



Avaliação da qualidade das bases de dados BRAPCI e PERI da área de Ciência da Informação

Quality evaluation of the BRAPCI and PERI databases of the Information Science area

Welze Rocha Arruda 

Graduada em Biblioteconomia
Universidade Federal do Rio Grande do Norte
welze.luna@gmail.com

Carla Beatriz Marques Felipe 

Doutoranda em Ciência da Informação
Instituto Brasileiro de Informação em Ciência e Tecnologia
felipecarla12@gmail.com

Raimunda Fernanda dos Santos 

Doutora em Ciência da Informação
Universidade Federal do Rio de Janeiro
nandaflorania@gmail.com

Resumo

Retrata aspectos relativos às bases de dados e a sua importância na disseminação e recuperação das produções científicas. Tem como objetivo avaliar duas bases de dados da área de Ciência da Informação, Base de Dados em Ciência da Informação (BRAPCI) e PERI, no que diz respeito aos critérios de qualidade estabelecidos na literatura. A partir disso, conceitua fontes de informação e descreve as suas tipologias, além de apresentar a definição e a função das bases de dados científicas. Descreve os sistemas de recuperação da informação analisados e demonstra sua relevância para a consolidação da área. Utiliza como metodologia a pesquisa bibliográfica, exploratória e descritiva, por meio da observação não participante. Conclui que as duas bases de dados analisadas se encontram dentro dos critérios propostos, exceto no item sobre a apresentação da Base de Dados PERI, em que foram apresentadas sugestões para a sua adequação.

Palavras-chave

Base de Dados em Ciência da Informação - BRAPCI. Critérios de qualidade - Base de Dados. PERI. Sistemas de Recuperação da Informação.

Abstract

Depicts aspects related to databases and their importance in the dissemination and recovery of scientific productions. Aims to evaluate two databases in the area of Information Science, Database in Information Science (BRAPCI) and PERI, with regard to the quality criteria established in the literature. Based on that, it conceptualizes sources of information and describes their typologies, in addition to presenting the definition and function of scientific databases. Describes the information retrieval systems analyzed and demonstrates their relevance for the consolidation of the area. Uses bibliographic, exploratory and descriptive research as a methodology, through non-participant observation. Concludes that the two databases analyzed are within the proposed criteria, except for the item on the presentation of the PERI Database, in which suggestions were made for their adequacy.

Keywords

Database. Database in Information Science - BRAPCI. Quality criteria. PERI. Recovery of Information System.



DOI: [10.28998/cirev.2020v7n1h](https://doi.org/10.28998/cirev.2020v7n1h)

Este artigo está licenciado sob uma [Licença Creative Commons 4.0](https://creativecommons.org/licenses/by/4.0/)

Submetido em: 08/08/2019

Aceito em: 09/04/2020

Publicado em: 15/05/2020

1 INTRODUÇÃO

Com a explosão informacional e o advento de novas formas de registrá-la surgiram ferramentas, técnicas e métodos para armazenamento, organização, representação e recuperação da informação em ambiente digital no intuito de satisfazer as necessidades informacionais de diferentes tipos de usuários.

As bases de dados, por exemplo, se configuram como sistemas de recuperação da informação cujo propósito é armazenar, representar e disponibilizar informações relevantes de acordo com as demandas de uma determinada comunidade, a partir de mecanismos eficientes de recuperação da informação. Diferentemente dos bancos de dados, os quais se configuram como repositórios de dados que são armazenados para finalidades específicas.

Seguindo essa linha de raciocínio, as bases de dados científicas têm como função principal armazenar e indexar publicações científicas de determinada(s) área(s) do conhecimento, visando satisfazer as necessidades informacionais da sua comunidade, bem como contribuir para a difusão e progresso da ciência a partir de fontes de informação científica.

Dessa forma, as bases de dados supracitadas se constituem como um marco para a difusão e compartilhamento do conhecimento científico em todo o mundo, pois são importantes fontes de informação para a comunicação e disseminação da ciência mediante acesso remoto. Com a proliferação de periódicos científicos e de bases de dados que os indexam surge também a necessidade de avaliá-las sob o ponto de vista de sua qualidade, além de enfatizar a importância da atuação dos profissionais da informação em sua estruturação, implementação e atualização.

Nesse sentido, esta pesquisa tem como objetivo avaliar, com base nos vetores propostos por Sayão (2000), a qualidade de duas bases de dados da área de Ciência da Informação, são elas: a Base de Dados em Ciência da Informação gerenciada pela Universidade Federal do Paraná (UFPR) em conjunto com a Universidade Federal do Rio Grande do Sul (UFRS), e a PERI, pertencente à Universidade Federal de Minas Gerais (UFMG).

Por serem duas bases de dados na área de Ciência da Informação e se encontrarem imersas em um contexto acadêmico, pressupõe-se que esses sistemas estejam adequados às normas mais confiáveis, no que diz respeito à qualidade da informação em critérios como disponibilidade, acessibilidade, armazenamento, representação e recuperação da informação.

Sob esse viés, a importância desta pesquisa decorre da necessidade de avaliar as principais características de bases de dados usualmente consultadas por pesquisadores da Ciência da Informação e áreas afins. Além disso, este estudo mostra-se como uma investigação útil para bibliotecários e cientistas da informação, podendo servir também como referência para pesquisas no contexto dos Sistemas de Recuperação da Informação.

Visando dar continuidade a essas considerações, a seção a seguir discorre sobre conceito e tipologias de fontes de informação.

2 FONTES DE INFORMAÇÃO

As fontes de informação surgiram a partir da necessidade de registrar e disseminar a informação por meio de diversos suportes visando atender os mais variados públicos. Para Cunha (2001, p. 8), “[...] o conceito de fontes de informação ou documento é muito amplo, pois pode abranger manuscritos e publicações impressas, além de objetos, como amostras minerais, obras de arte ou peças museológicas [...]”. Assim, são consideradas fontes de informação qualquer meio em que a informação possa ser encontrada. Tais fontes, por sua vez, são divididas

em três categorias, distribuídas de acordo com as suas respectivas funções: primárias, secundárias e terciárias.

De acordo com Pacheco e Valentim (2010) as fontes de informação primárias retratam a interferência direta do autor. As fontes secundárias, por sua vez, facilitam o acesso às fontes primárias, a partir de uma representação realizada de acordo com o seu arranjo ou finalidade. De outro modo disposto, as fontes terciárias viabilizam a recuperação das fontes primárias e secundárias.

As fontes de informação podem ser classificadas como informais e formais. Segundo Muller (2000), as primeiras utilizam os canais informais e contemplam, essencialmente, comunicações de caráter mais pessoal ou referem-se às pesquisas em andamento, dentre as quais: apresentações de estudos em andamentos, trabalhos publicados em eventos técnico-científicos que possuem características semelhantes, etc. Nas fontes informais, a informação é recente e é orientada a públicos restritos, proporcionando acesso limitado. Em linhas gerais, elas não se encontram armazenadas em repositórios, o que dificulta o processo de sua recuperação.

Ainda de acordo com essa autora, as fontes formais, por sua vez, utilizam os canais formais que viabilizam ampla divulgação das publicações científicas. Essas fontes de informação contam com todo o processo de tratamento e estocagem da informação, o que implica na facilidade do seu acesso por parte dos usuários. Os documentos institucionais/técnicos/científicos, livros, artigos publicados em periódicos científicos, obras de referência, etc. se configuram como exemplos de fontes formais.

Ambas as tipologias de fontes de informação (informais e formais) são importantes para a comunicação científica, uma vez que “[...] permitem aos cientistas identificarem os meios de se comunicarem e divulgarem suas produções para a comunidade, ampliando o leque de interação voltado para a produção do conhecimento.” (REZENDE, 2016, p. 24).

É importante ressaltar que a pesquisa bibliográfica é de suma importância para os pesquisadores, pois permite que fiquem cientes do que está ocorrendo de mais atual e não repliquem pesquisas, gastando tempo e recursos. Essa pesquisa só pode ser realizada a partir de fontes com credibilidade. Desse modo, a credibilidade advém principalmente de periódicos científicos, que possuem critérios em suas publicações, dentre eles: a periodicidade e a análise dos artigos por pares eminentes das áreas de conhecimento em que atuam.

As bases de dados estão presentes nessa cadeia como fontes formais e com tipologia primária, secundária ou terciária, dependendo das informações por elas armazenadas, bem como dos processos de busca e recuperação das produções científicas existentes nesse contexto.

As bases de dados agregam, de maneira sistemática, produções científicas das áreas do conhecimento em que fazem parte. Nesse contexto, podem ser indexados e organizados os artigos de periódicos, artigos apresentados em eventos científicos e profissionais, dissertações, teses e livros de autores que se destacam na área de atuação, facilitando a disseminação dos saberes produzidos nas diversas áreas do conhecimento.

3 BASE DE DADOS

A criação das bases de dados tem por finalidade atender ao crescente aumento ou *boom* informacional ocorrido na década de 1950, na grande área da indústria da informação, como enfatiza Cunha (1989). Nesse entendimento, é possível dizer que o seu crescimento está atrelado ao desenvolvimento das tecnologias utilizadas nos serviços de informação, onde são empregados serviços de indexação e análise bibliográfica que facilitam a divulgação de traba-

Ihos acadêmicos através das citações, por exemplo. Na visão de Campello, Cendón e Kremer (2000, p. 318):

As bases de dados são recursos que apresentam muitos modos de se pesquisar, com diversos pontos de acesso, possibilitando a busca por campos específicos, como palavras-chave, pelo(s) nome(s) do(s) autor(es), utilizando-se de lógica booleana, escolhendo o período de cobertura, e outros recursos que permitem buscas muito específicas, mais eficientes que os demais, como os índices impressos utilizados anteriormente.

Resumidamente, é possível dizer que as bases de dados estão presentes em locais físicos ou virtuais onde o material informacional (temático ou não) é armazenado para posterior uso. Podem conter a informação completa ou remeter a outra(s) base de dados e/ou fonte(s) de informação para a obtenção do material em sua íntegra. Essas bases podem utilizar diversos recursos que viabilizam a recuperação da informação, dentre eles as palavras chaves, operadores *booleanos*¹, buscas simples e avançadas.

O objetivo das bases de dados, segundo Albrecht e Ohira (2000, p. 133) é “[...] fornecer informação atualizada (recursos estruturais), precisa e confiável (não dar a informação pela metade) e de acordo com a demanda (oferecer o que o usuário necessita)”. Além disso, deve possuir um armazenamento eficiente e fornecer mecanismos igualmente eficientes de recuperação da informação.

Com o advento da internet, as bases de dados digitais se tornaram, segundo Sayão (2000, p.149), um “[...] simulacro da memória consensual do conhecimento humano, pois cada vez mais os saberes humanos estão sendo encapsulados por esse produto emblemático da indústria da informação.”. Essas bases de dados são sistemas cuja difusão das informações é mais rápida e ampla, o que promove aumento de visibilidade da produção científica. Diante disso, são fontes de informação primordiais para o avanço de pesquisas em todas as áreas do conhecimento.

Percebe-se, portanto, que as bases de dados viabilizam a ampliação do espaço de pesquisa, revelam especificidades da área e permitem ao pesquisador vislumbrar as iniciativas e novidades no campo científico. Portanto, para a sua criação/implementação/atualização faz-se necessário levar em consideração desde os critérios de coleta e seleção de objetos informacionais a serem armazenados na base de dados até os mecanismos de representação, busca e acesso à informação no sistema.

No contexto da Ciência da Informação, as bases de dados agregam as produções intelectuais dessa área, visando satisfazer as necessidades informacionais de pesquisadores e da comunidade científica. Desse modo, se faz necessária uma investigação sobre os pontos qualitativos de tais sistemas.

4 PROCEDIMENTOS METODOLÓGICOS

Os procedimentos metodológicos deste estudo foram empreendidos a partir de uma pesquisa bibliográfica para a sua fundamentação teórica. Nesse sentido, foram recuperadas produções científicas nacionais e internacionais sobre os seguintes assuntos: a) fontes de informação; b) bases de dados; c) critérios de avaliação de sistemas de recuperação; d) histórico,

¹ São conectivos cujo objetivo é definir como deve ser realizada a combinação entre os termos ou expressões de uma busca em um sistema de recuperação de informação, são eles: “and” (e), “or” (ou), “not” (não) (BAESA-YATES; RIBEIRO-NETO, 2013).

características e funcionalidades das bases de dados avaliadas. As buscas de tais assuntos foram realizadas em janeiro de 2020 em portais de pesquisa, repositórios institucionais, bibliotecas digitais, bem como nos endereços eletrônicos das bases de dados avaliadas.

Além disso, de acordo com os seus objetivos, esta pesquisa é de caráter exploratório e descritivo, uma vez que analisa e descreve os elementos presentes em duas bases de dados da área de Ciência da Informação. Outrossim, a pesquisa possui abordagem qualitativa, haja vista que observa e qualifica as bases de dados de acordo com os critérios propostos por Sayão (2000).

4.1 Universo da Pesquisa

No universo das bases de dados, foram selecionadas duas que realizam *harsveting* (coleta) de publicações periódicas na área de Ciência da Informação: a Base de Dados em Ciência da Informação (BRAPCI) e a PERI. A primeira é coordenada pela Universidade Federal do Paraná (UFPR) em conjunto com a Universidade Federal do Rio Grande do Sul (UFRS) e sua plataforma *online* nasceu de um projeto intitulado "Metodologia para criação de uma base de dados *online* de acesso público: modelizando práticas para a socialização de saberes", criado pela pesquisadora Leilah Santiago Bufrem e subsidiada pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (SANTOS, 2016, p. 93). Ainda de acordo com essa autora, a BRAPCI vem sendo consolidada e ampliada com o objetivo de recuperar e disponibilizar publicações científicas nacionais e internacionais na área de Ciência da Informação, contribuindo também para o acesso por parte de pesquisadores de áreas correlatas, dada às características interdisciplinares dessa área do conhecimento.

Atualmente, a BRAPCI disponibiliza 19.255 (dezenove mil duzentos e cinquenta e cinco) textos publicados em 68 (sessenta e oito) periódicos nacionais e 14 (quatorze) periódicos internacionais impressos e eletrônicos da área de Ciência da Informação. Dos periódicos disponíveis, 65 (sessenta e cinco) estão ativos e 17 (dezessete) são históricos (descontinuados). Além disso, a base de dados indexa anais de três eventos científicos da área de Ciência da Informação (UNIVERSIDADE FEDERAL DO PARANÁ, 2019).

De outro modo disposto, a PERI é uma base de dados referencial mantida pela Universidade Federal de Minas Gerais e vinculada à Biblioteca Professora Etelvina Lima. Seu objetivo consiste em divulgar as produções científicas nas áreas Biblioteconomia, Ciência da Informação, Arquivologia, Museologia e áreas afins. Essa base de dados foi criada em 1987 e apresenta referências de artigos de periódicos e trabalhos publicados em anais de eventos técnico-científicos. Essa base de dados também agrega documentos existentes no acervo da biblioteca supracitada e que são disponibilizados através da internet, apresentando resumos e referências. Atualmente a PERI contém 68 Periódicos indexados e 22 anais de eventos técnico-científicos (UNIVERSIDADE FEDERAL DE MINAS GERAIS, 2019).

4.2 Critérios selecionados para a análise dos dados

Nesta pesquisa foram selecionados os vetores para a avaliação da qualidade de bases de dados propostos por Sayão (2000). Esse autor apresenta vetores cujos aspectos contribuem para a análise de bases de dados de qualquer área, com vistas a identificar questões referentes ao conteúdo, armazenamento, processamento, distribuição e apresentação das informações em sistemas de recuperação da informação.

Os critérios de qualidade são divididos em cinco vetores, os quais estão especificados e descritos no Quadro 1.

Quadro 1 – Vetores e descrições

VETORES	DESCRIÇÃO
1 CONTEÚDO INFORMACIONAL	<p>1.1 INFORMAÇÃO Conceitos genéricos como exatidão, atualização, novidades e consistência. A confiança nas fontes geradoras da informação, as estratégias empregadas na coleta e na seleção desta informação.</p>
	<p>1.2 ESTRUTURA Deve refletir o universo de usuários reais aos quais as bases de dados se propõem. Representar o domínio da realidade que se propõe relatar.</p>
	<p>1.3 REPRESENTAÇÃO É estudada através de instrumentos como tesouros, esquemas de classificação, listas de autoridades, depende da maestria dos indexadores e é um item muito passível de críticas.</p>
2 ARMAZENAMENTO	<p>É o meio onde a informação é disponibilizada para o usuário. Inclui o modo como é armazenada e os métodos de acesso a ela. Devem ser analisadas a capacidade de armazenamento, a rapidez de acesso, a confiabilidade, a portabilidade e a longevidade.</p>
3 PROCESSAMENTO	<p>Analisa as transformações que ocorrem para o processamento da informação. No caso analisado, foram estudados os <i>softwares</i> agregados que permitem operações sobre o conteúdo disponibilizado, observando nas bases a capacidade de recuperação, habilidade na reformatação das informações, a impressão, a ordenação e a facilidade de atualização.</p>
4 DISTRIBUIÇÃO	<p>Esse vetor permite avaliar se o conteúdo informacional da base de dados é disseminado de modo a atingir no momento adequado todo o mercado potencial de usuários. São os fatores <i>timeliness</i> (tempo oportuno), velocidade, confiabilidade, garantia de chegada da informação até o usuário, e as operações de detenção e correção de erros.</p>
5 APRESENTAÇÃO	<p>Disposição do conteúdo informacional no site. Os vetores neste caso são ergonomia da interface e a facilidade de uso.</p>

Fonte: Sayão (2000, p. 161-166).

Sayão (2000) afirma que os fatores dependem das tecnologias vigentes para o acesso à informação e que as bases de dados não produzem a informação. Elas têm a função de agregar valor às informações já existentes (disseminação e reprodução da informação consolidada).

Visando dar segmento a essas considerações, a seguir são apresentados os resultados obtidos neste estudo, mediante as análises realizadas de acordo com os vetores supracitados.

5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

A partir dos critérios relacionados por Sayão (2000), foi realizada a avaliação das Bases de Dados da BRAPCI e PERI, atribuiu-se P (Possui) para a existência do elemento no vetor, NP (não possui) para a não existência do elemento no vetor e PP (Possui Parcialmente) para a existência parcial do elemento no vetor.

Quadro 2 – Resultado da análise dos vetores das Bases de Dados BRAPCI e PERI

VETORES		
1 Conteúdo informacional	BRAPCI	PERI
Confiança	P	P
Estratégia	P	P
Seleção	P	P
Estrutura	P	P
Representação	P	P
2 Armazenamento		
Capacidade	P	P
Rapidez de acesso	P	P
Confiabilidade	P	P
Portabilidade	P	P
Longevidade	P	P
3 Processamento		
Capacidade de recuperação	P	P
Habilidade nas reformatações da informação	Sem acesso	Sem acesso
Impressão	P	P
Ordenação	NP	P
Facilidade de atualização	P	P
4 Destruição		
Timeliness	P	P
Velocidade	P	P
Confiabilidade	P	P
Correção de erros	NP	NP
5 APRESENTAÇÃO		
Ergonomia da interface	PP	PP
Facilidade de uso	P	PP

Fonte: Elaborado pelas autoras (2019).

Visando complementar a descrição da análise sintetizada no quadro anterior, a seguir são apresentados os resultados da avaliação dos vetores em cada base de dados.

5.1 Base de dados BRAPCI

As análises qualitativas apresentadas estão organizadas em consonância com os elementos constituintes de cada vetor, a seguir:

a) Conteúdo informacional

Neste vetor, são contemplados aspectos relativos à informação contida na BRAPCI, bem como questões concernentes à sua estrutura. Além disso, é realizada a descrição de como ocorre a representação das produções científicas disponíveis nesse sistema.

1) Informação

A BRAPCI coleta, armazena e indexa publicações científicas da área de Ciência da Informação. A sua confiabilidade vem do fato de que o material disponibilizado é institucional (já avaliado pela academia) ou é proveniente de periódicos ou eventos que possuem normas rígidas para a publicação de artigos - os quais passam por avaliações às cegas pelos pares. Dessa forma, acredita-se que as informações apresentadas nessa base de dados são confiáveis.

A partir das análises realizadas, verificou-se que as publicações científicas disponíveis na BRAPCI, em sua maioria, são nacionais. Dos 82 (oitenta e dois) periódicos científicos, 14 (quatorze) são internacionais, provenientes dos seguintes países: Peru, México, Cuba, Costa Rica, Espanha, Argentina, Portugal e Colômbia.

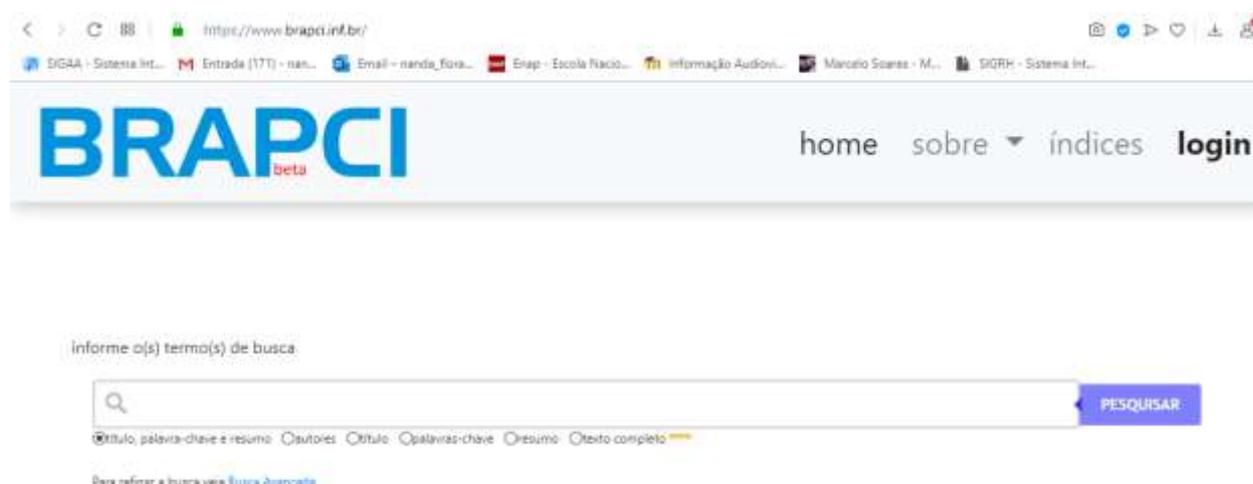
No que concerne à coleta das produções científicas e dos seus respectivos metadados, Santos (2016) enfatiza que essa atividade é realizada periodicamente por meio de *harvesting* nos periódicos eletrônicos da área de Ciência da Informação. Essa atividade objetiva detectar as novas publicações científicas disponíveis nesses periódicos para a sua inclusão na base de dados. Com isso, o sistema apresenta aos indexadores as novas edições coletadas pela base de dados, bem como todos os seus respectivos metadados para serem indexados/revisados.

2) Estrutura

No que diz respeito à sua estrutura, verifica-se que a BRAPCI visa atender, de maneira efetiva, às necessidades informacionais dos seus usuários. O sistema apresenta três módulos na sua configuração: o *módulo público*, para consultas realizadas pelos usuários; o *módulo de manutenção*, para revisão e correções da base; e o *módulo pesquisador*, para análises bibliométricas (UNIVERSIDADE FEDERAL DO PARANÁ, 2019)².

Ainda de acordo com as informações disponíveis no seu endereço eletrônico, a BRAPCI amplia o espaço documentário permitido ao pesquisador, facilitando a visão de conjunto da produção na área, ao mesmo tempo, em que revela especificidades do domínio científico. Atualmente disponibiliza textos publicados em periódicos nacionais impressos e eletrônicos da área de Ciência da Informação. A Figura a seguir, apresenta a imagem da página principal da base de dados.

Figura 1- Interface principal da BRAPCI



Fonte: UFPR (2019).

No âmbito da base de dados, é possível acessar o resumo, as referências utilizadas no artigo, além da transcrição da informação em MARC 21 e espaço para o RDA, já visualizando uma futura utilização.

² Informação disponível em: <https://www.brapci.inf.br/index.php/res/about>. Acesso em: 02 abr. 2020.

3) Representação

A representação consiste em uma operação que condiciona a recuperação da informação e deve ser realizada considerando os elementos físicos e temáticos dos objetos informacionais (textos, imagens, áudios, vídeos, etc.). No âmbito das bases de dados, essa operação é de extrema importância, uma vez que possibilita a identificação, recuperação e acesso às produções científicas nas interfaces de buscas. Consequentemente, a representação também pode influenciar na visibilidade e no impacto da comunicação científica.

Segundo Santos (2016), os metadados de descrição física e de conteúdo das produções científicas são coletados por meio de *harvesting* nos periódicos eletrônicos. De outro modo disposto, a indexação das produções científicas disponíveis na BRAPCI é norteadas por uma política de indexação que foi criada em 05 de novembro de 2011 e atualizada em 27 de setembro de 2014. Nesse documento são enfatizadas questões relativas aos seguintes elementos: cobertura do assunto, processos de indexação, vocabulário utilizado, agrupamento de termos (remissivas) e suas respectivas estatísticas.

Ainda de acordo com essa autora, a política de indexação da base de dados ressalta que o indexador deve fazer o uso dos mesmos termos informados pelos autores nas palavras-chave das produções científicas para descrever o conteúdo das publicações disponíveis no sistema. Nessa perspectiva, o processo de representação temática da informação disponível na base de dados é realizado de forma pós-coordenada (posteriormente), uma vez que os termos utilizados no momento da indexação de cada documento são decorrentes da coleta automática dos metadados incorporando os mesmos descritores que foram atribuídos pelos autores ou periódicos no ato da publicação dos artigos.

Santos (2016) constata que, em relação ao emprego de vocabulário controlado, a política de indexação da BRAPCI não orienta a utilização de instrumentos de controle terminológico como auxílio para a realização da indexação das produções científicas armazenadas nesse sistema. Nesse contexto, são utilizados apenas mecanismos para o processamento automático de descritores, embora essa base de dados possa incorporar diferentes tesouros em sua estrutura.

A ausência de controle terminológico pode implicar na incidência de ambiguidade terminológica em descritores no metadado “assunto” da base de dados. Portanto, sugere-se a adoção de instrumento(s) de controle terminológico para melhorias no processo de representação e, consequentemente, recuperação das produções científicas armazenadas nesse sistema.

b) Armazenamento

A partir das análises realizadas, verificou-se que a BRAPCI armazena uma quantidade significativa de publicações científicas (19.255 textos), incluindo publicações em periódicos científicos nacionais e internacionais, o que sinaliza a ampla cobertura da base de dados. As publicações científicas internacionais armazenadas nesse sistema são provenientes de periódicos dos seguintes países: Peru, México, Cuba, Costa Rica, Espanha, Argentina, Portugal e Colômbia.

Embora esse sistema armazene publicações científicas nacionais e internacionais, é possível constatar a predominância artigos produzidos e publicados em periódicos nacionais.

Pela volatilidade dos *links* dos periódicos e pela constante troca de endereços virtuais, a BRAPCI realiza a coleta não somente dos metadados, mas incorpora os artigos na íntegra com os PDF's, transformando o sistema não somente em uma base de referência. Por questões de direitos autorais, são disponibilizados apenas os arquivos que atendem à política de acesso

aberto, *Creative Commons*, ou que tenha autorização dos editores ou representantes legais da publicação (UNIVERSIDADE FEDERAL DO PARANÁ, 2019)³.

Com base nessa perspectiva, a base de dados utiliza duas estratégias para disponibilizar a informação: no próprio site ou pelos *links* que remetem à fonte original que publicou a produção científica consultada.

No que concerne à longevidade, a utilização nos sites de origem de identificadores persistentes, sendo o mais conhecido o DOI (*Digital Object Identifier* ou Identificador de Objeto Digital), demonstra a preocupação na garantia dessa informação na *web*. No caso da Base BRAPCI, de 65 (sessenta e cinco) periódicos estudados, 50 (cinquenta) possuem Identificador de Objeto Digital.

c) Processamento

Em relação ao vetor processamento, especificamente no que concerne à sua capacidade de recuperação, a base de dados dispõe do *módulo público*, contemplando a possibilidade de busca simples nos seguintes campos: “autor”, “título”, “palavra-chave”, “resumo”, “referência” e “ano”.

Além disso, é possível que o usuário realize uma busca avançada utilizando as seguintes estratégias: a) *busca por texto composto* (ex.: mineração de texto); b) pela *variação de termos* (busca pelo termo e sua variação mediante o uso de um asterisco, ex.: bibliomet*, *metria); c) pela *variação de uma letra dos termos* (utilizando o sinal de interrogação, ex.: bibliotecári?, indexad?r); d) pela *variação do termo* (busca de termos com início e fim iguais, ex.: bib*ca, nesse caso a base de dados retorna tudo o que começa com “bib” e termine com “ca”); e) pelo emprego de operadores *booleanos* (conectivos que delimitam os resultados da busca, são eles: “and”, “or”, “not”).

Considerando a facilidade de atualização do sistema, a BRAPCI também passou a utilizar o *ElasticSearch* como mecanismo de busca. Baseado no Lucene, esse servidor fornece um mecanismo de pesquisa de texto completo distribuído, com capacidade para vários usuários, com uma interface da Web HTTP e documentos JSON sem esquema. *Elasticsearch* é desenvolvido em Java e é lançado sob os termos da licença *Apache* (UNIVERSIDADE FEDERAL DO PARANÁ, 2019)⁴. Ainda no que diz respeito à sua atualização, observou-se que a coleta das produções científicas dos seus respectivos metadados para armazenamento na base de dados é realizada periodicamente.

O usuário consegue também baixar em formato PDF e realizar a impressão das produções científicas de maneira fácil e intuitiva no sistema. Identificou-se ainda que a BRAPCI não adota critérios de ordenamento dos resultados das buscas.

d) Distribuição

O fator *timeliness* diz respeito à pontualidade, ou seja, capacidade da base de dados completar uma tarefa necessária ou cumprir uma tarefa antes ou em um tempo esperado. Com isso, é possível associar esse fator à velocidade e capacidade de atualização de modo contínuo e beta contínuo no que concerne à recuperação e inclusão de novas produções científicas, conforme apresentado no vetor “armazenamento”. Além disso, nas buscas realizadas nesse sistema, é possível identificar produções científicas de 1972 até o período corrente.

³ Informação disponível em: <https://www.brapci.inf.br/index.php/res/about>. Acesso em: 02 abr. 2020.

⁴ Informação disponível em: <https://www.brapci.inf.br/index.php/res/about>. Acesso em: 02 abr. 2020.

De acordo com as informações apresentadas no sistema, a BRAPCI está realizando uma pesquisa *online* com os seus usuários com a finalidade de avaliar a base de pesquisa criada sob o ponto de vista do consultente. A finalidade dessa pesquisa com usuários está atrelada à possibilidade de implementar futuras melhorias de interface e de conteúdo, de acordo com as necessidades e com o nível de satisfação dos mesmos. Portanto, verifica-se que a base de dados busca a garantia de chegada da informação até os seus consultentes.

Mediante buscas realizadas nesse sistema, verificou-se que a BRAPCI não realiza operações de detecção e correção de erros, uma vez que ao tentar realizar consultas a partir de descritores com grafias erradas, não foram apresentadas sugestões de termos similares e/ou correções. Além disso, não foram recuperadas produções científicas cujos termos indexados apresentem similaridade sintática ou semântica.

e) Apresentação

Após as análises realizadas, verificou-se que a BRAPCI contém uma interface simples e intuitiva, o que pode facilitar o processo de busca e recuperação da informação. Em sua página principal são apresentados, de maneira clara, os metadados e funcionalidades que os usuários podem utilizar em suas estratégias de busca. Esses últimos foram descritos, de maneira detalhada, no vetor processamento.

Entretanto, verifica-se que a ergonomia está presente parcialmente na base de dados, uma vez que esse sistema não possui recursos de acessibilidade, podendo dificultar o seu acesso por usuários com deficiência visual, por exemplo. Dessa forma, considerando-se que a BRAPCI objetiva atender usuários pesquisadores e comunidade científica em geral, acredita-se que esse sistema não é completamente compatível com as limitações dos usuários potenciais, aspecto que pode ser considerado como sugestão para melhorias sob o ponto de vista da ergonomia e da arquitetura da informação. Pesquisas futuras também podem ser realizadas pela base de dados com os usuários, no intuito de identificar problemas de acessibilidade e usabilidade através de aplicações de testes.

5.2 Base de dados PERI

Nessa subseção, as análises qualitativas também estão organizadas em consonância com os elementos constituintes de cada vetor, a seguir:

a) Conteúdo informacional

Neste vetor, são contemplados aspectos relativos à informação contida na base de dados PERI, bem como são apontadas questões concernentes à sua estrutura.

1) Informação

A Base de Dados PERI contém referências de artigos de 68 (sessenta e oito) periódicos e trabalhos publicados em 22 (vinte e dois) anais de evento técnico-científicos, refletindo principalmente a literatura nacional da Ciência da Informação. Dentre os periódicos indexados, 23 (vinte e três) se encontram ativos.

Essa base de dados indexa um periódico internacional, o *Brazilian Journal of Information Science*, o que demonstra a predominância de produções científicas nacionais referenciadas nesse sistema. A PERI, embora seja uma base de dados referencial, também disponibiliza do-

cumentos existentes no acervo da Biblioteca Professora Etelvina Lima, os quais também estão disponibilizados na internet, apresentando os seus respectivos resumos e/ou texto completo.

A confiabilidade dessa base de dados está atrelada ao fato de que o material referenciado também é institucional, já avaliado pela academia, ou é proveniente de periódicos e eventos que possuem normas rígidas para a publicação de artigos. Dessa forma, acredita-se que as informações apresentadas na PERI também são confiáveis.

O sistema não apresenta informações sobre como é realizada a coleta das produções científicas e dos seus respectivos metadados, também não foram encontradas produções científicas que descrevem ou analisam essa base de dados. Todavia, acredita-se que a coleta dessas informações também é realizada periodicamente, uma vez que foram buscadas e recuperadas referências de produções científicas do período corrente na base de dados.

2) Estrutura

No que concerne à sua estrutura, verifica-se que a PERI apresenta dois módulos em sua configuração: o *módulo público*, para consultas realizadas pelos usuários; o *módulo de manutenção*, para revisão e correções da base.

A PERI disponibiliza, de modo predominante, referências e resumos de textos publicados em periódicos nacionais impressos e eletrônicos da área de Ciência da Informação. Nesse sentido, a base de dados não fornece na íntegra os artigos de periódicos, uma vez que o seu objetivo é direcionar o usuário para o periódico que publicou o trabalho (fonte primária), a partir do endereço eletrônico, conforme é possível visualizar na figura a seguir.

Figura 2- Interface de busca da base de dados PERI



Fonte: UFMG (2019).

Observa-se ainda que a PERI contém campos de busca simples e guiada, além de um dicionário de termos para facilitar o processo de recuperação da informação no sistema, aspectos a serem detalhados posteriormente.

3) Representação

Considerando que a representação condiciona a recuperação da informação, essa operação consiste em um dos principais vetores na avaliação da qualidade das bases de dados. No âmbito da PERI não foram encontradas informações relativas aos procedimentos e instrumentos utilizados para fins de representação das produções científicas referenciadas no sistema. Também não foi identificada uma política de indexação.

Entretanto, a partir das buscas e análises realizadas, verificou-se que os administradores fazem uso dos mesmos termos informados pelos autores nas palavras-chave das produções científicas para descrever o conteúdo das publicações referenciadas no sistema, aspecto que pode estar associado à sua política de indexação. O dicionário de termos, por sua vez, apresenta todos os descritores utilizados nessa indexação.

A ausência de controle terminológico pode implicar na incidência de ambiguidade terminológica e erros ortográficos em descritores no metadado “assunto” da base de dados. Portanto, recomenda-se a criação e implementação de uma política de indexação, bem como o emprego de instrumento(s) de controle terminológico para auxiliar na indexação e nas buscas realizadas pelos usuários desse sistema.

b) Armazenamento

A partir das análises realizadas, observou-se que a PERI referencia e indexa artigos de 68 (sessenta e oito) periódicos e trabalhos publicados em 22 (vinte e dois) anais de evento técnico-científicos, refletindo principalmente a literatura nacional da Ciência da Informação. Entretanto, dentre os periódicos indexados, 23 (vinte e três) se encontram ativos.

Embora essa base de dados também armazene referências de publicações científicas internacionais, é possível constatar a predominância de produções nacionais.

Com base nessa perspectiva, a base de dados disponibiliza a informação por meio de *links* que remetem à fonte original que publicou a produção científica ou apresenta o texto para a consulta no próprio site - no caso de publicações já disponíveis no acervo da biblioteca em que o sistema está subordinado.

Com relação aos *links* das fontes referenciadas na base de dados, não foi identificada nenhuma desatualização, limitação ou dificuldade para acesso. No que concerne à longevidade, a utilização nos sites de origem de identificadores persistentes, sendo o mais conhecido o DOI (*Digital Object Identifier* ou Identificador de Objeto Digital), demonstra a preocupação na garantia dessa informação na *web*. No caso da base PERI, de 23 (vinte e três) periódicos ativos, apenas dois não possuem DOI.

c) Processamento

Em relação ao vetor processamento, observa-se a dependência da base de dados em relação às fontes primárias de informação, uma vez que esse sistema é referencial. No acesso, foi analisado que há boa recuperação dos *links* disponibilizados na base, os quais encontram-se atualizados e direcionam o usuário para a fonte primária.

Especificamente no que concerne à sua capacidade de recuperação, a base de dados dispõe do *módulo público* que possibilita a realização de busca simples e guiada, além da utilização de um dicionário de termos.

No campo de busca simples é possível pesquisar pelo título ou assunto da produção científica. De outro modo disposto, no campo de busca guiada, o usuário consegue realizar a

busca por “periódico”, “autor”, “título”, “descriptor” e “resumo”. Além disso, é possível realizar estratégias de busca mediante o uso de operadores *booleanos*.

O dicionário de termos, por sua vez, permite listar todos os termos indexados na base, em ordem alfabética, objetivando facilitar a navegação e a recuperação da informação no sistema.

Quanto à ordenação dos resultados das buscas, a base PERI apresenta as publicações científicas recuperadas em ordem decrescente de ano, das mais recentes às mais antigas. Essa base de dados também permite ao usuário selecionar as referências que consideram relevantes e imprimi-las juntamente com os seus respectivos resumos.

d) Distribuição

Considerando que o fator *timeliness* diz respeito à capacidade da base de dados se atualizar no que concerne à recuperação e inclusão das referências de novas produções científicas, observa-se que o sistema contempla esse vetor. Todavia, a sua estrutura e as suas funcionalidades não são atualizadas de modo contínuo e beta contínuo.

Outrossim, não foram identificadas informações acerca de pesquisas de usuários realizadas pela base de dados com a finalidade de avaliar o sistema. A finalidade do estudo de usuários está atrelada à possibilidade de implementar futuras melhorias de interface e de conteúdo, considerando o nível de satisfação dos mesmos - aspecto que se configura como sugestões a serem consideradas pelos gestores da base de dados.

Mediante análises realizadas nesse sistema, constatou-se que a PERI não realiza operações de detecção e correção de erros, uma vez que ao tentar realizar buscas a partir de descritores com grafias erradas, não foram apresentadas sugestões de termos similares e/ou correções. Também não foram recuperadas produções científicas cujos termos indexados apresentam similaridade sintática ou semântica.

e) Apresentação

Após as análises realizadas, observou-se que a PERI apresenta os metadados e funcionalidades que podem ser utilizadas nas estratégias de busca. Entretanto, esse sistema possui uma interface escura e dispõe letras pequenas, podendo dificultar o acesso à informação por parte dos usuários que possuem deficiência visual.

Dessa forma, verifica-se que a ergonomia está presente parcialmente na base de dados, uma vez que esse sistema não possui recursos de acessibilidade, podendo dificultar o seu acesso por usuários.

Portanto, considerando-se que a PERI visa atender usuários pesquisadores e comunidade científica em geral, acredita-se que esse sistema também não é completamente compatível com as eventuais limitações dos seus usuários, aspecto que pode ser considerado como sugestão para melhorias sob o ponto de vista da ergonomia e da arquitetura da informação. Estudos futuros também podem ser realizadas com os usuários, no intuito de identificar problemas de acessibilidade e usabilidade mediante a aplicações de testes.

6 CONSIDERAÇÕES FINAIS

A informação científica consiste em um instrumento que contribui para o progresso e inovação no desenvolvimento de um país. Portanto, a mesma precisa ser organizada e proces-

sada de maneira eficiente no contexto das bases de dados para fins de sua recuperação pelos pesquisadores e comunidade em geral.

As bases de dados científicas, por sua vez, devem primar por qualidade e exatidão, objetivando contribuir para o processo de busca e recuperação da informação vinculada. Para tanto, esses sistemas devem considerar os critérios qualitativos que podem ajudar em sua melhoria, seja em contextos físicos ou em plataformas digitais.

Esta pesquisa teve como objetivo avaliar, com base nos vetores propostos por Sayão (2000), a qualidade de duas bases de dados da área de Ciência da Informação: a BRAPCI e a PERI. Esse autor apresenta vetores cujos aspectos contribuem para a análise de bases de dados de qualquer área, com vistas a identificar questões referentes ao conteúdo, armazenamento, processamento, distribuição e apresentação das informações em sistemas de recuperação da informação.

Na análise realizada utilizando os vetores supracitados, foi possível observar que 76% dos critérios de qualidade foram contemplados em ambos os sistemas, o que demonstra a qualidade das duas bases de dados como instrumento de divulgação científica para pesquisadores, docentes e comunidade científica em geral.

Em relação ao vetor “conteúdo informacional”, verificou-se que a BRAPCI possui ampla cobertura em relação às publicações da área de Ciência da Informação, contemplando uma quantidade significativa de periódicos nacionais e internacionais nessa área do conhecimento. Esse sistema realiza a coleta dos metadados e incorpora as produções científicas na íntegra em formato PDF, diferentemente da proposta da base de dados PERI cuja característica está atrelada a uma base de dados referencial.

No vetor “representação” identificou-se que apenas a BRAPCI possui política de indexação registrada e publicada. Embora possam incorporar tesouros em sua estrutura, verificou-se ainda que ambas as bases de dados não utilizam instrumentos de controle terminológico no processo de indexação e recuperação das produções científicas. Portanto, sugere-se a adoção dessas ferramentas com vistas a melhorar as práticas de representação e busca das publicações científicas nesse contexto.

No que se refere ao vetor “processamento”, ambas as bases de dados possibilitam que o usuário realize buscas simples e avançadas, a partir de estratégias realizadas de acordo com o modelo computacional utilizado pelo sistema. Ao apresentar diferentes possibilidades de busca para os usuários, essas bases de dados permitem que os mesmos delineiem estratégias compatíveis com as suas necessidades informacionais.

No atributo “*timeliness*”, por sua vez, constatou-se que ambas as bases de dados possuem capacidade de atualização de modo contínuo no que concerne à recuperação e inclusão de informações sobre novas publicações científicas. De outro modo disposto, no vetor “distribuição” identificou-se que a BRAPCI está realizando uma pesquisa *online* com os seus usuários para avaliar esse sistema de recuperação da informação. Portanto, verifica-se que essa base de dados busca a garantia de chegada da informação até o usuário, aspecto que pode vir a ser considerado pelos gestores da base de dados PERI.

No vetor “apresentação” de ambas as bases de dados, verificou-se que esses sistemas não estão completamente compatíveis com as ferramentas de acessibilidade. Desse modo, sugere-se a inclusão de melhorias no que concerne à ergonomia e à arquitetura da informação desses sistemas.

Em linhas gerais, é importante ressaltar que os vetores de avaliação têm valores individuais em bases de dados e são considerados em consonância com a perspectiva de cada usuário. Por exemplo, o usuário que busca informações atualizadas sobre “*Fake News*” valoriza o

atributo *timeliness* do vetor “distribuição” e o atributo “atualização” do vetor “conteúdo informacional” em sistemas de recuperação da informação.

Portanto, como sugestão para pesquisas futuras, propõe-se a realização de um estudo com usuários estudantes e/ou pesquisadores para a validação dos critérios de qualidade apresentados nesta pesquisa.

REFERÊNCIAS

ALBRECHT, R. F.; OHIRA, M. L. B. Base de dados: metodologia para seleção e coleta de documentos. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 5, n.5, 2000. Disponível em: <https://revista.acbsc.org.br/racb/article/download/347/411>. Acesso em: 02 abr. 2020.

BAESA-YATES, R.; RIBEIRO-NETO, B. **Recuperação de informação: conceitos e tecnologias das máquinas de busca**. 2.ed. Porto Alegre: Bookman, 2013.

CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M. **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: Editora UFMG, 2000.

CUNHA, M. B. Base de dados no Brasil: um potencial inexplorado. **Ciência da Informação**, Brasília, v. 18, n.1, p. 45-57, jan./jun. 1989. Disponível em: <http://repositorio.unb.br/handle/10482/5593>. Acesso em: 02 abr. 2020.

CUNHA, M. B. **Para saber mais: fontes de informação em ciência e tecnologia**. Brasília: Brinquet de Lemos Livros, 2001.

MUELLER, S. P. M. A ciência, o sistema de comunicação científica e a literatura científica. In: CENDÓN, B. V.; CAMPELLO, B. S.; KREMER, J. M. (Orgs). **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: Editora UFMG, 2000, p. 21-34.

PACHECO, C. G.; VALENTIM, M. L. P. Informação e conhecimento como alicerces para a gestão estratégica empresarial: um enfoque nos fluxos e fontes de informação. In: VALENTIM, M. (Org.). **Gestão, mediação e uso da informação**. São Paulo: Cultura Acadêmica, 2010. p. 319-341.

REZENDE, A. **Proposta de critérios de avaliação de fontes de informação na internet para pesquisadores da saúde**. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Pernambuco, Recife, 2016. Disponível em: <https://repositorio.ufpe.br/handle/123456789/17610>. Acesso em: 02 abr. 2020.

SANTOS, R. F. **Modelos colaborativos de indexação social e a sua aplicabilidade na base de dados referencial de artigos de periódicos em ciência da informação (BRAPCI)**. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Pernambuco, Recife 2016. Disponível em: <https://repositorio.ufpe.br/handle/123456789/17218>. Acesso em: 02 abr. 2020.

SAYÃO, L. F. Bases de Dados e suas qualidades. In: LUBISCO, Nídia M. L.; BRANDÃO, Lídia M. B (Org.). **Informação e Informática**. Salvador: EDUFBA, 2000. p.143-180.

UNIVERSIDADE FEDERAL DE MINAS GERAIS. **Base Peri**: literatura nacional nas áreas de Biblioteconomia, Ciência da Informação, Arquivologia, Museologia e outras interdisciplinares, Belo Horizonte, 2019. Disponível em: <http://biblio.eci.ufmg.br/produtos-e-servicos/bases-de-dados-eci/base-peri/>. Acesso em: 02 jan. 2020.

UNIVERSIDADE FEDERAL DO PARANÁ. **Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI)**, Curitiba, 2019. Disponível em: <https://www.brapci.inf.br/>. Acesso em: 02 jan. 2020.