e na integridade contextual dos objetos digitais.

3.6 Integração e conectividade

Esta dimensão analisa o grau de interligação de uma fonte com o ecossistema global de informação por meio do uso estratégico e sistemático de identificadores persistentes e padronizados, tais como o DOI (para objetos digitais), o ORCID (para identificação de pesquisadores), o *International Standard Serial Number* (ISSN) (para publicações seriadas) e o *Research Organization Registry* (ROR) (para afiliações institucionais). Esses identificadores moldam a infraestrutura de conexão essencial para garantir a preservação de longo prazo, a interligação de dados e metadados e a integração nativa dos recursos no paradigma *Linked Open Data* (Dados Abertos Interconectados) (Jacobsen *et al.*, 2020).

Contudo, a literatura aponta que a consolidação desse ecossistema interconectável ainda enfrenta desafios estruturais. A negligência em relação ao princípio FAIR I2 (que preconiza o uso de vocabulários controlados ou ontologias formais), combinada à ausência de implementações baseadas em sintaxe RDF (*Resource Description Framework*), indica que muitas fontes de informação científica na Web ainda operam sob o paradigma da Web de documentos tradicional. Esse cenário as distancia da interoperabilidade sistêmica e automatizada promovida pela Web Semântica (Alves *et al.*, 2022; Guandalini & Santos, 2018), impossibilitando o alcance dos estágios mais avançados de *Linked Data* (Berners-Lee, 2006) e divergindo das recomendações de Gilliland (2016) quanto ao rigor necessário nos padrões de valores de dados.

Sob essa ótica teórica, o modelo proposto fundamenta-se no princípio de que a conectividade depende de uma cadeia de conformidades interdependentes. A carência de vocabulários controlados, somada à omissão de Licenças de Uso expressas de forma padronizada nos metadados, configura barreiras severas que fragmentam o fluxo informacional e paralisam a conectividade automática entre sistemas de descoberta (Hahnel & Valen, 2020; Caplan & Arms, 1999).

Dessa forma, a análise da dimensão Integração e Conectividade visa adequar as fontes de informação às exigências da e-Ciência por meio da adoção de tecnologias e padrões semânticos capazes de tecer uma rede robusta, duradoura e legível por máquina de conjuntos de dados e metadados integrados.

4. Percurso metodológico

Para viabilizar a construção e a validação do modelo proposto, este estudo delineou um percurso metodológico estruturado em quatro etapas complementares, caracterizando-se como uma pesquisa aplicada, de abordagem quali-quantitativa, com

objetivos exploratórios e descritivos. A vertente qualitativa manifesta-se na análise compreensiva e interpretativa da literatura e das diretrizes institucionais. Por sua vez, o caráter quantitativo expressa-se na atribuição de escores numéricos (de 0 a 3) para mensurar o grau de aderência das fontes às regras e critérios estabelecidos.

O levantamento bibliográfico foi realizado por meio do Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), abrangendo as bases de dados internacionais Web of Science e Scopus, além da Base de Dados em Ciência da Informação (BRAPCI) e do Google Acadêmico, com delimitação nos campos da Ciência da Informação e da Ciência da Computação. O *corpus* documental recuperado nessas bases foi submetido de forma sistemática à Análise de Conteúdo, seguindo o aporte metodológico de Bardin (2016), o que permitiu a codificação, categorização e inferência dos dados que subsidiaram a construção do modelo proposto.

A primeira etapa do percurso metodológico consistiu na aplicação sistemática da Análise de Conteúdo (Bardin, 2016). Esse procedimento desdobrou-se em suas três fases fundamentais:

Fase 1 – Pré-análise: compreendeu-se a organização do *corpus* e o contato inicial com o conjunto de documentos recuperados na pesquisa bibliográfica. Realizou-se a leitura flutuante dos 59 documentos selecionados para sedimentar a familiarização com a literatura e validar sua aderência ao escopo do estudo. Essa etapa assegurou o cumprimento estrito das regras de exaustividade, homogeneidade e pertinência preconizadas por Bardin (2016).

Fase 2 – Exploração do material: consistiu nas operações de codificação e desmembramento analítico do material selecionado. Adotaram-se como unidades de registro os fragmentos textuais focados em critérios, métricas ou requisitos de avaliação de objetos digitais e seus metadados, utilizando os parágrafos circundantes como unidades de contexto. A categorização guiou-se pelo critério semântico por exclusão mútua, agrupando as unidades de registro conforme as convergências conceituais emergentes.

Fase 3 – Tratamento dos resultados, inferência e interpretação: os dados categorizados foram submetidos a análises reflexivas e confrontados com o referencial teórico adotado. A frequência e a densidade conceitual das categorias permitiram a inferência de padrões, viabilizando a conversão dos dados brutos em elementos reflexivos. Esse processo forneceu o arcabouço necessário para a modelagem e sustentação das dimensões analíticas propostas.

Este procedimento garantiu rigor científico ao desenho da pesquisa, demonstrando que as seis dimensões analíticas propostas não emergiram de forma intuitiva, mas sim como resultado de um processo indutivo de extração, categorização e interpretação teórica da literatura especializada. A aplicação prática desse fluxo operacional viabilizou a organização dos fragmentos textuais por aproximação conceitual, identificando núcleos

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

temáticos recorrentes na literatura, a exemplo de interoperabilidade, metadados, identificadores persistentes, o protocolo *Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH) e o licenciamento. O isolamento e a validação sistemática desses elementos subsidiaram a consolidação das seis dimensões analíticas que integram o modelo de avaliação proposto.

A segunda etapa metodológica consistiu na delimitação de um *corpus* empírico composto por 14 periódicos científicos e cinco repositórios institucionais *stricto sensu* em Museologia. A escolha desse domínio justifica-se pela complexidade de um ecossistema marcado pela heterogeneidade de objetos e pela necessidade de salvaguarda da memória digital (Lemos *et al.*, 2023). A forte relação da área com a cultura digital, a preservação patrimonial e a interoperabilidade de seus ativos de informação torna-a o cenário ideal para validar o modelo proposto e aferir a conformidade aos Princípios FAIR.

Os 14 periódicos que compõem o *corpus* desta pesquisa são: Anais do Museu Paulista, Anais do Museu Histórico Nacional, Arquivos do Museu de História Natural e Jardim Botânico da UFMG, Boletim do Museu de Biologia Mello Leitão, Boletim do Museu Paraense Emílio Goeldi. Ciências Naturais, Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas, Revista do Museu de Arqueologia e Etnologia MAE/UFBA, Cadernos do CEOM, Hawo, Humanidades em Revista, Museologia & Interdisciplinaridade, Patrimônio e Memória, Revista Fontes Documentais e Ventilando Acervos.

Já os cinco repositórios institucionais (RI) são: RI da Universidade Federal da Bahia (UFBA), RI da Universidade Federal de Pernambuco (UFPE), RI da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), RI da Universidade Federal do Rio Grande do Sul (UFRGS) e RI da Universidade de São Paulo (USP).

A terceira etapa metodológica concentrou-se na formulação dos critérios e das regras de avaliação, fundamentadas em regras implementadas pelo especialista do domínio, para cada uma das seis dimensões analíticas do modelo de avaliação. Essas diretrizes foram estruturadas a partir de padrões consolidados da *Web* de Dados, a exemplo do esquema de metadados *Dublin Core*, do protocolo OAI-PMH, *Digital Object Identifier* (DOI) e *Open Researcher and Contributor* ID (ORCID), de licenças *Creative Commons* (CC) e dos princípios de *Linked Open Data* (LOD). Para garantir o alinhamento operacional dessas regras, os parâmetros métricos do projeto *FAIRsFAIR* foram adotados como referência de calibração, evitando transposições diretas e assegurando adaptabilidade a diferentes ecossistemas. Essa modelagem normativa foi respaldada pelas diretrizes internacionais da GO FAIR, do *Committee on Data* (CODATA) e da *Research Data Alliance* (RDA), sendo refinada por meio de uma imersão técnica na documentação de metadados dos periódicos e repositórios institucionais e pela aplicação de pré-testes que validaram a consistência métrica do método proposto.

A quarta e última etapa compreendeu a operacionalização do Modelo de Avaliação Diagnóstica. Para a operacionalização do modelo, a extração e a organização dos metadados foram estruturadas em uma planilha de registro das verificações (conforme detalhado na Tabela 1). Nessa matriz de dados, as linhas representam as fontes de informação científicas auditadas, enquanto as colunas sistematizam o mapeamento das seis dimensões analíticas do modelo.

O arranjo desses metadados desdobrou-se em 13 critérios FAIR adaptados de Wilkinson *et al*. (2016) (F1, F2, F3, F4, A1.1, A1.2, A2, I1, I2, I3, R1.1, R1.2 e R1.3), distribuídos conforme sua aderência conceitual. Para fins de precisão na coleta, as categorias genéricas A1 e R1 foram desfragmentadas, organizando os metadados diretamente a partir de suas subdivisões específicas. A cada elemento extraído aplicaram-se regras de verificação previamente definidas, permitindo mensurar e tabular o grau de aderência das fontes. O recorte ilustrado na Tabela 1 exemplifica a organização lógica e as pontuações atribuídas aos metadados da dimensão Padrão de Dados e Indexação a partir de Metadados da fonte repositórios institucionais.

Tabela 1 - Pontuação atribuída às dimensões Padrão de Dados e Indexação a partir de Metadados da fonte repositórios institucionais

Fontes de informação na Internet	Padrão de Dados												Indexação a Partir de (meta)dados											
	F				A			I			R						A			I			R	
	F1	F2	F3	F4	A1.1	A1.2	A2	I1	I2	I3	R1.1	R1.2	R1.3	F1	F2	F3	F4	A1.1	A1.2	A2	I1	I2	I3	R1.1	R1.2	R1.3
BA - UFBA	2	2	N/A	N/A	2	N/A	N/A	1	0	N/A	0	N/A	1	2	1	2	2	3	N/A	N/A	3	0	2	0	N/A	1
PE - UFPE	2	1	N/A	N/A	2	N/A	N/A	1	0	N/A	2	N/A	2	2	1	2	2	3	N/A	N/A	3	0	1	0	N/A	2
RJ - UNIRIO	1	1	N/A	N/A	2	N/A	N/A	1	0	N/A	0	N/A	0	1	1	1	1	1	N/A	N/A	3	0	1	0	N/A	0
RS - UFRGS	2	1	N/A	N/A	2	N/A	N/A	1	0	N/A	1	N/A	1	2	1	2	0	3	N/A	N/A	3	0	1	0	N/A	1
SP - USP	3	1	N/A	N/A	3	N/A	N/A	1	0	N/A	0	N/A	1	3	1	3	1	3	N/A	N/A	3	0	1	0	N/A	1
CRITÉRIOS NÃO APLICÁVEIS NA DIMENSÃO																										

Fonte: Elaborada pelas autoras (2026).

A mensuração foi realizada por meio de uma escala de maturidade composta por quatro níveis (0 a 3) e pela opção Não Aplicável (N/A), utilizada quando determinado critério não era pertinente às características da fonte avaliada. Os níveis indicam:

- Nível 0 (Inexistente/Não aderente): ausência total do elemento avaliado ou desconformidade absoluta com os padrões exigidos;
- Nível 1 (Inicial/Baixa aderência): presença incipiente, fragmentada ou puramente textual do elemento, carecendo de automação ou padronização;
- Nível 2 (Intermediário/Aderência parcial): elemento presente, mas com limitações na legibilidade por máquina ou na integração sistêmica;

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

- Nível 3 (Consolidado/Alta aderência): plena conformidade técnica, semântica e estrutural, com dados integrados, automatizados e legíveis por máquina.

Visando mitigar a subjetividade e assegurar a consistência da auditoria, a atribuição dos níveis de conformidade pautou-se em regras operacionais estritas, associadas à presença, à disponibilidade e à resolubilidade de elementos estruturais, tais como identificadores persistentes, vocabulários controlados, protocolos de interoperabilidade e licenças de uso. A coleta e a avaliação empírica dos dados foram conduzidas de forma manual entre 1º e 28 de março de 2026. Na sequência, os dados obtidos foram processados de maneira automatizada para o cálculo das pontuações finais, com base na aplicação da Fórmula 1 apresentada a seguir.

$$i = \sum_{j=1}^{m=6} \left(\left(\sum_{k=1}^{f(j)} Valor_{i,j,k} \right) / n \right)$$

Pontuação por Dimensão Analítica (1)

Na Fórmula 1, “i” representa a fonte de informação, “j” indica as 6 dimensões analíticas, “k” corresponde ao critério específico avaliado na dimensão “j” (variando conforme o ecossistema da fonte por meio de “f(j)”), e “n” determina o total de critérios válidos aplicados. A pontuação de cada dimensão analítica é verificada a partir da média dos critérios válidos aplicados à fonte avaliada, desconsiderando-se aqueles classificados como Não Aplicáveis (N/A). Já a Pontuação Geral (FAIR*ness*) é determinada pela média das seis dimensões analíticas, constituindo o principal indicador para ordenação das fontes de informação científicas e geração de análises comparativas.

Este estudo centrou-se na auditoria rigorosa de metadados estruturais e descritivos, elementos determinantes para a inteligibilidade, a recuperação e o processamento automatizado de objetos digitais por sistemas computacionais. O exame exaustivo dos ecossistemas informacionais investigados, aliado ao cruzamento sistemático das variáveis coletadas, permitiu rastrear assimetrias precisas e diagnosticar empiricamente os níveis de maturidade das fontes de informação científicas analisadas. Os resultados gerados por esse diagnóstico, consolidados pela validação da arquitetura conceitual do modelo, oferecem uma base empírica que fundamenta as discussões aprofundadas na seção seguinte.

5. Resultados da pesquisa

Esta seção apresenta os resultados da investigação a partir da estrutura e validação empírica do Modelo de Avaliação Diagnóstica. Concebido como uma ferramenta de

auditoria conceitual e técnica, o modelo foi projetado sob uma abordagem gradativa e não binária da conformidade FAIR, permitindo fornecer um diagnóstico qualiquantitativo preciso sobre a localização, a acessibilidade, a interoperabilidade e o reúso de ativos digitais. Sua arquitetura avalia o desempenho informacional de fontes de informação científicas ao apontar os gargalos que impedem a livre circulação dos dados.

Cabe ressaltar que a engrenagem metodológica do modelo afere exclusivamente aspectos estruturais, semânticos e infraestruturais dos metadados. Desse modo, a avaliação concentra-se estritamente na adequação técnica aos Princípios FAIR, não contemplando qualquer análise sobre o mérito científico das publicações indexadas pelas fontes.

5.1 Modelo de Avaliação Diagnóstica FAIR

O modelo (Figura 1) apresenta os Princípios FAIR ao centro, a partir dos quais se fundamenta a avaliação das fontes de informação científicas. As seis dimensões analíticas que constituem a arquitetura do modelo (posicionadas de forma adjacente aos princípios) funcionam como os pilares estruturais dessa avaliação. Mapeados cientificamente por meio da aplicação sistemática da Análise de Conteúdo (Bardin, 2016) sobre a literatura especializada, esses eixos englobam: padrão de dados; indexação a partir de metadados; licença de uso; infraestrutura tecnológica e de comunicação; documentação e proveniência; e integração e conectividade, conforme fundamentados na Seção 3. Longe de serem categorias isoladas, esses pilares traduzem de forma operacional as quatro diretrizes fundamentais de localização, acessibilidade, interoperabilidade e reúso, consolidando uma avaliação multidimensional e integrada. A seguir, a Figura 1 apresenta conceitualmente o Modelo de Avaliação Diagnóstica.

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Figura 1- Modelo de Avaliação Diagnóstica



Fonte: Elaborado pelos autores (2026).

O diferencial proposto por este arranjo reside na sua capacidade de superar avaliações puramente generalistas ou binárias, comuns em outras ferramentas da literatura (González *et al.*, 2022; Candela *et al.*, 2024). Ao propor uma auditoria altamente granular, o modelo viabiliza uma verificação aprofundada da aderência dos objetos digitais e de seus respectivos metadados ao ecossistema FAIR, respeitando as especificidades de fontes heterogêneas. Para detalhar essa aplicação prática, o Quadro 1 apresenta, de forma sistemática, o foco de avaliação de cada dimensão analítica e as suas respectivas evidências verificáveis.

Quadro 1 – Síntese Estrutural das dimensões analíticas

Dimensão Analítica	Foco da Avaliação	Critérios e Evidências Verificáveis
Padrão de dados	Estruturas de representação e intercâmbio semântico.	Uso de esquemas de metadados estruturados (ex: Dublin Core); codificação em sintaxes *Web* (RDF/XML, JSON); adoção de vocabulários controlados.

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Indexação a partir de metadados	Potencial de recuperação e visibilidade sistêmica.	Riqueza do preenchimento dos campos descritivos; compatibilidade e exposição de dados para indexadores externos e *harvesters*.
Licença de uso	Segurança jurídica para o compartilhamento.	Presença de licenças de uso declaradas (ex: *Creative Commons*); inserção dessas licenças em formato legível por máquina nos metadados.
Infraestrutura Tecnológica e de Comunicação	Capacidade de transporte e coleta de dados.	Configuração das plataformas OJS e DSpace; conformidade técnica com o protocolo OAI-PMH; disponibilização de APIs de acesso.
Documentação e Proveniência	Confiabilidade, autoria e histórico dos objetos.	Histórico de modificações; atribuição clara de responsabilidade institucional; existência de políticas formais de preservação digital.
Integração e conectividade	Interligação e persistência em rede.	Atribuição de identificadores persistentes globais (DOI, ORCID, ISSN, ROR); conexão com grafos de dados do *Linked Open Data*.

Fonte: Elaborada pelas autoras (2026).

A camada subsequente apresenta os critérios e regras de avaliação respaldados pelas diretrizes internacionais da GO FAIR e pela imersão na literatura que trata das boas práticas do tratamento da informação. Para cada dimensão analítica foi elaborado um conjunto de critérios e regras através da análise minuciosa dos ecossistemas que essas fontes estão inseridas. O Quadro 2 apresenta o total de critérios e regras elaboradas por dimensão analítica.

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Quadro 2 - Total de critérios e regras por dimensão analítica

Dimensão Analítica	Critérios	Regras
Padrão de dados	14	56
Indexação a partir de metadados	20	80
Licença de uso	10	40
Infraestrutura Tecnológica e de Comunicação	10	40
Documentação e Proveniência	16	64
Integração e conectividade	18	72

Fonte: Elaborada pelas autoras (2026).

O conjunto completo dos critérios (88) e regras (352) encontra-se disponível para consulta no Repositório Institucional da Universidade Federal do Espírito Santo (Passos, 2026). Esse conjunto elaborado através da imersão nos ecossistemas e validados pela literatura constitui o diferencial proposto na pesquisa: a avaliação única de cada fonte de informação científica. Ressalta-se que a avaliação não tem o objetivo de ser punitiva ou puramente binária; pelo contrário, o modelo apoia-se em uma lógica de aderência gradativa. Dessa forma, as regras funcionam como um roteiro de melhoria contínua, permitindo que gestores de repositórios e editores de periódicos identifiquem gargalos específicos e ajustem seus fluxos editoriais e computacionais à medida que atendem, progressivamente, às exigências normativas do ecossistema FAIR.

5.1 Aplicação do modelo em periódicos e repositórios

A validação empírica do modelo ocorreu por meio de sua aplicação diagnóstica no *corpus* selecionado, permitindo testar a sensibilidade da ferramenta diante das particularidades de cada ecossistema informacional. Para tanto, a auditoria consistiu no exame minucioso nos sites dos periódicos e repositórios institucionais, com foco na análise da documentação pública orientadora para submissão de artigos ou depósito de teses e dissertações, combinada à verificação direta da presença e consistência dos elementos estruturais nos metadados, conforme estabelecido pelas regras operacionais.

5.2 Aplicação em periódicos científicos

A análise dos periódicos revelou a predominância do *Open Journal Systems* (OJS) em 85% do *corpus*, seguido pelo *WordPress*, que responde pelos 15% restantes das fontes

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

visitadas. A avaliação demonstrou uma sólida maturidade estrutural no que tange ao uso do padrão *Dublin Core* (85%) e à capacidade de exposição de metadados por meio do protocolo OAI-PMH ativo (78%).

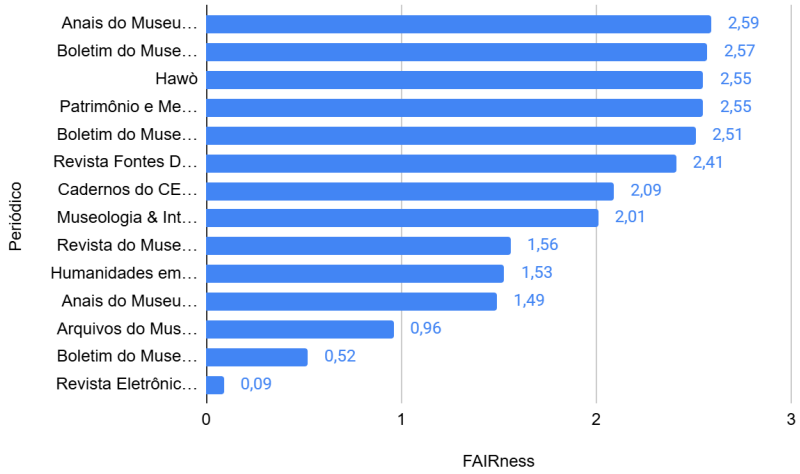
Contudo, o diagnóstico revelou fragilidades acentuadas na dimensão de indexação e controle terminológico. Embora os campos de palavras-chave estivessem preenchidos, observou-se uma quase total ausência de orientações institucionais ou travas técnicas que exigissem o uso de vocabulários controlados e tesauros da área, resultando em uma indexação livre que prejudica a recuperação precisa. Ademais, embora as páginas textuais das revistas mencionassem as licenças *Creative Commons*, essa informação por vezes falhava em aparecer codificada de forma explícita nos metadados (*dc.rights.uri*) de cada artigo individual, limitando a coleta automatizada.

Diante desse panorama, o Gráfico 1 sintetiza o ranking geral resultante da avaliação dos periódicos avaliados. Os dados revelam que o melhor desempenho foi alcançado pela revista Anais do Museu Paulista: História e Cultura Material, com 2,59 pontos, seguida de perto pelo Boletim do Museu de Biologia Mello Leitão (2,57 pontos), e pelos títulos Hawo e Patrimônio e Memória, ambos com 2,55 pontos. Nesses periódicos de alta performance, verificou-se o uso consolidado do ecossistema OJS integrado ao padrão Dublin Core, o preenchimento consistente de metadados de licenciamento, além da adoção sistemática de DOIs e identificadores de autoria.

Em contrapartida, os menores índices de conformidade foram registrados pela Revista Eletrônica Ventilando Acervos, com apenas 0,09 pontos, e pelo Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas, com 0,52 pontos. Nesses cenários de baixa pontuação, constatou-se a utilização de plataformas não nativas para a gestão editorial, como o WordPress, o não alinhamento ao padrão Dublin Core, a disponibilização de arquivos em formatos restritos e a ausência crônica de licenças de uso explícitas.

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Gráfico 1 - *Ranking* de avaliação dos periódicos



Fonte: Elaborado pelas autoras (2026).

A avaliação do indicador persistente DOI demonstrou sua aplicação em 64% dos periódicos. Em contrapartida, os metadados revelaram um cenário de baixa adesão para os identificadores de autoria e filiação institucional, com o ORCID registrando 50% de presença e o ROR apenas 21%. O ID Lattes apresentou o cenário mais crítico, com apenas uma ocorrência identificada de forma isolada no corpo do PDF.

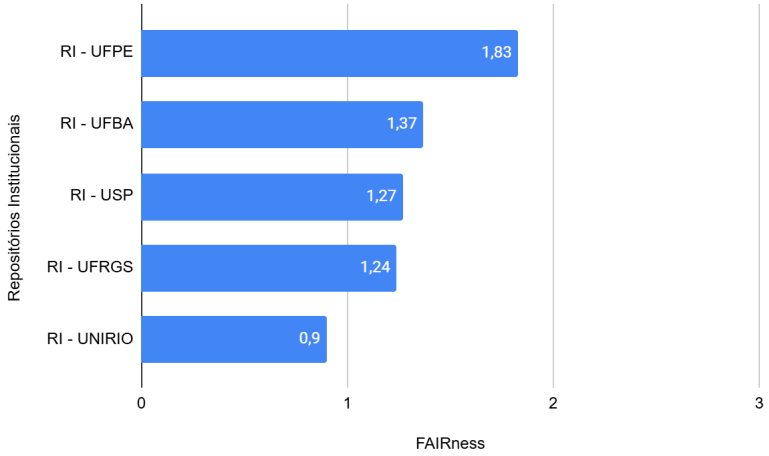
5.3 Aplicação em repositórios de teses e dissertações

Nos repositórios que aplicam *software* DSpace (80% do *corpus*) e estruturados sob o padrão *Dublin Core* (100%), a infraestrutura de comunicação mostrou-se robusta e altamente padronizada, reflexo do amadurecimento de políticas institucionais de acesso aberto consolidadas e orientadas pelo Ibict. Complementarmente, o protocolo de interoperabilidade OAI-PMH apresentou-se ativo em 100% do *corpus* investigado. O Gráfico 2 apresenta o desempenho dos repositórios em função das dimensões analíticas.

Observa-se que os repositórios institucionais apresentaram um desempenho geral inferior ao dos periódicos científicos em todas as dimensões analisadas. O índice mais elevado de conformidade FAIR foi registrado pelo RI UFPE, devido à utilização do *software* DSpace, padrão Dublin Core associado o protocolo OAI-PMH ativo, à presença do identificador persistente Handle plenamente resolvível, e à inclusão sistemática das credenciais ORCID e Lattes nos metadados, além da indicação clara do tipo de acesso e da licença de uso. Em contrapartida, o menor desempenho foi verificado no RI UNIRIO.

Esse cenário de baixa aderência decorre da persistência do protocolo HTTP em detrimento do HTTPS, associada à ausência de licenças de uso estruturadas e de identificadores persistentes tanto para os objetos digitais quanto para a autoria apesar de utilizar o *software* DSpace, padrão Dublin Core associado o protocolo OAI-PMH ativo.

Gráfico 2 - *Ranking* de avaliação dos repositórios institucionais



Fonte: Elaborado pelas autoras (2026).

O modelo de avaliação evidenciou como principal gargalo a carência de vocabulários controlados, somada a baixos índices de preenchimento para elementos essenciais: o ORCID e as licenças de uso figuram em apenas 20% dos registros (esta última mapeada pelo metadado *dc.rights.uri*), enquanto o ID Lattes atinge 40%. Diante disso, observa-se uma desconexão entre os requisitos exigidos na documentação de depósito e a real integridade dos metadados apresentados.

6. Discussão: contribuição das dimensões analíticas para o campo

O Modelo de Avaliação Diagnóstica proposto e validado nesta pesquisa afasta-se de abordagens puramente mecânicas, concebendo a avaliação da conformidade FAIR não como um *checklist* computacional binário, mas como uma engrenagem de auditoria analítica. A principal contribuição do modelo reside na capacidade de estabelecer uma ponte epistemológica indispensável entre os novos mandamentos da *e*-Ciência, ou Ciência Orientada por Dados, e as bases teóricas e históricas da Ciência da Informação e da Biblioteconomia (Borgman, 2015).

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Ao decompor a *FAIRness* nas seis dimensões analíticas, demonstra-se que as demandas contemporâneas por "localização" e "reúso" são, fundamentalmente, reconfigurações tecnológicas de práticas clássicas da nossa disciplina, tais como:

- A organização do conhecimento;
- A representação descritiva e temática;
- O controle terminológico rigoroso;
- A curadoria digital e a governança de dados.

Adicionalmente, o modelo demonstrou sensibilidade empírica ao ser aplicado aos ecossistemas de periódicos e repositórios institucionais, justamente por reconhecer as assimetrias, restrições e finalidades de cada ambiente informacional. O principal objetivo do modelo de avaliação é proporcionar uma avaliação diagnóstica qualiquantitativa que aponte caminhos viáveis de aprimoramento do tratamento da informação dentro dessas fontes.

Nesse contexto, as seis dimensões propostas constituem um conjunto de critérios aplicável à avaliação de fontes informacionais heterogêneas, por contemplarem aspectos transversais da interoperabilidade, independentemente da natureza ou da finalidade de cada ambiente. As dimensões Padrão de Dados e Indexação a partir de Metadados concentram-se na qualidade da representação da informação, na adoção de padrões e na organização dos dados, favorecendo sua descoberta, interpretação e recuperação. As dimensões Licença de Uso e Documentação e Proveniência avaliam os mecanismos de governança, transparência, rastreabilidade e condições de reutilização dos dados. Por sua vez, Infraestrutura Tecnológica e de Comunicação e Interconectividade examinam os requisitos tecnológicos que possibilitam o acesso, a integração e a troca de informações entre diferentes sistemas, considerando a disponibilidade de protocolos, serviços e mecanismos que viabilizam a interoperabilidade em ambientes distribuídos.

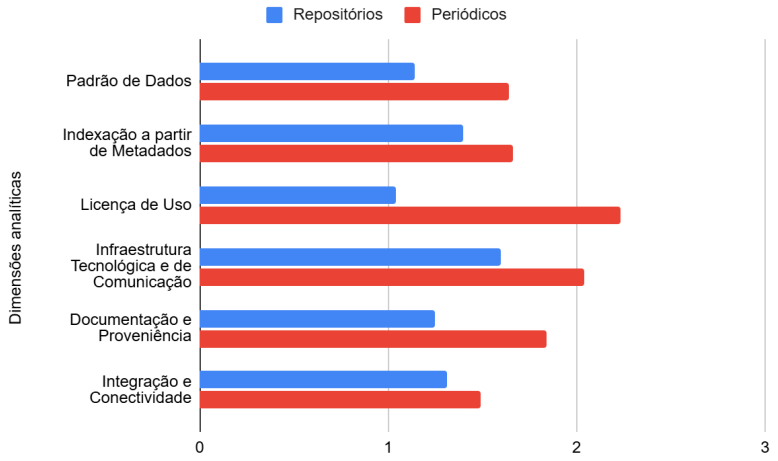
Sua aplicação em periódicos e repositórios científicos apontou uma melhor FAIR*ness* por parte dos periódicos que já se declaram *Open Access* e contemplam maior visibilidade pela indexação em plataformas internacionais. Já os achados no ecossistema repositórios vão ao encontro das ações de melhorias aplicadas pela CAPES onde, medidas de melhorias adotadas pelo GOPG No Brasil, a CAPES lançou o Guia de Interoperabilidade (2025) para integrar objetos digitais entre repositórios institucionais e a Plataforma Sucupira. Esse esforço sinaliza que as melhorias recaem sobre os repositórios, que deixam de ser meros depósitos para se tornarem ambientes vivos de informação científica voltados a análises e políticas públicas.

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Assim, o modelo ao propor a avaliação por dimensão dimensão analítica

O Gráfico 3 apresenta o desempenho comparativo das fontes de informação científicas avaliadas frente às seis dimensões analíticas.

Gráfico 3 - Desempenho comparativo das fontes de informação avaliadas frente às dimensões analíticas



Fonte: Elaborado pelas autoras (2026).

Dessa forma o Modelo de Avaliação Diagnóstica se apresenta como um instrumento a ser empregado em ações de curadoria, por gestores, profissionais da informação e por fontes que busquem o aprimoramento do tratamento da informação sob sua custódia visando a atuação em *Web* semântica.

7. Considerações finais

Este artigo cumpriu seu objetivo ao propor e aplicar um Modelo de Avaliação Diagnóstica FAIR estruturado em seis dimensões analíticas aplicadas a fontes de informação científicas na *Web*. A validação empírica confirmou a hipótese de que, embora as infraestruturas de periódicos e repositórios operem sobre bases técnicas consolidadas (OJS, DSpace e Dublin Core), subsistem lacunas em metadados de licenciamento, uso de vocabulários controlados e preenchimento de identificadores persistentes, o que compromete o pleno potencial de reúso e circulação do conhecimento científico.

Adv. Know. Repr.; 2026; volume 6; number 2. DOI: 10.5281/zenodo.22099744. Dossiê temático: O impacto das tecnologias emergentes na organização da informação e na organização do conhecimento

Como principal limitação, reconhece-se que a aplicação do modelo de forma manual demanda tempo e esforço analítico consideráveis dos profissionais da informação, o que pode restringir auditorias em larga escala.

Como desdobramento natural e horizonte futuro de pesquisa, aponta-se que as regras de verificação e os critérios de maturidade desenhados neste estudo possuem a plasticidade necessária para evoluir para uma ferramenta tecnológica semiautomatizada ou totalmente automatizada de avaliação, uma *FAIR Assessment Tool,* voltada a fontes científicas. O desenvolvimento desse *software* permitirá que editores de periódicos e gestores de repositórios realizem diagnósticos em tempo real, elevando o patamar de qualidade, visibilidade e perenidade da ciência aberta em ambientes digitais.

Referências

ALVES, L. S. S.; TARTAROTTI, R. C. D.; FUJITA, M. S. L. Avaliação do uso de vocabulário controlado em repositórios institucionais. *Informação@ Profissões*, Londrina, v. 11, n. 1, p. 52-77, 2022. Disponível em: https://ojs.uel.br/revistas/uel/index.php/infoprof/article/view/45158.

ARAUJO, P.; FÜHR,F.; LIMA, K. C. R.; ZULPO, S. As práticas de ciência aberta dos periódicos científicos correntes da biblioteca digital de periódicos da universidade federal do paraná. *Cadernos BAD (Portugal)*, v. 5, n., 2021. Disponível em: https://doi.org/10.48798/cadernosbad.2801.

BAHIM, C. *et al.* The FAIR data maturity model: an approach to harmonise FAIR assessments. *Data Science Journal*, v. 19, 2020. Disponível em: https://datascience.codata.org/articles/dsj-2020-041.

BARDIN, Laurence. *Análise de conteúdo*. Lisboa: Edições 70, 2016.

BERNERS-LEE, T. *Linked Data: Design Issues*. [S.l.]: W3C, 2006. Disponível em: https://www.w3.org/DesignIssues/LinkedData.html.

BORGMAN, C. L. *Big Data, Little Data, No Data: Scholarship in the Networked World*. Cambridge, Londres: TheMIT Press, 2015. 416 pp.

CANDELA, L. *et al*. The FAIR assessment conundrum: reflections on tools and metrics. *Data Science Journal*, v. 23, 2024. Disponível em: https://datascience.codata.org/articles/10.5334/dsj-2024-033.

CAPLAN, P. & ARMS, W. Y. Reference linking for journal articles. *D-Lib Magazine*, v.5, n.4, 1999. Disponível em: https://www.dlib.org/dlib/july99/caplan/07caplan.html.

CHALHUB, T. Análise das iniciativas para implementação do acesso livre à produção científica em repositórios de países americanos e europeus. *In*: PINHEIRO, L. V. R.; OLIVEIRA, H. C. P. (Org.). *Múltiplas facetas da comunicação e divulgação científica.* Brasília: Ibict, 2012. p. 293-319. Disponível em: https://livroaberto.Ibict.br/Handle/123456789/711.

CRUZ, M. C. A; FERNEDA, E.; FUJITA, M. S. L. A disponibilização de vocabulário controlado aos usuários para a recuperação da informação. *RICI: Revista Ibero-americana de Ciência da Informação*, v. 15, n. 1, 2022. Disponível em: https://periodicos.unb.br/index.php/RICI/article/view/42464/33460.

DEVARAJU, A. & HUBER, R. F-UJI - An Automated FAIR Data Assessment Tool. *Data Science Journal*, v.2, 2020. Disponível em: https://zenodo.org/records/4063720.

DEVARAJU, A. *et al*. From conceptualization to implementation: FAIR assessment of research data objects. *Data Science Journal*, v. 20, n. 1, 2021. Disponível em: https://doi.org/10.5334/dsj-2021-004.
DIMITROVA, M. *et al*. Semantic publishing enables text mining of biotic interactions. *Biodiversity Information Science and Standards*, v. 4, 2020: e59036. Disponível em: https://doi.org/10.3897/biss.4.59036.

EMONET, V. *et al*. Towards an extensible FAIRness assessment of FAIR Digital Objects. *Research Ideas and Outcomes (RIO)*, Bulgária, v. 8, e94988, p. 1-3, 2022. Disponível em: https://doi.org/10.3897/rio.8.e94988.

GILLILAND, A. J. Setting the stage. *In*: BACA, Murta (ed.). *Introduction to metadata.* 3rd ed. Los Angeles: Getty Research Institute, 2016. E-book. Disponível em: https://www.getty.edu/publications/intrometadata/setting-the-stage/.

GONZALEZ, E. *et al*. Challenges for FAIR digital object assessment. *Research Ideas and Outcomes (RIO)*, Bulgária, v. 8, p. 1-5, 2022. DOI: 10.3897/rio.8.e 95943. Disponível em: https://riojournal.com/article/95943/.

GUANDALINI, C. A.; SANTOS, A. A. Linked Open Data: conceito, relações e importância na Era da Informação. *Múltiplos Olhares em Ciência da Informação*, Belo

Horizonte, v. 8, n. 2, 2018. Disponível em: https://periodicos.ufmg.br/index.php/moci/article/view/16911.

GUANDALINI, C.; FURNIVAL, A. C. M.; ARAKAKI, A. C. S. Boas práticas científicas na elaboração de planos de gestão de dados. *Revista Digital Biblioteconomia e Ciência da Informação*, Campinas, v. 17 p. 1-20, 2018. Disponível em: https://www.SciELO.br/j/rdbci/a/GZ3b86PHsZJXXrhtQZn9fRD/?format=pdf&lang=pt.

HAHNEL, M. & VALEN, D. How to (easily) extend the FAIRness of existing repositories. *Data Intelligence*, v. 2, n. 1-2, p. 192-198, 2020. Disponível em: https://doi.org/10.1162/dint_a_00041.

HENNING, P. *et al*. Desmistificando os princípios FAIR: conceitos, métricas, tecnologias e aplicações inseridas no ecossistema dos dados FAIR.*Tendências da Pesquisa Brasileira em Ciência da Informação*, São Paulo, v. 11, n. 1. 2018. Disponível em: https://revistas.ancib.org/index.php/tpbci/article/view/458/443.

HOLANDA, P. M. C. & LOURENÇO, C. A. Treinamento e implementação da RDA: olhares e perspectivas. *RDBCI: Revista Digital de Biblioteconomia e Ciência da Informação*, Campinas, SP, v.19, e021016, 2021. Disponível em: https://brapci.inf.br/v/162369.

JACOBSEN, A. *et al*. FAIR principles: Interpretations and implementation considerations. *Data Intelligence*, Beijing, v. 2, p. 10-29, 2020. DOI: 10.1162/dint_r_00024. Disponível em: https://doi.org/10.1162/dint_r_00024 .

JUTY, N. *et al*. (2020). Unique, persistent, resolvable: identifiers as the foundation of FAIR. *Data Intelligence*, v. 2, n. 1-2, p. 30-39, 2020. Disponível em: https://doi.org/10.1162/dint_a_00025.

KAMM, S. *et al*. A survey on machine learning based analysis of heterogeneous data in industrial automation. *Computers in Industry*, v. 149, 2023. Disponível em: https://www.sciencedirect.com/science/article/pii/S0166361523000805.

LEMOS, D. L. S. & COELHO JÚNIOR, A. Qualidade de dados em acervos do patrimônio cultural: uma avaliação diagnóstica semiautomática nos objetos culturais sob gestão do Instituto Brasileiro de Museus. *Encontros Bibli*, Florianópolis, v. 28, 2023. Disponível em: https://periodicos.ufsc.br/index.php/eb/article/view/90510/52571.

LEMOS, D. L. S. *et al*. Organização e representação da informação e do conhecimento em contextos informacionais: uma proposta de um modelo teórico-conceitual para a qualidade de objetos culturais digitais. *Fronteiras da Representação do Conhecimento*, v. 3, n. 2, 2023. Disponível em: https://zenodo.org/records/8350312.

MARQUES, G. S. *et al*. Padronização de metadados adotados no portal agregador de conteúdo científico Oasisbr. *Ciência da Informação*, Brasília, v.54, n. 2, p. 1-11, 2025. Disponível em: https://revista.Ibict.br/ciinf/article/view/7223/7344.

MACHADO, R. B. & PEREIRA, A. M. Análise do padrão RDA: um estudo aplicado em teses e dissertações em literatura e cinema. *RDBCI: Revista Digital Biblioteconomia e Ciência da Informação*. Campinas, São Paulo. v.15, n.1, p.130-147, jan./abr., 2017. Disponível em: https://brapci.inf.br/v/40021. Acesso em: 25 jan. 2026.

MARÍN-ARRAIZA, P. & HEREDIA, A. FAIR PIDs: O papel da ORCID no fortalecimento dos Princípios FAIR. *In*: SALES, L. F. *et al.* (org.). *Princípios FAIR aplicados à gestão de dados de pesquisa.* Rio de Janeiro: Ibict, 2021. p.15-26. Disponível em: https://api.arca.fiocruz.br/api/core/bitstreams/22ee6376-eabd-4c9e-9a7f-aff923ee7a4c/content.

MARTINS, D. L. & FERREIRA, S. M. S. P. Protocolo OAI-PMH e Sistemas Federados de Informação: fundamentos de arquitetura da informação para análise de dados do portal de produção científica da área de Ciências da Comunicação Univerciencia.org. *Liinc em Revista*, v.8, n.2, 2012. Disponível em: https://www.researchgate.net/publication/314195943_Protocolo_OAI-PMH_e_Sistemas_Federados_de_Informacao.

MORAES, L. S. & SALES, L. F. Uso das redes sociais pelos repositórios institucionais de acesso aberto. *Ciência da Informação*, Brasília, D. F., v. 48, n. 3 (Supl.), p. 140-146, set./dez. 2019. Disponível em: https://brapci.inf.br/v/136344.

PASSOS, H. G. S. 2026. *Modelo de Avaliação Diagnóstica fundamentado nos Princípios Fair para fontes de informação na web.* 2026. [Dissertação de mestrado não publicada] , Universidade Federal do Espírito Santo. Repositório Institucional da Universidade Federal do Espírito Santo.

OLIVEIRA, M. C. A. *Proposta de modelo de metadados baseado no padrão Dublin Core para a descrição de patentes em repositórios institucionais*. 2023. 128 f. Dissertação (Mestrado em Ciência da Informação) - Instituto Brasileiro de Informação

em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2023. Disponível em: https://ridi.Ibict.br/Handle/123456789/1362.

PEREIRA, J. P. S.; RODRIGUES, R. S.; SANTOS, S. M.. Periódicos científicos com indexação descontinuada: a Coleção SciELO Brasil. *Transinformação*, v. 32, 2020. Disponível em: https://www.SciELO.br/j/tinf/a/ZJpwWc58yfcLXn6xFC3Ch7s/.

SANTOS, A. S. & ARAKAKI, F. A. A. Aplicação do RDA em instituições latino-americanas: um panorama do processo de implementação. *Revista Brasileira de Biblioteconomia e Documentação*, São Paulo, v. 18, p. 01-18, 2022. Disponível em: https://brapci.inf.br/v/229137.

SANTOS, L. O. B. S. *et al*. FAIR Data Point: A FAIR-Oriented Approach for Metadata Publication. *Data Intelligence*, 2023, v. 5, n. 1, p. 163–183. Disponível em doi: https://doi.org/10.1162/dint_a_00160.

SAYÃO, L. F. Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes - urn, purl, doi, Handle system, crossref e openurl. *Transinformação*, v. 19, n. 1, 2007. Disponível em: https://brapci.inf.br/v/115890.

TARIQ, Z.; ZAINAB, B.; HUSSAIN, M. Z.. Evaluating the effectiveness and resilience of SSL/TLS, HTTPS, IPSec, SSH, and WPA/WPA2 in safeguarding data transmission. **UCP Journal of Engineering and Technology**, v. 1, n.2, 2024. Disponível em: https://doi.org/10.24312/ucp-jeit.01.02.144.

TOMAÉL, M. I. & ALCARÁ, A. R. *Fontes de informação digital*. Londrina: Eduel, 2016. Disponível em: https://books.google.com.br/books?id=bLg0EAAAQBAJ&printsec=frontcover&hl=pt-BR&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false.

VIDOTTI, S. A. B. G.; TORINO, E.; CONEGLIAN, C. S. #SejaJUSTOeCUIDADOSO: princípios FAIR e CARE na gestão de dados de pesquisa. *In*: SALES, L. F et al., (Orgs.) *Princípios FAIR aplicados à gestão de dados de pesquisa.* Ibict, Rio de Janeiro, 2021.

WILKINSON, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, [*S.l.*], v. 3, n. 1, p. 1-9, 2016. Disponível em: Disponível em: https://www.nature.com/articles/sdata201618.

ZENG, M. L. Interoperability. *Knowledge Organization*, v. 46, n. 2, p. 122-146, 2019. Disponível em: https://article.imrpress.com/journal/KO/46/2/10.5771/0943-7444-2019-2-122/71f20dff13a617263334a524e8a2120e.pdf.

ZENG, M. L. Knowledge organization systems (KOS). *Knowledge Organization*, v.35, n.2/3,p.160-82, 2008.Disponível em: https://www.academia.edu/26672820/Knowledge_Organization_Systems_KOS_.