

ISSN 3085-5624

Eixo Temático 2 - Informação, Comunicação e Processos Tecnológicos

**INDEXAÇÃO AUTOMÁTICA E ENCONTRABILIDADE DA INFORMAÇÃO:
como pilares para a eficiência dos repositórios institucionais*****AUTOMATIC INDEXING AND INFORMATION FINDABILITY:
pillars for the efficiency of institutional repositories***

Cíntia Holanda – Universidade Federal de Pernambuco (UFPE) – cmsholanda@gmail.com –
Orcid: <https://orcid.org/0009-0005-7217-3546>

Remi Lapa – Universidade Federal de Pernambuco (UFPE) – remi.lapa@ufpe.br – Orcid:
<https://orcid.org/0009-0000-7333-0121>

Modalidade: Trabalho completo

Resumo: O estudo tem como objetivo analisar a relação entre a indexação automática e a encontrabilidade da informação em repositórios institucionais. A metodologia envolve uma revisão bibliográfica sobre encontrabilidade e indexação automática, utilizando bases de dados nacionais e internacionais. Os resultados destacam a importância dos metadados na ligação entre o sujeito informacional e a informação necessária, e na descrição do conteúdo dos documentos a serem indexados, melhorando a precisão da recuperação e o acesso à informação em ambientes digitais.

Palavras-chave: indexação automática; encontrabilidade da informação; repositórios institucionais; recuperação da informação.

Abstract: *The study aims to analyze the relationship between automatic indexing and the findability of information in institutional repositories. The methodology involves a literature review on findability and automatic indexing, using national and international databases. The results highlight the importance of metadata in linking the informational subject to the necessary information and in describing the content of the documents to be indexed, improving the accuracy of retrieval and access to information in digital environments.*

Keywords: *automatic indexing; information findability; institutional repositories; information retrieval.*

1 INTRODUÇÃO

A crescente produção e disponibilidade de dados digitais têm redefinido o acesso e a gestão do conhecimento. Vechiato e Vidotti (2014) destacam a relação direta entre os avanços nas Tecnologias de Informação e Comunicação (TICs) e a ampliação das oportunidades de recuperação, acesso e uso da informação, medida pela ligação e o relacionamento entre dados e informações em diferentes dispositivos e sistemas, bem como a colaboração dos

sujeitos informacionais na produção e organização da informação e do conhecimento em ambientes informacionais digitais.

Nessa perspectiva, em meio à efervescente produção acadêmica, científica e cultural, os Repositórios Institucionais (RIs) se configuram como elementos cruciais no que tange à salvaguarda e disseminação de dados e informações. Os RIs, como interfaces digitais, desempenham um papel estratégico ao possibilitar que instituições acadêmicas e culturais possam acumular, estruturar e fornecer uma ampla diversidade de materiais, abrangendo desde teses e dissertações até expressões artísticas e criações científicas (Borges et al., 2019).

A definição de repositório por Romani, Fusco e Santos (2010), compreende um banco de dados que oferece serviços cruciais de divulgação, disponibilização e preservação de documentos, independentemente de estarem em formato digital ou digitalizado. Tais plataformas abrigam uma diversidade de obras em diferentes formatos, com o acesso viabilizado por meio de redes computacionais. Esta estrutura desempenha um papel fundamental ao facilitar a gestão e a avaliação da expressão intelectual das organizações, contribuindo de maneira substantiva para a promoção de suas atividades de pesquisa e ensino. A integração eficaz da gestão de conteúdo digital com os objetivos acadêmicos e científicos fortalece a relevância dessas instituições em um cenário cada vez mais voltado para o conhecimento.

A disseminação dos RIs no Brasil, principalmente nas Instituições de Ensino Superior (IES), consolidou-se a partir de 2002, impulsionada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), que promoveu ações para prover ferramentas tecnológicas às IES. A visibilidade da produção científica foi ampliada, evidenciando o papel dos RIs no armazenamento, na preservação e na disponibilização, por meio do acesso livre, da produção técnico-científica formada pelos docentes, técnicos-administrativos e discentes da instituição competente (Custódio; Vechiato, 2017).

Contudo, o aumento exponencial de informações nos Repositórios Institucionais apresenta o desafio da representação precisa do conteúdo, impactando a Encontrabilidade da Informação (EI) e a Recuperação da Informação (RI). Ambos os processos compartilham elementos similares, como: representação, organização, armazenamento e acesso (Roa-Martínez, 2019). Assim, é crucial investigar abordagens inovadoras e eficazes para superar os desafios decorrentes do aumento constante de dados e informações.

É relevante destacar que a Encontrabilidade da Informação pode ser percebida como uma etapa subsequente à Recuperação da Informação. A simples recuperação de conteúdo na *web* não assegura o efetivo encontro da informação desejada. A verdadeira essência da encontrabilidade reside na capacidade de localizar a informação desejada, atender às necessidades informacionais e, sobretudo, garantir que a assimilação da informação ocorra, viabilizando a criação de novos conhecimentos (Sanchez; Vechiato; Vidotti, 2019).

Na Ciência da Informação (CI), uma vertente de pesquisa compromete-se a otimizar o acesso e a utilização da informação, concentrando-se em aspectos de Usabilidade e Encontrabilidade. É uma abordagem que busca tornar a informação mais acessível e compreensível para atender às demandas informacionais dos usuários (Vidotti *et al.*, 2017). De acordo com essa visão, são desenvolvidas estratégias, técnicas e sistemas que visam facilitar a interação dos usuários com a informação, garantindo uma experiência informacional mais eficiente e satisfatória. São esforços que não apenas promovem o acesso, mas contribuem, também, para a eficácia na utilização da informação em ambientes digitais.

A Encontrabilidade da Informação trata da habilidade de localizar informações relevantes de maneira rápida e precisa dentro de vastos conjuntos de dados. É um conceito, relativamente recente, que se refere à qualidade com que a informação pode ser localizável ou navegável, é o grau no qual um determinado objeto é facilmente encontrado ou localizado e o grau no qual um sistema ou ambiente suporta a navegação e a recuperação, sendo um dos elementos cruciais para a eficácia na recuperação da informação (Morville, 2005 e Vechiato; Vidotti, 2014). Segundo Morville (2005), a Encontrabilidade manifesta-se quando se procura por informações através da navegação em um site (*web site*) ou por meio das estratégias de pesquisa empregadas em um mecanismo de busca (*search engine*).

A indexação é um processo fundamental que consiste em atribuir termos a documentos, facilitando sua recuperação e organização em sistemas de informação, independentemente de serem físicos ou digitais. Na CI, existem diversos modelos de indexação, abrangendo desde a tradicional indexação manual até abordagens mais contemporâneas que exploram Tecnologias de Informação e Comunicação, como a Indexação Automática (IA) e semiautomática (Santos, 2017). Segundo Martins, Monard e Matsubara (2003), as abordagens mais avançadas de indexação empregam técnicas de aprendizado de máquina para analisar e extrair automaticamente informações relevantes dos documentos,

aprimorando assim o processo de recuperação.

Neste estudo, a Indexação destaca-se como ferramenta indispensável para a eficiente localização e utilização de informações nos Repositórios Institucionais, especialmente no controle terminológico, gestão e acesso a recursos acadêmicos e de pesquisa. Ao considerar a Encontrabilidade da Informação (EI) e os conceitos inerentes à Indexação, a aplicação da Indexação Automática emerge como fator significativo para a melhoria da precisão da recuperação, redução da sobrecarga dos usuários e aprimoramento do acesso à informação em ambientes informacionais digitais, especialmente nos RIs. Essa abordagem é percebida como facilitadora essencial, proporcionando aos usuários acesso mais eficaz aos documentos técnico-científicos e aprimorando a recuperação desses materiais armazenados nos RIs. Dessa maneira, orientada pela perspectiva da Indexação Automática, esta pesquisa busca refletir sobre a influência da IA no aprimoramento da Encontrabilidade nos RIs, destacando seu papel essencial no processo de Recuperação da Informação.

2 ENCONTRABILIDADE DA INFORMAÇÃO EM AMBIENTES DIGITAIS: EXPLORANDO CONCEITOS E PRÁTICAS

A pesquisa sobre “Encontrabilidade da Informação” se origina das contribuições de Morville (2005) sobre *findability*. Esse termo está vinculado à reflexão sobre como “encontrar informações” na *Web*, buscando aprimorar a forma como as informações são localizadas, apresentadas, organizadas e compreendidas pelos usuários (Campos; Sousa; Oliveira, 2021; Camossi; Rodas, 2023). O conceito científico integrado à Ciência da Informação, no Brasil, foi introduzido por Vechiato (2013), que apresentou os Atributos de Encontrabilidade da Informação (AEI), o Modelo de Encontrabilidade da Informação (MEI) e as Recomendações de Encontrabilidade da Informação (REI) (Camossi; Rodas, 2023), que são elementos que constituem a base fundamental para o estudo e aprimoramento da experiência informacional.

Os AEI são conceituados como características que fomentam as possibilidades de que usuários encontrem a informação, em um sistema ou ambiente informacional e que passam por inúmeras etapas, transitando pelo ciclo da informação, como a produção, a representação e da disseminação ao armazenamento e à preservação. Abrangem desde aspectos relacionados à representação da informação, organização e mediação com os sujeitos informacionais até os elementos de ubiquidade e usabilidade nos ambientes informacionais

(Vechiato; Vidotti, 2014; Vidotti *et al.*, 2017).

O Quadro 1 apresenta, de forma sistematizada, os AEI, oferecendo uma visão clara e compreensiva de como eles contribuem para a efetividade do processo de Encontrabilidade da Informação em diferentes configurações informacionais.

Quadro 1 - Atributos de Encontrabilidade da Informação (AEI)

Taxonomias Navegacionais	Estruturas <i>top-down</i> que organizam categorias para facilitar a navegação e descoberta de informações, com critérios como categorização coerente, controle terminológico, relacionamento hierárquico e multidimensionalidade.
Instrumentos de Controle Terminológico	Envolve vocabulários controlados, como tesouros e ontologias, para apoiar a representação de recursos informacionais.
Folksonomias	Relacionadas à organização social da informação, permitem a classificação de recursos informacionais, sendo utilizadas em estruturas <i>bottom-up</i> , ampliando as possibilidades de acesso.
Metadados	Representação dos recursos informacionais armazenados em bancos de dados para recuperação da informação.
Mediação dos Informáticos	Associação ao desenvolvimento de sistemas, dispositivos e interfaces usando linguagens computacionais.
Mediação dos Profissionais da Informação	Envolvimento de sujeitos institucionais na seleção, estruturação e disseminação da informação.
Mediação dos Sujeitos Informacionais	Ações infocomunicacionais realizadas pelos sujeitos em sistemas e ambientes informacionais.
Affordances	Incentivos e pistas nos objetos que possibilitam aos sujeitos realizar ações na interface do ambiente, relacionadas à orientação, localização, encontrabilidade, acesso e descoberta de informações.
Wayfinding	Associado à orientação espacial para facilitar a localização, encontrabilidade e descoberta de informações na interface do ambiente.
Descoberta de Informações	Condicionada aos atributos de Encontrabilidade, refere-se às facilidades oferecidas pela interface (navegação/mecanismos de busca) para encontrar informações adequadas às necessidades do sujeito.
Acessibilidade e Usabilidade	Relacionados à capacidade de o sistema permitir acesso equitativo à informação (acessibilidade) e facilidades inerentes ao uso da interface (usabilidade).
Intencionalidade	Enfatiza a importância das experiências e habilidades dos sujeitos informacionais no projeto de ambientes e sistemas de informação.
Mobilidade, Convergência e Ubiquidade	Associados ao meio ambiente externo aos sistemas informacionais, dinamizando e potencializando as possibilidades dos sujeitos em encontrar informações por meio de diferentes dispositivos e em diferentes contextos e situações.

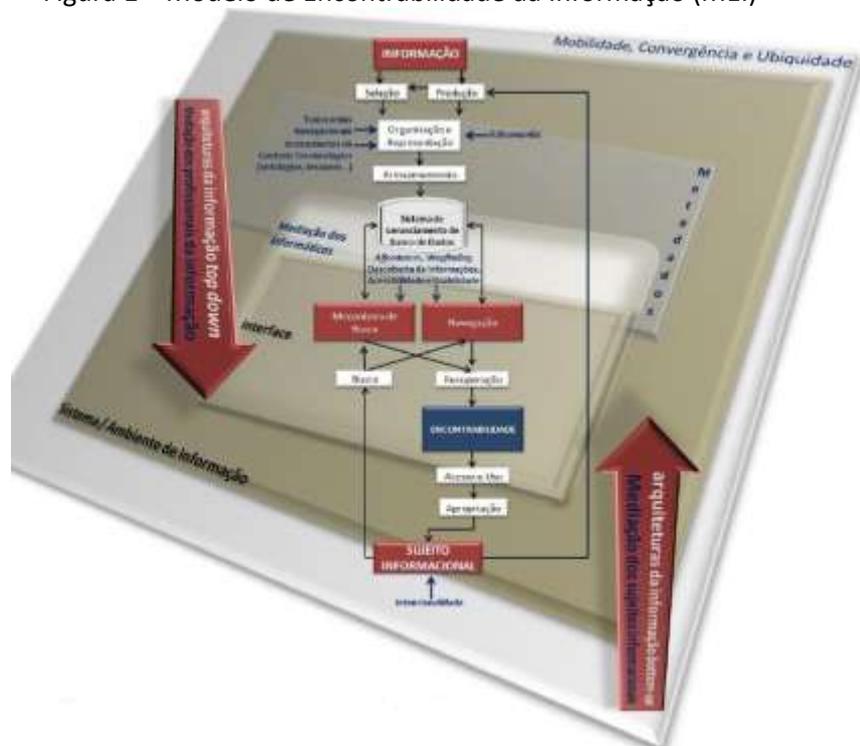
Fonte: Adaptado de Vechiato e Vidotti (2014).

A interação essencial entre usuário e informação é solidificada por meio desses atributos, destacando a importância de considerar cuidadosamente cada elemento durante o planejamento e otimização de ambientes informacionais. Essa abordagem estratégica, pode

não apenas oferecer perspectivas valiosas, mas também potencializar o desenvolvimento e o aprimoramento desses ambientes, resultando em uma troca ainda mais eficaz entre os sujeitos informacionais e a vasta quantidade de informações disponíveis.

O Modelo de Encontrabilidade da Informação representa os seguintes atributos da Encontrabilidade da Informação: taxonomias navegacionais; instrumentos de controle terminológico; folksonomias; metadados; mediação dos informáticos; *affordances*; *wayfinding*; descoberta de informações; acessibilidade e usabilidade; mediação dos profissionais da informação; mediação dos sujeitos informacionais; intencionalidade; e mobilidade, convergência e ubiquidade (Vechiato; Vidotti, 2014). Conforme delineado por Campos, Sousa e Oliveira (2021), o enfoque no sistema ou ambiente de informação é concebido para contemplar tanto uma arquitetura da informação 'de cima para baixo' (*top-down*) quanto uma arquitetura da informação 'de baixo para cima' (*bottom-up*). Essa dualidade é significativa em termos estratégicos para a organização e o processamento da informação (Figura 1).

Figura 1 – Modelo de Encontrabilidade da Informação (MEI)



Fonte: Vechiato e Vidotti (2014, p. 172).

O MEI pode ser aplicado em diversos sistemas de informação, ambientes

informacionais ou arquiteturas da informação, sendo influenciado pelo contexto sociotécnico emergente que conecta a intencionalidade dos sujeitos ao desenvolvimento tecnológico. Tal convergência altera substancialmente as formas de mediação de todos os atores, desde que devidamente considerada. A elaboração do MEI proporcionou a definição das Recomendações de Encontrabilidade da Informação, representando um suporte valioso para o desenvolvimento de protótipos (Vechiato, 2013) (Quadro 2).

Quadro 2 – Recomendações de Encontrabilidade da Informação (REI)

Recomendações	
1	Elaborar taxonomias navegacionais coerentes com a terminologia utilizada pelos sujeitos.
2	Implementar recursos de classificação social (folksonomia) e de navegação por meio das tags atribuídas pelos sujeitos.
3	Investir na mediação infocomunicacional dos sujeitos.
4	Ampliar as possibilidades dos sujeitos em encontrar a informação por meio de diferentes ambientes e dispositivos.
5	Investigar os comportamentos, as competências, as experiências e a Intencionalidade dos sujeitos.
6	Aplicar recomendações e avaliações de acessibilidade e de usabilidade.
7	Investir na utilização de affordances para orientar os sujeitos no espaço (wayfinding) no decorrer da navegação.
8	Investir na utilização de affordances para a query e os resultados de busca.
9	Proporcionar busca pragmática.

Fonte: Vechiato (2013).

As mediações dos sujeitos institucionais, sejam informáticos ou profissionais da informação, têm como foco facilitar a recuperação de informações, alinhando-se à necessidade do sujeito informacional de encontrar ou descobrir informações relevantes nos ambientes digitais. A conexão entre esses tipos de mediação e a intencionalidade do sujeito ocorre por meio da navegação via interface ou mecanismos de busca. Ao abordar, nesse contexto, o conceito de Encontrabilidade da Informação, refere-se a um processo subjetivo centrado no sujeito informacional, impulsionado por uma necessidade específica. Esse processo ocorre durante a navegação, destacando-se a interface com atributos como taxonomias navegacionais, *affordances*, *wayfinding*, acessibilidade e usabilidade, e durante a busca, utilizando a ferramenta de busca do ambiente ou mecanismos externos, destacam-se atributos como instrumentos de controle terminológico, folksonomias e metadados, todos vinculados às mediações dos sujeitos institucionais e informacionais (Torino; Vidotti; Vechiato, 2020).

com uma base conceitual e prática ajustada para a Ciência da Informação, ancorada na mediação infocomunicacional e, ao considerar os aspectos sociotécnicos envolvidos no projeto desses ambientes, bem como a intencionalidade dos sujeitos, é possível promover melhorias significativas na recuperação, no acesso e na apropriação da informação. Dessa maneira, para promover a Encontrabilidade da Informação, é preciso pensar nas possibilidades de sua produção, organização, representação, armazenamento e preservação, com ênfase no acesso, noutro e na apropriação (Vechiato, 2013; Vidotti *et al.*, 2017).

Compreende-se, portanto, que os desafios relacionados à Encontrabilidade abrangem a necessidade de se adaptar continuamente às modernizações, integrar abordagens interdisciplinares para lidar com a complexidade informacional e promover práticas inclusivas. É a forma de assegurar que a informação não apenas se ajuste, mas que também atenda de maneira eficaz e abrangente às diversas necessidades dos usuários, especialmente em um ambiente digital que está em constante evolução e transformação.

3 CONTRIBUIÇÕES DA INDEXAÇÃO AUTOMÁTICA PARA A EFICIÊNCIA DA ENCONTRABILIDADE DA INFORMAÇÃO NOS REPOSITÓRIOS INSTITUCIONAIS

Os Repositórios Institucionais surgiram a partir de iniciativas e movimentos internacionais iniciados na década de 1990, como a *Open Archives Initiative* (OAI) e o *Open Access* (OA), que tinham como objetivo promover e fomentar discussões sobre o acesso livre às produções acadêmico-científicas. Em meio a essas discussões, foram estabelecidas políticas e estratégias para consolidar o movimento de acesso aberto, sendo duas dessas estratégias a criação de revistas de acesso aberto e o autoarquivamento das produções em repositórios digitais de acesso aberto, conhecidos como Via Dourada (*Gold OA*) e Via Verde (*Green OA*), respectivamente. Assim, a implementação de ambientes informacionais, como os repositórios institucionais, tem sido fortalecida, especialmente pelas instituições de ensino e pesquisa (Sanchez; Vechiato; Vidotti; 2019).

Os RIs agregam toda a produção técnico-científica da comunidade acadêmica, composta por docentes, técnicos administrativos e alunos de pós-graduação, visando armazenar, preservar e disponibilizar textos completos de forma gratuita. De acordo com Custódio e Vechiato (2017), cada repositório institucional estabelece sua política específica alinhada aos

objetivos da instituição à qual está vinculado. No entanto, todos compartilham o mesmo propósito fundamental de armazenar, preservar e disseminar informações.

Vechiato e Vidotti (2014) estabelecem treze atributos para a encontrabilidade da informação em ambientes informacionais, contudo, conforme Custódio e Vechiato (2017), apenas dez desses atributos são aplicáveis em repositórios institucionais: taxonomias navegacionais, folksonomias, instrumentos de controle terminológico, metadados, mediação dos informáticos, mediação dos profissionais da informação, mediação do sujeito informacional e intencionalidade, *affordances*, *wayfinding* e a descoberta de informações.

Ao considerar os princípios fundamentais delineados até o momento nesta pesquisa, torna-se pertinente refletir sobre o impacto da rápida evolução tecnológica na Encontrabilidade da Informação em RIs. As emergentes inovações tecnológicas exercem uma influência direta na eficácia dos sistemas de informação, destacando, assim, a necessidade de estratégias robustas de indexação e categorização. Nesse contexto, a efetiva adoção de taxonomias navegacionais e ferramentas de controle terminológico assume um papel determinante, pois pode assegurar uma organização eficiente e uma recuperação mais precisa.

A indexação automática é conceituada como um procedimento automatizado em que o computador realiza a extração de palavras, expressões ou radicais do texto, a fim de selecionar descritores que representam integralmente o conteúdo textual, prescindindo da intervenção direta do ser humano (Santos; Ribeiro, 2003). Pode originar-se por meio de três perspectivas: uma indexação manual assistida por programas informáticos, que auxiliem o profissional no armazenamento dos termos de indexação identificados pelo homem; uma indexação semi-automática realizada pela máquina, mas com a validação dos termos por um profissional e; uma indexação automática realizada completamente pela máquina, sem a interferência do profissional (Gil-Leiva; Ortuño; Corrêa, 2022).

Segundo Lancaster (2004), há duas formas de se realizar a indexação automática, por extração ou por atribuição. A primeira acontece na ocasião em que um programa de computador indexa o documento por meio de critérios de frequência, posição e contexto das palavras que aparecem no decorrer do texto, para na sequência extraí-las e apresentá-las como descritores. Silva e Correa (2020) esclarecem que a segunda consiste em associar termos autorizados de um vocabulário controlado a um perfil formado por um conjunto de termos alternativos ou expressões equivalentes que ocorrem com frequência nos documentos, os

termos autorizados que possuem ocorrências em um documento são ponderados e posteriormente são selecionados os mais relevantes.

Gil-Leiva, Ortuño e Corrêa (2022) e Suominen, Inkinen e Lehtinen (2022) apontam a aplicação da indexação automática como uma opção factível diante do desafio que os profissionais da informação têm em realizar uma indexação com uniformidade, homogeneidade e a qualidade necessária para garantir o acesso posterior à informação por conta do aumento significativo das publicações e do crescente número de documentos processados e divulgados. Sob essa ótica, a automação do processo de indexação não apenas otimiza a eficácia, mas também contribui para a consistência e a precisão na representação da informação.

Borges, Maculan e Lima (2008) destacam a importância crucial da semântica na indexação automática, ressaltando que ela viabiliza ao *software* a identificação não apenas de termos isolados, mas também de estruturas e significados complexos nos documentos. O viés semântico contribui significativamente para a melhoria da exatidão e do desempenho do processo de indexação automática, uma vez que permite uma compreensão mais aprofundada e contextualizada do conteúdo informacional.

Janaite Neto (2018) e Santos (2017) são pesquisadores que propõem a aplicação da indexação automática ou semiautomática como uma ferramenta facilitadora para o acesso e recuperação de documentos armazenados em repositórios. São estudos que reconhecem a busca por aprimoramentos no processo de indexação dos documentos e que oferecem suportes valiosos na recuperação da informação. Além disso, a capacidade da indexação automática, de ser atualizada automaticamente com a adição de novos documentos ao repositório, pode garantir aos usuários o acesso a informações sempre atualizadas.

Percebe-se que a indexação automatizada constitui uma esfera dinâmica, com a finalidade de simplificar o processamento e a organização da informação. Em ambientes de Repositórios Institucionais, esse intento pode ser concretizado mediante a identificação e representação do conteúdo presente nos documentos informacionais. Nesse âmbito, compreende-se que a Indexação Automática (IA) configura-se como uma ferramenta potencialmente contributiva para o processo da Encontrabilidade da Informação. Essa contribuição se dá mediante a apresentação de um conjunto de descritores ou metadados, os quais têm a capacidade de recuperar e viabilizar o uso informacional dos elementos que

compõem o repositório.

Na visão de Torino, Vidotti e Vechiato (2020), os metadados desempenham um papel crucial ao estabelecerem a relação entre o sujeito informacional e a informação desejada, mesmo quando distribuída em variados repositórios informacionais. Adicionalmente, os metadados apresentam a habilidade de organizar dados compreensíveis tanto para indivíduos quanto para agentes computacionais, conferindo-lhes a característica de serem interoperáveis e suscetíveis à integração em diferentes contextos e sistemas. Essa versatilidade dos metadados não apenas promove uma coesão efetiva entre as informações, mas também contribui substancialmente para a eficiência do processo de Encontrabilidade da Informação.

A IA desempenha um papel de relevância fundamental na otimização da EI em RIs, desempenhando uma função determinante na eficiência do processamento de dados e no aprimoramento da recuperação da informação. Com base nas análises e reflexões realizadas, neste estudo, torna-se possível destacar, superficialmente, algumas contribuições de grande significado para a Encontrabilidade da Informação em Repositórios Institucionais. Cada ponto abordado apresenta aspectos específicos que exercem impacto positivo, reforçando, assim, a importância da indexação automática nesse contexto, conforme ilustrado no Quadro 3.

Quadro 3 - Contribuições da Indexação Automática para a Encontrabilidade da Informação

Eficiência na Organização e Estruturação de Dados	O sistema computacional extrai elementos linguísticos, como palavras, expressões e radicais, com o intuito de selecionar descritores, facilitando a organização e estruturação dos dados, promovendo a eficácia na representação e recuperação das informações.
Desambiguação de Termos Homógrafos e Polissêmicos	Visa mitigar recuperações incorretas de informações, promovendo a precisão na representação dos conteúdos nos repositórios.
Enfoque Semântico	Possibilita que o software identifique não apenas termos isolados, mas também compreenda estruturas complexas e significados nos documentos, aprimorando a representação semântica nas bases de dados.
Aplicabilidade e Facilitação	A aplicação da indexação automática configura-se como uma ferramenta estratégica, facilitando o acesso e a recuperação eficiente de documentos em repositórios institucionais, otimizando a eficácia dos processos de busca e recuperação da informação.
Atualização Automática	Quando novos documentos são adicionados aos repositórios, assegura aos usuários acesso a informações sempre atualizadas, promovendo a eficácia contínua nos processos de recuperação da informação.
Contribuição dos Metadados	Os descritores estabelecem uma ligação entre o sujeito informacional e a informação, transcendendo as fronteiras de diferentes repositórios, favorecendo a eficiência nos processos de recuperação.

Fonte: Os autores (2023).

Desde o início dos anos 2000, as publicações científicas foram impregnadas com a Web Semântica (que utiliza XML [*eXtensible Markup Language*], RDF [*Resource Description Framework*] e OWL [*Ontology Web Language*]), com a adição, em 2012, da norma JATS (*Journal Article Tag Suite*) para a descrição do conteúdo formal, textual e gráfico dos artigos científicos (Gil-Leiva; Ortuño; Corrêa, 2022). Contudo, há uma escassez de estudos que relacionem o processo de indexação automática com a encontrabilidade da informação. Um possível motivo para isso é indicado por Roa-Martinez (2019), ao observar que a Recuperação da Informação é extensamente discutida tanto na Ciência da Computação quanto na Ciência da Informação, ao passo que a Encontrabilidade se tornou foco de estudos nos últimos anos.

A área da Encontrabilidade da Informação é considerada relativamente recente, oferecendo várias possibilidades de estudos. No que concerne à sua interconexão com a Indexação Automática, percebe-se que a utilização estratégica desta última pode ser usada para aprimorar a Encontrabilidade. Embora estejamos diante de um estágio inicial e embrionário, este trabalho representa uma iniciativa pioneira na análise dessa intrincada relação, estabelecendo as bases para investigações mais profundas e buscando contribuir de maneira substancial para o desenvolvimento de pesquisas futuras nesse promissor campo de estudo.

4 CONSIDERAÇÕES FINAIS

A interação entre a Indexação Automática e a Encontrabilidade exerce um papel crucial na eficácia dos Repositórios Institucionais em várias dimensões. Uma indexação eficiente e uma encontrabilidade robusta elevam a visibilidade do conteúdo dos RIs. A capacidade aprimorada de localizar documentos, através de motores de busca e outros mecanismos, amplia a abrangência das pesquisas e a disseminação do conhecimento, proporcionando benefícios significativos tanto para os autores quanto para a comunidade acadêmica como um todo.

A utilização da indexação automática emerge como uma ferramenta valiosa na organização estruturada do conteúdo dos Repositórios Institucionais, permitindo a categorização e classificação eficazes dos documentos com base em critérios relevantes, como tópicos e autores. Essa abordagem otimiza a navegação e a busca no repositório,

simplificando a localização de informações desejadas de forma ágil. Adicionalmente, possibilita a atribuição de palavras-chave, metadados e descrições pertinentes aos documentos, simplificando o acesso às informações. Como resultado, os usuários podem empregar consultas de pesquisa para encontrar documentos com base em termos-chave relacionados, tornando o processo mais eficiente e intuitivo.

A usabilidade dos RIs está intimamente ligada à encontrabilidade da informação. Quando os documentos são indexados de maneira eficiente, os usuários experimentam uma navegação mais satisfatória ao buscar o conteúdo desejado. Esse aspecto é particularmente crucial para pesquisadores, professores e estudantes que utilizam os RIs para acessar materiais acadêmicos e científicos. Uma alta encontrabilidade e uma indexação automática eficaz funcionam como estímulo para que os autores contribuam com suas pesquisas nos RIs, confiantes de que seus trabalhos serão organizados de maneira apropriada e facilmente localizados por outros pesquisadores. Esse ciclo virtuoso tem o potencial de impulsionar o crescimento e a relevância contínua do repositório ao longo do tempo.

A colaboração entre indexação automática e encontrabilidade pode ser aplicada para personalizar a experiência do usuário nos repositórios. Ao analisar os metadados e os termos de pesquisa, os sistemas podem sugerir conteúdos relevantes, tornando a pesquisa mais eficiente e agradável, o que torna os RIs fontes eficazes de acesso ao conhecimento acadêmico e científico. Essa relação é fundamental para assegurar que os documentos armazenados sejam facilmente encontrados, estejam bem organizados e, como resultado, tenham um impacto positivo na disseminação e na utilização do conhecimento.

Em resumo, a indexação automática apresenta-se como uma ferramenta promissora para aprimorar a encontrabilidade da informação em repositórios institucionais. Ao otimizar a eficácia na busca e recuperação de informações, ela amplia a acessibilidade e o valor dos conteúdos para a comunidade acadêmica e a sociedade em geral. Em meio ao constante crescimento da produção de conhecimento, a indexação automática assume um papel crucial na organização e preservação desse conhecimento, assegurando sua relevância contínua e acesso para todos.

REFERÊNCIAS

BORGES, G. S. B.; MACULAN, B. C. M. S.; LIMA, G. A. B. O. Indexação automática e semântica: Siti, Maceió, v. 6, e181, 2024

estudo de análise do conteúdo de teses e dissertações. **Informação & Sociedade**: Estudos, João Pessoa, v. 18, n. 2, p. 181-193, maio/ago. 2008. Disponível em: <https://brapci.inf.br/index.php/res/v/91214>. Acesso em: 11 nov. 2023.

BORGES, L. C.; CASTRO, A. G.; SILVA, D. M.; VASCONCELLOS, B. B. B.; VITIELLO, B. C. Potencialidades dos repositórios institucionais das universidades federais brasileiras: apontamentos sobre *software*, equipe, manual, tutorial e política. **Perspectivas em Gestão & Conhecimento**, João Pessoa, v. 9, n. 2, p. 245-265, maio/ago. 2019. Disponível em: <http://dx.doi.org/10.21714/2236-417X2019v9n2p245>. Acesso em: 22 nov. 2023.

CAMOSSI, G.; RODAS, C. M. Contribuições das técnicas de search engine optimization para encontrabilidade da informação. **Encontros Bibli**, Florianópolis, v. 28, 2023. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/93371/53431>. Acesso em: 27 out. 2023.

CAMPOS, A.; SOUSA, M. R. F. de; OLIVEIRA, H. P. C. de. Encontrabilidade da informação e arquitetura da informação: possíveis relações teóricas. **Encontros Bibli**: revista eletrônica de Biblioteconomia e Ciência da INFORMAÇÃO, v. 26, p. 01-19, 2021. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/77624>. Acesso em: 10 dez. 2023.

CUSTÓDIO, N. C.; VECHIATO, F. L. Encontrabilidade da informação em repositórios institucionais: uma proposta de instrumento de avaliação. **Informação na Sociedade Contemporânea**, Natal, v. 1, p. 1-17, 2017. Edição especial. Disponível em: <https://brapci.inf.br/index.php/res/download/106617>. Acesso em: 27 out. 2023.

GIL-LEIVA, I.; ORTUÑO, P. D.; CORRÊA, R. F. Indización automática de artículos científicos sobre Biblioteconomía y Documentación con SISA, KEA y MAUI. **Revista Española de Documentación Científica**, [S. l.], v. 45, n. 4, p. e338, 2022. Disponível em: <https://redc.revistas.csic.es/index.php/redc/article/view/1371>. Acesso em: 19 out. 2023.

JANAITE NETO, J. **Recuperação de informação baseada em ontologia**: uma proposta utilizando o modelo vetorial. 2018. 105 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Marília, SP, 2018. Disponível em: https://bdtd.ibict.br/vufind/Record/UNSP_f5570a550c83884d60ea4b9267da5ebc. Acesso em 10 nov. 2023.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. ver. atual. Brasília: Briquet de Lemos, 2004.

MARTINS, C. A.; MONARD, M. C.; MATSUBARA, E. T. Reducing the Dimensionality of Bag-of-words Text Representation used by Learning Algorithms. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND APPLICATIONS, 3., 2003, Benalmádena (ES). **Anais [...]**. Benalmádena (ES): IASTED: Acta Press, 2003. Disponível em: https://www.researchgate.net/publication/228983901_Reducing_the_dimensionality_of_bag-of-words_text_representation_used_by_learning_algorithms. Acesso em: 21 nov. 2023.

MORVILLE, P. **Ambient findability**. Sebastopol: O'Really, 2005.

ROA-MARTÍNEZ, S. M. **Da Information Findability à Image Findability: aportes da polirrepresentação, recuperação e comportamento de busca.** 2019. 235 f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista (Unesp), Faculdade de Filosofia e Ciências, 2019. Disponível em: <https://acesse.dev/Xqyfo>. Acesso em: 18 nov. 2023.

ROMANI, L. S.; FUSCO, E.; SANTOS, P. L. V. A. C. Análise e implantação de repositório digital utilizando software livre DSPACE. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO (SBSI), 6., 2010, Marabá. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2010. p. 250-261. Disponível em: <https://sol.sbc.org.br/index.php/sbsi/article/view/14717>. Acesso em: 21 out. 2023.

SANCHEZ, F. A.; VECHIATO, F. L.; VIDOTTI, S. A. B. G. Recomendações de encontrabilidade da informação para repositórios institucionais. In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 8., 2019, Londrina. **Anais [...]**. Londrina: SECIN: UEL, 2019. Disponível em: <http://www.uel.br/eventos/cinf/index.php/secin2019/secin2019/paper/view/577/406>. Acesso em: 22 nov. 2023.

SANTOS, A. F. dos. **A atividade indexação apoiada por plataformas computacionais.** 2017. 153 f. Dissertação (Mestrado em Gestão da Informação) – Universidade do Estado de Santa Catarina. Florianópolis, 2017. Disponível em: https://www.udesc.br/arquivos/faed/id_cpmenu/1440/dissertacao_alessa_15689006572758_1440.pdf. Acesso em: 10 nov. 2023.

SANTOS, G. C.; RIBEIRO, C. M. **Acrônimos, siglas e termos técnicos:** arquivística, biblioteconomia, documentação, informática. Campinas, SP: Editora Átomo, 2003.

SILVA, S. R. B.; CORREA, R. F. Sistemas de indexação automática por atribuição: uma análise comparativa. **Encontros Bibli:** revista eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 25, p. 01-15, 2020. Disponível em: <https://brapci.inf.br/index.php/res/download/142235>. Acesso em: 21/11/2023.

SUOMINEN, O.; INKINEN, J.; LEHTINEN, M. Annif and Finto AI: developing and implementing automated subject indexing. **JLIS.it**, v. 13, n. 1, jan. 2022. Disponível em: <https://www.jlis.it/index.php/jlis/article/view/437>. Acesso em: 9 nov. 2023.

TORINO, E; VIDOTTI, S. A. B. G.; VECHIATO, F. L. Contribuições do atributo Metadados para a Encontrabilidade da Informação. **Em Questão**, Porto Alegre, v. 23, n. 2, p. 437-457, maio/ago., 2020. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/93072>. Acesso em: 9 nov. 2023.

VECHIATO, F. L. **Encontrabilidade da informação:** contributo para uma conceituação no campo da Ciência da Informação. 2013. 206f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2013. Disponível em: Acesso em: 3 nov. 2023.

VECHIATO, F. L.; VIDOTTI, S. A. B. G. **Encontrabilidade da informação.** São Paulo: Cultura Siti, Maceió, v. 6, e181, 2024

Acadêmica, 2014. 196 p. Disponível em:

<http://www.culturaacademica.com.br/catalogo/encontrabilidade-da-informacao/>. Acesso em: 22 nov. 2019.

VECHIATO, F. L.; VIDOTTI, S. A. B. G. Encontrabilidade da informação: atributos e recomendações para ambientes informacionais digitais. **Informação & Tecnologia (ITEC)**, Marília/João Pessoa, v.1, n. 2, p. 42-58, jul./dez., 2014. Disponível em: <https://brapci.inf.br/index.php/res/download/53982>. Acesso em: 28 out. 2023.

VIDOTTI, S. A. B. G.; ROA-MARTÍNEZ, S. M.; CONEGLIAN, C. S.; FERREIRA, A. M. J. F. C.; VECHIATO, F. L. As contribuições das heurísticas de usabilidade para a encontrabilidade da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais [...]**. Marília, 2017. Disponível em: <https://brapci.inf.br/index.php/res/v/104849>. Acesso em: 5 nov. 2023.