26º ENANCIB
Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação
A Amazônia na sociedade global em rede: assimetrias e potencialidades informacionais, tecnológicas e ambientais
Belém, Pará, 09 a 13 de novembro de 2026
Universidade Federal do Pará

XXVI ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - XXVI ENANCIB

ISSN 2177-3688

GT 2 –Organização e Representação do Conhecimento

ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO NA CIÊNCIA ABERTA: UMA AVALIAÇÃO FAIR NO REPOSITÓRIO MULTIDISCIPLINAR DEPOSITA DADOS

INFORMATION ORGANIZATION AND REPRESENTATION IN OPEN SCIENCE: *A FAIR ASSESSMENT OF THE MULTIDISCIPLINARY REPOSITORY DEPOSITA DADOS*

Raimunda Fernanda dos Santos – Universidade Federal do Rio Grande do Norte (UFRN)
Carla Beatriz Marques Felipe – Universidade Federal do Rio de Janeiro (UFRJ)

Modalidade: Trabalho Completo

Resumo: Estuda questões concernentes à organização e representação da informação no contexto dos repositórios de dados de pesquisa, em especial no Deposita Dados. Tem como objetivo avaliar a conformidade das práticas de representação da informação e a qualidade dos metadados no repositório multidisciplinar brasileiro Deposita Dados à luz dos Princípios FAIR (*Findable, Accessible, Interoperable, Reusable*). Utiliza como metodologia a pesquisa bibliográfica e documental, de natureza descritiva. Adota uma abordagem qualitativa, delineada por meio de um estudo de caso focado no repositório Deposita Dados. Para a análise do repositório com base nos Princípios FAIR, foi utilizada a ferramenta F-UJI (*FAIRsFAIR Research Data Object Assessment Service*) para analisar uma amostra de dez conjuntos de dados. Os resultados revelaram um cenário promissor e dual: o repositório atingiu um índice global de 92% de aderência aos princípios, alcançando nível avançado nos eixos de Encontrabilidade e Acessibilidade devido ao uso eficiente de identificadores persistentes (DOIs e *Handles*) e indexação. Os eixos de Interoperabilidade e Reutilização situaram-se no nível moderado, evidenciando lacunas na adoção de vocabulários controlados específicos e na captura de metadados de proveniência. As considerações finais apontam que, embora o repositório cumpra com distinção seu papel de custódia democrática, carece de camadas de mediação semântica. Conclui-se que a governança da plataforma deve evoluir para a incorporação de Perfis de Aplicação flexíveis e tecnologias da Web Semântica. Tal prática permitirá conectar a infraestrutura geral às especificidades das comunidades científicas, superando o paradoxo da "disponibilidade sem acessibilidade" e garantindo a efetiva reprodutibilidade da Ciência Aberta no país.
Palavras-chave: organização da informação; representação da informação; repositórios de dados de pesquisa; acesso aberto; Deposita dados.

Abstract: This study examines issues concerning information organization and representation within the context of research data repositories, with a specific focus on Deposita Dados. It aims to assess the compliance of information representation practices and metadata quality in the Brazilian multidisciplinary repository Deposita Dados in light of the FAIR Principles (Findable, Accessible, Interoperable, Reusable). The methodology comprises descriptive, bibliographic, and documentary research. It adopts a qualitative approach outlined through a case study focused on the Deposita Dados repository. For the repository analysis based on the FAIR Principles, the F-UJI automated tool (FAIRsFAIR Research Data Object Assessment Service) was employed to evaluate a sample of ten datasets stored on the platform. The results revealed a promising yet dual scenario: the repository achieved a 92% global adherence index to the principles, reaching an advanced level in the Findability and Accessibility axes due to the efficient use of persistent identifiers (DOIs and Handles) and global indexing. Conversely, the Interoperability and Reusability axes were situated at a moderate level, evidence of gaps in the adoption of specific controlled vocabularies and the capture of provenance metadata. The final considerations point out that although the repository fulfills its role of democratic custody with distinction, it lacks semantic mediation layers. It is concluded that the platform's governance must evolve to incorporate flexible Application Profiles and Semantic Web technologies. Such a practice will allow connecting the general infrastructure to the specificities of scientific communities, overcoming the paradox of "availability without accessibility" and guaranteeing the effective reproducibility of Open Science in the country.
Keywords: information organization; information representation; research data repositories; open access; Deposita Dados.

1 INTRODUÇÃO

Os dados de pesquisa constituem a essência da investigação científica, atuando como a evidência fundamental que sustenta a validade das conclusões de um estudo. Para além de sua função comprobatória, tais dados são os pilares da reprodutibilidade, visto que sua disponibilidade permite tanto a validação de resultados quanto o fomento a novas pesquisas revisadas por pares. Nesse sentido, a eficácia da transparência e da colaboração no ambiente acadêmico torna-se intrinsecamente dependente da forma como esses dados são organizados e disponibilizados.

Sob essa perspectiva, os dados de pesquisa caracterizam-se como registros processados e coletados durante a investigação científica. Eles abrangem uma variedade de tipos e formatos que vão desde planilhas estruturadas e transcrições de entrevistas até registros visuais, sonoros e computacionais (Santos; Felipe, 2025). Mais do que meros subprodutos do processo experimental, esses dados constituem objetos informacionais complexos que, para serem integrados efetivamente ao ecossistema da Ciência Aberta, demandam uma organização que preserve seu valor probatório e sua integridade semântica ao longo do tempo.

A crescente produção de dados tem impulsionado a criação de repositórios digitais especializados. Essas plataformas têm como objetivo armazenar, organizar, preservar e disseminar dados de pesquisa de forma acessível, viabilizando seu reuso em futuras investigações científicas (Bonetti; Arakaki, 2022).

Nesses repositórios, os dados são representados por meio de metadados, que são informações estruturadas que descrevem um recurso informacional. Para Alves (2010, p. 47), eles são “atributos que representam uma entidade (objeto do mundo real) em um sistema de informação”. Dessa maneira, os metadados funcionam como uma espécie de "carteira de identidade" dos dados, fornecendo detalhes essenciais sobre seu conteúdo, contexto, autoria e formato.

Em consonância com Santos e Felipe (2025), entende-se que a representação adequada dos dados por meio de metadados padronizados e descrições claras é crucial para sua recuperação e reuso em novas pesquisas.

Uma boa representação não apenas contribui para que os usuários encontrem os dados mais relevantes para suas necessidades, mas também garante as condições para que eles sejam compreendidos e aplicados em diferentes contextos. Sem uma descrição precisa, um conjunto de dados torna-se praticamente inútil, mesmo que esteja tecnicamente disponível online. Além disso, a padronização na representação contribui significativamente para a interoperabilidade entre sistemas e plataformas.

Para que o compartilhamento de dados vá além da mera disponibilidade técnica e alcance plena utilidade científica, é imperativo que a gestão de dados seja norteada por diretrizes normativas rigorosas. Nesse contexto, os Princípios FAIR emergem como um conjunto de diretrizes estratégicas projetadas para garantir que os dados sejam efetivamente encontráveis (*Findable*), acessíveis (*Accessible*), interoperáveis (*Interoperable*) e reutilizáveis (*Reusable*).

No contexto brasileiro, observa-se uma disparidade na aplicação de padrões de metadados e diretrizes de governança de dados, o que pode comprometer a visibilidade e o impacto da produção científica nacional. Diante desse panorama, o repositório Deposita Dados destaca-se como infraestrutura estratégica para a custódia e o compartilhamento de dados de pesquisa; contudo, permanecem lacunas quanto à extensão em que suas práticas descritivas e operacionais atendem aos rigorosos critérios internacionais de maturidade em Ciência Aberta.

Dessa forma, consolida-se o seguinte problema de pesquisa: Em que medida as práticas de representação da informação e a qualidade dos metadados do repositório Deposita Dados atendem aos princípios FAIR, e qual é o nível de maturidade dessa plataforma na promoção da Ciência Aberta no contexto brasileiro?

Diante do exposto, este estudo tem como objetivo avaliar a conformidade das práticas de representação da informação e a qualidade dos metadados no repositório Deposita Dados, com base nos princípios FAIR, a fim de identificar o nível de maturidade da plataforma na promoção da Ciência Aberta no contexto brasileiro.

A relevância deste estudo justifica-se pela necessidade premente de superar o paradoxo da "disponibilidade sem acessibilidade". Embora o volume de dados depositados em repositórios tenha crescido exponencialmente, a falta de rigor técnico na descrição desses ativos pode transformá-los em "dados órfãos", ou seja, arquivos que, apesar de armazenados, não podem ser localizados ou interpretados por humanos ou máquinas. Ao avaliar o repositório Deposita Dados sob a ótica dos Princípios FAIR, esta pesquisa busca determinar se a infraestrutura de informação do Brasil está, de fato, cumprindo seu papel de facilitar a descoberta científica ou se permanece apenas como um repositório passivo de arquivos.

Além disso, a realização desta pesquisa é essencial para o fortalecimento das políticas de Ciência Aberta no Brasil. A análise da qualidade dos metadados e do nível de maturidade da plataforma oferece subsídios estratégicos para gestores de informação e agências de fomento, permitindo a implementação de melhorias na curadoria de dados. Justifica-se, portanto, pela oportunidade de estabelecer parâmetros de excelência que garantam o retorno do investimento em pesquisa pública, assegurando que o conhecimento gerado seja um bem comum efetivamente reutilizável, interoperável e transparente para toda a sociedade.

Do ponto de vista da Organização e Representação da Informação e do Conhecimento, esta pesquisa justifica-se pela necessidade de investigar como as estruturas de representação atuam como mediadoras semânticas em ambientes digitais complexos. Ao analisar o repositório Deposita Dados, este estudo contribui para o corpo teórico da área ao discutir como a padronização e a exaustividade da representação da informação influenciam na construção de redes de conhecimento interoperáveis.

2 REFERENCIAL TEÓRICO

De acordo com Gäal e Martins (2022), o Acesso Aberto atua como um mecanismo de comunicação científica voltado à democratização do conhecimento, operando na remoção de barreiras financeiras e jurídicas aos resultados de pesquisa. É imperativo ressaltar, contudo, que a abertura não é absoluta: em repositórios de dados, o acesso pode ser limitado por salvaguardas relacionadas à confidencialidade, à privacidade ou à propriedade intelectual. Superadas essas restrições éticas e legais, o modelo potencializa a circulação da informação entre diferentes campos, fomentando práticas interdisciplinares que permitem a pesquisadores consultar e aplicar dados provenientes de diversas áreas do saber.

Nesse contexto, o compartilhamento de dados fundamenta-se na premissa de que tais ativos possuem valor potencial de reuso, tanto em seu campo disciplinar de origem quanto em investigações de caráter interdisciplinar (Borgman, 2010). Contudo, sob a ótica da Organização da Informação e do Conhecimento, para que essa utilidade se concretize, a mera disponibilidade técnica de um arquivo é insuficiente. É imprescindível que os dados passem por um processo rigoroso de representação descritiva e temática, gerando metadados robustos que garantam sua interpretabilidade e inteligibilidade. Sem essa ancoragem contextual e documental, os conjuntos de dados convertem-se em estruturas opacas, compreensíveis apenas para seus criadores originais, o que inviabiliza o fluxo da comunicação científica.

Por conseguinte, o ecossistema da Ciência Aberta exige que o dado de pesquisa deixe de ser interpretado como um elemento estático e isolado e passe a ser tratado como um objeto informacional dinâmico e complexo. Para tanto, a governança desses recursos deve contemplar políticas de preservação digital de longo prazo, visando mitigar os riscos de obsolescência tecnológica e assegurar a integridade, a acessibilidade, a interpretação e o potencial de reutilização das informações ao longo do tempo (Sayão; Sales, 2012). Na perspectiva da Organização da Informação, essa governança se materializa na criação de ecossistemas capazes de preservar não apenas a estrutura física dos arquivos, mas a integridade das redes de significados que os tornam úteis para a ciência.

Para que o compartilhamento e a recuperação da informação ocorram de forma eficiente nesses ambientes digitais, faz-se necessário adotar arquiteturas normativas e padronizadas. É nesse cenário que se posicionam os Princípios FAIR, um conjunto de diretrizes essenciais que balizam a encontrabilidade, a acessibilidade, a interoperabilidade e a reutilização

dos dados (Diniz; Dias; Silva, 2018). O acrônimo estrutura os requisitos mínimos para que os recursos informacionais sejam efetivamente processados tanto por agentes humanos quanto por algoritmos: Findable (Encontrável); Accessible (Acessível); Interoperable (Interoperável); Reusable (Reutilizável).

Cabe ressaltar que a encontrabilidade, no contexto dos princípios FAIR, refere-se a uma propriedade intrínseca do objeto digital (frequentemente aplicada a dados de pesquisa) de ser localizado, como ocorre por meio do uso de identificadores persistentes. Por outro lado, no âmbito da Ciência da Informação, a encontrabilidade da informação constitui um conceito mais amplo, associado à capacidade de um recurso ser localizado e percebido em um ambiente informacional complexo. Desse modo, a fim de evitar sobreposições conceituais e imprecisões terminológicas, o termo Localizável será empregado nesta pesquisa para traduzir o atributo de encontrabilidade do primeiro princípio FAIR.

Diferente de uma visão puramente instrumental, a correspondência entre os princípios FAIR e os fundamentos da Organização da Informação (OI) e Organização do Conhecimento (OC) revela que cada pilar do acrônimo depende diretamente da qualidade dos instrumentos de representação. Os processos de catalogação de dados (*data curation*) e a aplicação de esquemas de metadados padronizados atuam como pontes de mediação semântica. Ao estabelecer Identificadores Persistentes (PIDs), mapear a proveniência e aplicar Sistemas de Organização do Conhecimento (SOC) (tais como tesauros, taxonomias, ontologias e esquemas de classificação), os repositórios cumprem a função clássica da área: transformar dados dispersos em conhecimento estruturado, localizável e intercambiável.

Aprofundando essa arquitetura, as exigências de cada faceta do acrônimo demonstram o papel central da representação, conforme Wilkinson *et al.* (2016):

- **Encontrabilidade (*Findable*)**: Alinha-se diretamente com a função de busca e recuperação da informação. Exige que dados e metadados recebam identificadores persistentes, únicos e globais (como DOI ou *Handle*), e que a descrição documental seja detalhada o suficiente para permitir a indexação inequívoca em catálogos e motores de busca.
- **Acessibilidade (*Accessible*)**: Trata das condições de manuseio do objeto informacional. Determina a recuperação por meio de protocolos de comunicação abertos e padronizados (como HTTP/HTTPS). Sob o prisma da preservação documental, impõe que o registro de metadados permaneça perenemente disponível no repositório,

garantindo a memória institucional da pesquisa mesmo que o dado original precise ser descontinuado ou protegido por sigilo.

- **Interoperabilidade (*Interoperable*)**: Representa o núcleo semântico da OC. Exige que a informação seja expressa por meio de linguagens formais, compartilhadas e semanticamente ricas (como RDF, XML e JSON-LD). Para que plataformas distintas dialoguem, os metadados devem utilizar vocabulários controlados e estruturas conceituais que também sigam os preceitos FAIR, mitigando ambiguidades terminológicas entre diferentes comunidades científicas.
- **Reutilização (*Reusable*):** É o ápice do processo de organização, onde o dado atinge sua máxima utilidade. Demanda uma descrição contextual exaustiva (pluralidade de atributos precisos), o mapeamento detalhado da proveniência histórica (ciclo de vida do dado) e a vinculação a licenças jurídicas explícitas (como *Creative Commons*). É a qualidade da representação que dita a capacidade de um cientista alheio ao projeto original compreender os limites, vieses e aplicabilidade daquele conjunto de dados em novos cenários.

Desse modo, a convergência entre os Princípios FAIR e a Organização da Informação e do Conhecimento demonstra que a consolidação da Ciência Aberta transcende a infraestrutura de TI; trata-se de um imperativo político, metodológico e epistemológico. Como apontam Albagli e Maciel (2009), a democratização do conhecimento científico, especialmente no cenário do Sul Global, está intrinsecamente atrelada ao desenvolvimento de infraestruturas de informação robustas e bem desenhadas. São essas estruturas de representação que impedem que o dado de pesquisa se torne "ruído" ou se perca na imensidão digital, garantindo que o investimento público se reverta em um patrimônio científico comum, verdadeiramente acessível, integrável e emancipador para a sociedade.

3 PROCEDIMENTOS METODOLÓGICOS

A presente investigação caracteriza-se como um estudo de caso, de natureza descritiva e abordagem quali-quantitativa, estruturado a partir de pesquisa bibliográfica e documental centrada no repositório Deposita Dados.

No que tange à análise da organização e representação em repositórios de dados de pesquisa, selecionou-se como objeto de estudo o Deposita Dados1. Este Repositório Comum de Dados de Pesquisa tem como propósito reunir e disponibilizar conjuntos de dados gerados por pesquisadores brasileiros vinculados a instituições que ainda não dispõem de infraestrutura própria para o depósito de dados (Deposita Dados, 2026).

A escolha do Deposita Dados como campo empírico desta pesquisa justifica-se por seu papel como repositório dedicado preservação e disseminação de dados de pesquisa. A plataforma reúne conjuntos de dados provenientes de diferentes áreas do conhecimento, oferecendo um ambiente estruturado para pesquisadores e instituições que não possuem repositórios, mas que querem contribuir com a Ciência aberta, compartilhando os seus dados.

Para a análise do repositório com base nos Princípios FAIR, foi utilizada a ferramenta F-UJI (FAIRsFAIR Research Data Object Assessment Service), cujo objetivo é avaliar programaticamente a conformidade com os princípios FAIR de objetos de dados de pesquisa, com base em métricas desenvolvidas pelo projeto FAIRsFAIR (Devaraju; Huber, 2020). A ferramenta permite que a avaliação seja realizada por meio de uma URL ou DOI de um conjunto de dados.

O ' F ' significa JUST e ' UJI ' significa 'Teste' em malaio. Portanto, F-UJI é uma ferramenta de teste JUSTA (Devaraju; Huber, 2020). A ferramenta efetua avaliações automáticas com bases nos metadados disponíveis em cada conjunto de dados. Para tal, ela emprega informações constantes na página descritiva dos conjuntos de dados em provedores de identificadores persistentes como DataCite2, e em outros serviços, tais como re3data3 e FAIRsharing4. Cabe frisar que é uma ferramenta que ainda está em desenvolvimento.

Para a consecução desta pesquisa, a amostragem foi constituída por dez conjuntos de dados extraídos do repositório investigado, selecionados a partir do critério de recência dos depósitos. A escolha dos 10 dez conjuntos de dados teve caráter exploratório e preliminar, visando à necessidade de efetuar uma análise inicial da implementação dos princípios FAIR no

[1] Disponível em: https://depositadados.ibict.br/.
[2] O *DataCite* é uma organização internacional que tem como principal objetivo promover recuperação da informação, especialmente em conjuntos de dados científicos. Encontra-se disponível em: https://datacite.org/.
[3] O *re3data* é um buscador global de repositórios de dados de pesquisa. Encontra-se disponível em: https://www.re3data.org/.
[4] O *FAIRsharing* é um registro online de padrões, bancos de dados e políticas de dados utilizados em diferentes áreas da pesquisa científica. Disponível em: https://fairsharing.org/.

repositório investigado. O recorte amostral baseou-se nos dez conjuntos de dados mais recentes disponíveis durante o período de realização da pesquisa, sem a intenção de gerar uma amostra estatisticamente representativa do conjunto de dados do repositório, mas de permitir uma primeira avaliação de suas características de organização, representação e disponibilização dos dados.

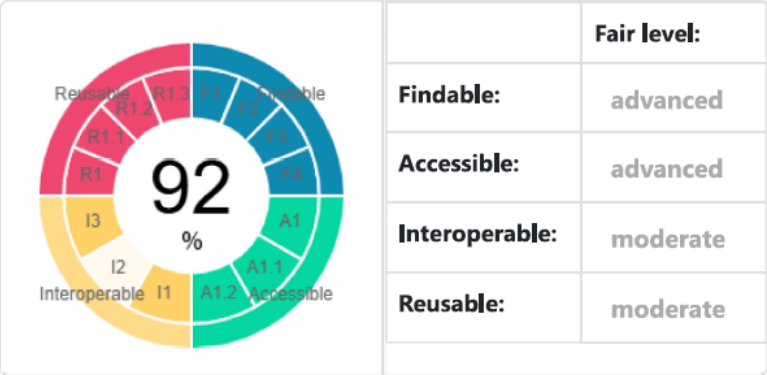
Para minimizar os impactos relacionados a utilização da ferramenta que ainda está em desenvolvimento, a etapa de coleta de dados foi realizada em maio de 2026, tomando-se esse período como referência para a coleta e análise dos dados. Consequentemente, os resultados apresentados refletem as condições e os critérios disponíveis na ferramenta na ocasião da avaliação.

Esse procedimento visou garantir a análise de registros atualizados e representativos das práticas contemporâneas de organização e representação da informação na plataforma. Ademais, por se tratar de um repositório institucional de caráter multidisciplinar, a amostra selecionada reflete essa heterogeneidade, contemplando os conjuntos de dados mais recentes disponibilizados pelo repositório. Essa diversidade abrange áreas como Computação, Ciências Sociais, Ciência da Informação, Engenharia, Artes e Humanidades, e Saúde e Ciências da Vida, com temáticas que variam de estudos bibliométricos e repositórios digitais até Internet das Coisas, pesquisas com insetos, produções artísticas e investigações sobre doenças genéticas graves.

4 RESULTADOS E DISCUSSÃO

No que tange à aderência do repositório aos Princípios FAIR, a avaliação automatizada dos dez conjuntos de dados selecionados revelou um índice global de conformidade de 92%. A homogeneidade desse resultado entre todas as amostras analisadas evidencia a maturidade e a consistência das políticas de curadoria digital, organização e arquitetura de metadados implementadas pela instituição. A Figura 1 apresenta a síntese da avaliação realizada pela ferramenta F-UJI.

Figura 1 – Síntese da avaliação realizada pela ferramenta F-UJI



	Fair level:
Findable:	advanced
Accessible:	advanced
Interoperable:	moderate
Reusable:	moderate

Fonte: F-UJI (2026).

A ferramenta de diagnóstico empregada segmenta cada pilar analítico (*Findable, Accessible, Interoperable* e *Reusable*) em quatro níveis progressivos de maturidade: incompleto, iniciante, moderado e avançado. Os resultados obtidos nesta modelagem diagnóstica apontam que o repositório Deposita Dados atingiu o nível **avançado** nos eixos **Localizável** e **Acessível**, chancelando a alta eficácia da infraestrutura tecnológica na indexação e na disponibilidade dos objetos digitais via protocolos universais. Em contrapartida, os eixos de **Interoperabilidade** e **Reutilização** situaram-se no nível **moderado**.

Embora tais eixos reflitam a incorporação sistemática de boas práticas documentais, eles revelam fraturas latentes que demandam o refinamento de semânticas estruturadas e o adensamento de dados sobre o ciclo de vida dos recursos, passos fundamentais para potencializar a integração e a reciclagem do conhecimento científico.

Alinhada aos objetivos desta investigação, foca-se a análise detalhada nos subprincípios que regem as atividades de representação e a qualidade dos metadados. Estes elementos conectam-se de forma umbilical aos eixos *Findable*, *Interoperable* e *Reusable*, constituintes da tríade essencial para a recuperação e o uso compartilhado de dados.

4.1 EIXO FINDABLE (LOCALIZÁVEL)

O pilar **Localizável** é desdobrado pela ferramenta em critérios rigorosos de identificação e descrição. Os subprincípios **F1-01** (atribuição de identificadores globais, únicos e persistentes aos metadados) e **F1-02** (atribuição de PIDs aos dados propriamente ditos) demonstraram plena conformidade no Deposita Dados. Essa aderência, mediada pelo uso sistemático de DOIs e *Handles*, é o primeiro requisito para mitigar o link rot (degradação de links) e assegurar a perenidade dos objetos na teia científica global.

Quanto ao subprincípio **F2**, que preconiza a descrição dos dados por meio de metadados ricos e detalhados (título, autoria, data de publicação, palavras-chave, dentre outros), verificou-se conformidade integral em toda a amostra. Sob o prisma da Organização da Informação, este resultado atesta a presença de descrições informacionais densas, essenciais para reduzir a ambiguidade conceitual e expandir a descoberta orgânica.

Essa consistência estende-se ao subprincípio **F3**, que atesta a vinculação explícita entre a instância dos metadados e o identificador do arquivo que eles descrevem. Por fim, o subprincípio **F4** demonstrou pleno alinhamento, visto que o Deposita Dados encontra-se formalmente indexado no *Re3data* (catálogo global de repositórios de dados de pesquisa). Esse arranjo garante que os dados transcendam as fronteiras da instituição e se tornem visíveis a agregadores globais.

4.2 EIXO INTEROPERABLE (INTEROPERÁVEL)

O trânsito pelo princípio da Interoperabilidade revela o primeiro gargalo estrutural da plataforma. O repositório obteve êxito no subprincípio **I1**, que afere o uso de linguagens formais, acessíveis e amplamente aplicáveis para a representação do conhecimento. A plataforma emprega esquemas de metadados padronizados (como o *Dublin Core* e estruturas derivadas em XML/JSON), garantindo uma leitura sintática eficiente tanto por agentes humanos quanto por sistemas computacionais. O subprincípio **I3**, que prevê a inclusão de referências qualificadas (links semânticos) para outros metadados e recursos correlatos, também foi validado como uma prática recorrente, gerando redes incipientes de contextualização.

Contudo, a desconformidade manifestou-se no subprincípio **I2**, que exige que os metadados utilizem, eles próprios, vocabulários, ontologias ou terminologias controladas que sigam os preceitos FAIR. O Deposita Dados não pontuou neste critério, sendo este o fator determinante para o rebaixamento do pilar de Interoperabilidade para o nível moderado. Formula-se a hipótese de que essa limitação decorre diretamente da natureza **multidisciplinar (não temática)** do repositório. Implementar e mapear dinamicamente terminologias altamente específicas para domínios que variam da Computação às Artes representa um desafio complexo na ausência de perfis de aplicação flexíveis ou de ferramentas de conversão semântica integradas à interface de depósito.

3.3 EIXO REUSABLE (REUTILIZÁVEL)

No domínio do Reuso, o subprincípio **R1** subdivide-se em camadas avaliativas fundamentais. O critério **R1.01**, focado em verificar se os metadados especificam adequadamente o conteúdo para permitir a triagem inicial pelo usuário, foi plenamente atendido. O repositório oferece campos suficientes para que a relevância do conjunto de dados seja aferida. Da mesma forma, o desdobramento **R1.1-01**, responsável por avaliar a presença explícita de licenças de uso e termos de restrição jurídica (como as atribuições *Creative Commons*), registrou conformidade unânime, conferindo segurança legal indispensável para a reciclagem dos dados.

Por outro lado, o subprincípio **R2**, focado no registro de dados de proveniência (*provenance metadata*), foi classificado como apenas **parcialmente atendido**. A análise empírica revelou que, embora o repositório documente com precisão o momento do depósito inicial e a autoria, ele falha em registrar o histórico de ciclo de vida do dado, as etapas de processamento técnico, as versões anteriores ou modificações subsequentes. Para a ferramenta F-UJI, essa lacuna obscurece a linhagem do dado (*data lineage*), comprometendo a confiabilidade necessária para a reprodução experimental.

Por fim, o subprincípio **R3** evidenciou a clássica fricção entre generalismo e especificidade. No desdobramento **R3.01-01** (uso de padrões de metadados recomendados por comunidades científicas específicas), o repositório não obteve pontuação integral por adotar um esquema de metadados abrangente e generalista em detrimento de esquemas de domínio (como o *DDI* para Ciências Sociais ou *Darwin Core* para a Biodiversidade). Paradoxalmente, o critério **R3.01-02** (disponibilização dos dados em formatos de arquivo recomendados pelas respectivas comunidades, como a preferência por formatos abertos e não proprietários como: .csv, .txt ou. json) foi atendido integralmente. O repositório demonstra, portanto, um esforço deliberado em curar a tipologia dos arquivos depositados, embora careça de sofisticação na camada descritiva que os recobre.

3.4 REFLEXÕES SOB A PERSPECTIVA DA ORGANIZAÇÃO DO CONHECIMENTO

Em síntese, a classificação geral do repositório Deposita Dados como **avançado** (graças ao seu forte desempenho global de 92%) não deve camuflar as assimetrias identificadas nos eixos de Interoperabilidade e Reuso. A opção por padrões de metadados generalistas garante a "custódia democrática", permitindo que pesquisadores de qualquer área submetam seus dados

sem barreiras técnicas intransponíveis. Contudo, essa facilidade de entrada transfere para o usuário final ou para sistemas externos o esforço cognitivo de desambiguação terminológica.

No âmbito da Organização do Conhecimento, essa realidade impõe a necessidade de transição de uma postura passiva de armazenamento para uma postura ativa de curadoria baseada em **Perfis de Aplicação**. Para que a maturidade migre do patamar moderado para o avançado em sua totalidade, a representação da informação na plataforma não pode ser estática. É imperativo incorporar tecnologias da Web Semântica (como mapeamentos em *Simple Knowledge Organization System* (SKOS) e grafos de conhecimento interconectados) capazes de cruzar os metadados gerais do repositório com ontologias de domínio específicas, transformando strings de texto em vetores de dados conectados (*Linked Open Data*).

Com o objetivo de sintetizar as discussões, o Quadro 1 correlaciona os princípios, níveis de maturidade observados, subprincípios validados e as respectivas notas analíticas.

Quadro 1 – Síntese de aderência aos princípio FAIR no repositório Deposita Dados

Princípio	Nível De Maturidade	Subprincípios Atendidos	Observações
Findable **(Localizável)**	**Avançado**	**F1, F2, F3 e F4**	Conformidade plena. Uso robusto e institucionalizado de identificadores persistentes (PIDs). Alta densidade descritiva nos esquemas de metadados, alinhada a redes globais de indexação (Re3data).
Acessible **(Acessível)**	**Avançado**	**Critérios gerais**	Alta eficiência. Dados e metadados são recuperáveis por meio de protocolos de comunicação abertos e padronizados (HTTP/HTTPS). Os metadados permanecem perenemente acessíveis, atuando como registro institucional permanente.
Interoperable **(Interoperável)**	**Moderado**	**I1 e I3**	Conformidade parcial. O repositório obtém êxito no uso de linguagens formais legíveis por máquina (XML/JSON). Contudo, falha no critério I2 devido à ausência de sistemas de organização do conhecimento padronizados (tesauros/ontologias), reflexo direto de sua natureza multidisciplinar.
Reusable **(Reutilizável)**	**Moderado**	R1.01, R1.1-01 e R3.01-02	Conformidade parcial. A descrição do conteúdo e o licenciamento jurídico (Creative Commons) são bem estruturados. No entanto, o critério R2 é frágil (metadados de proveniência e versionamento insuficientes) e o R3.01-

			01 falha devido à dependência de templates de metadados generalistas em detrimento de perfis específicos de domínio.

Fonte: Elaborado pelas autoras.

Em suma, os dados empíricos sintetizados no Quadro 1 revelam que o repositório Deposita Dados caminha a passos largos rumo à consolidação das diretrizes da Ciência Aberta no cenário nacional, embora ainda enfrente as dores de crescimento típicas de plataformas multidisciplinares. O expressivo índice de 92% de conformidade global atesta o acerto da instituição ao investir em uma infraestrutura tecnológica capaz de blindar a recuperação e a acessibilidade dos seus ativos. Contudo, sob o prisma da Organização da Informação e do Conhecimento, o diagnóstico deixa claro que o futuro da plataforma depende da superação dos gargalos semânticos e de proveniência observados nos eixos de interoperabilidade e reuso. Ao desvelar esse cenário, esta investigação não apenas mapeia o nível de maturidade atual da infraestrutura brasileira, mas fornece os subsídios necessários para que gestores de informação possam arquitetar perfis de aplicação mais robustos, assegurando que o dado depositado hoje permaneça um insumo científico vivo, inteligível e passível de novas descobertas amanhã.

5 CONSIDERAÇÕES FINAIS

Esta investigação teve como propósito avaliar a conformidade das práticas de representação da informação e a qualidade dos metadados no repositório Deposita Dados à luz dos Princípios FAIR. Diante do crescimento exponencial da produção científica, o estudo buscou responder ao desafio de superar o paradoxo da "disponibilidade sem acessibilidade", identificando em que medida a infraestrutura comum de dados de pesquisa no Brasil atua como um agente facilitador de descobertas ou como um mero depósito passivo de arquivos digitais.

Os resultados obtidos por meio da ferramenta diagnóstica F-UJI revelaram um cenário promissor e, ao mesmo tempo, desafiador. O índice global de 92% de aderência aos Princípios FAIR atesta que o Deposita Dados cumpre com distinção o seu papel de "custódia democrática". Ao atingir o nível avançado nos eixos de **Localizável** (*Findable*) e Acessibilidade *(Accessible*), o repositório garante que os dados gerados por pesquisadores de instituições sem infraestrutura própria recebam identificadores persistentes e universais (DOIs e *Handles*), sejam indexados globalmente no Re3data e fiquem protegidos contra a obsolescência de links e a opacidade documental.

Contudo, as análises nos eixos de Interoperabilidade (I*nteroperable*) e Reuso (*Reusable*), ambos situados no nível moderado, revelam os limites práticos de uma abordagem descritiva generalista. Sob a ótica da Organização da Informação e do Conhecimento, evidenciou-se que a natureza multidisciplinar do repositório gera dificuldades intrínsecas para a adoção de Sistemas de Organização do Conhecimento específicos, como ontologias e tesauros (falha no subprincípio I2), bem como no emprego de padrões de metadados de domínio (falha no R3.01-01). Adicionalmente, a fragilidade na captura de metadados de proveniência (critério R2) reduz a transparência sobre o ciclo de vida e a linhagem dos dados (*data lineage*), criando ruídos que podem comprometer a reprodutibilidade científica.

Conclui-se que o Deposita Dados não é um repositório passivo, mas uma infraestrutura robusta que carece de camadas de mediação semântica. Para migrar do nível moderado para o avançado em sua totalidade, a governança da plataforma precisa evoluir para a incorporação de Perfis de Aplicação flexíveis e tecnologias da Web Semântica (como o mapeamento em SKOS e *Linked Open Data*). Essas soluções permitiriam conectar a base descritiva geral do repositório às especificidades terminológicas de cada comunidade científica, desonerando o usuário final do esforço de desambiguação.

Como limitações deste estudo, aponta-se o escopo amostral centrado em dez conjuntos de dados e o caráter ainda em desenvolvimento da ferramenta F-UJI, que pode apresentar restrições na leitura de especificidades regionais de sistemas não temáticos. Como recomendações para pesquisas futuras, sugere-se o desenvolvimento e a testagem de perfis de aplicação de metadados voltados a repositórios multidisciplinares, bem como investigações que avaliem a percepção dos pesquisadores brasileiros quanto às barreiras no preenchimento de metadados ricos no momento do depósito.

Em última análise, este estudo contribui para o *corpus* teórico da Ciência da Informação ao demonstrar que a consolidação da Ciência Aberta no Brasil transcende a engenharia de tecnologia da informação. Trata-se de um imperativo metodológico e político, onde a qualidade da representação da informação e do conhecimento assume o papel central de transformar arquivos armazenados em bens comuns efetivamente integráveis, reutilizáveis e emancipadores para toda a sociedade.

REFERÊNCIAS

ALBAGLI, Sarita; MACIEL, Maria Lucia. Novas condições de circulação e apropriação da informação e do conhecimento: questões no debate contemporâneo. *In*: ENCONTRO NACIONAL DE PESQUISA E OÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 10., 2009, João Pessoa. **Anais** [...]. João Pessoa: ANCIB, 2009. Disponível em: https://ridi.ibict.br/bitstream/123456789/314/1/ALBAGLIENANCIB2009.pdf. Acesso em: 18 jun. 2026.

ALVES, Rachel Cristina Vesú. **Metadados como elementos do processo de catalogação**. 2010. 132 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010. Disponível em: https://repositorio.unesp.br/handle/11449/103361. Acesso em: 18 jun. 2026.

BONETTI, Letícia Guarany; ARAKAKI, Ana Carolina Simionato. A importância das boas práticas em repositórios de dados de pesquisa. *In*: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 22., 2022, Porto Alegre. **Anais** [**...**]. Porto Alegre: ANCIB, 2022. Disponível em: https://enancib.ancib.org/enancib/pt_BR/article/view/24/24. Acesso em: 20 jun. 2026.

BORGMAN, Christine L. Research Data: who will share what, with whom, when and why. **RatSWD Working Paper**, [*s. l.*], n. 161, out. 2010. Disponível em: Disponível em: https://www.konsortswd.de/wp-content/uploads/RatSWD_WP_161.pdf. Acesso em: 18 jun. 2026.

DEPOSITA DADOS. [Página institucional]. Brasília: IBICT, 2026. Disponível em: https://depositadados.ibict.br/about.xhtml. Acesso em: 18 jun. 2026.

DEVARAJU, Anusuriya; HUBER, Robert. **F-UJI - An Automated FAIR Data Assessment Tool**. [S. l.]: Zenodo, 2020. Disponível em: https://doi.org/10.5281/zenodo.6361400. Acesso em: 18 jun. 2026.

DINIZ, Bárbara Carvalho; DIAS, Guilherme Ataíde; SILVA, Alzira Karla Araújo da. Os princípios FAIR como diretrizes de maximizar o uso e (re)uso dos dados de pesquisa: traçando um paralelo com a gestão de documentos arquivísticos digitais. *In*: WORKSHOP DE INFORMAÇÃO, DADOS E TECNOLOGIA, 2., 2018, João Pessoa. **Anais** [...]. João Pessoa: IBICT, 2018. Disponível em: https://widat.ibict.br/index.php/widat2023/article/view/142/123. Acesso em: 18 jun. 2026.

GÄAL, Lígia Pereira Muniz; MARTINS, Márcio Souza. Acesso aberto no contexto da pesquisa em Ciência da Informação. **Transinformação**, Campinas, v. 34, e220016, 2022. Disponível em: https://www.scielo.br/j/tinf/a/s9z43WHqpXbncwvzrgmDKFj/?lang=pt. Acesso em: 18 jun. 2026.

SANTOS, Raimunda Fernanda dos; FELIPE, Carla Beatriz Marques. Repositórios de dados de pesquisa à luz da organização e representação da informação. *In*: CONGRESSO DA SOCIEDADE INTERNACIONAL PARA A ORGANIZAÇÃO DO CONHECIMENTO, 8., 2025, Canela. **Anais** [**...**].

Canela: IKSO BRASIL, 2025. Disponível em: https://isko.org.br/ojs/index.php/iskobrasil/article/view/70. Acesso em: 17 jun. 2026.

SAYÃO, Luís Fernando; SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**, João Pessoa, v. 22, n. 3, 2012. Disponível em: https://periodicos.ufpb.br/index.php/ies/article/view/12224. Acesso em: 17 jun. 2026.

WILKINSON, Mark D.; DUMONTIER, Michel; AALBERSBERG, IJsbrand Jan; APPLETON, Gabrielle *et al*. The FAIR guiding principles for scientific data management and stewardship. **Scientific Data**, [*s. l.*], v. 3, p. 160018, 2016. Disponível em: https://www.nature.com/articles/sdata201618. Acesso em: 17 jun. 2026.