

10, 11 e 12 de novembro de 2025

POLITÉCNICO DO PORTO / ISCAP
PORTO - PORTUGAL

**FORMAÇÃO E INVESTIGAÇÃO EM
CIÊNCIA DA INFORMAÇÃO**

OPORTUNIDADES E DESAFIOS



BASE DE DADOS EM ARQUIVÍSTICA EM PERSPECTIVA:

TRATAMENTO TEMÁTICO E RECUPERAÇÃO DA INFORMAÇÃO EM FOCO

João Marcello de Lima Marques, Universidade de Brasília, 0009-0005-7132-5333, Brasil,
joaomunb@gmail.com

Katia Isabelli Melo, Universidade de Brasília, 0000-0002-0559-3891, Brasil, isabelli@unb.br

Eixo: Organização da informação

1 Introdução

A sociedade como um todo vive um momento de grande crescimento informacional. A comunidade científica devido aos grandes avanços nas tecnologias e nas constantes mudanças percebidas vem produzindo cada vez mais trabalhos em todas as suas formas, artigos, monografias, teses e dissertações, dentre outros.

A escrita científica contribui para a resolução de problemas e para a proposta de melhorias, permitindo que a sociedade continue evoluindo. Através da investigação e das análises realizadas no âmbito científico, é possível identificar novos modos de agir, pensar e perceber várias coisas, aplicando esses conhecimentos de forma prática na sociedade.

A Arquivologia evoluiu significativamente com a implementação das novas tecnologias presentes no mercado. O trabalho dos profissionais de arquivos passou a incorporar tecnologias da informação, automatizando processos e integrando o meio digital ao cotidiano, com o objetivo de otimizar as tarefas técnicas relacionadas à gestão documental.

Com as mudanças percebidas, o fluxo de produção científica envolvendo a área da Arquivologia teve um crescimento perceptível (Silva, 2023). Começou uma movimentação para investigar não apenas o uso de

tecnologias da informação na área, mas também outros temas que envolvem o fazer arquivístico.

Outro motivo para o aumento da produção científica na área da Arquivologia consiste no crescente número de eventos científicos que permeiam assuntos relevantes para a área, impulsionando pesquisadores antigos e novos, de todos os níveis, a engajarem com a comunidade científica.

Com a percepção do crescimento da produção científica na área de Arquivologia, em 2019, identificou-se a necessidade de construir um repositório digital, que reunisse a produção científica e técnica, de autores brasileiros, relacionados à Arquivologia¹.

Surgiu a Base de Dados em Arquivística (BDA), ferramenta disponibilizada para a sociedade, em 2021. No cadastramento elege-se, majoritariamente, os itens produzidos no idioma português. No entanto, uma parcela atende as produções científicas e técnicas nos idiomas inglês, francês, espanhol refletindo contribuições de profissionais convidados que participaram de eventos científicos ou que contribuíram em edições dos periódicos seriados.

A BDA está em contínuo progresso, com o constante cadastro de novos itens e materiais, juntamente com a implementação de ajustes destinados a aprimorar a eficácia da plataforma tanto para seus colaboradores

quanto para seus usuários.

A presente pesquisa tem como objetivo analisar como se processa a organização das produções científicas cadastradas na BDA considerando as técnicas de Tratamento Temático da Informação (TTI) adotadas. O objetivo é analisar as técnicas de Tratamento Temático da Informação (TTI) usadas na BDA, justificando-se pela necessidade de organização eficiente devido ao crescimento da produção científica em Arquivologia. Para atender o objetivo maior, traçou-se os seguintes objetivos específicos: investigar como é feito o processo de TTI na BDA; identificar possíveis problemáticas no uso do TTI na BDA; analisar a relação entre TTI e Recuperação da Informação (RI), e averiguar como o uso de metodologias de tratamento da informação influencia na RI na BDAⁱⁱ.

A pesquisa se justifica pela crescente produção científica na área de Arquivologia e esta demanda sistemas de organização eficientes para garantir a acessibilidade e a recuperação da informação (Silva, 2023).

Uma das contribuições para o crescimento da produção científica na área da Arquivologia está ligada à criação de novos cursos e ao exponencial crescimento de discussões relacionadas a área dentro do território brasileiro.

2 Referencial Teórico

O referencial teórico da pesquisa baseou-se numa vasta literatura, com destaque para os estudos de Taylor & Joudrey (2004), Silva (2011), Santos & Corrêa (2017), (Coelho (2020), Melo (2021), Wang (2023), Melo, Souza, & Paiva (2023).

O aumento do uso da internet para disseminação de informações atingiu seu ponto máximo após a Segunda Guerra Mundial, impulsionado pelo imenso volume de dados e informações produzidos diariamente, sem contar com tecnologias adequadas para processá-los (Oliveira *et al*, 2019).

Uma das soluções tecnológicas encontradas

para atender à demanda de processamento, organização e armazenamento de informações foram as bases de dados. O surgimento e a evolução dessa ferramenta ocorreram de forma rápida, transformando-a em um repositório capaz de reunir informações de qualidade.

No âmbito da Arquivologia, no início dos estudos sobre base de dados o conceito se confundia com o de banco de dados (Albrecht & Ohira, 2000). Cunha e Cavalcanti no Dicionário de Biblioteconomia e Arquivologia definem base de dados como:

[...] coleção de valores de dados inter-relacionados de tal natureza que, de acordo com o sistema de gerenciamento de base de dados, os arquivos que contêm os dados podem integrar-se temporariamente em uma única estrutura conectada ou integrar-se somente por ocasião da consulta. (Cunha & Cavalcanti, 2008, p. 43).

O estudo de base de dados existe no âmbito tanto da Ciência da Computação, focado em dados, quanto na C.I., que tem como uma de suas lentes o cadastro, gerenciamento e acesso de produções bibliográficas. Na Arquivologia temos a base de dado referencial bibliográfica definida por Cunha e Cavalcanti como:

[...] base de dados que contém registros automatizados, relativos a documentos e itens bibliográficos. Pode, ou não, conter resumos e é formada por "uma série de registros bibliográficos ligados entre si, onde cada um em geral apresenta uma combinação dos seguintes componentes: número do documento; título; autor; referência da fonte; resumos; texto integral; termos ou expressões de indexação; citações ou quantidade de referências; instituição de origem do documento, ou endereço do autor, ou ambos; língua do documento-fonte; informação de uso interno, como números de classificação ou localização (Cunha & Cavalcanti, 2008, p. 44)

Um repositório digital pode ser classificado de acordo com o seu uso. Quando o conteúdo de um repositório é focado em um assunto, ele é considerado um repositório temático e quando o conteúdo reflete o acervo de produção científica de uma instituição, ele é chamado de repositório institucionalⁱⁱⁱ (Pinheiro & Ferrez, 2014).

As bases de dados bibliográficas fazem parte do cotidiano de um profissional da Ciência da Informação, como o Arquivista e o Bibliotecário, sendo uma importante fonte de pesquisa sobre diversos temas. Elas podem ser de conhecimentos gerais, como o *Google Scholar*, que indexa produções bibliográficas de todas as áreas do conhecimento, ou de conhecimentos temáticos, como a BDA, estudo de caso retratado, que indexa apenas produções científicas sobre a área da Arquivologia, de autores brasileiros.

3 Procedimentos Metodológicos

O percurso metodológico deste trabalho se deu como um estudo de caso descritivo e qualitativo da BDA. A pesquisa se iniciou com a navegação da estrutura da Base de Dados, permitindo uma análise detalhada do funcionamento e dos métodos de organização empregados.

O estudo de caso se caracteriza como um estudo profundo e exaustivo de um ou alguns objetos trazendo conhecimento sobre o que foi analisado (Yin, 2001). É considerada aplicada pois busca a aplicação prática para a solução de problemas identificados no objeto de estudo. (Boaventura, 2004).

De acordo com Prodanov e Freitas (2013), a pesquisa qualitativa busca interpretar fenômenos a partir da compreensão de significados, contextos e experiências vivenciadas, sendo especialmente útil quando se deseja investigar relações e dinâmicas que não podem ser expressas apenas em dados numéricos. Como o foco deste estudo está na descrição e análise de práticas relacionadas ao uso de taxonomias, categorização e indexação, uma abordagem quantitativa não

seria suficiente para alcançar os objetivos propostos. A pesquisa qualitativa, por sua vez, permite interpretar as funcionalidades da BDA e as escolhas técnicas a partir de uma perspectiva contextual e crítica.

Quanto ao processo metodológico, o estudo de caso é o mais apropriado por permitir uma investigação profunda de uma unidade específica dentro de seu contexto real. Segundo Prodanov e Freitas (2013, p. 127), o estudo de caso “é uma investigação empírica que investiga um fenômeno contemporâneo dentro de seu contexto da vida real, especialmente quando os limites entre fenômeno e contexto não estão claramente definidos”. No presente trabalho, essa definição aplica-se perfeitamente, já que o funcionamento da BDA e suas práticas de organização estão diretamente relacionadas ao seu contexto digital e institucional.

O estudo de caso permite, ainda, a coleta de múltiplas fontes de evidência, como a observação direta da plataforma, a análise dos documentos cadastrados, e a revisão bibliográfica especializada. Essa triangulação metodológica proporciona maior confiabilidade e profundidade aos resultados obtidos, permitindo identificar não apenas as técnicas empregadas, mas também suas limitações e possibilidades de aprimoramento.

Assim, tanto a abordagem qualitativa quanto o estudo de caso são fundamentais para alcançar os objetivos da pesquisa, pois permitem compreender o fenômeno investigado em sua complexidade e propor melhorias efetivas para o tratamento e a recuperação da informação na BDA.

O presente trabalho consiste em coletar e analisar informações sobre como a BDA organiza os materiais técnico científicos cadastrados. Buscou-se identificar se foram utilizadas técnicas/métodos de TTI para organizar as informações depositadas.

Ademais da metodologia indicada, para alcançar os objetivos foram adotados os seguintes procedimentos: levantamento bibliográfico pertinente ao assunto, tendo

como fontes de pesquisa a BDA, a Base de Dados em Ciência da Informação (BRAPCI) e a base de dados *Google Scholar*.

4 Tecnologia da Informação na Arquivologia

No último século a sociedade se modificou em uma velocidade acelerada devido aos grandes avanços na tecnologia. Muitas áreas do conhecimento tiveram que adaptar a sua realidade com o surgimento de novas tecnologias e demandas. A Arquivologia foi uma das áreas afetadas pelo surgimento de novas tecnologias, principalmente devido a percepção que a Arquivologia é uma ciência interdisciplinar (Vieira *et al*, 2015).

A interdisciplinaridade da Arquivologia com a área da Tecnologia da Informação está relacionada com a prática do tratamento, armazenamento, acesso e recuperação de informação e documentos, de acordo com Vieira *et al* (2015). O uso de *Softwares* como os Sistemas Informatizados de Gestão Arquivística de Documentos (SIGAD), o Archivematica, o ICA-AtoM e as bases de dados auxiliam o trabalho do Arquivista, diariamente.

O Conselho Nacional de Arquivos (CONARQ, 2022) no e-ARQ Brasil^{IV} define o SIGAD como um conjunto de procedimentos e operações técnicas, característicos de gestão arquivística de documentos, processado por computador. Esse *Software* é utilizado na produção, classificação e no gerenciamento dos documentos produzidos e recebidos por determinada entidade.

Em seu Guia de Usuários Archivematica, Costa *et al* (2016) definem Archivematica como um sistema de preservação digital de *Software* livre, gratuito e de código aberto, responsável por manter os dados baseados em padrões de preservação digital e possibilitar o acesso a longo prazo das coleções de objetos digitais.

O Access to Memory (AtoM) é definido como um *Software* de código aberto programado para *Web*, baseado em padrões desenvolvidos pelo International Council on Archives (ICA) (Artefactual Systems, 2015). Utilizado no

processo de descrição, acesso, e difusão dos documentos depositados nos fundos arquivísticos de uma entidade, formando um Repositório Arquivístico Digital Confiável.

Os *Softwares* são utilizados em conjunto no processo de gestão documental proposto pelo modelo *Open Archival Information System* (OAIS), criado pelo *Consultative Committee for Space Data System* (CCSDS, 2012). O modelo é responsável por especificar os principais critérios que devem ser seguidos em um projeto de preservação digital.

Para além do mundo da gestão de documentos e dos repositórios arquivísticos digitais confiáveis, dentro da área da Ciência da Informação (C.I.) e da Arquivologia existem os repositórios digitais de conhecimento, como as bases de dados, que são um importante meio de consulta científica, podendo elas serem de conhecimentos gerais ou temáticas.

Podemos ter uma visão inicial de como a área da Arquivologia e a área da Tecnologia da Informação estão interligadas. É de extrema importância que os profissionais das duas áreas estejam em constante colaboração para elaborar e utilizar corretamente *softwares* arquivísticos e repositórios digitais de conhecimento, como as bases de dados, sem prejudicar a entidade e os usuários. -

Ao que tange a C.I., esta é uma ciência interdisciplinar (Almeida Junior, 2020) responsável por fornecer base teórica e metodológica para as áreas que a compõem, como a Arquivologia e a Biblioteconomia (Rabelo & Cé, 2023). E isso se dá, pois, essas duas áreas têm como objeto principal de estudo a informação registrada em algum suporte.

Mesmo tendo a informação como fonte de estudo, o conceito de documento para essas áreas possui especificidades. Por serem áreas correlatas são perceptíveis os limites e trocas relacionados aos objetos, técnicas e teorias fundamentais desses campos, como fica evidenciado pela estreita relação entre os conceitos discutidos sobre os repositórios digitais de documentos tanto na Arquivologia

quanto na Biblioteconomia, dentro do escopo da C.I.

O Dicionário de Biblioteconomia e Arquivologia escrito por Cunha e Cavalcanti (2008, p. 321) define repositório como: “Lugar físico ou digital onde objetos são armazenados [...]”, “Conjunto ou base de dados com informações importantes. Biblioteca digital”. No Tesauro Brasileiro de Ciência da Informação, um repositório digital é definido como um meio para armazenar, gerenciar e preservar conteúdos informacionais no formato eletrônico (Pinheiro & Ferrez, 2014).

O Re却tório digital é formado por “coleções digitais de documentos de interesse para a pesquisa científica e, no caso dos institucionais, representam a sua memória científica.” (Pinheiro & Ferrez, 2014, p. 195.). Se destaca no caso dos repositórios digitais a possibilidade de compartilhamento dos metadados do seu conteúdo através de modelos de dados, como o Dublin Core^{v3}, e protocolos como OAI-PMH^{vi}, que garantem a interoperabilidade entre ferramentas, transformando assim o re却tório digital em um arquivo aberto (Martins *et al*, 2017).

O tratamento técnico aplicado aos documentos bibliográficos e documentos arquivísticos é feito de forma distinta. Essa disparidade de tratamento também se reflete nos repositórios digitais, devido às singularidades de tratamento de cada categoria documental. Ao analisar a categoria de repositórios na Arquivologia percebe-se a existência de vários tipos de repositórios, como os repositórios arquivísticos digitais confiáveis (RDC-Arq) e os repositórios de dados, conhecidos como base de dados.

4.1 Estudo de caso: a Base de Dados em Arquivística (BDA)

Na contemporaneidade houve um grande crescimento na produção científica na área de Arquivologia devido a busca por novos conhecimentos e soluções para problemas encontrados no cotidiano de trabalho de um Arquivista e pessoas que trabalham com

arquivos (Silva, 2023). A produção científica desempenha um papel fundamental no avanço de todas as áreas do conhecimento, permitindo a descoberta de novas ideias e soluções e na Arquivologia esse fenômeno não foi diferente.

A construção da BDA reflete uma necessidade vista em muitas áreas do conhecimento. De acordo com Sayão,

Quando um pesquisador, diante de um microcomputador ligado a um banco de dados [...] à procura de informações que definam, completem ou estabeleçam as fronteiras do seu trabalho de pesquisa, ele repete o mesmo gesto de quem mergulha na memória de seu grupo para reconstruir as lembranças comuns e dessa forma manter íntegra a sua comunidade. (1996, pp. 314).

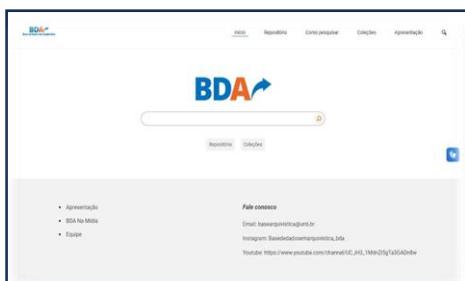
Uma das inspirações para esse projeto veio do *Centro de Información Documental de Archivos* (CIDA), um setor vinculado a *Subdirección General de Los Archivos Estatales* que é subordinado ao *Ministerio de Cultura y Deporte* da Espanha. O CIDA é responsável por reunir, difundir e disponibilizar a literatura sobre Arquivologia produzida no território espanhol e em países Iberoamericanos (Melo, Souza & Paiva, 2022).

A BDA nasce de uma necessidade de reunir as produções técnicas e científicas da área da Arquivologia do Brasil, majoritariamente em língua portuguesa, em um só lugar^{vii}. A produção da interface da base de dados foi feita a partir da ferramenta Tainacan, que é um plugin de código aberto para *Wordpress*, utilizado na criação de repositórios digitais, que tem como foco a gestão e publicização de acervos digitais de diversas áreas, responsável por organizar e expor coleções através do uso de metadados (Brasiliiana Museus, 2024).

O *Wordpress* foi desenvolvido em 2003 como um *Software* de código aberto. Inicialmente era utilizado para criação de blogs, orientado e utilizado através de uma comunidade ativa de usuários que contribuem, continuamente, com novos códigos para aumentar e melhorar

suas funcionalidades (Messenlehner & Coleman, 2014). A Figura 1, a seguir, ilustra a página inicial da BDA.

Figura 1: Página Inicial da Base de Dados em Arquivística (BDA)



Fonte Base de Dados em Arquivística (2024).

O projeto da BDA foi dividido em etapas que se iniciaram com o projeto de criação e implementação da ferramenta. Além disso, foram desenvolvidas outras etapas dedicadas à manutenção e melhorias da usabilidade da base de dados, além do cadastramento das produções científicas e técnicas (Melo, 2020). Atualmente, a Base de Dados conta com, aproximadamente, 10.024 produções científicas cadastradas, incluindo artigos científicos, eventos, livros e outras contribuições acadêmicas. Com a implementação das taxonomias na BDA se tornou necessária a vinculação de palavras-chave e keywords aos itens cadastrados para auxiliar na recuperação da informação.

O uso de taxonomias também impactou o processo de categorização das produções cadastradas, seja considerando-os em sua totalidade, como no caso de eventos científicos, livros, entre outros, ou como itens vinculados a uma categoria geral, como Anais de eventos científicos, capítulos de livros e artigos publicados nos Anais.

A visibilidade da BDA impactou diretamente a comunidade científica que debate sobre Arquivologia. A Base conseguiu se tornar um ponto de referência em pesquisas nacionais e internacionais (Melo & Paiva, 2023), e como objeto de pesquisas como apresentado por Silva (2023).

Com o crescimento da importância da BDA como fonte de referência e os impactos causados por ela, se viu necessário revisar o

uso das taxonomias dentro da plataforma, para identificar possíveis problemas ligados ao seu uso e com isso poder estudar propostas para a melhoria delas.

As produções científicas e técnicas a serem cadastradas na BDA passam por um processo de tratamento da informação antes de serem disponibilizadas, tendo como objetivo proporcionar a melhor experiência possível ao usuário da ferramenta.

Os profissionais da C.I., incluindo Arquivistas e Bibliotecários, desempenham um papel crucial na organização e tratamento da informação. Seus objetivos consistem em proporcionar uma experiência eficaz e eficiente para os usuários, ao mesmo tempo em que consideram o impacto dessas atividades em seu trabalho diário.

Um dos primeiros passos para o tratamento informacional é definir o objeto informacional (OI), que é definido por Taylor & Joudrey (2004) como uma unidade informacional organizável, sendo registrada em vários formatos como texto, imagem, registro sonoro, etc. No caso das bases de dados, o objeto informacional a ser tratado são as produções científicas e técnicas, como os artigos, manuais e monografias.

O tratamento da informação permite que ela seja organizada de forma padronizada, facilitando a recuperação e a disseminação (Hillesheim & Fachin, 2016). Nesse contexto, Dias define que:

Nos sistemas de informação e de recuperação da informação, o tratamento da informação é definido como a função de descrever os documentos, tanto do ponto de vista físico (características físicas dos documentos) quanto do ponto de vista temático (ou de descrição do conteúdo). (Dias, 2001, pp. 3).

Nessa perspectiva, Brascher e Monteiro (2010) destacam também a importância de analisar as informações do ponto de vista tanto físico, quanto temático na organização da informação em repositórios digitais. O tratamento temático auxilia na construção e

adoção de taxonomias para categorizar as informações a serem cadastrados nas bases de dados.

5 Tratamento Temático da Informação

O tratamento da informação é composto por várias subáreas, uma delas considerada fundamental para a organização e representação do conhecimento, se chama Tratamento Temático da Informação (TTI) (Tartarotti, Dal'edove & Fujita, 2015). O TTI está ligado com o estudo do conteúdo do documento a ser analisado.

O estudo do conteúdo utiliza um conjunto de instrumentos, processos e produtos que permite que sejam explorados dois aspectos, um relacionado à determinação do assunto e o segundo à suas especificações (Brascher & Guimarães, 2018). O TTI também pode ser chamado de análise documental, análise temática, análise de assuntos, descrição de conteúdo (Café & Sales, 2010).

O TTI na cadeia de tratamento documental é utilizado para fazer a mediação entre a produção e o uso da informação (Guimarães, Sales & Grácio, 2012). Sua análise é voltada para extrair informações pertinentes aos atributos temáticos, tendo como base de seus estudos a análise do assunto e do conteúdo dos documentos. Em resumo, abrange as áreas de análise, descrição e representação do conteúdo dos documentos tendo como objetivo sua organização e recuperação (Guimarães, Ferreira & Freitas, 2012).

O TTI pode ser entendido também como o ato de produzir informação sobre a informação analisada (Vieira, Oliveira & Cunha, 2017). Seu uso é capaz de gerar produtos através das análises realizadas em cima de um documento.

Um dos produtos gerados através do uso do TTI são as taxonomias, termos descritores, utilizadas para categorizar ou indexar informações em Repositórios Digitais como as bases de dados. A categorização e a indexação são técnicas que auxiliam na organização e na recuperação de informações em ambientes

informacionais.

O uso de taxonomias ficou conhecido a partir da classificação de seres humanos utilizada na Biologia ao categorizar os seres vivos em classes e subclasses (Mendes & Pinto, 2019). Essa prática rompeu barreiras e começou a ser utilizada em outras áreas do conhecimento como na Arquivologia. Nessa área de acordo com Novo (2010), as taxonomias são utilizadas principalmente nos estudos de recuperação da informação e na experiência do usuário, pois são responsáveis por melhorar e facilitar o acesso à informação nelas organizadas e classificadas.

Com o crescimento do uso de tecnologias da informação e do grande número de informações criadas, a recuperação de informações ficou bastante prejudicada. Tendo em vista essa problemática, se viu necessário promover estudos para encontrar métodos para melhorar a organização e a recuperação de informação. Segundo Terra e Bax (2003) esses estudos visam simplificar a procura e a navegação, desenvolvendo estruturas de informação e categorias que sejam facilmente compreendidas pelos usuários que utilizam o sistema.

Segundo Sousa e Araújo Júnior,

A taxonomia pode ser definida como sistema de classificação que apoia o acesso à informação, permitindo classificar, alocar, recuperar e comunicar informações em um sistema de maneira lógica. (2013, pp. 140).

Tem por finalidade permitir agregação de informação e dados, possibilitando acesso através de navegação. Nessa prática as informações são construídas por palavras e traduzidas por conceitos responsáveis por agrupar informações que possuam conteúdo parecidos.

Nessa perspectiva, Miranda e Simeão (2005, s. p.) descrevem que “na análise dos documentos são exigidos princípios, métodos e técnicas que permitem examinar, distinguir e separar cada uma das partes do documento para determinar a categoria, estrutura formal,

propriedades e significado de seu conteúdo temático". A análise feita de cada documento permite a extração e criação de conceitos que serão utilizados para organizar as informações, criando assim as taxonomias.

A estrutura taxonômica é responsável por representar através de conceitos o mapeamento de um domínio de conhecimento e geralmente são apresentadas em uma forma hierárquica, onde são estabelecidas classes e, dentro delas, subclasses para melhor organizar a informação. (Novo, 2010).

Através de investigação em literatura de trabalhos de autores da Arquivologia foi possível perceber que não existem critérios metodológicos para a elaboração de taxonomias, sendo os desenvolvedores responsáveis por estabelecê-los no momento da criação (Viana, 2022).

No artigo de Wang (2023), a autora destaca a importância da taxonomia e da recuperação da informação no contexto do crescente número de pesquisas retratadas, apontando que os sistemas de informação atuais enfrentam desafios significativos para lidar com esse fenômeno. Wang argumenta que a taxonomia, como estrutura classificatória de conhecimento, é essencial para organizar, rastrear e distinguir obras confiáveis daquelas comprometidas, facilitando uma recuperação mais precisa e contextualizada da informação. A autora propõe que os sistemas de recuperação devem incorporar novas funcionalidades, como a capacidade de sinalizar retratações e de oferecer interações mais transparentes aos usuários, contribuindo para um ecossistema informacional mais confiável e ético.

O uso de taxonomias tem se mostrado essencial para a recuperação eficiente da informação em ambientes digitais, especialmente em bases de dados voltadas à ciência da informação. Segundo Coelho (2020), as taxonomias atuam como estruturas organizacionais que facilitam a navegação, o acesso e a localização de conteúdos, promovendo uma melhor experiência

informacional aos usuários.

A autora destaca que, quando bem construídas, as taxonomias permitem a representação lógica e sistemática do conhecimento, favorecendo a indexação precisa dos documentos e, consequentemente, aumentando a relevância e a precisão dos resultados recuperados em sistemas de busca.

Além disso, Coelho (2020) aponta que as chamadas taxonomias navegacionais facetadas possibilitam múltiplos caminhos de acesso à informação, permitindo ao usuário realizar buscas a partir de diferentes perspectivas e critérios.

Esse aspecto é particularmente importante em repositórios digitais como a BDA, onde a diversidade de tipos documentais e de temáticas exige uma estrutura flexível, porém coerente, para garantir a eficácia na recuperação. Dessa forma, o uso de taxonomias não apenas organiza o acervo, mas também potencializa o acesso ao conhecimento, contribuindo para a democratização da informação e para a valorização dos repositórios especializados.

O uso de taxonomias é um grande aliado no processo de organização e recuperação da informação, estas são criadas através da análise documental e utilizadas no processo de criação de categorias para reunir as informações, e no processo de descrição da informação, através da indexação dos documentos por assunto.

Ao que tange a categorização, Silva (2011, p. 272) a define como "o processo cognitivo de compreensão das características dos objetos por critérios de similitude ou dessemelhança". Essa técnica é muito utilizada na Arquivologia pois ajuda na organização e na recuperação de documentos.

A categorização é uma atividade cognitiva, pois é um processo totalmente ligado ao pensar e ao agir dos seres humanos. É uma prática que ocorre quando dois ou mais objetos, seres ou eventos distintos são reunidos devido a uma certa semelhança ou afinidade entre eles (Silva & Lima, 2011).

É uma necessidade humana criar categorias para classificar os inúmeros elementos da natureza. No entanto, dentro dessas categorias existem subclasses que refinam a categorização inicial, transformando em menos geral, para mais específico. Quanto mais subclasses são definidas, menos abrangente se torna a classificação. (Marques, 2021).

Nas bases de dados referenciais muitas vezes se utiliza essa técnica para melhor organizar os documentos depositados, podendo ser categorizados por assunto, ou por modelo de documento, como artigos, periódicos, monografias, livros entre outros tipos de produção científica, além de auxiliar no processo de recuperação da informação. Outro método muito utilizado para organizar é a indexação.

Santos e Corrêa (2017) definem a indexação como:

[...] consiste na descrição do conteúdo de um documento, de forma concisa e condensada, por meio do emprego de termos - também denominados como palavras-chave ou descritores - que exercem a função de pontos de acesso mediante os quais um documento pode ser identificado e recuperado. (Santos & Corrêa, 2017, p. 249).

Em complemento, o procedimento de indexação, de acordo com Souza e Hillesheim, é assim definido:

[...] resulta na representação temática dos documentos por meio de índices e resumos, possibilitando a recuperação dos assuntos principais [...] a indexação descreve seu conteúdo utilizando vários termos de indexação, usualmente selecionados de algum tipo de vocabulário controlado. (Souza & Hillesheim, 2014, p. 85).

A NBR 12.676 (1992), norma que versa sobre métodos para análise de documentos, define três etapas para o processo de indexação: o exame do documento e estabelecimento do

assunto do seu conteúdo, a identificação dos conceitos presentes no assunto e a tradução desses conceitos nos termos de uma linguagem de indexação.

Essa mesma norma, prevê o uso de instrumentos auxiliares para definir os termos para uso da indexação, como os tesouros, códigos de classificação e cabeçalhos de assuntos. Esses instrumentos auxiliam na escolha e definição dos termos a serem utilizados de forma padronizada.

A padronização dos termos de indexação é importante, pois a partir dessa técnica é possível identificar homônimos, eliminar ambiguidades, escolher o melhor termo para representar uma palavra ou sentença que possuem várias formas de serem escritas, como por exemplo: arquivista e arquivologista.

Para a indexação de materiais escritos como os presentes em base de dados, a NBR 12.676 (1992), define que deve ser feita uma leitura técnica do documento a ser analisado, isso elimina a necessidade da leitura por completo do material, os termos devem ser extraídos com a leitura do título, subtítulo, resumo, sumário e introdução. Após a análise das seções indicadas é possível ter informação suficiente para a extração dos termos para realizar a indexação.

O trabalho do indexador está diretamente ligado ao processo de organização e recuperação da informação, pois ao padronizar e conhecer os termos utilizados para indexar as obras ali representadas, é possível fazer pesquisas com uma maior relevância e entregar resultados melhores ao usuário. O uso da indexação é uma das formas de ter uma melhor recuperação da informação.

O Dicionário de Biblioteconomia e Arquivologia de Cunha e Cavalcanti (2008, p. 307) define a R.I. como: "Restituição dos dados constantes do sistema, para obtenção de informações específicas ou genéricas" e "A restituição, ou recuperação, abrange o processo total de identificação, busca, encontro e extração da informação

armazenada".

A recuperação da informação tem como primeira atividade a montagem de uma estratégia de busca. O objetivo da estratégia de busca é entender e delimitar qual é a lacuna informacional do usuário, definir os melhores termos para realizar a busca, visando a extração de materiais e documentos pertinentes com o objeto de pesquisa (Marques, 2021).

Vieira e Garrido (2011, s. p.) definem o objetivo principal da recuperação da informação como:

[...] recuperar mais documentos que alcancem a necessidade informacional dos usuários e recuperar menos documentos irrelevantes, é o objetivo central da recuperação da informação. No entanto, tudo está interligado: a recuperação efetiva da informação relevante depende tanto das tarefas dos usuários, quanto da visão lógica dos documentos adotados pelos sistemas de recuperação de informação. [...] A princípio, o objetivo de um sistema de recuperação de informação é apenas atender a necessidade de informação do usuário, e não recuperar com exatidão todos os documentos requeridos. A recuperação de dados diferencia-se da recuperação da informação, uma vez que a sua precisão e exatidão é essencial e imprescindível. [...]

Os operadores booleanos^{viii} auxiliam no processo de busca em Base de Dados. Sua estrutura básica é composta pelos termos *and*, *or*, *not*. O *And* é utilizado para unir dois termos distintos. Por exemplo “indexação *and* Arquivologia”. Ao utilizar este operador, vão ser pesquisadas e recuperadas na base, documentos que possuem em seu escopo os dois termos. O *or* é utilizado como forma elegível, busca por um termo ou outro, como no caso de “veículo *or* automóvel”. O *not* é utilizado para excluir um ou mais termos. (Picalho, Lucas & Amorim, 2022).

Ingwersen (1992) traz em sua obra três pontos de vista para pesquisa e desenvolvimento da

recuperação da informação: *Aboutness*, considerado o tema ou o assunto a ser pesquisado; tipos de representação da informação; a relevância e avaliação dos resultados recuperados.

Utilizar os operadores booleanos refinam o processo de busca e melhoram a recuperação e extração de materiais e informações pertinentes para uso. Além dos operadores booleanos, as técnicas de tratamento temático da informação, como a categorização e a indexação, também auxiliam no processo de recuperação da informação.

Neste trabalho foram abordadas temáticas sobre o uso de tecnologias da informação na Arquivologia, repositórios digitais, base de dados, métodos e técnicas da área de representação temática da informação e a recuperação da informação. No segmento, demonstrou-se como elas são aplicadas na BDA.

A discussão teve como objeto de estudos a identificação de técnicas de organização da informação na BDA para propor melhorias e analisar a relação entre TTI e RI.

6 Análise e discussão dos resultados

Para realizar a investigação de como é feito o TTI dentro da BDA, o primeiro passo adotado foi a leitura de bibliografia pertinente ao assunto. O procedimento permitiu a identificação de métodos e técnicas de organização da informação que poderiam ser adotados pela base de dados para a organização de seus documentos, como a taxonomia, categorização, indexação, entre outros métodos.

O segundo passo adotado foi navegar na própria plataforma para identificar quais dos métodos/técnicas foram utilizadas na BDA. Foi possível identificar três metodologias adotadas, as taxonomias, a categorização e a indexação, utilizadas para organizar e auxiliar na recuperação da informação.

Na BDA, empregam-se taxonomias para estabelecer a nomenclatura dos metadados a serem preenchidos durante o cadastramento

dos documentos, por exemplo: ano, autores, coleções, edições, editoras, *key words*, locais, natureza, organizadores, palavras-chave. Também adotou-se a taxonomia “categoria” para classificar os objetos cadastrados em categorias que os englobam.

A Base de Dados por ser desenvolvida através do Tainacan, utiliza de taxonomias para se criar vocabulários de categorização em seu acervo digital (Luna, 2023). As taxonomias são responsáveis por dar nome às coleções presentes na BDA.

O Tainacan não utiliza um tesouro padrão para definir o nome de suas coleções. Devido ao grande número de áreas do conhecimento que utilizam essa ferramenta para gerenciar as suas coleções, a escolha da nomenclatura fica no encargo da instituição gestora. (Silva & Segundo, 2019). A Figura 2, a seguir, demonstra as coleções presentes na BDA.

Figura 2: Página de Coleções da Base de Dados em Arquivística



Fonte: Base de Dados em Arquivística (2024)

As coleções estão organizadas com base em dois tipos principais de categorização: Obras e Arquivística, totalizando 10.036 itens cadastrados. A coleção Obras reúne, aproximadamente, 890 itens, incluindo livros, cadastro principal dos periódicos e anais de eventos em geral.

Por sua vez, a coleção Arquivística conta atualmente com 9146 itens, entre os quais se destacam artigos científicos, capítulos de livros e artigos apresentados nos eventos científicos. Esses registros podem, ou não, estar associados a materiais cadastrados na coleção Obras.

A segunda categoria das coleções, Arquivística, pode ser entendida como subcategoria da categoria Obras, sendo que os documentos depositados são categorizados por tipo de produção bibliográfica. Na coleção de eventos científicos, foram cadastrados os Anais, cadernos de resumos, programação dos eventos científicos da área de Arquivologia. Na coleção de periódicos foram cadastradas as revistas científicas seriadas, impressas e digitais, e outros tipos de periódicos. Já na coleção de monografias, são depositados os livros, capítulos de livros e manuais relacionados à área de Arquivologia.

Ao navegar pela página da BDA, podemos observar o uso da categorização, uma importante ferramenta para a organização e recuperação da informação. Com o aprofundamento da navegação dentro dos materiais depositados na Base de Dados, também se consegue perceber o uso de outra técnica que auxilia na organização e recuperação da informação, que é a indexação.

O Tainacan permite a criação e a utilização de palavras-chave para descrever os documentos depositados. Os descritores utilizados nas obras podem ser considerados como termos de indexação. A extração dos termos para a indexação é realizada através da análise do documento ali depositado. A imagem a seguir, Figura 3, retirada da BDA, exemplifica como é feita essa descrição:

Figura 3: Página de documento cadastrado na Base de Dados em Arquivística



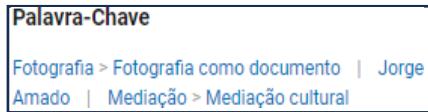
Fonte: Base de Dados em Arquivística (2024)

Na página é possível identificar os metadados preenchidos com as informações relevantes sobre o item como o resumo, paginação, observações entre outros. As informações aliadas às palavras-chave auxiliam na

identificação e na descrição do documento cadastrado na base.

A imagem a seguir, Figura 4, traz os termos utilizados para indexar um documento cadastrado na BDA, onde são apresentados os termos gerais (pai) e seus termos específicos.

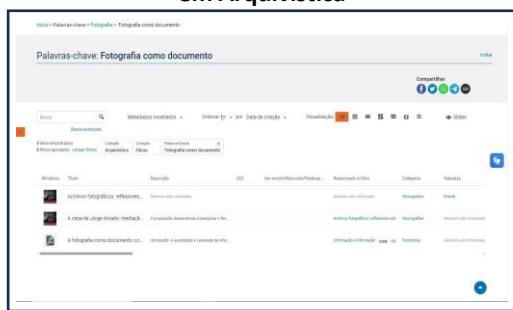
Figura 4: Termos descritores de uma obra cadastrada na BDA



Fonte: Base de Dados em Arquivística (2024)

A indexação realizada pela BDA não só descreve os documentos cadastrados por meio de taxonomias, mas também reúne e organiza todos os documentos que tratam do mesmo assunto em um único local. Ao clicar em um termo descritor, se consegue recuperar todos os documentos indexados com o termo conforme ilustrado na imagem a seguir, Figura 5.

Figura 5: Página da Palavra-chave “fotografia como documento” cadastrada na Base de Dados em Arquivística



Fonte: Página da Base de Dados em Arquivística (2025)

O uso da indexação, utilizada na representação e organização da informação, permite uma melhor usabilidade de uma base de dados, como visto na BDA, e auxilia na recuperação da informação dos documentos cadastrados.

6.1 Problemáticas identificadas no uso do tratamento temático da informação

Foram identificadas algumas problemáticas ligadas ao uso do tratamento temático das informações adotadas na BDA. Os problemas

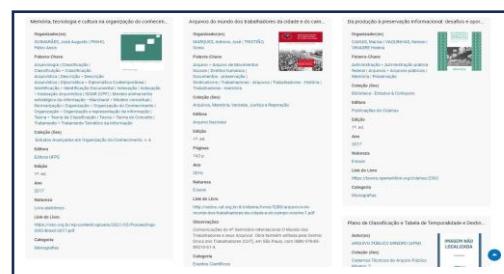
identificados estão relacionados à criação das categorias para categorização e no processo de indexação dos documentos cadastrados na Base de Dados.

Quando foi proposta, a BDA apresentava três categorizações: Periódicos, Eventos científicos e Monografias. Num segundo momento, a fim de unificar os procedimentos para o cadastramento, foram definidas duas categorias. A primeira, “Obras”, inclui o material bibliográfico geral, ou seja, a denominação do(s) evento(s) científico(s), o título do(s) periódico(s) e o título do(s) livro(s) e manual(ais). Na segunda “Arquivística”, são preenchidos, com detalhamento, os metadados referentes aos artigos apresentados nos eventos científicos, artigos publicados nos periódicos e, no caso dos livros, os capítulos que refletem temáticas relacionadas à arquivística.

Ainda que as nomenclaturas das categorias principais sejam genéricas, os vários tipos de materiais bibliográficos estão identificados. Para isso, o pesquisador precisa compreender o percurso durante as buscas, o que pode ocasionar um tempo mais longo na resposta.

Na imagem a seguir, Figura 6, temos exemplos de diferentes tipos de materiais cadastrados na categoria “Obras”, onde são apresentadas duas monografias e um evento científico.

Figura 6: Página da categoria “Obras”



Fonte: Base de Dados em Arquivística (2025)

Uma proposta de solução para as problemáticas ligadas à categorização seria extinguir as categorias “Obras” e “Arquivística” e retomar apenas a categorização ligada aos tipos de produção. A técnica permite a criação de categorias mais específicas para melhor ordenar os documentos.

Na categoria de “Eventos Científicos” seriam cadastrados os eventos como um todo e os itens ligados aos eventos seriam cadastrados em subcategorias como “Anais de evento”^{ix}, “Programação”, “Livro de resumos”^x e por fim criar mais uma subcategoria para categorizar os artigos apresentados no evento e publicados nos Anais, denominado “Artigo dos Anais”.

Na categoria de “Monografias” a nomenclatura seria substituída por “Livros e Manuais”, onde seriam cadastrados o inteiro teor dos documentos e os itens ligados aos livros e manuais em subcategorias como “Capítulo de livro” e “Capítulo de manual”. Na categoria de “Periódicos” seriam cadastrados os periódicos como um todo, e os itens ligados aos periódicos em uma subcategoria denominada “Artigo de periódico”.

Ao adotar categorias gerais e outras mais específicas, a recuperação dos itens cadastrados tornaria a pesquisa mais rápida. Essa técnica, categorias gerais e específicas, colabora com a melhora na experiência do usuário ao utilizar a BDA.

Ao navegar nas categorias e nos itens presentes na BDA, nota-se a presença de termos de indexação, representados pela nomenclatura de palavras-chave, utilizados como pontos de acesso e recuperação da informação. Aprofundando no uso da técnica foi possível identificar três pontos a serem considerados.

O primeiro é que nem todos os itens cadastrados estão indexados, por exemplo, em vários artigos, sobretudo os mais antigos, estão ausentes o resumo, as palavras-chave, o ano de produção entre outras informações relevantes para a extração dos termos de indexação. Isso deve-se ao fato de que os textos publicados anteriormente não registravam os resumos e as palavras-chave, ocasionando lacunas nos metadados a serem preenchidos.

Quanto ao segundo é a ausência do uso de uma linguagem documentária estruturada para padronizar e organizar os termos utilizados para indexação. Os dois pontos

levantados são importantes, pois prejudicam o usuário no processo de busca e recuperação da informação.

A ausência de atribuição das palavras-chave ocorre devido à ausência de informação na produção científica a ser cadastrada.

Há a possibilidade de os usuários realizarem pesquisas apenas utilizando os termos atribuídos como palavras-chave. Dada essa opção de navegação, é crucial que todos os itens cadastrados estejam devidamente indexados. O usuário pode empregar essa função da plataforma para recuperar todos os itens cadastrados que abordem uma palavra-chave específica. Quando documentos não são indexados, isso prejudica significativamente a recuperação da informação.

O outro ponto a ser observado é na criação dos termos a serem utilizados no momento da indexação. Ao navegar no item que possui a atribuição do termo, é possível notar a estrutura de uma linguagem documentária, apresentando a relação hierárquica dos termos, indo do mais geral para o mais específico, conforme a Figura 7, a seguir:

Figura 7: Palavras-chave atribuídas a um item cadastrado na Base de Dados

Palavra-Chave
Acesso > Acesso à informação Arquivo > Arquivo Universitário Memória > Memória Institucional Preservação > Preservação de Documentos

Fonte: Base de Dados em Arquivística (2024)

Para a criação de termos de indexação, geralmente, o profissional que trabalha com esse tipo de técnica extrai os termos do documento em sua linguagem natural. Com o auxílio de uma linguagem documentária, como um tesouro ou um vocabulário controlado básico, ele transforma esses termos da linguagem natural em uma linguagem documental. O termo traduzido para uma linguagem documental será utilizado para indexar todos os outros itens que tratam do mesmo assunto.

Visando a melhoria dos termos utilizados para a indexação dos itens cadastrados e por se

tratar de uma base de dados temática em Arquivologia, a BDA poderia utilizar o Tesauro Brasileiro de Ciência da Informação (2014), criado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT, o Dicionário de Biblioteconomia e Arquivologia, escrito por Cunha e Cavalcanti (2008) e o Dicionário Brasileiro de Terminologia Arquivística (2005), criado pelo Arquivo Nacional. Os instrumentos são fontes de pesquisa para a revisão dos termos já criados e extração de novos termos a serem utilizados para a indexação dos itens.

Outra alternativa é utilizar esses documentos como fonte de pesquisa e criar um tesauro ou um Vocabulário Controlado Básico (VCB), especializado em Arquivologia, para ser utilizado na Base de Dados. Criando uma linguagem documentária, além de auxiliar e melhorar o trabalho de cadastramento e indexação, também é possível contribuir com outros repositórios que indexem documentos que abordam a Arquivologia.

Após a revisão e criação dos termos, é possível selecionar os itens já cadastrados na Base de Dados que ainda não possuem palavras-chave atribuídas e, em seguida, designar adequadamente os termos que melhor descrevem cada item. Esse processo contribuirá significativamente para a organização e a recuperação eficiente da informação pelos usuários, que poderão utilizar as palavras-chave atribuídas aos itens para realizar buscas de maneira mais eficaz. Ressalta-se que estão previstos projetos para melhoria da funcionalidade da BDA e adequação do vocabulário controlado.

Através da literatura das áreas de C.I e de Arquivologia, foi possível compreender a relação entre o uso de metodologias de TTI com a RI. Com o uso correto das técnicas, as informações a serem depositadas são organizadas de maneira eficiente, o que impacta diretamente na recuperação da informação ali registrada.

Na BDA, foram empregadas três metodologias de TTI. A primeira consiste no uso de taxonomias para criar e alimentar os metadados de catalogação dos itens. A

segunda envolve a categorização, que organiza os itens cadastrados em grupos que melhor os representam. Por fim, a terceira metodologia é a indexação, utilizada para descrever o conteúdo dos itens cadastrados.

O uso de taxonomias e a categorização desempenham um papel fundamental na recuperação da informação, especialmente no contexto de bases de dados em Ciência da Informação, conforme discutido por Santos e Duque (2023). As taxonomias funcionam como estruturas ontológicas que organizam conceitos e categorias de forma hierárquica e sistemática, facilitando a classificação e a representação do conhecimento. Essa organização permite que os sistemas de recuperação da informação possam mapear e relacionar termos e conceitos de maneira mais precisa, melhorando a precisão e a relevância dos resultados encontrados.

A categorização e a indexação estão diretamente ligadas ao processo de recuperação da informação. Através da categorização, é possível organizar e recuperar itens que compartilham características em comum, enquanto a indexação possibilita a organização e a recuperação de itens com conteúdo similares.

A utilização de metodologias de organização e descrição da informação enriquece e facilita a experiência do usuário ao acessar a base de dados como fonte de pesquisa e extração de informações relevantes. Além disso, essas metodologias auxiliam e padronizam o processo de cadastramento de novos itens.

7 Considerações finais

A pesquisa foi realizada para identificar as técnicas e metodologias de tratamento temático da informação utilizadas na BDA, propor soluções e melhorias e analisar a relação entre TTI e RI. O desenvolvimento da pesquisa se deu a partir de leituras referentes à organização da informação, focando no tratamento temático de materiais

bibliográficos e na navegação na base de dados.

Por meio da revisão de literatura e da navegação na plataforma, foi possível iniciar um estudo de caso, identificando os métodos e técnicas de organização da informação utilizados na (BDA), como o uso de taxonomias, categorização e indexação. A identificação permitiu observar algumas problemáticas relacionadas ao uso e à aplicação das metodologias no repositório.

Os pontos observados para melhoria estão relacionados à escolha das taxonomias para categorização, à forma de agrupamento dos materiais nas categorias, à falta de padronização dos termos de indexação e aos materiais cadastrados na plataforma que não receberam termos indexadores.

Para a melhoria e soluções dos problemas relacionados a indexação foi proposta uma revisão dos termos já cadastrados, utilizando documentos auxiliares, como o Tesauro Brasileiro da Ciência da Informação (2014), o Dicionário de Biblioteconomia e Arquivologia (2008) e o Dicionário Brasileiro de Terminologia Arquivística (2005), para revisar e extrair novos termos para serem utilizados na indexação dos itens cadastrados.

O uso dos documentos proporciona a oportunidade de criar um Tesauro ou um Vocabulário Controlado Básico (VCB) específico para a área de Arquivologia, a ser implementado para a criação e padronização dos termos de indexação dos itens cadastrados na Base de Dados.

Após a revisão e criação dos novos termos, a próxima etapa proposta foi identificar as produções científicas e técnicas que não possuem palavras-chave atribuídas e, posteriormente, atribuí-las por meio de uma análise prévia dos itens cadastrados, visando identificar as expressões que melhor descrevem o conteúdo dos documentos depositados.

A partir da leitura dos materiais da área de C.I. e Arquivologia foi possível identificar que o tratamento temático da informação não apenas influencia na organização da informação, mas também na recuperação da informação. Através do uso da categorização, é possível organizar e recuperar itens que compartilham características em comum, enquanto a indexação possibilita a organização e a recuperação de itens com conteúdo similares.

Para estudos futuros sugere-se a análise de outras estruturas da BDA, promovendo investigações constantes do funcionamento e da naveabilidade, para que ela continue sendo uma fonte de pesquisa eficiente e de qualidade para os usuários interessados na área de Arquivologia.

Portanto, a pesquisa a partir do entendimento da organização dos itens cadastrados na BDA, propôs possíveis alternativas de melhorias para serem aplicadas no uso das taxonomias, no tratamento temático da informação, que vão gerar impacto na recuperação da informação. Buscou-se, dessa maneira, possibilitar ajustes na navegação e usabilidade da plataforma por parte dos usuários.

Pretende-se que, a partir dos resultados da pesquisa, a BDA continue em constante evolução e crescimento, trazendo maiores benefícios para a comunidade acadêmica da Arquivologia e áreas afins. Ademais, as propostas apresentadas possibilitam uma melhor naveabilidade e usabilidade do repositório por seus usuários e por seus colaboradores, se transformando em um importante aliado para as pesquisas voltadas para a temática arquivística.

8 Referências

Albrecht, R. F., & Ohira, M. L. B. (2000). Bases de dados: metodologia para seleção e coleta de documentos. *Revista Acb*, 5(5), 131-144.

Almeida Júnior, O. F. de. (2020). Prefácio. In L. S. Freitas et al. (Orgs.), *O PPGCI/UFF nos Enancib: Trabalhos premiados* (pp. 9–13). São Paulo: Tikinet. <http://ppgci.uff.br/wp->

content/uploads/sites/86/2021/04/O-PPGCI-UFF-nos-Enancib_Completo_Final-1.pdf

Artefactual Systems. (2015). *AtoM*.
<https://www.accesstomemory.org/pt-br/>

Associação Brasileira de Normas Técnicas. (1992). *NBR 12676: Métodos para análise de documentos – determinação de seus assuntos e seleção de termos de indexação*.
<https://www.semas.pa.gov.br/wp-content/uploads/2018/04/NBR-12676-INDEXACAO.pdf>

Boaventura, E. M. (2004). *Metodologia da pesquisa: Monografia, dissertação e tese*. Atlas.

Bräscher, M., & Guimarães, J. A. C. (2018). Tratamento temático da informação (TTI): influência dos paradigmas físico, cognitivo e social em artigos de revisão de literatura no período de 1966-1995. *Liinc Em Revista*, 14(2).
<https://doi.org/10.18617/liinc.v14i2.4347>

Bräscher, M., & Monteiro, F. de S. (2010). Organização da informação em repositórios digitais. *Encontros Biblio: revista eletrônica de biblioteconomia e ciência da informação*, 15(29).
<https://periodicos.ufsc.br/index.php/eb/article/view/13705>

Brasil. Casa Civil. Secretaria-Executiva. Arquivo Nacional. Conselho Nacional de Arquivos. (2022). *E-arq Brasil: modelo de requisitos para sistemas informatizados de gestão arquivística de documentos*. CONARQ.
<https://www.gov.br/conarq/pt-br/centrais-de-conteudo/publicacoes/EARQV203MAI2022.pdf>

Brasil. Ministério da Cultura. Brasiliana Museus. (s.d.). *Testes de Boques Tainacan*.
<https://brasiliana.museus.gov.br/testes-de-bloques-tainacan/>

Café, L. M. A., & Sales, R. de. (2010). Organização da informação: conceitos básicos e breve fundamentação teórica. In J. Robredo & M. Bräscher (Orgs.), *Passeios pelo bosque da informação: estudos sobre representação e organização da informação e do conhecimento* (pp. 115–129). IBICT.

Coelho, A. I. R. C. G. (2020). *A produção científica sobre taxonomias navegacionais facetadas: Estudo a partir de bases de dados da ciência da informação (2010–2018)* [Dissertação de mestrado, Universidade de Coimbra].

Consultative Committee for Space Data Systems (CCSDS). (2012, June). *Reference model for an open archive information system (OAI/S)* (CCSDS 650.0-M-2). CCSDS Secretariat.
<http://public.ccsds.org/publications/archive/650x0m2.pdf>

Costa, M., et al. (2016). *Guia do usuário Archivematica*. Instituto Brasileiro de Informação em Ciência e Tecnologia.
<http://livroaberto.ibict.br/123456789/1063/4/Manual-Archivematica.pdf>

Crow, R., et al. (2002). *The case for institutional repositories: a SPARC position paper*.
<https://sparcopen.org/wp-content/uploads/2016/01/instrepo.pdf>

Cunha, M. B. da, & Cavalcanti, C. R. de O. (2008). *Dicionário de Biblioteconomia e Arquivologia*. Briquet de Lemos.
<http://www.realp.unb.br/jspui/handle/10482/34113>

Dias, E. W. (2001). Contexto digital e tratamento da informação. *DataGramZero: Revista de Ciência da Informação*, 2(5), 1-10.

Guimarães, J. A. C., Ferreira, G. M., & Freitas, M. F. M. (2012). Correntes teóricas do tratamento temático da informação: uma análise de domínio da presença da catalogação de assunto e da indexação nos congressos de ISKO-España. In *Congreso ISKO Capítulo Español, 10., 2011. Actas* (pp. 181–194). Ferrol. https://iskoiberico.org/wp-content/uploads/2014/09/dd282-181-194_guimaraes.pdf

Guimarães, J. A. C., Sales, R., & Grácio, M. C. C. (2012). A dimensão interdisciplinar da análise documental nos contextos brasileiro e espanhol no âmbito da organização do conhecimento. *DataGramZero*, 13(6).
<https://cip.brapci.inf.br/download/45874>

Hillesheim, A. I. de A., & Fachin, G. R. B. (2016). Prática de tratamento da informação: relato de experiência docente. *Revista ACB*, 21(3), 627–638.
<https://revista.acbsc.org.br/racb/article/view/1204>

Ingwersen, P. (1992). *Information retrieval interaction*. Taylor Graham.
https://peteringwersen.info/publications/0060_ingroupersen_iri.pdf

Luna, M. M. (2023). Tainacan Beta 0.20.3 – Gerenciando e exibindo termos de taxonomias. Tainacan.
<https://tainacan.discourse.group/t/tainacan-beta-0-20-3-gerenciando-e-exibindo-termos-de-taxonomias/1436>

Marques, J. M. L. (2021). *Nível de relevância dos resultados de busca apresentados pela Netflix: um estudo de caso com consumidores de conteúdos LGBTQI+* (Trabalho de Conclusão de Curso, Universidade de Brasília).
<https://bdm.unb.br/handle/10483/28601>

Martins, D. L., Silva, A. S., Santarém Segundo, J. A., & Albuquerque, E. M. (2017). Repositório digital com o software livre Tainacan: Revisão da ferramenta e exemplo de implantação na área cultural com a revista Filme Cultura. Anais do XVIII Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ENANCIB).
<https://www.researchgate.net/publication/338045811>

Melo, K. I. (2021). BASE DE DADOS EM ARQUIVÍSTICA: A PRODUÇÃO CIENTÍFICA BRASILEIRA. In *Arquivo, documento e informação em cenários híbridos: anais do Simpósio Internacional de Arquivos* (Vol. 8). Eventus.
<https://www.even3.com.br/anais/simposiointernationalearquivos/292347-BASE-DE-DADOS-EM-ARQUIVISTICA--A-PRODUCAO-CIENTIFICA-BRASILEIRA>

Melo, K. I., Souza, S., & Paiva, D. (2023). Construção da base de dados em arquivística: uma ferramenta de pesquisa. *Participação*, 1(38).
<https://periodicos.unb.br/index.php/participacao/article/view/46805>

Mendes, I., & Pinto, V. B. (2019). Taxonomia nas áreas da biblioteconomia e da ciência da informação: uma revisão sistemática. *Páginas A&B, Arquivos e Bibliotecas (Portugal)*, (12).
<https://cip.brapci.inf.br//download/127650>

Miranda, A., & Simeão, E. (2005). Estrutura da informação e modelo extensivo: uma abordagem para a Ciência da Informação. In A. Miranda & E. Simeão (Orgs.), *Informação e tecnologia: conceitos e recortes* (pp. 177–199). Universidade de Brasília. (Série Comunicação da Informação Digital, v. 1).

Novo, H. F. (2010). A taxonomia enquanto estrutura classificatória: uma aplicação em domínio de conhecimento interdisciplinar. *PontodeAcesso*, 4(2), 131–156.
<https://periodicos.ufba.br/index.php/revistaici/article/view/4103>

Messenlehner, B., & Coleman, J. (2014). *Criando aplicações web com WordPress: WordPress como um framework de aplicações*. Novatec Editora.
<https://books.google.com.br/books?hl=pt-BR&lr=&id=1KPDBAAQBAJ&oi=fnd&pg=PA293&dq=conceito+de+wordpress>

Oliveira, R. R. de, & Carvalho, C. L. de. (2009). Implementação de interoperabilidade entre repositórios digitais por meio do protocolo OAI-PMH (pp. 3–9). Universidade Federal de Goiânia, Instituto de Informática.
https://www.researchgate.net/profile/Renan-Oliveira/publication/266459173_Implementacao_de_Interoperabilidade_entre_Repositorios_Digitais_por_meio_do_Protocolo_OAI-PMH/links/55bf3d9b08aed621de122906/Implementacao-de-Interoperabilidade-entre-Repositorios-Digitais-por-meio-do-Protocolo-OAI-PMH.pdf

Oliveira, E. D., et al. (2019). Base de dados e suas vertentes: uma organização das emissoras de rádio da região de Campinas. In *Anais do 28º Congresso Brasileiro de Biblioteconomia, Documentação e Ciência da Informação - FEBAB*.
<https://portal.febab.org.br/cbbd2019/article/view/2390>

Picalho, A. C., Lucas, E. R. de O., & Amorim, I. S. (2022). Lógica booleana aplicada na construção de expressões de busca. *AtoZ: Novas Práticas em Informação e Conhecimento*, 11.
<https://brapci.inf.br/#/v/193823>

Pinheiro, L. V. R., & Ferrez, H. D. (2014). *Tesauro Brasileiro de Ciência da Informação*. Instituto Brasileiro de Informação em Ciência e Tecnologia.
http://sitehistorico.ibict.br/publicacoes-e-institucionais/tesauro-brasileiro-de-ciencia-da-informacao-1/copy_of_TESAUROCOMPLETOFINALCOMCAPA24102014.pdf

Prodanov, C. C., & Freitas, E. C. (2013). *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico* (2ª ed.). Editora Feevale.

Rabelo, N. B., & Cé, G. (2023). Repositórios digitais: especificidades e proximidades sob a ótica da ciência da informação. *Brazilian Journal of Information Science*, 17.
<https://brapci.inf.br/#/v/220255>

Rocha, M. M. V. da. (2021). *Um olhar sobre os cursos de bacharelado em arquivologia no Brasil à luz do regime de informação* (Tese de doutorado, Universidade Federal da Paraíba).
https://repositorio.ufpb.br/jspui/handle/123456789/20883?locale=pt_BR

Santos, R. F. dos, & Corrêa, R. F. (2017). Organização da informação em repositórios

digitais: uma abordagem sobre a política de indexação da Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação (BRAPCI). In *Congresso Brasileiro em Organização e Representação do Conhecimento (ISKO-BRASIL)*, Recife.

<https://cip.brapci.inf.br//download/122077>

Santos, L. S., & Duque, C. G. (2023). *Estudos ontológicos aplicado ao contexto: Base de dados em Ciência da Informação - BRAPCI*.

Sayão, L. F. (1996). Bases de dados: a metáfora científica. *Ciência da Informação*, 25(3).

<https://doi.org/10.18225/ci.inf.v25i3.629>

<https://revista.ibict.br/ciinf/article/view/629>

Silva, A. R. da, & Lima, G. Â. B. de O. (2011). As categorias e a categorização cognitiva contemporânea: enfoque sobre os olhares da biblioteconomia e ciência da informação e das ciências cognitivas. *Revista Interamericana de Bibliotecologia*, 34(3).

<https://cip.brapci.inf.br//download/83778>

Silva, A. R. da. (2011). Estudo dos princípios de categorização na biblioteconomia e ciência da informação: Ranganathan - entre a teoria clássica e a abordagem cognitiva contemporânea. *Perspectivas em Ciência da Informação*, 16(4).

<https://brapci.inf.br/#/v/32784>

Silva, M. E. dos S. (2023). Bases de dados em arquivologia no Brasil: estudo da usabilidade na Base de Dados em Arquivística (BDA) e na Base de Dados Pesquisas Arquivísticas Brasileiras (PAB) (Trabalho de Conclusão de Curso, Universidade Federal da Paraíba, João Pessoa).

<https://repositorio.ufpb.br/jspui/bitstream/123456789/28930/1/MESS13112023.pdf>

Silva, L. C., & Segundo, J. E. S. (2019). Componentes de representação da informação em ambientes de informação digital: estudo do sistema de organização do software Tainacan. In *ENANCIB 2019*.

<https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1333/912>

Sousa, R. T. B. de, & Araújo Júnior, R. H. de. (2015). A classificação e a taxonomia como instrumentos efetivos para a recuperação da informação arquivística. *Ciência da Informação*, 42(1).

<https://revista.ibict.br/ciinf/article/view/1400>

Souza, F. P. de, & Hillesheim, A. I. de A. (2014). Tratamento da informação e o uso das tecnologias da informação e comunicação. *Biblionline*, 10(2).

<https://brapci.inf.br/#/v/16519>

Souza, M. I. F., Vendrusculo, L. G., & Melo, G. C. (2000). Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. *Ciência da Informação*, 29, 93–102.

<https://www.scielo.br/j/ci/a/tcW3q4WvNBQNTqTylK8qfFF/>

Tartarotti, R. C. D., Dal'Evedove, P. R., & Fujita, M. S. L. (2015). Atuação bibliotecária na análise de assunto de teses nas dimensões teóricas da catalogação de assunto e indexação. In J. A. C. Guimarães & V. Dodebe (Orgs.), *Organização do conhecimento e diversidade cultural* (pp. 329–337).

ISKO-Brasil; FUNDEPE. <https://cip.brapci.inf.br/download/135506>

Taylor, A. G., & Joudrey, D. N. (2004). *The organization of information*. Libraries Unlimited.

https://books.google.com.br/books?hl=pt-BR&id=hK_2EAAAQBAJ

Terra, J. C. C., & Bax, M. P. (2003). Portais corporativos: instrumento de gestão de informação e de conhecimento. In I. Paim (Org.), *A gestão da informação e do conhecimento* (pp. 33–53). Belo Horizonte.

Viana, J. Q., et al. (2022). Metodologia para a construção de taxonomia corporativa facetada.

<https://repositorio.ufmg.br/handle/1843/44638>

Vieira, A. A. N., et al. (2015). Novas perspectivas e diálogos para a arquivologia: tecnologias e ciências jurídicas. *Informação & Tecnologia*, 2(1).

<https://brapci.inf.br/#/v/41340>

Vieira, A. P. F., Oliveira, L. P. de, & Cunha, T. M. (2017). Incursões sobre o tratamento temático da informação: estudo da política de indexação em bibliotecas universitárias goianas. *Informação em Pauta*, 2(1), 28–49.

<http://www.periodicos.ufc.br/informacaoempauta/article/view/6542/30021>

Viera, A. F. G., & Garrido, I. S. (2011). Folksnomia como uma estratégia para recuperação colaborativa da informação. *DATAGRAMAZERO*, 12(2).

<https://brapci.inf.br/#/v/7336>

Wang, P. (2023, March). Rising of retracted research works and challenges in information systems: Need new features for information retrieval and interactions. In *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval* (pp. 69–82).

Yin, R. K. (2001). *Estudo de caso: planejamento e métodos* (2. ed.). Bookman.

<https://edisciplinas.usp.br/pluginfile.php/659841>

6/mod_resource/content/1/Livro%20Robert%20Y
in.pdf

NOTAS

ⁱ A Base de Dados em Arquivística (BDA), desde a sua criação, é coordenada pela Profa. Katia Isabelli Melo. A proposta inicial surgiu como um Projeto de Iniciação Científica (ProIC), da Universidade de Brasília.

ⁱⁱ A pesquisa resulta de um estudo desenvolvido ao término do curso de graduação em Arquivologia, da Universidade de Brasília.

ⁱⁱⁱ Conforme Crow (2002), repositórios institucionais são coleções que capturam e preservam a produção intelectual de uma instituição ou comunidade.

^{iv} O e-ARQ Brasil é um modelo de requisitos para a gestão arquivística de documentos digitais, desenvolvido pelo Conselho Nacional de Arquivos (CONARQ) no Brasil. O objetivo do e-ARQ Brasil é fornecer diretrizes e normas para a criação, manutenção, preservação e acesso a documentos arquivísticos digitais, garantindo a autenticidade, confiabilidade e integridade desses documentos ao longo do tempo. (CONARQ, 2022).

^v Conforme registrado pelos autores, “Dublin Core pode ser definido como sendo o conjunto de

elementos de metadados planejado para facilitar a descrição de recursos eletrônicos.” (Souza, Vendrusculo & Melo, 2000, p. 93).

^{vi} Segundo os autores, “O OAI-PMH (Open Archives Initiative— Protocol for Metadata Harvesting) é um protocolo que possibilita aos participantes da iniciativa OAI compartilhar seus metadados, para aplicações externas que se interessem na coleta desses dados.” (Oliveira & Carvalho, 2009, p. 7).

^{vii} Integram a equipe da BDA discentes de graduação, bolsistas e voluntários, vinculados ao Projeto de Iniciação Científica e ao Projeto de Extensão, coordenados pela Profª. Katia Isabelli Melo, docente do Curso de Arquivologia, da Universidade de Brasília.

^{viii} “[...] a álgebra booleana consiste em um conjunto lógico, binário e bivalente criado pelo matemático George Boole na metade do século XIX.” (Picalho, Lucas & Amorim, 2022, p. 3).

^{ix} Utilizado para cadastrar os Anais como um todo.

^x Como há eventos para os quais os Anais não estão disponíveis, são cadastradas as programações ou cadernos de resumos relacionados a esses eventos.