Relatório final de atividades
Projeto: SAVITS: Sistema de Avaliação e Validação de Impactos em Tecnologias Sociais

Coordenador do Projeto:	**Washington Luís Ribeiro de Carvalho Segundo**		
Coordenação:	**Coordenação-Geral de Informação Científica e Técnica - CGIC**		
Nome do(a) pesquisador(a):	**Gabriel Marcos Batista Guimarães**		
Vigência da Bolsa de Pesquisa:	**01/05/2026**	**Até:**	**30/10/2026**
Encerramento em:	**30/10/2026**		
Ref. Fundação:	**32725**	**Ref. Ibict:**	**0008/2026-81**

Descrição das Atividades da Bolsa de Pesquisa:

1. Realizar estudo conceitual e benchmarking sobre repositórios digitais, acessibilidade e IA aplicada. 2. Entrega de relatório das atividades desenvolvidas.

1 INTRODUÇÃO

Os repositórios digitais constituem importantes instrumentos para a organização, preservação e disseminação da produção científica e acadêmica, contribuindo para ampliar o acesso à informação e fortalecer os processos de comunicação científica. Nesse contexto, os repositórios institucionais assumem papel relevante na disponibilização da produção intelectual das instituições de ensino e pesquisa (Lynch, 2003).

No contexto da democratização do acesso à informação, a acessibilidade digital apresenta-se como um aspecto fundamental para garantir que conteúdos e funcionalidades disponibilizados na Web possam ser utilizados por diferentes perfis de usuários. Nesse sentido, as Diretrizes de Acessibilidade para Conteúdo Web (WCAG) estabelecem recomendações voltadas à redução de barreiras de acesso, considerando aspectos relacionados à percepção, operação, compreensão e compatibilidade dos conteúdos digitais (W3C, 2023).

Paralelamente às questões de acessibilidade, a aplicação de Inteligência Artificial (IA) tem se consolidado como um recurso tecnológico promissor para elevar a eficiência operacional dos repositórios, constituindo outro eixo central desta pesquisa. A utilização de IA nesses ambientes pode abranger atividades como a criação e o enriquecimento de metadados, a extração automática de informações e o apoio à organização e recuperação de conteúdos. Essas aplicações apresentam potencial para otimizar processos de gerenciamento e facilitar a descoberta de informações pelos usuários (Silva, 2023).

Diante dessas transformações, torna-se relevante investigar as características dos repositórios digitais e as possibilidades de aplicação de recursos de acessibilidade e Inteligência Artificial nesses ambientes. Assim, este estudo tem como objetivo realizar um levantamento conceitual e um *benchmarking* de repositórios digitais, buscando identificar características, boas práticas e recursos que possam contribuir para o aprimoramento desses ambientes de informação.

2 MÉTODO

Para o desenvolvimento da atividade, foi realizada uma pesquisa de caráter exploratório e descritivo, com abordagem qualitativa. Conforme Prodanov e Freitas (2013), a pesquisa exploratória tem como objetivo proporcionar maior familiaridade com o problema, enquanto a pesquisa descritiva observa, registra e analisa as características de um determinado fenômeno. Essa estrutura alinha-se ao objetivo de identificar e analisar conceitos, práticas e recursos relacionados a repositórios digitais, acessibilidade e Inteligência Artificial. O trabalho foi estruturado a partir de levantamento bibliográfico e análise comparativa de plataformas.

Inicialmente, foi realizado um levantamento bibliográfico em fontes acadêmicas e institucionais. Esse procedimento, segundo Prodanov e Freitas (2013), é desenvolvido a partir de material já elaborado, constituído principalmente de artigos científicos e documentos técnicos, sendo indispensável para fundamentar a pesquisa em bases teóricas sólidas. A busca visou identificar referências relacionadas aos repositórios digitais, à acessibilidade em ambientes virtuais e às possibilidades de aplicação de Inteligência Artificial na organização da informação.

Em seguida, foram definidos critérios para a realização do benchmarking, estruturado aqui como uma técnica de análise comparativa voltada à identificação das melhores práticas do setor para conduzir a um desempenho superior dos acervos. A avaliação considerou três dimensões principais: características dos repositórios digitais, acessibilidade e aplicação de Inteligência Artificial. Para os repositórios, foram observados mecanismos de busca, organização das informações e metadados. Na dimensão de acessibilidade, avaliaram-se a navegação e a apresentação dos conteúdos, tendo como referência normativa as Diretrizes de Acessibilidade para Conteúdo Web (W3C, 2023). Quanto à Inteligência Artificial, o foco recaiu sobre recursos de geração e enriquecimento de metadados, recuperação da informação e automação de processos.

Posteriormente, foram selecionados repositórios digitais para a aplicação dos critérios estabelecidos. As plataformas foram analisadas de forma comparativa, buscando identificar características em comum, diferenças, boas práticas e recursos relevantes para o aprimoramento de ambientes de informação digital institucionais.

Por fim, as informações obtidas foram organizadas em quadros, tabelas e figuras, permitindo sistematizar os resultados do levantamento e do benchmarking. A análise qualitativa dos dados possibilitou identificar tendências e oportunidades relacionadas à acessibilidade e à aplicação de IA em repositórios digitais, fornecendo os subsídios necessários para as discussões apresentadas na seção de resultados.

3 RESULTADOS ALCANÇADOS

A partir dos procedimentos descritos na seção anterior, foram obtidos resultados relacionados às características dos repositórios digitais, aos aspectos de acessibilidade e às possibilidades de aplicação de Inteligência Artificial nesses ambientes. O levantamento bibliográfico permitiu identificar conceitos e práticas relevantes para a compreensão do cenário atual dos repositórios, enquanto o benchmarking possibilitou comparar diferentes plataformas a partir dos critérios estabelecidos.

Inicialmente, o levantamento permitiu identificar que os repositórios digitais apresentam funcionalidades voltadas principalmente à organização, armazenamento, preservação e disseminação de objetos digitais. Entre os recursos observados, destacam-se mecanismos de busca e recuperação da informação, organização dos conteúdos por comunidades ou coleções e utilização de metadados para descrição dos documentos.

Para sistematizar os principais aspectos considerados na análise, o Quadro 1 apresenta as dimensões e os critérios utilizados no benchmarking.

Quadro 1 - Dimensões e critérios utilizados no benchmarking

Dimensão	Critérios analisados
Repositórios digitais	Busca, organização dos conteúdos e metadados.
Acessibilidade	Navegação, apresentação dos conteúdos e interação.
Inteligência Artificial	Geração e enriquecimento de metadados, recuperação da informação e automação.

Fonte: Autoria própria (2026).

Assim, a aplicação desses critérios permitiu uma avaliação padronizada das soluções estudadas. Os resultados obtidos em cada uma dessas dimensões são discutidos nos tópicos subsequentes.

3.1 Repositórios Digitais

Para a realização deste benchmarking, foram analisadas quatro plataformas referenciais: DSpace (versões recentes baseadas em Angular), Zenodo, Dataverse e EPrints. O Quadro 2 consolida as características operacionais e o potencial tecnológico avaliado nesses sistemas.

Quadro 2 - Características estruturais e potencial de inovação (IA)

Repositório	Mecanismo de Busca	Organização Interna	Metadados	Integração com IA
DSpace	Filtros facetados (Apache Solr)	Comunidades e Coleções	Dublin Core	Busca semântica (RAG) e extração inteligente de metadados
Zenodo	Elasticsearch	Comunidades temáticas	JSON Schema / DataCite	Automação preditiva de curadoria e metadados (Projeto AIRDEC)
Dataverse	Apache Solr	Dataverses e Datasets	Padrão DDI	Assistentes de linguagem natural (AutoSage/DataChat) e APIs de LLMs
EPrints	Busca booleana simples	Árvore de assuntos	Esquemas legados	Inexistente nativamente (exige plugins e scripts externos em Perl)

Fonte: Autoria própria (2026).

A avaliação demonstrou que as plataformas mais modernas (DSpace e Zenodo) possuem os melhores mecanismos de busca, sustentados por indexadores robustos como o Apache Solr e Elasticsearch, que permitem filtros facetados em tempo real e a recuperação precisa por identificadores persistentes (como o DOI).

Quanto à organização dos documentos, o DSpace e o EPrints seguem a hierarquia institucional clássica (Comunidades, Coleções e Itens), enquanto o Zenodo adota uma abordagem mais flexível baseada em comunidades temáticas abertas. O Dataverse, por sua vez, diferencia-se por ser otimizado para a gestão de conjuntos de dados de pesquisa (datasets), organizando os conteúdos de forma a facilitar a reprodutibilidade científica. Todos trabalham rigorosamente com metadados para garantir a interoperabilidade (especialmente via protocolo OAI-PMH), mas enquanto o DSpace foca no padrão universal Dublin Core, o Zenodo utiliza o DataCite e o Dataverse prioriza esquemas descritivos voltados para dados (DDI).

Ilustrando a aplicação prática dessas características estruturais, a Figura 1 apresenta a interface de busca do DSpace, com seu front-end baseado no framework Angular, destacando recursos de navegação e a disposição dos filtros laterais, com facetas geradas pelo motor Apache Solr, fundamental para recuperação precisa de informações.

Figura 1 - Interface de busca facetada e organização de itens no DSpace



Fonte: Captura de tela (2026).

Dessa maneira, os filtros visuais e a estrutura de navegação mostram na prática o atendimento aos requisitos do sistema, facilitando a busca e permitindo que qualquer pessoa explore o acervo de forma simples e intuitiva.

3.2 Acessibilidade

A análise da usabilidade inclusiva teve como parâmetro as recomendações das Diretrizes de Acessibilidade para Conteúdo Web (WCAG), resultando na avaliação sintetizada no Quadro 3.

Quadro 3 - Avaliação de critérios de acessibilidade (WCAG)

Critério	DSpace	Zenodo	Dataverse	EPrints
Navegação por teclado	Atende	Atende	Atende	Parcial
Contraste e Texto Alt.	Atende	Atende	Parcial (Falhas em listagens)	Baixo (Falhas de contraste)
Estrutura semântica	Parcial (devido a ajustes ARIA)	Parcial	Atende	Não atende

Fonte: Autoria própria (2026).

A comparação evidencia que a acessibilidade está diretamente atrelada à modernidade da arquitetura de front-end das plataformas. Repositórios como Zenodo e as versões atuais do DSpace entregam alta conformidade, pois suas interfaces já são desenvolvidas com bibliotecas de componentes semanticamente estruturados. Eles atendem aos requisitos de navegação por teclado, contraste e suporte a tecnologias assistivas (leitores de tela). Em contrapartida, o EPrints e instâncias não personalizadas do Dataverse apresentam barreiras, especialmente na navegação sem mouse e na ausência de atributos estruturais profundos, demandando customizações manuais para atingirem plena conformidade.

3.3 Inteligência Artificial Aplicada

A partir da coluna de inovação mapeada no Quadro 2, nota-se que a aplicação de IA nesses ecossistemas ainda ocorre de forma assimétrica. A prática mais madura observada concentra-se no tratamento do back-end informacional, especificamente na extração automática de metadados de arquivos depositados (recurso em expansão no Dataverse e em instâncias customizadas do DSpace) e em motores de recomendação de

publicações (Zenodo). Essa automação minimiza falhas humanas de digitação durante o autoarquivamento e eleva a eficiência da gestão institucional.

3.4 Síntese do Benchmarking

Em suma, mostra-se que o aprimoramento contínuo dos repositórios está diretamente atrelado à renovação de sua infraestrutura. O DSpace e o Zenodo destacaram-se nos resultados, demonstrando que a transição de interfaces antigas para modelos orientados a APIs resolve barreiras técnicas de acessibilidade. Além disso, essa modernização prepara as instituições para incorporar, de forma padronizada e segura, as novas ferramentas de Inteligência Artificial, assegurando o cumprimento pleno da democratização do conhecimento.

4 CONSIDERAÇÕES FINAIS

O presente estudo, juntamente com o benchmarking realizado, possibilitou uma análise do cenário atual dos repositórios digitais. Os resultados mostraram que a evolução dessas plataformas não depende apenas da modernização de suas interfaces, mas também das tecnologias que sustentam seu funcionamento. Nesse contexto, arquiteturas baseadas em APIs, presentes em versões recentes do DSpace e no Zenodo, contribuem para a integração entre sistemas, o uso de metadados padronizados e uma recuperação mais eficiente das informações.

Em relação à acessibilidade digital, observou-se que a adoção das diretrizes WCAG está relacionada à estrutura tecnológica das plataformas. Sistemas mais recentes tendem a integrar recursos como navegação por teclado e estrutura semântica adequada, enquanto plataformas mais antigas podem exigir intervenções para eliminar barreiras de acesso. A Inteligência Artificial também se apresenta como um recurso relevante, principalmente na curadoria e na extração automática de metadados, contribuindo para tornar o gerenciamento e a recuperação da produção científica mais eficientes.

Dessa forma, conclui-se que a integração entre tecnologias modernas, acessibilidade e Inteligência Artificial pode contribuir para o fortalecimento dos repositórios digitais. Além de preservar e organizar a produção científica, essas plataformas devem garantir o acesso à informação de forma inclusiva e eficiente, acompanhando as necessidades tecnológicas e dos diferentes públicos que utilizam esses sistemas.

REFERÊNCIAS

CERN; OPENAIRE. Zenodo. Genebra: CERN, 2013-. Disponível em: https://zenodo.org/. Acesso em: 7 set. 2026.

ELASTIC. Elasticsearch: o mecanismo de busca e análise distribuído. Mountain View: Elastic, [202-?]. Disponível em: https://www.elastic.co/pt/elasticsearch/. Acesso em: 8 set. 2026.

IQSS. The Institute for Quantitative Social Science. The Dataverse Project. Cambridge: Harvard University, [202-?]. Disponível em: https://dataverse.org/. Acesso em: 7 set. 2026.

LYNCH, Clifford A. Institutional repositories: essential infrastructure for scholarship in the digital age. portal: Libraries and the Academy, v. 3, n. 2, p. 327-336, 2003. DOI: 10.1353/pla.2003.0039.

LYRASIS. DSpace. [S. l.]: Lyrasis, [202-?]. Disponível em: https://dspace.lyrasis.org/. Acesso em: 12 set. 2026.

PRODANOV, C. C.; FREITAS, E. C. de. Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico. 2. ed. Novo Hamburgo: Feevale, 2013.

SILVA, P. N. Recuperação de informação na Ciência da Informação: produção acadêmico-científica brasileira (2012-2021). Transinformação, Campinas, v. 35, e237336, 2023. Disponível em: https://doi.org/10.1590/2318-0889202335e237336. Acesso em: 7 set. 2026.

THE APACHE SOFTWARE FOUNDATION. Apache Solr. Forest Hill: The Apache Software Foundation, [202-?]. Disponível em: https://solr.apache.org/. Acesso em: 12 set. 2026.

UNIVERSITY OF SOUTHAMPTON. EPrints: digital repository software. Southampton: University of Southampton, [202-?]. Disponível em: https://www.eprints.org/. Acesso em: 7 set. 2026.

W3C. World Wide Web Consortium. Web Content Accessibility Guidelines (WCAG) 2.2. W3C Recommendation, 5 out. 2023. Disponível em: https://www.w3.org/TR/WCAG22/. Acesso em: 7 set. 2026.

Documento assinado digitalmente
GABRIEL MARCOS BATISTA GUIMARAES
Data: 22/10/2026 14:38:35-0300
Verifique em https://validar.iti.gov.br

Gabriel Marcos Batista Guimarães

Pesquisador

De acordo:

Documento assinado digitalmente
WASHINGTON LUIS RIBEIRO DE CARVALHO SEG
Data: 23/10/2026 10:05:08-0300
Verifique em https://validar.iti.gov.br

Washington Luís Ribeiro de Carvalho Segundo
Coordenador do projeto