

ABISMO SOCIOTÉCNICO: BARREIRAS À INTELIGÊNCIA ARTIFICIAL E DADOS LIGADOS EM REPOSITÓRIOS

SOCIOTECHNICAL ABYSS: BARRIERS TO ARTIFICIAL INTELLIGENCE AND LINKED DATA IN REPOSITORIES

XIV WICI 2026
WORKSHOP INTERNACIONAL
EM CIÊNCIA DA INFORMAÇÃO

FERNANDA CHRISTINA DA COSTA FIGUEIREDO

Universidade Federal de Minas Gerais (UFMG)

ORCID: <https://orcid.org/0000-0003-4036-6923>

ROR: <https://ror.org/0176yjwt32>

MODALIDADE: RESUMO EXPANDIDO

Resumo: A comunicação científica enfrenta o desafio de transformar repositórios institucionais em ecossistemas inteligentes. O objetivo deste estudo é analisar as barreiras estruturais, instrumentais, políticas e interdisciplinares que inviabilizam a padronização de dados e o uso de Inteligência Artificial (IA) e Dados Abertos Conectados (*Linked Open Data* - LOD). A metodologia caracteriza-se como uma pesquisa de abordagem qualitativa, com objetivo descritivo e de natureza aplicada, fundamentada em revisão bibliográfica e análise de conteúdo. Os resultados indicam que a estagnação decorre da ausência de políticas obrigatórias, obsolescência de sistemas, falhas na indexação e distanciamento sociotécnico entre equipes. Conclui-se que superar esse cenário exige letramento algorítmico, representação colaborativa e governança efetiva.

Palavras-chave: Repositórios Institucionais; Inteligência Artificial; Linked Open Data.

Abstract: Scientific communication faces the challenge of transforming institutional repositories into intelligent ecosystems. The objective of this study is to analyze the structural, instrumental, political, and interdisciplinary barriers that hinder data standardization and the use of Artificial Intelligence (AI) and Linked Open Data (LOD). The methodology is characterized as a qualitative research, with a descriptive objective and applied nature, based on bibliographic review and content analysis. The results indicate that stagnation stems from the lack of mandatory policies, system obsolescence, indexing failures, and the sociotechnical gap between teams. Overcoming this scenario requires algorithmic literacy, collaborative representation, and effective governance.

Keywords: Institutional Repositories; Artificial Intelligence; Linked Open Data.



1 INTRODUÇÃO

A comunicação científica brasileira evoluiu com o Acesso Aberto (Targino; Garcia; Paiva, 2014), impulsionada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) na adoção de Repositórios Institucionais (RIs) (Medeiros; Ferreira, 2014; Shintaku; Vechiato, 2018). Contudo, o modelo engessou-se no armazenamento passivo de textos, chocando-se com a *e-Science*, que exige o reúso de dados brutos e computacionais (Rodrigues; Dias, 2025).

Ao analisar a infraestrutura desses ambientes, diagnostica-se uma estagnação tecnológica: os RIs operam como “depósitos digitais” isolados (Rodrigues; Dias, 2025; Sousa, 2025). Embora modelos de linguagem externa (LLMs) possam acessar e processar documentos textuais superficialmente, a arquitetura interna desses repositórios permanece passiva. Especificamente, os RIs brasileiros estão distantes da aplicação nativa de subáreas da Inteligência Artificial (IA) voltadas à Organização do Conhecimento, como o Processamento de Linguagem Natural (PLN) para indexação semiautomática e o *Machine Learning* para o desenvolvimento e atualização de ontologias (Oliveira; Santos, 2024).

A integração entre essas subáreas de IA e o *Linked Open Data* (LOD) representa uma inovação ainda pouco explorada na prática (Paula; Simão; Barbosa, 2020): enquanto a IA atua na modelagem e padronização temática dos metadados extraídos dos documentos, o LOD utiliza padrões da Web Semântica (como o *framework* RDF e URIs) para interligar esses metadados estruturados a grafos de conhecimento globais (Segundo, 2015).

Diante disso, o problema de pesquisa questiona: Quais barreiras estruturais, instrumentais, políticas e interdisciplinares bloqueiam a implementação de IA (focada em indexação) e LOD nos RIs brasileiros? O objetivo é mapear essas falhas que impedem a maturidade semântica necessária para a mineração nativa por máquinas. A justificativa pauta-se na urgência de romper o “mito da neutralidade técnica” (Sousa, 2025). Integrar IA e LOD exige dados rigorosamente estruturados, políticas mandatórias (Medeiros; Ferreira, 2014) e letramento algorítmico (Sousa, 2025), transformando os RIs em infraestruturas epistêmicas ativas.



2 DESENVOLVIMENTO

A metodologia deste estudo consiste em uma pesquisa de abordagem qualitativa, com objetivo descritivo e de natureza aplicada, baseada em revisão bibliográfica e documental. O *corpus* compreende publicações da Ciência da Informação focadas na implantação, indexação e transição de RIs para *e-Science* e LOD. Os dados foram tratados via análise de conteúdo interpretativa para identificar padrões. A partir dessa leitura transversal, as barreiras à inovação tecnológica foram categorizadas em três eixos.

A análise revela que a dificuldade de inserir os RIs brasileiros no ecossistema de IA e LOD decorre de fraturas históricas. No eixo político-estrutural, a implementação via indução *top-down* pelo IBICT não foi sustentada por políticas locais efetivas. A ausência de Políticas Institucionais de Informação (PIIs) mandatórias causa o esvaziamento dos RIs, pois o depósito voluntário falha diante da resistência dos pesquisadores (Medeiros; Ferreira, 2014). Somado ao desinteresse das instâncias superiores, muitos RIs sequer constam nos Planos de Desenvolvimento Institucional (PDI) (Jesus *et al.*, 2021; Targino; Garcia; Paiva, 2014). Sem mandatos, faltam dados volumosos, essenciais para treinar algoritmos locais.

No eixo sociotécnico, o cenário é marcado pela obsolescência do *software* livre hegemônico, o *DSpace* (Sousa; Campos; Amaro, 2024). Esse fenômeno é causado pela falta de letramento técnico cruzado entre as equipes de Biblioteconomia e Tecnologia da Informação (TI), o que paralisa atualizações (Jesus *et al.*, 2021). Ademais, há uma cultura nociva de não compartilhamento de códigos-fonte; as equipes de TI que desenvolvem extensões não as socializam, inviabilizando a evolução em rede (Shintaku; Vechiato, 2018).

No eixo metodológico, a transição para a Web Semântica é bloqueada por um déficit agudo na indexação: a esmagadora maioria dos RIs não possui políticas de indexação visíveis e usa termos livres, impedindo a formação de ontologias (Oliveira; Santos, 2024). Consequentemente, os RIs estagnam no nível de 3 estrelas do LOD, publicando PDFs isolados e falhando em adotar padrões semânticos (como o *framework* RDF e URIs) imprescindíveis para interligar dados a bases globais (Paula; Simão; Barbosa, 2020; Segundo, 2015)



3 CONSIDERAÇÕES FINAIS

A pesquisa diagnosticou que o bloqueio à IA (como PLN e *Machine Learning*) e ao LOD nos RIs brasileiros decorre de um profundo abismo sociotécnico. A ausência de mandatos (Medeiros; Ferreira, 2014), o distanciamento entre Biblioteconomia e TI (Jesus *et al.*, 2021) e o déficit na indexação (Oliveira; Santos, 2024) mantêm os RIs invisíveis para a mineração semântica (Paula; Simão; Barbosa, 2020).

Para superar a estagnação rumo à *e-Science* (Rodrigues; Dias, 2025), propõe-se um modelo estratégico em três fases: 1) Governança e Letramento: Aprovação imediata de PIs mandatórias para povoamento e promoção do letramento algorítmico das equipes interdisciplinares (Sousa, 2025); 2) Representação Híbrida: Adoção de ferramentas de IA (Processamento de Linguagem Natural e *Machine Learning*) na base do sistema para apoiar a indexação semiautomática e a atualização dinâmica de ontologias, superando o uso exclusivo da linguagem natural (Oliveira; Santos, 2024); 3) Interligação Semântica: Superação da obsolescência do *software* (Sousa; Campos; Amaro, 2024) para implementar padrões da Web Semântica (RDF/URIs), elevando o RI a 5 estrelas no LOD (Segundo, 2015), permitindo que a inteligência da máquina cruze os dados localmente e com a nuvem global.

Conclui-se que o sucesso futuro na comunicação científica não deve ser medido apenas pelo volume de PDFs, mas pela capacidade estrutural da infraestrutura em gerenciar o reuso de dados brutos e garantir sua conectividade algorítmica e global.



REFERÊNCIAS

JESUS, Deise Lourenço de; SOUSA, Angelica; DETONI, Juliana; CUNHA, Murilo. Barreiras no processo de desenvolvimento de repositórios institucionais nos institutos federais do Brasil. **Brazilian Journal of Information Studies**, [S. l.], v. 15, publicação contínua, p. 1-26, 2021. DOI: <https://doi.org/10.36311/1981-1640.2021.v15.e02111>. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/11634>. Acesso em: 10 fev. 2026.

MEDEIROS, Simone Assis; FERREIRA, Patrícia Aparecida. Política pública de acesso aberto à produção científica: um estudo sobre a implementação de repositórios institucionais em instituições de ensino superior. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 4, n. 2, p. 195-217, jul./dez. 2014. DOI: <https://doi.org/10.22478/ufpb.1981-0695.2015v10n1.23266>. Disponível em: <https://periodicos.ufpb.br/index.php/pbcib/article/view/23266>. Acesso em: 10 fev. 2026.

OLIVEIRA, Thamires Nascimento de; SANTOS, Raimunda Fernanda dos. Indexação e recuperação em repositórios institucionais brasileiros: recomendações de melhorias para a difusão do conhecimento científico. **Biblos: Revista do Instituto de Ciências Humanas e da Informação**, Rio Grande, v. 38, n. 1, p. 139-160, jan./jun. 2024. DOI: <https://doi.org/10.63595/biblos.v38i1.17881>. Disponível em: <https://periodicos.furg.br/biblos/article/view/17881>. Acesso em: 08 fev. 2026.

PAULA, Luciano Bernardes de; DEUS, William Simão de; BARBOSA, Ellen Francine. Análise de repositórios de REAs em relação ao uso dos padrões de Linked Open Data. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO (SBIE), 31., 2020. **Anais [...]**. [S. l.]: CBIE, 2020. p. 312-321. DOI: <https://doi.org/10.5753/cbie.sbie.2020.312>. Disponível em: <https://sol.sbc.org.br/index.php/sbie/article/view/12787>. Acesso em: 21 jan. 2026.

RODRIGUES, Adriana Alves; DIAS, Guilherme Ataíde. Repositórios de dados abertos de pesquisa em plataformas de universidades brasileiras. **Ciência da Informação em Revista**, [S. l.], v. 12, p. 1-12, 2025. DOI: [10.28998/cirev.2025v12e17799](https://doi.org/10.28998/cirev.2025v12e17799). Disponível em: <https://periodicos.ufal.br/cir/article/view/17799>. Acesso em: 27 jul. 2026.

SANTAREM SEGUNDO, Jose Eduardo. Web semantica, dados ligados e dados abertos: uma visão dos desafios do brasil frente às iniciativas internacionais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, [S. l.], v. 8, n. 2, p. 219-239, 2015. Disponível em: <https://revistas.ancib.org/tpbci/article/view/359>. Acesso em: 27 jul. 2026.

SHINTAKU, Milton; VECHIATO, Fernando Luiz. Histórico do uso do DSpace no Brasil com foco na tecnologia. **Revista Informação na Sociedade Contemporânea**, Natal, v. 2, n. 1, p. 1-16, jan./jun. 2018. DOI: <https://doi.org/10.21680/2447-0198.2018v2n0ID13097>. Disponível em: <https://periodicos.ufrn.br/informacao/article/view/13097>. Acesso em: 27 jul. 2026.

SOUSA, Juliana Araujo Gomes de; CAMPOS, Phillipe de Freitas; AMARO, Bianca. Evolução dos repositórios institucionais brasileiros construídos com o DSpace: contexto histórico. *In*: CONFERÊNCIA LUSÓFONA DE CIÊNCIA ABERTA (ConfOA), 15., 2024. **Anais** [...]. [S. l.], 2024. DOI: <https://doi.org/10.18225/ci.inf.v53i3.7220>. Disponível em: <https://brapci.inf.br/v/313780>. Acesso em: 29 jul. 2026.

SOUSA, Nuno Miguel Teixeira. Academic libraries as hubs of artificial intelligence: a critical and epistemic approach. **Discover Artificial Intelligence**, [S. l.], v. 5, n. 221, p. 1-11, 2025. Disponível em: <https://link.springer.com/article/10.1007/s44163-025-00490-8>. Acesso em: 29 jul. 2026.

TARGINO, Maria das Graças; GARCIA, Joana Coeli Ribeiro; PAIVA, Maria José Rodrigues. Repositórios institucionais brasileiros: entre o sonho e a realidade. **Revista FSA**, Teresina, v. 11, n. 1, p. 117-133, jan./mar. 2014. Disponível em: <http://www4.unifsa.com.br/revista/index.php/fsa/article/view/320>. Acesso em: 29 jul. 2026.