
A utilização de linguagens de indexação por bibliotecas universitárias: mapeamento sistemático da literatura

El uso de lenguajes de indización en las bibliotecas universitarias: mapeo sistemático de la literatura

The use of indexing languages by university libraries: systematic literature mapping

Jessica Beatriz TOLARE, Mariângela Spotti Lopes FUJITA

Universidade Estadual Paulista – Unesp, Campus de Marília, Dep. de Ciência da Informação, Av. Hygino Muzzi Filho, 737, Campus Universitário, 17525-900 – Marília, SP (Brasil). [jessica.tolare | mariangela.fujita] @unesp.br

Resumen

Se analizan los estudios sobre lenguajes de indización en las bibliotecas universitarias para verificar las similitudes y diferencias, los métodos, las propuestas, las contribuciones y las recomendaciones, a través del mapeo sistemático de la literatura en las bases BRAPCI, LISTA, WoS y Scopus entre 2002 y 2022. Las variables que pueden afectar el proceso de indización son: uso de vocabulario controlado e inapropiado de otras áreas del conocimiento; atribución de la indización por parte del autor; el profesional indizador; falta de política de indización y lenguajes y la consonancia entre ellos; falta de método en la construcción del lenguaje de indización; indisponibilidad del lenguaje de indización por parte de las bibliotecas para los usuarios; falta de un manual de indización; comprensión de los profesionales que realizan la mediación del lenguaje de indización y su complejidad. Cada estudio analizado presenta, al menos en algunos aspectos, similitudes y diferencias entre sí, en métodos, propuestas, contribuciones y recomendaciones. Utilizar solamente las directrices de la política de indización y el manual no es suficiente para resolver todos los problemas que involucran el proceso de indización, es necesario realizar una combinación de multifactores: el indizador, la calidad del lenguaje de indización en su construcción, el instrumento utilizado, los sistemas y el contexto de la biblioteca universitaria.

Palabras clave: Indización. Lenguajes de indización. Bibliotecas universitarias.

1. Introdução

Existem diferentes tipos de bibliotecas como públicas, especializadas, escolares, particulares, comunitárias e universitárias. Cada biblioteca possui sua própria especificidade e seu público-alvo, desenvolvendo sua identidade com produtos e serviços que irá oferecer.

Quando se trata de uma biblioteca universitária, seu objetivo consiste em atender à instituição ao qual ela está vinculada, especificamente, os alunos, professores, funcionários, pesquisadores e, às vezes, a comunidade local. Por isso, oferece

Abstract

The studies on indexing languages in university libraries are analyzed to verify similarities and differences, methods, propositions, contributions and recommendations, through the systematic mapping of literature in the databases BRAPCI, LISA, WoS and Scopus, in the period 2002-2022. The variables that can affect the indexing process are: use of controlled and inappropriate vocabulary from other areas of knowledge; attribution of indexing by the author; the professional indexer; lack of indexing policy and languages and the consonance between them; lack of method in the construction of the indexing language; unavailability of the indexing language by libraries to users; lack of an indexing manual; understanding of the professionals who perform the mediation of the indexing language and its complexity. These variables showed how the indexing process can be affected, involving aspects of the indexer, the system, the elaborated guidelines and how the indexing language construction occurs. We concluded that each study analyzed presented, at least in some aspects, similarities and differences between each other, methods, propositions, contributions and recommendations. Using only the indexing policy guidelines and the manual is not enough to solve all the problems involving the indexing process, it is necessary to perform a combination of multifactors, which are: the indexer, the quality of the indexing language in its construction, the instrument used, the systems and the context of the university library.

Keywords: Indexing languages. Indexing. Academic libraries.

um tipo de serviço de cunho específico e, conseqüentemente, precisa possuir uma organização adequada e própria para tornar a informação acessível, tanto em seu acervo online quanto físico.

A biblioteca universitária tem necessidade de sistematizar processos e condutas de indexação e, nessa tarefa, terá que definir uma política de indexação cujo desenvolvimento depende de planejamento, elaboração e implantação dos procedimentos, técnicas, normas e manual com o obje-

tivo de organizar e recuperar a informação inserida nesse ambiente (Fujita, Favato e Zaniboni & Fagundes, 2016, p. 23). Nesse sentido, a indexação em uma biblioteca universitária possui particularidades em seu processo envolvendo aspectos como acervo especializado e público-alvo específico. O processo de indexação nesse ambiente requer que os termos da linguagem de indexação atribuídos, estejam de acordo com a temática no qual o acervo abarca, pois caso isso não ocorra, pode causar discrepâncias em sua recuperação e prejudicar o usuário. Logo, esse processo se difere de outros tipos de bibliotecas, como públicas e comunitárias, pois, devido a sua especificidade, é necessário envolver especialistas para realizar esse procedimento.

O uso inadequado de uma política de indexação, para Fujita e Gil Leiva (2010), reflete diretamente no uso de linguagens de indexação pelo bibliotecário, da mesma forma que as linguagens de indexação exercem influência na política de indexação tendo em vista os resultados da recuperação.

Ao longo dos anos, pesquisas foram desenvolvidas sobre o uso de linguagens de indexação em bibliotecas, especialmente, em universitárias, caracterizando a problemática desta pesquisa: como a literatura publicada aborda as linguagens de indexação em bibliotecas universitárias?

Para isso, esta investigação propõe analisar estudos sobre linguagens de indexação em bibliotecas universitárias mediante mapeamento sistemático de literatura para realizar pesquisa bibliográfica em base de dados com o objetivo de analisar semelhanças e diferenças, métodos, proposições, contribuições e recomendações.

Espera-se que a análise da literatura publicada sobre o uso de linguagens de indexação em bibliotecas universitárias possa constatar como a variável linguagem de indexação influencia nesse processo de indexação, pois são determinantes para a representação no processo de indexação e podem interferir diretamente na recuperação de recursos informacionais dentro desse ambiente (Arboit, 2017).

2. Linguagem de indexação

A indexação é definida pela norma ISO 25.964-1 (International Standardization for Organization, 2011, p. 5) como uma “[...] análise intelectual do assunto de um documento, com o objetivo de identificar os conceitos representados, a fim de possibilitar a recuperação da informação.”

Neste estudo, a indexação é vista como um processo composto por fases e etapas para sua execução. Cada corrente teórica segue uma quantidade de fases e mesmo havendo divergência,

suas fases podem ser resumidas em: *exame do documento e estabelecimento do assunto de seu conteúdo*, etapa do exame do documento por meio da leitura documentária, para conhecer o assunto (Fujita, 2017); *identificação e seleção dos conceitos presentes no assunto*, fase que possui o intuito de extrair os conceitos que traduzam a essência do documento (Dias & Naves, 2007); e *tradução dos conceitos*, em que os conceitos são traduzidos por termos definidos pela linguagem de indexação.

A linguagem de indexação é definida por Silva e Brito (2018, p. 96) como um “conjunto de termos padronizados, que representam um conceito da linguagem natural, com vistas a uma padronização dos sistemas de indexação”.

A norma internacional *American National Standards Institute/National Information Standards* (ANSI/NISO, 2010) define linguagem de indexação como

Um vocabulário controlado ou um sistema de classificação com regras para sua aplicação. Uma linguagem de indexação é utilizada para a representação dos conceitos tratados nos documentos [objeto de conteúdo] e para a recuperação de tais documentos [objeto de conteúdo] para um armazenamento de informações e de um sistema de recuperação.

A linguagem de indexação é considerada um componente-chave para os sistemas de informação, sobretudo quando a informação é profissional ou acadêmica, devido a ser um contexto associado com informações complexas e de alta qualidade. Por isso, ela precisa ser extensa, para conseguir abordar adequadamente qualquer tipo de conhecimento, principalmente, para as ciências humanas, que pode possuir uma variedade de perspectivas diferentes (Vállez, Pedraza-Jiménez, Condia, Blanco & Rovira, 2015; Golub Tyrkkö, Hansson & Ahlström, 2020).

O uso da linguagem de indexação, dentro de um sistema de informação, ocorre em dois momentos: o primeiro consiste na entrada, momento em que se insere no sistema, no qual a representação da informação é realizada por intermédio do controle definido pela linguagem de indexação; o segundo momento corresponde à saída, quando o usuário realiza a busca pela informação, expressando sua necessidade e utilizando termos do vocabulário controlado (Cintra, Tálamo, Lara & Kobashi, 2002).

Para Gil Urdiciain (2004) y Cruz, Santos e Fujita (2016), a linguagem de indexação é construída por um conjunto de regras preestabelecidas em nível sintático, semântico e pragmático, baseando-se em teorias e metodologias oriundas da terminologia, da linguística textual, da semântica

estrutural e da sintaxe, tendo como princípio o de realizar o controle terminológico e aperfeiçoar o acesso à informação.

Moura, Silva e Amorim (2002), consideram que existe uma complexidade no desenvolvimento de linguagens de indexação, pois ela requer uma compreensão de variáveis que podem influenciar no processo. Essa complexidade se deve ao fato de que é tomado como princípio a garantia literária, a qual ocorre quando um vocabulário controlado é capaz de representar conceitos que realmente se encontram na literatura de assunto. Por isso, esse princípio tende a inibir algumas interfaces da busca entre os usuários e os sistemas de informação, como, por exemplo, o usuário iniciante, que não possui domínio geral dos assuntos e das categorias em que eles estão; e os profissionais, pois as categorias de assunto não foram inseridas no documento.

3. Procedimentos metodológicos

A pesquisa se caracteriza como exploratória e descritiva, de cunho qualitativo. O método utilizado para o seu desenvolvimento foi o mapeamento sistemático da literatura, pois possibilitou investigar os problemas relatados em pesquisas anteriores sobre uma determinada temática.

O mapeamento sistemático garante uma confiabilidade no trabalho por conter etapas bem definidas e tratar-se de um estudo menos aprofundado, mas que ocorre de forma mais abrangente. Esse método é composto por três etapas: planejamento da revisão para identificação da real necessidade da pesquisa; definição do protocolo da revisão; e escrita dos resultados. O protocolo especifica questões da pesquisa e estratégias utilizadas (Kitchenham & Charters, 2007; Ramires, 2022). Os protocolos são compostos por diferentes passos, e, para observar o seu uso, adaptou-se a tabela de Ramires (2022).

De todos os trabalhos apresentados nos resultados (152) nas buscas das bases de dados nacionais e internacionais, foram selecionados 45 estudos para o desenvolvimento da pesquisa. Os critérios estabelecidos para se chegar aos resultados dessa seleção consistiram na premissa de que os estudos abordassem a linguagem de indexação em bibliotecas universitárias, especificamente, apresentando as temáticas de pesquisa em pelo menos um dos locais do texto: no título, nas palavras-chave (tabela 2) ou o seu desenvolvimento nos resumos dos trabalhos selecionados.

Depois de selecionado os 45 estudos, foram realizadas leituras parciais do título, resumo, pala-

avras-chave, da introdução, resultados e considerações finais, a fim de observar se os autores realmente abordaram e desenvolveram o uso de linguagem de indexação em bibliotecas universitárias.

<i>Passos do protocolo</i>	
<i>Passo 1: Mapeamento</i>	
<i>Objetivo</i>	Analisar artigos sobre linguagens de indexação em bibliotecas universitárias, com objetivo de comparar semelhanças e diferenças, métodos, proposições, contribuições e recomendações.
<i>Pergunta norteadora</i>	Como a literatura publicada aborda as linguagens de indexação em bibliotecas universitárias?
<i>População</i>	Artigos científicos
<i>Intervenção</i>	Publicações científicas em português e inglês
<i>Controle</i>	Publicações científicas em texto completo
<i>Resultados</i>	Seleção e análise dos dados das publicações científicas recuperadas para identificação de pesquisas, autores, datas das pesquisas, problemas, procedimentos e desenvolvimentos
<i>String de busca</i>	Linguagem de indexação; linguagem de indexação AND biblioteca universitária; linguagem de indexação AND biblioteca acadêmica; indexing language; indexing language AND University language; indexing language AND academic library
<i>Passo 2: Definição de critérios de seleção de fontes</i>	
<i>Critério de busca</i>	Publicações em forma de artigos científicos em texto completo
<i>Idioma</i>	Português e inglês
<i>Métodos de seleção</i>	Delimitação do espaço temporal no levantamento bibliográfico (2002-2022) Busca bibliográfica em bases de dados do assunto do primeiro item; Análise e seleção de textos da pesquisa bibliográfica com base no assunto determinado; Leitura parcial da introdução, resultado, considerações finais e resumo que abordaram a temática delimitada para seleção final da análise; Leitura aprofundada dos trabalhos selecionados; Fichamentos dos textos selecionados para realização da análise
<i>Passo 3: Definição de fontes</i>	
<i>Bases de dados</i>	BRAPCI, LISTA, WoS e Scopus.
<i>Passo 4: Critérios de seleção de documentos</i>	
<i>Critérios de inclusão</i>	Publicações científicas em forma de artigos disponíveis em texto completo que contemplam as strings de buscas citados. A inclusão de textos terá como critério a leitura do título, resumo, palavras-chave, resumo, introdução, resultados e considerações finais.
<i>Critérios de exclusão</i>	Ser artigos de periódicos fora do espaço temporal delimitado; ser de outro idioma do selecionado; não ter nenhuma relação com a temática proposta e não ser a tipologia documental definida.
<i>Tipos de documentos</i>	Artigos de periódicos

Tabela 1. Passos do protocolo do mapeamento sistemático

Após essa leitura parcial, dos 45 estudos, foram selecionados 8, nos quais foram feitas leituras

aprofundadas: (Menard, 2010; Tartarotti, 2014; Cruz, Fujita & Santos, 2017; Sunny & Angadi, 2017; Fujita, Moreira, Santos e Cruz & Ribas, 2018; Cruz, 2019; Fujita, Cruz, Patrício & Branco, 2019; Golub, Tyrkkö, Hansson & Ahlström, 2020).

Os critérios de exclusão se deram ao perceber que as bases de dados apresentaram trabalhos com temáticas que eram diferentes das pesquisadas; fora do espaço temporal delimitado; não estarem no idioma ou na tipologia definida. Houve trabalhos em que foi necessário excluir da seleção, pois tinham como objetivo desenvolver as seguintes temáticas: “elementos da linguística em indexação”, “aplicabilidade do tesouro em biblioteca digital”, “pensamento dialógico da indexação social”, “teoria do conceito”, “indexação jurídica”, “indexação automática”, “ontologias”, “tradução intersemiótica dos mecanismos de buscas” entre outros assuntos diferentes. Nesses estudos, a linguagem de indexação foi apenas citada brevemente, por isso, a exclusão ocorreu por não terem relação de fato com os objetivos propostos.

4. Análise dos resultados

As buscas em bases de dados possibilitaram encontrar em seus resultados trabalhos que abordaram linguagem de indexação em bibliotecas universitárias. Por isso, a análise de resultados foi subdividida em: (I) busca e recuperação de estudos em bases de dados e (II) análise de estudos sobre linguagem de indexação em bibliotecas universitárias.

4.1. Busca e recuperação de estudos em bases de dados

No Gráfico 1 estão os resultados da busca feita a partir do termo composto “linguagem de indexação” em bases de dados nacionais e internacionais.

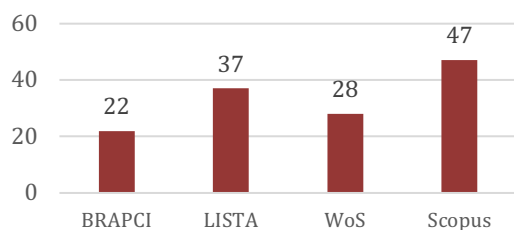


Gráfico 1. Resultados a partir dos termos “linguagem de indexação” em bases de dados nacionais e internacionais

O Gráfico apresenta a quantidade de trabalhos encontrados nos resultados em bases de dados nacionais e internacionais, a partir da busca pelo

termo composto “Linguagem de indexação”. Pode-se observar que a base com maior número de trabalhos foi a Scopus, seguido pela LISTA, WoS e BRAPCI.

A Scopus apresentou 47 pesquisas, no qual 2 abordavam o tema do estudo. Na LISTA foram encontrados 37 trabalhos, no qual apenas 1 possui a temática. A WoS teve 28 estudos em seus resultados, no qual 2 apresentaram relevância em seu assunto. Por último, a BRAPCI apresentou 22 trabalhos, no qual 5 foram considerados relevantes. A partir desses dados, pode-se analisar que houve uma quantidade muito pequena de trabalhos que de fato abordaram o uso de linguagens de indexação em bibliotecas universitárias, principalmente, quando comparado com a quantidade de resultados apresentados.

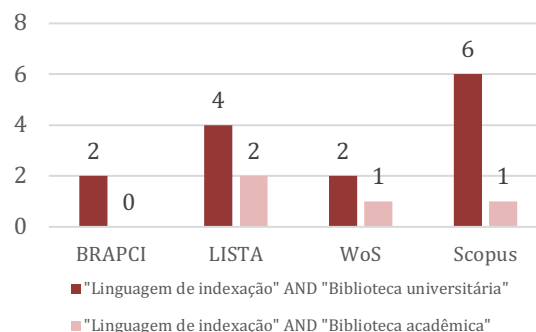


Gráfico 2. Resultados a partir dos termos “linguagem de indexação” AND “biblioteca universitária” em bases de dados nacionais e internacionais

No Gráfico 2 são apresentados os resultados da busca combinando dois termos compostos: “linguagem de indexação” AND “biblioteca universitária” e “linguagem de indexação” AND “biblioteca acadêmica”. A base de dados com a maior quantidade de estudos apresentados em seus resultados foi a Scopus, com 6, sendo 5 relevantes; seguido pela LISTA, com 4 pesquisas, no qual apenas 1 aborda a temática. A WoS e a BRAPCI apresentaram 2 trabalhos em cada, em que, na primeira base foi encontrado apenas 1 estudo e na segunda base os dois trabalhos foram considerados relevantes. Para um trabalho ser considerado relevante, eles tinham que abordar o uso de linguagem de indexação em bibliotecas universitárias. Pode-se observar nele uma quantidade muito pequena de estudos encontrados, no qual sua variação ocorreu entre nenhum trabalho, na base da BRAPCI, até 2 estudos, na base da LISTA. As bases WoS e Scopus apresentaram um trabalho em cada.

Por fim, cada base de dados apresentou uma quantidade diferente de trabalhos em seus resul-

tados; mesmo assim, notou-se que houve trabalhos que se repetiram em mais de uma base de dados. A base de dados que mais apresentou resultados foi Scopus, com 54 pesquisas; seguido pela base LISTA com 43 estudos e pela WoS com 31 trabalhos. A base de dados da BRAPCI ficou em último lugar com 24 pesquisas. Observou-se que o uso de estratégias de buscas como refinamento, operadores booleanos e aspas otimizaram as buscas e deixaram os resultados mais selecionados e consistentes com a temática pesquisada.

Ao analisar as bases de dados, pode-se perceber falhas ao recuperar os trabalhos, podendo gerar questionamento acerca da qualidade da indexação dos documentos, devido aos resultados apresentados após às buscas. As informações disponíveis sobre o processo de indexação nas bases de dados mostraram que os documentos são implementados por meio de indexadores e pelos próprios autores, não deixando claro se há uso da indexação automática ou vocabulário controlado na atribuição de termos. Nesse sentido, utilizar estratégias na hora de realizar a busca por publicações em bases de dados é algo imprescindível e uma forma de otimizar a recuperação nos resultados. Na análise foi possível observar como os resultados se alteraram com o uso de diferentes tipos de estratégias e como uma pesquisa pode ser afetada em sua qualidade dependendo dos tipos de buscas realizadas.

Essas falhas podem ocorrer por diferentes variáveis como, problemas na indexação dos sistemas, erros na atribuição de termos representativos pelos próprios autores ou pelo indexador e falta de conhecimento sobre o uso da linguagem de indexação para padronizar o processo.

4.2. Análise de estudos sobre linguagem de indexação em bibliotecas universitárias

Os 8 trabalhos selecionados para compor a análise estão relacionados na Tabela 1. A temática abordada neles coincide e se relaciona com o tema desenvolvido nesta pesquisa. Essa relação ocorre pelos trabalhos abordarem o uso de linguagem de indexação em bibliotecas universitárias, através da representação e recuperação da informação em repositórios, uso de vocabulários controlados e avaliação da indexação e do profissional indexador. Por outro lado, analisar uma quantidade menor de trabalhos possibilitou realizar análises mais aprofundadas em estudos mais pertinentes ao tema de investigação.

A análise terá início com a observação das palavras-chave atribuídas aos trabalhos selecionados e, na sequência, cada trabalho será individu-

almente apresentado a partir do objetivo, metodologia, resultados e principais conclusões. Ao final, a análise conjunta e comparada dos resultados dos trabalhos.

A tabela II apresenta as palavras-chave dos trabalhos que abordaram a linguagem de indexação. A necessidade em analisá-las ocorreu devido a importância em observar como as diferentes pesquisas atribuíram as palavras-chave para representar seus próprios estudos. Essa análise é pertinente e essencial para o trabalho, pois mostra uma perspectiva que pode interferir na representação e, conseqüentemente, na recuperação, bem como na seleção dos trabalhos para a análise. Pode-se observar uma grande diversidade de palavras-chave, que variam entre assuntos mais gerais até os mais específicos.

Autor(es)	Palavras-chave
Menard (2010)	Não apresenta palavras-chave no texto
Tartarotti (2014)	"Tratamento Temático da Informação"; "Bibliotecas universitárias"; "Protocolo Verbal"; "Avaliação da Indexação"; "Ciência, Tecnologia e Sociedade"
Cruz, Fujita e Santos (2017)	"Linguagem de indexação"; "Política de indexação"; "Indexação"; "Bibliotecas universitárias"
Sunny e Angadi (2017)	"Systematic literature review"; "Evaluation studies"; "Thesaurus"; "Digital information retrieval systems"; "Information retrieval effectiveness"; "Usability studies"
Fujita, Moreira, Santos, Cruz e Ribas (2018)	"Indexing languages"; "Hierarchical"; "Engineering"; "Subject heading lists"
Cruz (2019)	"Linguagem de indexação"; "Controle de vocabulário"; "Bibliotecas universitárias"
Fujita, Cruz, Patrício e Branco (2019)	"Organização do conhecimento"; "Indexação"; "Linguagens de indexação"; "Política de indexação"; "Bibliotecas universitárias"
Golub, Tyrkkö, Hansson, Ahlström (2020)	"Digital libraries"; "Digital humanities"; "Institutional repositories"; "Humanities"; "Knowledge organization"; "Bibliographic databases"; "Subject indexing"

Tabela II. Palavras-chave dos 8 trabalhos selecionados

Observa-se que em quatro estudos as palavras-chave com maior frequência foram: "bibliotecas universitárias" e "linguagens de indexação" / "indexing languages".

Houve trabalhos que apresentaram palavras-chave como, humanidades, engenharia, estudo de usabilidade, estudos de avaliação entre outros. Esses termos, por si só, não são adequados para representar os trabalhos e os que poderiam representar em algum nível não possuem relação direta com a temática abordada.

Percebeu-se, no geral, trabalhos que atribuíram palavras-chave adequadamente, mas também

trabalhos que utilizaram palavras-chave sem relação com a linguagem de indexação. Isso afeta a recuperação, pois mesmo quando o estudo é recuperado, a falta de representação pelas palavras-chave pode afetar a leitura do usuário, fazendo-o questionar se o texto será utilizado. Nessa condição, o texto, poderá ser descartado antes mesmo da leitura, pois, com o tempo limitado e a grande quantidade de texto para ser lido, o usuário apenas lê as informações principais como título, resumo e palavras-chave.

Os trabalhos selecionados serão analisados e discutidos pela seguinte ordem cronológica abaixo: Ménard (2010), Tartarotti (2014), Cruz, Fujita e Santos (2017), Sunny e Angadi (2017), Fujita, Moreira, Santos, Cruz e Ribas. (2018), Fujita, Cruz, Patrício e Branco (2019), Cruz (2019), Golub, Tyrkkö, Hansson, Ahlström (2020). Nesses trabalhos, foram analisados como os autores desenvolveram o uso da linguagem de indexação voltado para a biblioteca universitária, observando os resultados e conclusões apontados em seus estudos acerca dessa temática.

Ménard (2010) abordou em seu estudo a comparação entre dois vocabulários de indexação na recuperação de imagens. A autora concluiu que a análise dos termos de indexação demonstrou que ambas as abordagens de indexação, com o vocabulário controlado e não controlado, são inerentemente diferentes uma da outra. No entanto, em termos de linguagem de indexação, a análise dos termos confirma que as características das abordagens são bastantes semelhantes, dentro da mesma indexação, independentemente da linguagem adotada. Ainda de acordo com a autora, se existirem diferenças, elas são manifestadas mais na indexação utilizando vocabulário não controlado.

Dependendo das circunstâncias, as duas abordagens podem coexistir, ressaltando-se que a linguagem não controlada não deve necessariamente ser considerada como uma alternativa ou uma solução para substituir a indexação que emprega a linguagem controlada. Os resultados obtidos sugeriram que os sistemas de informação podem, em um futuro próximo, se beneficiar do uso combinado do vocabulário controlado e não controlado, recorrendo, por exemplo, à indexação colaborativa por parte dos usuários e por parte de pesquisadores no assunto.

Na investigação realizada por Tartarotti (2014) em sua dissertação acerca do profissional no tratamento temático da informação (TTI), foi observado que o catálogo com as bases de dados de áreas científicas especializadas permite a visibilidade informacional das bibliotecas universitárias, ressaltando que a determinação de assunto,

nesses sistemas, envolve um profissional indexador que, geralmente, não é especialista na área científica especializada em que atua. Por isso, é responsabilidade dele determinar de forma precisa o conteúdo do documento. A autora ressalta que esse processo pressupõe um nível elevado de complexidade, porque existem problemas relacionados à terminologia do documento e da própria área e a influência das variáveis leitor-indexador, texto e contexto do processo de leitura documentária dos profissionais.

No trabalho desenvolvido por Cruz, Fujita e Santos (2017, p. 222) sobre linguagem de indexação em bibliotecas universitárias, chegou-se à conclusão de que muitas instituições oferecem uma variedade de cursos e, com isso, a cobertura de assuntos da biblioteca é abrangente. Por isso, foi recomendado pelas autoras no investimento da construção de uma linguagem de indexação, já que a “[...] linguagem em consonância com a política de indexação trabalharia em função do sistema em questão ao buscar uma padronização mais consistente [...]”. Mesmo assim, as autoras constataram a necessidade em realizar a manutenção periódica da linguagem, devido à constante evolução das ciências, e para que a representação de assuntos possa ser adequada e específica aos objetivos e à realidade da biblioteca, desde que observada a capacidade do software de gestão da linguagem em inserir novos termos ou a retirada dos que não serão mais utilizados.

Sunny e Angadi (2017) desenvolveram uma pesquisa cuja observação ocorreu em estudos de avaliação publicados na língua inglesa e indexados nas bases de dados da LISA e Scopus, sobre a eficácia de tesouros na recuperação da informação em sistemas de informação digital. De acordo com os autores, um tesouro pode melhorar a eficácia da recuperação, se os documentos forem pré-indexados e se a busca for processada com os termos do tesouro. O estudo percebeu que a eficácia da recuperação também pode ser melhorada com a aplicação do tesouro por meio de outras técnicas, como modelo vetorial ou medida de similaridade de Jaccard para classificar os resultados.

Fujita, Moreira, Santos, Cruz e Ribas (2018) relataram o trabalho de uma equipe de catalogadores e pesquisadores engajados em construir uma estrutura hierárquica de uma linguagem de indexação para um catálogo online nas áreas do conhecimento de uma rede de bibliotecas universitárias. Para o desenvolvimento da linguagem de indexação, foram utilizadas estruturas hierárquicas, definindo categorias e subcategorias que formaram uma macroestrutura, usando parâmetros da Biblioteca do Congresso e Terminologia

da Biblioteca Nacional e do Vocabulário do Sistemas de Bibliotecas da Universidade de São Paulo. Os resultados possibilitaram a avaliação das estruturas hierárquicas dos idiomas utilizados na organização de termos superordenados e subordinados, o que contribuiu para a sistematização dos procedimentos operacionais contidos em um manual de linguagem de indexação para catálogos online de bibliotecas.

Fujita, Cruz, Patrício e Branco (2019) realizaram estudo analítico sobre linguagens de indexação em bibliotecas universitárias por meio da aplicação de questionários. Os resultados obtidos revelaram novos cenários, com avanços, problemas e lacunas no uso da linguagem pela biblioteca universitária. Um dos desafios está relacionado ao problema de falta de compreensão pelos profissionais que efetuam a mediação da linguagem de indexação na representação e recuperação da informação. Desse modo, foram realizadas buscas na interface de 24 bibliotecas, na tentativa de verificar a especificidade na representação, porém, não obtiveram o alcance da representatividade na estratégia de busca pelo usuário na recuperação, mostrando que nenhuma das bibliotecas analisadas, nesse aspecto, disponibilizam algum tipo de linguagem para consulta dos usuários. Foi levantada a hipótese da complexidade que seria, para o usuário, consultar mais de uma linguagem para elaborar a sua estratégia de busca.

Ainda no trabalho desenvolvido por Fujita, Cruz, Patrício e Branco (2019, p. 220), notou-se que houve também um avanço considerável em 20 bibliotecas (um terço do total de 60 bibliotecas da amostra), as quais desenvolveram linguagem própria, enquanto 10 bibliotecas declararam pretensão em construir. As autoras consideraram esse dado relevante, pois demonstra uma “[...] tendência da gestão do tratamento temático com política de indexação que define ações equilibradas entre a representação e a recuperação da informação [...]”. Já a lacuna evidenciada pelas autoras está relacionada a esse avanço, porque diz respeito ao desconhecimento de softwares de gestão e manutenção de linguagens de indexação e à possibilidade de construção compartilhada de linguagens de indexação pela interoperabilidade. O único software relatado foi o Tema-Tres por apenas uma biblioteca, o qual dispõe de dispositivos que facilitam o seu uso articulado aos softwares de gestão de bibliotecas.

A pesquisa desenvolvida por Cruz (2019), teve como proposta investigar a aplicação da linguagem de indexação em bibliotecas universitárias brasileiras, através da perspectiva do uso de ferramentas de controle do vocabulário. Como método a autora aplicou questionários, através do

Google Form, em bibliotecas universitárias brasileiras e comparou os resultados com os do estudo de Fujita, Cruz, Patrício e Branco (2019), no qual os autores fizeram um estudo com temática semelhante e abrangente. A autora verificou que todas as bibliotecas universitárias possuem profissionais dedicados à tarefa de realizar a indexação e que 84,8% utilizam a linguagem de indexação no TTI, mas, quando não encontram o termo na linguagem empregada, as bibliotecas usam a linguagem natural, podendo acarretar inconsistências no catálogo da biblioteca e disparidades na recuperação da informação. Ela constatou que metade das bibliotecas disponibilizam a linguagem de indexação no catálogo aos usuários e possui uma política de indexação documentada (Cruz, 2019).

Golub Tyrkkö, Hansson e Ahlström (2020, p. 1210, tradução nossa) procederam a um estudo sobre a indexação de assunto em humanidades, fazendo a comparação entre o repositório de uma universidade pública na Suécia e um serviço bibliográfico internacional. O estudo apontou que os objetivos bibliográficos fixados para garantir o acesso do assunto para artigos de periódicos em humanidades não são adequadamente suportados no repositório da universidade e na Scopus, considerada pelos autores como o maior banco de dados de resumos e citação comercial do mundo. Ainda de acordo com os autores, as políticas de indexação nos dois serviços parecem não atender às necessidades dos acadêmicos de humanidades para termos de indexação de assunto. Foi constatado que nenhum vocabulário controlado é utilizado no repositório da universidade pública da Suécia, em virtude, segundo os autores, do fato de que as categorias do repositório se destinam, principalmente, à análise estatística e não à recuperação, enquanto na Scopus, há o impedimento da recuperação eficaz por meio da omissão de seus vocabulários controlados. Os resultados também indicaram que existe falta de mapeamento entre vocabulários que geram duplicações e impedem os usuários de usar termos de um vocabulário em todos os recursos.

O trabalho desenvolvido por Golub Tyrkkö, Hansson e Ahlström (2020, tradução nossa) concluiu que as palavras-chave incluídas pelo autor, no momento de depositar uma dissertação ou tese no repositório, são complementares às palavras-chave do vocabulário controlado, em virtude de os autores não serem treinados para esse processo de indexação e nem receber diretrizes de como implementar o processo. Para sanar esse problema é sugerido o fornecimento de treinamento e orientações para os autores, com o intuito de aprimorar e agilizar o processo.

Em uma síntese dos estudos analisados pode-se observar uma variedade de abordagens do uso de linguagens de indexação por bibliotecas universitárias. No estudo de Ménard (2010), foi abordada a necessidade de coexistência entre a linguagem de indexação e a linguagem natural (ou não controlada). A justificativa ocorreu ao apontar que ela poderá ser utilizada no futuro como uma indexação colaborativa, ocorrendo por partes dos usuários e pesquisadores no assunto. A indexação colaborativa abordada para aproximar o usuário da indexação e possibilitar que a representação da informação receba diferentes perspectivas em seu processo na perspectiva de contribuir substancialmente para a indexação. Ambas as perspectivas poderão ser contempladas, entretanto, esse processo possui muita complexidade na representação dos conteúdos de recursos informacionais o que poderá gerar discrepâncias entre o usuário e o indexador.

Na dissertação de Tartarotti (2014), percebeu-se a existência de circunstâncias e variáveis de leitura documentária que influenciam no processo de indexação e no uso de linguagem de indexação. Essas variáveis foram comentadas por Fujita, Cruz, Patrício e Branco (2019) e Golub, Tyrkkö, Hansson e Ahlström (2020), no relato de problemas, lacunas e avanços da linguagem de indexação no âmbito da biblioteca universitária, bem como a falta do uso de linguagem de indexação ou o uso inadequado para controlar o vocabulário. Para Tartarotti (2014), o indexador pode influenciar no processo de indexação. A autora apontou que existem aspectos do profissional que podem interferir nesse processo como suas experiências, pré-conceitos e preconceitos, formação acadêmica e conhecimento em diferentes níveis do mundo. Enquanto, para Fujita, Cruz, Patrício e Branco (2019), também existem outros aspectos que podem interferir no processo de indexação como os próprios sistemas e o ambiente em que se inserem; a elaboração de diretrizes de política de indexação; e o uso da linguagem de indexação para controle de vocabulário. Ambos os trabalhos, em conjunto, demonstraram diferentes aspectos e em diferentes níveis que podem interferir no processo de indexação.

5. Considerações finais

Através do aprofundamento teórico, da observação das bases de dados e da análise dos estudos selecionados pode-se chegar a algumas considerações sobre o uso de linguagem de indexação em bibliotecas universitárias.

O desenvolvimento da fundamentação teórica, com base no levantamento da literatura e análise dos resultados permitiram retomar o objetivo de

analisar estudos sobre linguagens de indexação, por meio da pesquisa bibliográfica, com o intuito de analisar semelhanças e diferenças, métodos, proposições, contribuições e recomendações.

As semelhanças encontradas na literatura analisada abordam, principalmente, a influência do autor e do indexador no processo de indexação devido à diferentes interpretações e perspectivas que afetam a representação e recuperação da informação (Ménard, 2010; Tartarotti, 2014; Golub, Tyrjö, Hansson & Ahlström, 2020).

As diferenças observadas na literatura analisada referem-se a variáveis que interferem no processo de indexação tais como as relacionadas ao Indexador. Nesse sentido, sua formação acadêmica, pré-conceitos e preconceitos, conhecimento em diferentes níveis de mundo e suas experiências podem interferir no processo de indexação, pois se trata de um ser humano que possui personalidade e pensamentos que difere do outro ser humano (Tartarotti, 2014). Enquanto, para outros autores, a variável interferente é o sistema, com a falta de elaboração de políticas e manual de indexação (Fujita, Cruz, Patrício & Branco, 2019). Considera-se que ambas as variáveis estão ligadas inerentemente e interferem no processo de indexação.

Como proposição na literatura analisada, destaca-se que a combinação de vocabulário controlado e não controlado pode, em um futuro próximo, ser benéfica para o processo de indexação, caracterizando, dessa maneira, em um sistema de indexação colaborativa (Ménard, 2010). Ainda nesse sentido, o uso de tesouro, por exemplo, pode melhorar a eficácia na recuperação da informação, se os documentos forem pré-indexados e se a busca for processada com os termos do tesouro, mas nesse ponto envolve o conhecimento do usuário acerca do uso de tesouro no sistema (Sunny & Angadi, 2017).

Uma das contribuições relevantes da literatura analisada consistiu na sistematização dos procedimentos contidos em um manual de linguagem de indexação para catálogos online em uma rede de bibliotecas (Cruz, 2019). O método para a construção da linguagem de indexação, apresenta-se inovador porque consistiu no uso de estruturas hierárquicas para definir categorias e subcategorias, que formaram uma macroestrutura, utilizando parâmetros da *Library of Congress Subject Headings*, Terminologia da Biblioteca Nacional e o Vocabulário do Sistemas de Bibliotecas da USP (Fujita, Moreira, Santos, Cruz & Ribas, 2018). O método utilizado para a construção da linguagem de indexação interfere diretamente no seu funcionamento e, conseqüentemente, em seus resultados. Por isso, necessita

de cuidado e atenção, principalmente, nos detalhes de sua construção.

Com a observação da literatura analisada ressalta-se, também, a recomendação de treinamento e orientações para autores e indexadores sobre o processo de indexação de forma mais eficiente. A importância do treinamento decorre do impacto que a falta dele pode gerar na atuação do bibliotecário. A falta de atualização acerca de como realizar o processo de indexação, especialmente, no uso da linguagem de indexação, atribuição de termos, controle de vocabulário e do próprio sistema pode causar impacto na qualidade da indexação, prejudicando todo o procedimento e a recuperação da informação pelo usuário. Por isso, a necessidade de manter o indexador sempre atualizado com treinamentos e cursos (Ménard, 2010; Tartarotti, 2014; Golub; Tyrkö; Hansson e Ahlström, 2020); e, por outro lado, em investimentos na construção, manutenção e avaliação periódica da linguagem de indexação (Cruz, Fujita e Santos, 2017).

As ideias apresentadas pelos autores envolveram indexação compartilhada, necessidade de treinamento de autores e indexadores para adequada atribuição de termos com consulta da linguagem de indexação e a avaliação e manutenção constante da linguagem.

A partir do desenvolvimento dos trabalhos observou-se que a política de indexação precisa incluir a combinação de multifatores, de forma a garantir uma representação e recuperação da informação adequada ao sistema e seus diferentes atores. Os multifatores e as variáveis considerados que podem interferir nesse processo são: o indexador, a qualidade da linguagem de indexação, o instrumento utilizado, sistemas e o próprio contexto da biblioteca universitária.

Esse contexto requer um sistema de representação e recuperação da informação eficaz e, em função do emprego intensivo de informações com características comuns e como forma de assegurar esse processo, devem ser feitas a elaboração e a implantação da linguagem de indexação. A linguagem de indexação atua nos sistemas de informação, orientando o catalogador sobre quais os melhores termos para a representação de um assunto, da mesma maneira como auxilia os usuários na elaboração de estratégias de buscas por informações no sistema. Em sua ausência pode ocorrer problemas relacionados ao controle de vocabulário no sistema, afetando diretamente a representação e recuperação como resultados não condizentes com a busca realizada devido aos excessos de materiais de outras temáticas e falta dos documentos no assunto pesquisado. Enquanto a política de indexação é

um guia para tomada de decisões e o estabelecimento de diretrizes como: linguagem de indexação utilizada, como será feito o processo de indexação e recursos materiais e financeiros, atendimento das necessidades dos usuários entre outros procedimentos necessários (Moura, Silva & Amorim, 2002; White, 2013; Tejeda-Lorent et al., 2014; Váñez; Pedraza-Jiménez, Condia e Blanco & Rovira 2015).

Recomenda-se que haja continuidade em trabalhos sobre linguagens de indexação, devido ao desenvolvimento constante da tecnologia e da ciência, para que possibilite se aperfeiçoar, acompanhar e evoluir na temática, sem deixar se tornar obsoleto e causar perdas dos dados inseridos nos sistemas utilizados.

Referencias

- American National Standards Institute/National Information Standards Organization (ANSI/NISO). Z39.19-2005 (R2010). Guidelines for the construction format, and management of monolingual controlled vocabularies. Bethesda, Maryland: NISO Press, 2010.
- Arboit, A. E. (2017). É possível uma organização e representação do conhecimento sem partidos?. // P2P & Inov. Rio de Janeiro. 3:1 (set./mar. 2017).
- Cintra, A. M. M.; Tálamo, M. de F. G. M.; Lara, M. L. G.; Kobashi, N. Y. (2002). Para entender as linguagens documentárias. Rio de Janeiro: Polis, 2002.
- Cruz, M. C. A. (2019). Linguagens de indexação em bibliotecas universitárias: estudo analítico em território nacional. 2019. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, 2019.
- Cruz, M. C. A.; Santos, L. B. P.; Fujita, M. S. L. Linguagens de indexação em bibliotecas universitárias brasileiras: diagnóstico preliminar das regiões sul e sudeste. // X Asociación de Educación e Investigación en Ciencia de la Información de Iberoamérica y el Caribe (X EDICIC), Belo Horizonte, 2016.
- Cruz, M. C. A.; Fujita, M. S. L.; Santos, L. B. P. (2017). Linguagem de indexação no contexto da política de indexação: estudo em bibliotecas universitárias. // Pinho, F. A.; Guimarães, J. A. C. (Orgs.). Memória, tecnologia e cultura na organização do conhecimento. Recife: UFPE, 2017, 217-224.
- Dias, E. W.; Naves, M. M. L. (2007). Análise de assunto: teoria e prática. Brasília: Thesaurus, 2007.
- Fujita, M. S. L. (2017). Manual de política de indexação para bibliotecas universitárias da Unesp. // São Paulo: Unesp (2017) 36.
- Fujita, M. S. L.; Gil Leiva, I. (2010). As linguagens de indexação em bibliotecas nacionais, arquivos e sistemas de informação na América-latina. // Seminário Nacional de Bibliotecas Universitárias. 2. 2010.
- Fujita, M. S. L.; Moreira, W.; Santos, L. B. P.; Cruz, M. C. A.; Ribas, R. R. de B. (2018). Construction and evaluation of hierarchical structures of indexing languages for online catalogs of libraries: an experience of the São Paulo State University (Unesp). // Knowledge Organization. 45:3. 2018.
- Fujita, M. S. L.; Cruz, M. C. A.; Patrício, B. O. M.; Branco, L. B. P. R. (2019). Linguagens de indexação em bibliotecas universitárias: estudo analítico. // Informação & Informação. 24:1 (jan./abr. 2019) 190-225.

- Fujita, M. S. L.; Favato, V. A. M.; Zaniboni, M. M.; Fagundes, S. A. (2016). A elaboração da política de indexação para bibliotecas universitárias da Unesp. // Fujita, M. S. L. (Org.). Política de indexação para bibliotecas: elaboração, avaliação e implantação. Marília: Cultura Acadêmica, 2016.
- Fujita, M. S. L.; Santos, L. B. P. (2016). A estrutura lógico-hierárquica de linguagens de indexação utilizadas por bibliotecas universitárias. // Scire. 22:2 (jul./dez. 2016) 37-46
- Gil Urdiciain, B. (2004). Manual de lenguajes documentales. // Gijón: Trea. 2004.
- Golub, K; Tyrkkö, J; Hansson, J; Ahlström, I. (2020). Subject indexing in humanities: a comparison between a local university repository and international bibliographic service. // Journal of Documentation. 76:6 (2020) 1193-1214.
- International Organization for Standardization (ISO) 25.964-1. (2011). Information and documentation – thesauri and interoperability with other vocabularies – part 1: thesauri for information retrieval. // Genebra. 2011.
- Kitchham, B; Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. // Version 2.3. Keele: Keele University/School of Computer Science and Mathematics/Software Engineering Group. 65 p.
- Menard, E. (2010). Ordinary image retrieval in multilingual context: a comparison of two indexing vocabularies. // Aslib Proceedings: New Information Perspective. 62: 4/5 (2010) 428-437.
- Moura, M. A; Silva, A. P; Amorim, V. R. (2002). A concepção e o uso das linguagens de indexação face às contribuições da semiótica e da semiologia. // Informação & Sociedade: estudos. 12:1. 2002.
- Ramires, A. Q. (2022). Estratégias de leitura no ensino fundamental para formação de leitores profissionais em indexação. 2022. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Marília, 2019.
- Silva, R. C; Brito, J. F. (2018). Proposta de um manual de indexação para bibliotecas universitárias. // Informação@Profissões. 7:1 (jan./jun. 2018) 92-113.
- Sunny, S. K; Angadi, M. (2017). Evaluation the effectiveness of thesauri In digital information retrieval systems. // The electronic library. 36:1 (2018) 55-70.
- Tartarotti, R. C. D. E. (2014). Atuação bibliotecária no tratamento temático da informação em unidades de informação: estudo comparativo qualitativo. Dissertação (Mestrado em Ciência, Tecnologia e Sociedade) – Universidade Federal de São Carlos – UFSCar, 2014.
- Tejeda-Lorente, A.; Porcel, C; Peis, E; Sanz, R; Herrera-Viedma, E. (2014). A quality-based recommender system to disseminate information in a university digital library. Information Science. :261. 52-69.
- Vállez, M; Pedraza-Jiménez, R; Condia, L; Blanco, S; Rovira, C. (2015). Updating controlled vocabularies by analyzing query logs. // Online Information Review. 39:7 (2015) 870-884.

Enviado: 2022-04-10. Segunda versão: 2022-06-08.
Aceptado: 2022-06-10.
