

A disponibilização de vocabulário controlado aos usuários para a recuperação da informação

Maria Carolina Andrade e Cruz

Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Departamento de Ciência da Informação, Marília, SP, Brasil
maria.andrade@unesp.br

Edberto Ferneda

Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Departamento de Ciência da Informação, Marília, SP, Brasil
edberto.ferneda@unesp.br

Mariângela Spotti Lopes Fujita

Universidade Estadual Paulista, Faculdade de Filosofia e Ciências, Departamento de Ciência da Informação, Marília, SP, Brasil
mariangela.fujita@unesp.br

DOI: <https://doi.org/10.26512/rici.v15.n1.2022.42464>

Recebido/Recibido/Received: 2022-01-04

Aceitado/Aceptado/Accepted: 2022-03-31

ARTIGOS

Resumo

Os vocabulários controlados normalmente são utilizados na indexação, no processo de padronização dos conceitos para a representação da informação, atuando como agente mediador entre a linguagem natural contida nos documentos e a linguagem do sistema de informação. Contudo, os vocabulários controlados também podem favorecer a busca de informação a partir de sua disponibilização ao usuário, considerando que isso contribuiria para sua autonomia no processo recuperação da informação, tendo em vista que as bibliotecas possam disponibilizar esse recurso no catálogo *online* onde se realizam buscas a partir da iniciativa e necessidade dos usuários. Com isso, resgata-se pesquisa anterior em que foram aplicados questionários sobre o uso de vocabulários controlados em bibliotecas universitárias e destaca-se a questão relacionada à presença do vocabulário no catálogo online para uso do usuário, na qual 39 bibliotecas afirmam fazer uso de vocabulários controlados no processo de indexação e 18 delas disponibilizam o recurso no catálogo *online*. **Objetivo:** Com isso, o objetivo proposto é investigar os vocabulários controlados presentes nos catálogos das bibliotecas participantes e analisar a sua disposição, ou seja, como se apresentam aos usuários. **Metodologia:** Pesquisa exploratória a partir da análise em catálogos *online* de bibliotecas universitárias. **Resultados:** A pesquisa reconheceu a presença de vocabulários controlados em 16 catálogos distintos, em que foram encontradas lista de autoridades (11 bibliotecas), tesouros (4) e lista de cabeçalhos de assuntos (1). **Conclusão:** Tendo em vista as 39 bibliotecas que dispõem de vocabulários controlados para indexação, foram poucas as bibliotecas que disponibilizam algum para o usuário, apenas 16. Entretanto, para essas bibliotecas ou para outras além das analisadas nesta pesquisa, sugere-se melhor apresentação dos vocabulários nos catálogos, pois estes podem passar despercebidos por não estarem facilmente às vistas do usuário. Além disso, proporcionar aos usuários tutoriais de uso para que possam conhecer o recurso e aprimorar as buscas realizadas no catálogo.

Palavras-Chave: Vocabulário controlado. Recuperação da informação. Busca da informação. Catálogo em linha.

**La disponibilidad del vocabulario controlado para los usuarios para la recuperación de información
Resumen**

Los vocabularios controlados se utilizan normalmente en la indexación, en el proceso de estandarización de conceptos para la representación de la información, actuando como agente mediador entre el lenguaje natural contenido en los documentos y el lenguaje del sistema de información. Sin embargo, los vocabularios controlados también pueden favorecer la búsqueda de información a partir de su disponibilidad para el usuario, considerando que esto contribuiría a la autonomía del usuario en el proceso de recuperación de la información, teniendo en cuenta que las bibliotecas pueden poner a disposición este recurso en el catálogo en línea donde las búsquedas se realizan a partir de la iniciativa y necesidad de los usuarios. Así, se rescata una investigación previa en la que se aplicaron cuestionarios sobre el uso de vocabularios controlados en las bibliotecas universitarias y se destaca la pregunta relacionada con la presencia del vocabulario en el catálogo en línea para uso del usuario, en la que 39 bibliotecas afirman que utilizan vocabularios controlados en el proceso de indización y 18 de ellas ponen el recurso a disposición en el catálogo en línea. **Objetivo:** Así, el objetivo propuesto es investigar los vocabularios controlados presentes en los catálogos de las bibliotecas participantes y analizar su disponibilidad, es decir, cómo se presentan a los usuarios. **Metodología:** Investigación exploratoria a partir del análisis en catálogos online de bibliotecas universitarias. **Resultados:** La investigación reconoció la presencia de vocabularios controlados en 16 catálogos distintos, en los que se encontró la lista de autoridades (11 bibliotecas), el tesoro (4) y la lista de encabezamientos de materia (1). **Conclusión:** Considerando las 39 bibliotecas que tienen vocabularios controlados para la indización, fueron pocas las que ponen algunos a disposición del usuario, sólo 16. Sin embargo, para estas bibliotecas o para otras además de las analizadas en esta investigación, se sugiere una mejor presentación de los vocabularios en los catálogos, ya que pueden pasar desapercibidos por no ser fácilmente vistos por el usuario. Además, se recomienda proporcionar a los usuarios tutoriales sobre su uso para que puedan conocer el recurso y mejorar las búsquedas realizadas en el catálogo.

Palabras clave: Vocabulario controlado. Recuperación de información. Búsqueda de información. Catálogo en línea.

The availability of controlled vocabulary to users for information retrieval

Abstract

The controlled vocabularies are normally used in indexing, in the process of standardization of concepts for the representation of information, acting as a mediating agent between the natural language contained in documents and the language of the information system. However, controlled vocabularies can also favor the search for information from their availability to the user, considering that this would contribute to the user's autonomy in the information retrieval process, considering that libraries can make this resource available in the online catalog where searches are performed based on the initiative and needs of users. Thus, previous research on the use of controlled vocabularies in university libraries is rescued, and we highlight the question related to the presence of the vocabulary in the online catalog for the user's use, in which 39 libraries state that they use controlled vocabularies in the indexing process and 18 of them make the resource available in the online catalog. **Objective:** With this, the proposed objective is to investigate the controlled vocabularies present in the catalogs of the participating libraries and analyze their availability, that is, how they are presented to the users. **Methodology:** Exploratory research from the analysis in online catalogs of university libraries. **Results:** The research recognized the presence of controlled vocabularies in 16 distinct catalogs, in which authority list (11 libraries), thesaurus (4) and subject heading list (1) were found. **Conclusion:** Considering the 39 libraries that have controlled vocabularies for indexing, only 16 libraries (16) made some available to the user. However, for these libraries or for others besides those analyzed in this research, we suggest a better presentation of the vocabularies in the catalogs, since they may go unnoticed because they are not easily visible to the user. In addition, users should be provided with tutorials on how to use them so that they can get to know the resource and improve the searches performed in the catalog.

Keywords: Controlled vocabulary. Information retrieval. Information search. Online catalog.

1 Introdução

A área de Recuperação da Informação reflete a grande demanda na área de Ciência da Informação pelo interesse em estudos no desenvolvimento de métodos e técnicas para

aprimorar as formas de organizar a produção informacional da sociedade. Para entender o papel da recuperação da informação no contexto da Ciência da Informação, deve-se recordar as raízes de seu surgimento.

Segundo Borko (1968), a Ciência da Informação é a disciplina que investiga todos os aspectos que permeiam a informação, como suas propriedades, comportamentos, forças que governam seu fluxo e o processamento da informação para otimizar seu acesso e usabilidade. Além disso, o autor aponta como foco da disciplina “a origem, coleta, organização, estocagem, recuperação, interpretação, transmissão, transformação e uso de informação” (BORKO, 1968, p. 3).

Saracevic (1996) que acompanhou a evolução dos estudos e questões relacionadas a área, a definiu como:

A Ciência da Informação é um campo dedicado às questões científicas e à prática profissional voltadas para os problemas da efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação. No tratamento destas questões são consideradas de particular interesse as vantagens das modernas tecnologias informacionais (SARACEVIC, 1996, p. 47).

Percebe-se que a definição de Borko (1968) explicita a pesquisa, a investigação dos aspectos, propriedades e comportamentos relacionados à informação. Saracevic (1996) estabelece que a Ciência da Informação é um campo dedicado às questões científicas, mas também inclui a prática profissional, ressaltando os esforços realizados diante de questões ou problemas a serem resolvidos por essa ciência.

Segundo Saracevic (1999, p.1051), afirma que a Recuperação da Informação (RI) se apresenta como um dos principais ramos da Ciência da Informação. Segundo ele:

A história de qualquer campo é a história de algumas ideias poderosas. Eu sugiro que a Ciência da Informação tem três ideias poderosas, até agora. Essas ideias lidam com o processamento de informações de uma maneira radicalmente diferente da que era feita anteriormente ou em outro lugar. A primeira e original ideia, surgida na década de 1950, é a recuperação da informação, proporcionando o processamento da informação com base na lógica formal. A segunda, surgindo logo em seguida, é a relevância, orientando e associando diretamente o processo às necessidades e avaliações humanas de informação. O terceiro, derivado de outro lugar cerca de duas décadas depois, é a interação, permitindo trocas diretas e feedback entre sistemas e pessoas envolvidas em processos de RI (SARACEVIC, 1999, p.1052).¹

¹ A history of any field is a history of a few powerful ideas. I suggest that information science has three such powerful ideas, so far. These ideas deal with processing of information in a radically different way than was done previously or elsewhere. The first and the original idea, emerging in 1950s, is information retrieval, providing for processing of information based on formal logic. The second, emerging shortly thereafter, is relevance, directly orienting and associating the process with human information needs and

Para Ferneda (2012, p. 129), o processo de recuperação de informação consiste em identificar em um acervo documental os documentos que venham a atender a uma determinada necessidade de informação. A recuperação se efetiva por meio da comparação entre as representações dos documentos e a representação da necessidade do usuário. Um documento é recuperado se a sua representação coincidir total ou parcialmente com a representação da necessidade do usuário. Portanto, a eficiência do processo de recuperação de informação é dependente da qualidade da representação dos documentos e das buscas realizadas pelos usuários. Tais representações são geralmente formalizadas textualmente, caracterizando a recuperação de informação como um processo linguístico cuja eficiência depende de um controle adequado da linguagem de representação dos itens de informação e das requisições de seus usuários.

Na Ciência da Informação, os vocabulários controlados geralmente são utilizados pelos indexadores com preocupação na padronização e consistência dos conceitos descritivos do conteúdo informacional dos documentos. Em um sistema de recuperação de informação, o vocabulário controlado pode também ser disponibilizado aos usuários a fim de auxiliá-los na elaboração de suas estratégias de busca.

Em pesquisa realizada por Cruz (2019), foi aplicado a bibliotecários universitários um questionário com um conjunto de questões relacionadas à disponibilização do vocabulário controlado (linguagem de indexação) aos usuários. Tendo em vista as respostas obtidas nessas questões, esse trabalho tem como objetivo analisar as respostas obtidas na referida pesquisa, verificando a forma como esse recurso está sendo oferecido nos sistemas de informação. Para alcançar esses objetivos, foram relacionadas as respostas obtidas por meio do questionário, verificando as bibliotecas que afirmaram disponibilizar o vocabulário controlado aos usuários. A partir dessa lista de bibliotecas, realizou-se uma investigação no catálogo online de cada uma delas.

Na próxima seção desenvolve-se o referencial teórico a respeito das funções dos vocabulários controlados, as recomendações normativas da International Organization for Standard (ISO 95964) e a disponibilização da ferramenta para o usuário.

2 vocabulário controlado: função mediadora

As bibliotecas universitárias, como unidades de informação que visam a socialização do conhecimento, buscam através de técnicas e estratégias formas eficazes e viáveis de aprimorar

assessments. The third, derived from elsewhere some two decades later, is interaction, enabling direct exchanges and feedback between systems and people engaged in IR processes.

os meios de recuperação da informação. De acordo com Hjørland (2021) o objetivo do campo da recuperação da informação é facilitar o acesso a documentos, conhecimento e informação. As bibliotecas estão acompanhando os sinais dos tempos e buscam aprimorar o sistema de informação às quais pertencem. Os catálogos online são realidade para a maioria das bibliotecas e, na esfera universitária, os repositórios institucionais estão presentes para abranger a produção científica da organização.

Sobre o papel das universidades, a presença dos repositórios institucionais e catálogos online, Fujita, Tolare (2019) apresentam:

Em universidades, os repositórios são institucionais e contêm toda a produção científica das áreas do conhecimento abrangidas pelas atividades acadêmicas à semelhança de um catálogo bibliográfico online. Essa circunstância é própria do ambiente universitário onde a produção científica abrange diferentes áreas do conhecimento e são especializadas com inovação científica e tecnológica. Ser precursora em geração de conhecimentos é condição natural da existência da universidade e significa que todos os seus sistemas de informação precisam acompanhar esta evolução na organização do conhecimento para sua gestão (FUJITA; TOLARE, 2019, p. 96).

Esta pesquisa tem como universo de pesquisa o catálogo *online* de bibliotecas universitárias brasileiras. Mesmo assim, percebemos a semelhança dita pelas autoras acima entre os catálogos *online* e os repositórios institucionais. Ambos contribuem para a organização da informação da instituição, além de proporcionar ao usuário independência para fazer sua pesquisa e poder recuperar a informação desejada. O diferencial de ambientes como os sistemas de informação de uma biblioteca para um sistema de busca de dados, por exemplo, é a preocupação com a recuperação consistente da informação a partir de esforços para o tratamento dos documentos, priorizando o uso de ferramentas elaboradas, políticas internas com interesses em proporcionar ao usuário sucesso em sua busca. O sistema busca ser eficiente para quem o utiliza, evitando disparidades e perdas.

Sendo assim, o sucesso da recuperação da informação é reflexo dos esforços realizados pelas bibliotecas no tratamento e organização da informação. A demanda informacional dos usuários deve ser compreendida e acatada para traçar planos para solucionar qualquer tipo de lacuna no sistema de organização.

A indexação é um procedimento, no qual cabe a função de descrever o conteúdo dos documentos com coerência aos interesses organizacionais e da comunidade de usuários e, principalmente, concentrando-se na recuperação da informação. O processo de indexação compreende a análise de assunto mediante a leitura técnica do documento pelo indexador, a seleção de termos que sejam identificados como chaves para compreender o documento analisado e a tradução desses termos, por meio de um vocabulário controlado, buscando consistência no sistema de informação, como no catálogo online ou em repositórios.

A consistência na tradução dos termos refere-se ao controle de vocabulário, ou seja, à garantia de uma uniformidade dos descritores utilizados na indexação a partir do uso de um vocabulário controlado previamente escolhido e/ou elaborado. A norma ISO 25964-2 (2013, p. 15, tradução nossa) expressa que o controle de vocabulário é o “gerenciamento de um vocabulário para eliminar a ambiguidade e restringir a forma dos termos e limitar o número de conceitos e termos disponíveis para indexação”.

O vocabulário controlado consiste em uma ferramenta que visa o controle dos termos utilizados em um sistema de informação, ou seja, há uma estrutura construída para representar conceitos atribuídos aos documentos. Ressalta-se que os termos são expressões linguísticas responsáveis por sintetizar e representar conceitos (DAHLBERG, 1978). Dito isso, Zhang *et al.* (2015) apresenta o vocabulário controlado como “sistema semântico” utilizado para a organização e recuperação de informação, capaz de compreender o escopo de um determinado domínio.

Ressalta-se que há variações terminológicas para o conceito de “vocabulário controlado”, encontram-se na literatura termos como linguagem documentária, linguagem de indexação, linguagens descritoras, linguagens de recuperação da informação, linguagens de informação, linguagem documental e etc. (BARITÉ, 2011; DODEBEI, 2002). Além dessas, outra variação encontrada, sendo mais atual na literatura, é sistemas de organização do conhecimento (SOCS), termo amplo empregado na tentativa de contemplar as ferramentas de indexação e de classificação, também na busca em sanar questões de excesso de termos, como descrito acima, e suas delimitações conceituais (BARITÉ, 2011). Por conseguinte, esses termos representam uma variação terminológica do mesmo domínio. Assim, justifica-se o uso do termo “vocabulário controlado” neste trabalho, pois se pretende seguir em conformidade com a literatura normativa, como a ISO 25964-1 (2011) e a ISO 25964-2 (2013), que utilizam “*controlled vocabularies*”, mas sem descartar as outras formas.

Antes da sistematização computacional das unidades de informação, os vocabulários controlados eram mais simplificados em sua estrutura, sendo uma ferramenta destinada aos indexadores, no formato de listas de termos. Com a automatização dos sistemas de informação, os vocabulários controlados seguem acompanhado os aprimoramentos para gerir toda a informação a ser representada, passando a ser um recurso utilizado também pelos usuários no momento da pesquisa pela informação.

De acordo com Golub, Schmiede e Tudhope (2019) os sistemas de organização do conhecimento, assim como os vocabulários controlados, cumprem uma função essencial no gerenciamento de informações no meio digital e em demais aplicações, como em sua atuação como ferramenta de representação da informação, na indexação, base de conhecimento para

pesquisadores, mapas semânticos para domínios e disciplinas, uma estrutura conceitual que pode servir como base para sistemas fundamentados em conhecimento.

Os autores listam uma série de usabilidade para os vocabulários controlados, contudo, além disso incluímos a função de recuperação da informação a partir do uso dessas ferramentas à disposição do usuário.

Na pesquisa de Cruz (2019) foi realizado um estudo a respeito do uso de linguagens de indexação, ou seja, de vocabulários controlado pesquisa mais precisos. No estudo, destaca-se ainda o alinhamento que se obtém a partir do método de representação do documento com a forma de recuperação da informação.

Sendo assim, não somente o vocabulário controlado é uma ferramenta de semântica para a representação e recuperação de documentos, como também consiste em um elemento mediador entre a linguagem do indexador, documento e usuário. É por meio dessa mediação que a recuperação da informação será alcançada.

As bibliotecas podem optar por utilizar um vocabulário controlado elaborado por outra instituição ou organização e até mesmo utilizar mais de um vocabulário, dependendo da abrangência de seu acervo. Ressalta-se que ao utilizar mais de um vocabulário é necessário discriminar na política de indexação para quais assuntos eles serão utilizados, buscando evitar a presença de diferentes termos para representar um único conceito. Há também as bibliotecas que decidem elaborar um vocabulário controlado próprio, de acordo com sua política de indexação, específico aos assuntos contidos no acervo da coleção. A construção dessa ferramenta exige planejamento para enfrentar desafios como custo, usabilidade e manutenção.

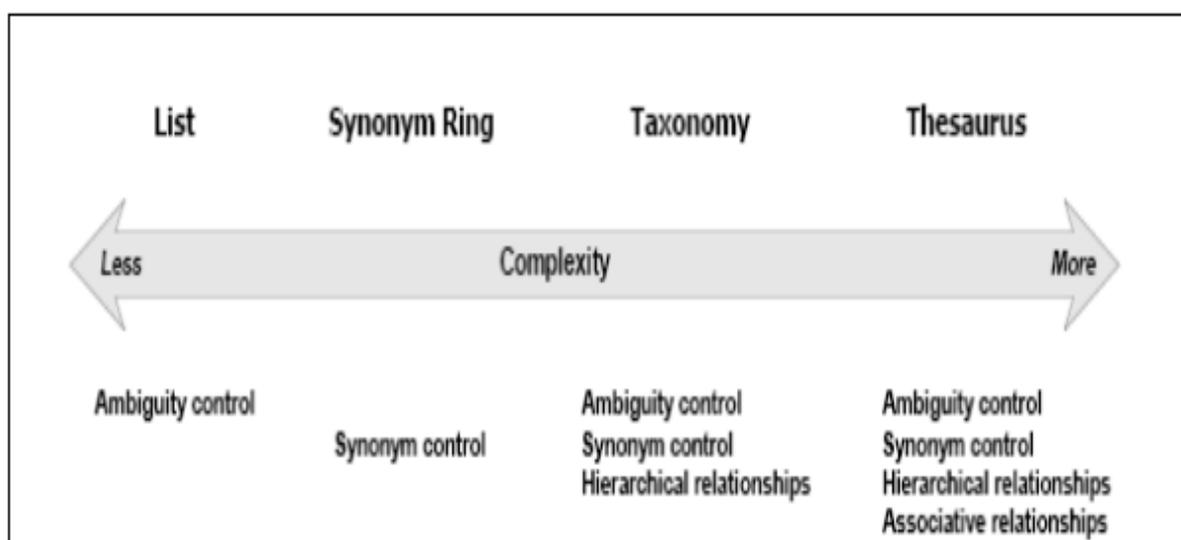
Diante disso, para amparar a construção de um vocabulário controlado existem normas específicas destinadas a isso; aliás, não somente para as no âmbito das bibliotecas universitárias. A análise feita pela autora destaca que ao oferecer esse recurso ao usuário, possibilita-se maior autonomia na busca, tendo em vista que o processo se torna mais consistente e poderá tornar os resultados de construção do vocabulário como também para o seu gerenciamento e manutenção. A norma mais atualizada sobre vocabulários controlados atualmente é a ISO 25964 (2011), na qual foi publicada em duas partes: “Information and documentation: part 1: thesauri and interoperability with Other vocabularies”, de 2011 e a “Information and documentation–thesauri and interoperability with Other vocabularies - part 2: Interoperability with Other vocabularies”, publicada em 2013.

De acordo com a norma, os vocabulários controlados são projetados para aplicações úteis a fim de identificar cada conceito com um rótulo consistente ao classificar documentos, indexá-los e/ou pesquisá-los (ISO 25964-1, 2011).

A norma que antecedeu a ISO 25964-1 (2011) foi a norma americana Z39:10 - *Guidelines for the construction, format, and management of monolingual controlled vocabularies*, desenvolvida pela American National Standards Institute (ANSI) da National Information Standards Organization (NISO), de 2005. Essa norma, como o próprio nome já diz, também compõe diretrizes para vocabulários controlados e demonstra, por meio da Figura 1, o grau de complexidade entre os vocabulários de acordo com sua estrutura. A norma expõe as listas, os anéis de sinônimos, as taxonomias e os tesouros. A partir dessa figura pode-se perceber que os vocabulários listados, em sua maioria, apresentam o controle de ambiguidade como constante e nos anéis de sinônimos, apresenta-se o controle de sinônimos. Contudo, a característica principal dessas ferramentas é o controle de vocabulário independente do seu grau de complexidade e estrutura.

Zeng (2008) apresenta quatro grupos de vocabulários controlados, sendo que utiliza o termo sistemas de organização do conhecimento (SOCS):

Figura 1 - Complexidade estrutural crescente entre vocabulários controlados.



Fonte: ANSI/NISO Z39.19 (2005, p. 17).

- 1- **Lista de termos:** glossários/ dicionários, listas de seleção, anéis de sinônimos.
- 2- **Modelos como metadados:** Lista de autoridades, diretórios e dicionários geográficos.
- 3- **Classificação e categorização:** Cabeçalho de assuntos, esquemas de categorização, taxonomias e esquemas de classificação.

4- Modelos de relacionamentos: Tesouros, redes semânticas e ontologias.

Na Figura 2 pode-se observar que a autora apresenta as categorias contemplando os diferentes tipos de SOCS, retratando por meio de uma visão geral suas estruturas e funções. No eixo vertical, são relacionadas as estruturas em que as categorias pertencem, apresentando-as como: 1. Plana; 2. Duas dimensões; 3. Múltiplas dimensões. No eixo horizontal, a autora relaciona as funções de cada recurso, diferenciando, na parte inferior da figura, suas principais finalidades, sendo elas: eliminar ambiguidade, controle de sinônimos, estabelecimento de relações: hierárquicas e associativas, e apresentar propriedades.

Figura 2- An overview of the structures and functions of KOS

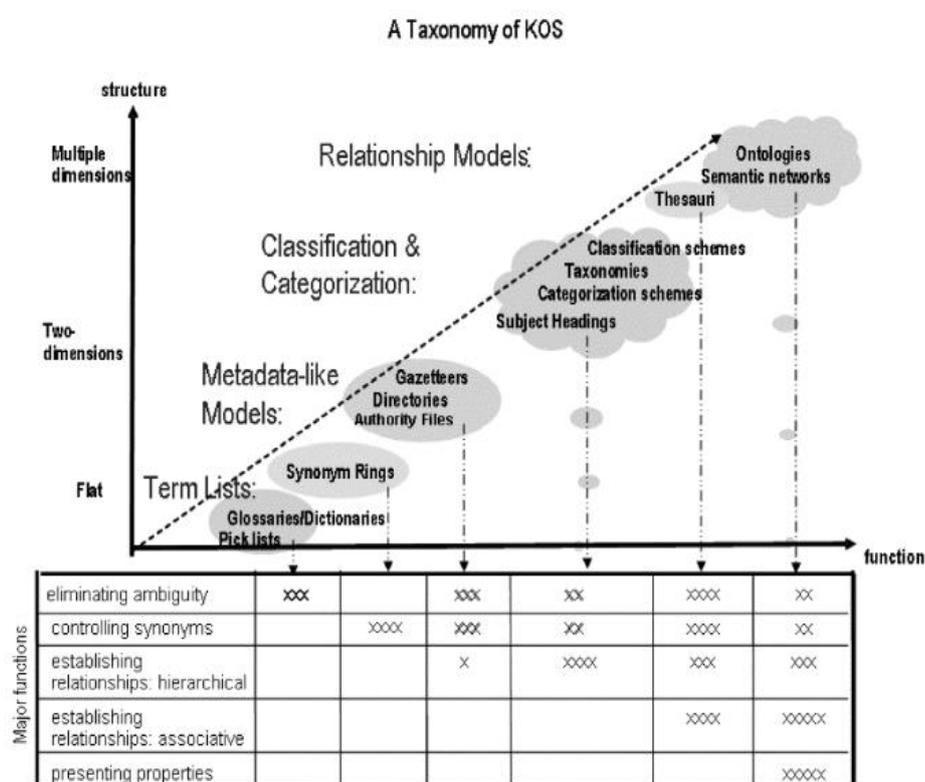


Figure 1. An overview of the structures and functions of KOS

Fonte: Zeng (2008).

A norma ISO 25964-2 (2013) compreende e aborda os seguintes vocabulários controlados: tesouros, listas de cabeçalhos de assunto, lista de autoridades, sistemas de classificação, taxonomias, ontologias, terminologia e anéis de sinônimos.

De acordo com a norma supracitada, os vocabulários são conceituados sinteticamente da seguinte forma:

- Lista/esquema de cabeçalho de assuntos: vocabulário estruturado, no qual são utilizados termos ou frases de modo conciso para descrever conceitos no processo de indexação. Contém regras para a combinação dos termos de forma pré-coordenada. Como exemplo de listas de cabeçalho de assuntos conhecidos apontam-se a *Library of Congress Subject Headings* (LCSH) e *Medical Subject Headings* (MeSH).
- Anéis de sinônimos: conjunto de termos sinônimos ou quase sinônimos listados de várias formas possíveis e são utilizados para se referir a um conceito particular.
- Lista de autoridades: conjunto de nome de entidades, são capazes de unir diferentes formas de se referir a um objeto de forma única e consistente, podendo contemplar pessoas, organizações e lugares, por exemplo.
- Sistemas de Classificação: organizam conceitos por meio de classes ou categorias, muito utilizado na biblioteconomia para prover a organização dos materiais bibliográficos nas prateleiras. Seguem combinações pré-coordenadas e utilizam notações para a classificação, por exemplo o Dewey Decimal Classification (DDC).
- Terminologia: Conjunto de designações que representam conceitos pertencentes a uma linguagem diferenciada, ou seja, podem ser por meio de termos, sinais, denominações ou símbolos.
- Taxonomias: esquema de categorias e subcategorias, usualmente de forma hierárquica, que podem ser usadas para classificação de itens, são comuns na organização da disposição de informações em sites.
- Tesouros: vocabulário controlado estruturado, os conceitos são representados por termos organizados de forma que as relações entre esses conceitos sejam tornadas explícitas e os termos preferidos sejam acompanhados por entradas iniciais para sinônimos ou quase-sinônimos.
- Ontologias: possui diferentes significados de acordo com sua área de atuação, contudo, especificamente no contexto de modelagem de dados, as ontologias referem-se “ao uso de uma linguagem formal para estabelecer uma representação formalizada de um domínio de conhecimento.” Configuram-se em “[...] um conjunto de fatos (afirmações sobre indivíduos) juntos formam uma base de conhecimento” (ISO 25964, 2013, p. 72, tradução nossa).

Independente da nomenclatura e de sua composição estrutural, os vocabulários controlados se diferenciam na sua estrutura e se complementam no objetivo em comum de favorecer o tratamento da informação por meio de ferramentas desenvolvidas para representação da informação

A partir do que foi visto, no item seguinte, serão discutidos os aspectos metodológicos para amparar os objetivos de pesquisa traçados.

3 Procedimentos metodológicos

Essa pesquisa caracteriza-se como exploratória com análise de dados em sites dos catálogos online de bibliotecas universitárias brasileiras. A pesquisa exploratória é definida pela capacidade de investigar novos fenômenos e trazer maior familiaridade com o objeto de

pesquisa. Cervo, Bervian e Silva (2007) define a pesquisa exploratória como ideal para descrever situações e identificar as relações existentes entre elementos. É uma pesquisa quali-quantitativa por apresentar análises diante dos dados estatisticamente relacionados.

Sendo assim, esse trabalho foi desenvolvido a partir da pesquisa iniciada em Cruz (2019) como forma de continuar com questões levantadas pertinentes aos objetivos aqui propostos em: analisar as respostas obtidas na pesquisa de Cruz (2019) sobre a disponibilização dos vocabulários no catálogo online das bibliotecas e verificar a forma em que está sendo apresentado esse recurso ao usuário.

Na pesquisa de Cruz (2019) foram recuperados 46 questionários referentes à linguagem de indexação e o seu uso em bibliotecas universitárias brasileiras. Houve a preocupação de recuperar respostas de todo o Brasil, contemplando todas as regiões brasileiras. O questionário contou com diversas perguntas com o objetivo de contribuir com os estudos “sobre o controle de vocabulário ao realizar estudo analítico sobre o uso de linguagens de indexação no Brasil”. Sendo assim, o questionário apresentou três categorias em sua composição: “Indexação na biblioteca”; “Linguagem de indexação”; “Política de indexação e linguagem de indexação”.

Contudo, nesta pesquisa a questão está centrada na disponibilização da linguagem de indexação, ou seja, dos vocabulários controlados aos usuários no catálogo online da biblioteca.

Observa-se que a região que mais teve retorno dos questionários foi a região Sudeste (19 bibliotecas respondentes) e Sul (13) por apresentarem maior quantidade de universidades. Em seguida o Nordeste (8), Norte (6) e o Centro-oeste (2).

Foi constatado que das 46 respostas reunidas, 39 bibliotecas afirmam utilizar linguagem de indexação para fins de tratamento da informação. Em sequência, houve a questão sobre a disponibilização da linguagem de indexação no catálogo *online* da biblioteca e verificou-se que: 38 respondentes (uma biblioteca não respondeu) utilizam linguagem de indexação, sendo que 18 bibliotecas disponibilizam o recurso no catálogo e 20 não.

Em sua análise, Cruz (2019) relata que são vinte bibliotecas que não disponibilizam a linguagem em seus catálogos. Porém, praticamente a metade das bibliotecas, dezoito delas, afirmam dispor desse recurso. Entende-se com esse resultado que as bibliotecas estão com foco na participação do usuário e em sua capacitação em realizar a busca com maior precisão na escolha dos termos.

Ao disponibilizar o vocabulário controlado ao usuário permitimos a ele maior autonomia de pesquisa, uma vez que o processo de busca se tornará mais eficiente e os resultados mais precisos. Isso só seria possível com a combinação do processo de indexação com vocabulário padronizado pela instituição mais a disponibilidade do recurso para o usuário. Os vocabulários controlados fazem um processo de mediação entre a linguagem do usuário, a linguagem do

sistema e as linguagens consultadas pelo indexador em outros sistemas de bibliotecas. Posto isso, é visto a importância de informar sobre o papel desse instrumento que é utilizado pelo bibliotecário e deve ser parte do processo de busca do usuário. Dessa forma, há um alinhamento no método de representação do conteúdo dos documentos com a forma de recuperação da informação.

A partir desses resultados apresentados, foi realizado uma análise dos dados encontrados a partir da busca no site das bibliotecas que afirmaram disponibilizar a linguagem de indexação/vocabulário controlado na área de busca pela informação de seus sistemas, ou seja, no catálogo online de cada biblioteca.

4 Apresentação e discussão dos resultados

A partir da pesquisa realizada nas 18 bibliotecas que afirmaram disponibilizar o vocabulário controlado nos catálogos online para sua consulta pelo usuário, verificou-se a presença do recurso em 16 catálogos, sendo que em duas delas não foi encontrado.

Na verificação dos catálogos das bibliotecas, nas 16 páginas, encontrou-se vocabulários como:

- Lista de autoridades – 11 bibliotecas
- Tesouro – 4 bibliotecas
- Lista de cabeçalho de assuntos – 1 biblioteca

Todos os catálogos que apresentam as listas de autoridades são dispostos pelo software de gestão de bibliotecas Pergamum, na qual a maioria das bibliotecas estudadas o utilizam. O tesouro, mencionado pelas outras 4 bibliotecas, são da mesma instituição universitária que elaborou uma macroestrutura com ampla cobertura de assuntos para atender a comunidade interna e externa. Nota-se que, neste caso, houve o cuidado em documentar e apresentar o histórico e metodologia utilizada no desenvolvimento da macroestrutura, a equipe de implementação e a atual gestora, além de um tutorial de uso do vocabulário. Ao final, uma biblioteca dispõe no catálogo o cabeçalho de assuntos.

Todas as bibliotecas que apresentam como vocabulário as listas de autoridades em seu catálogo, o exibe da seguinte forma:

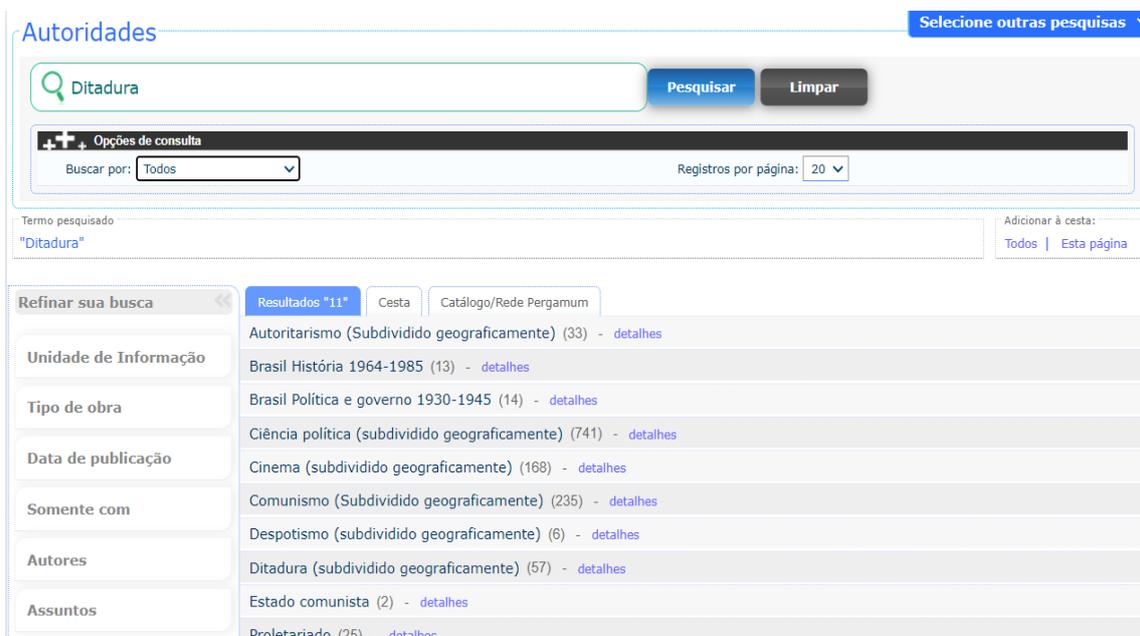
Figura 4 –Exibição das autoridades no catálogo *online*



Fonte: *Software* de gestão de bibliotecas Pergamum.

Os catálogos que apresentam as listas de autoridades as listas de cabeçalhos de assuntos são exibidas por meio de pesquisa, ou seja, a “lista” não é apresentada de forma integral ao usuário, e sim por meio do resultado de pesquisa de acordo com o termo buscado. Como visto na figura 4, há a opção em selecionar outras formas de pesquisa, ao escolher as autoridades, o resultado aparece no seguinte formato:

Figura 5 -Resultado de pesquisa com opção de busca pelas autoridades



Fonte: *Software* de gestão de bibliotecas Pergamum.

Nos resultados obtidos, verificam-se os termos relacionados com a busca realizada, a quantidade de itens disponíveis no catálogo sobre o tema e a especificação da subdivisão geográfica quando presente.

Segundo Llanes Padrón, Fujita e Bastos (2014) os registros de autoridades compreendem um dos pontos de acesso mais importantes para a recuperação da informação, seu controle e normalização dos cabeçalhos permitem a recuperação, o intercâmbio, a difusão e o uso da informação.

No catálogo online que trabalha com a lista de cabeçalho de assuntos, observa-se nos resultados de busca do termo “ditadura”, por exemplo, o número de registros/materiais bibliográficos do assunto e suas variações.

Figura 6 - Resultado de pesquisa com busca por lista de cabeçalhos

Minerva - Busca por Listas de Cabeçalhos

Informar palavra ou expressão: Exibir entradas contendo a palavra ou frase

Selecionar lista para percorrer: ▼

Base para busca: ▼

Percorrer Lista: Todos os Campos

No. Regs.	Entrada
611	ditadura
1	ditadura-
30	ditaduras
1	ditaduta
1	dital
122	ditalia
27	ditalie
3	ditalien
10	ditam
1	ditamaraca

Fonte: Software de gestão de bibliotecas Minerva.

Diante dos resultados obtidos por meio da verificação dos catálogos das bibliotecas, nota-se que o recurso dos vocabulários controlados pode ter ficado fora da visão dos usuários fazendo com que não o encontrem ou não saibam de sua função de contribuição para que a busca seja mais eficiente.

As listas de cabeçalho de assuntos, antes da digitalização dos sistemas, eram feitas manualmente, no papel, hoje pode-se perceber a facilidade de incluir os cabeçalhos automaticamente. Como visto em Zeng (2008) os cabeçalhos de assuntos são sistemas classificatórios e de categorização, favorecem o controle de sinônimos, eliminar ambiguidades e estabelecer relação hierárquica dos conjuntos de termos.

Por meio da observação do catálogo online das quatro bibliotecas que utilizam a macroestrutura, o tesouro da instituição, percebeu-se a importância da disponibilização desse recurso, na qual é possível visualizar todas as ramificações das áreas trabalhadas no tesouro e suas relações. O tutorial de uso do vocabulário também é um agente facilitador ao usuário e ao pesquisador que busca informação.

5 Considerações finais

Os vocabulários controlados são recursos fundamentais na indexação e também podem ser considerados dessa forma para a recuperação da informação pelo usuário. A biblioteca universitária representa um elemento fundamental na vida acadêmica dos usuários, tanto dos discentes como dos docentes, e por meio de medidas coordenadas pode oferecer o recurso dos vocabulários controlados para consulta em seus catálogos online.

Diversas bibliotecas não possuem um vocabulário controlado construído especialmente para a realidade de sua instituição, por isso, deve-se levar em consideração a relevância dos vocabulários desenvolvidos. Há bibliotecas que utilizam vários tipos de sistemas de organização do conhecimento para o tratamento da informação por possuir um acervo diversificado, comum em ambientes universitários, contudo, isso pode dificultar o entendimento do usuário caso seja disponibilizado uma quantidade maior de vocabulários para a realização da busca.

Na pesquisa realizada no catálogo *online* das bibliotecas selecionadas, especialmente nos que utilizam as listas de autoridades e os cabeçalhos de assuntos, percebeu-se que esse recurso fica fora do campo de visão do usuário. Apenas pessoas familiarizadas com essa ferramenta, como bibliotecários, têm maior facilidade de identificar o local no site para acessar os vocabulários. Da mesma forma, considera-se fundamental, além de colocar de forma mais visível os vocabulários controlados, salientar sua função e usabilidade.

Sendo assim, com foco no sucesso da recuperação da informação, os bibliotecários cientes do uso e funcionalidade dos vocabulários controlados podem disponibilizar treinamentos de apresentação e uso do recurso, além de tutoriais com linguagem acessível na página da biblioteca e/ou mesmo no catálogo online, se houver possibilidade.

Referências

AMERICAN NATIONAL STANDARDS INSTITUTE. NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.19** Guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005.179 p.

BARITÉ, M. Sistemas de organización del conocimiento: una tipología actualizada. **Informação & Informação**, Londrina, v. 16, n. 3, p. 122-139, 2011. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/9952> Acesso em: 3 maio 2019.

BORKO, H. Information science: what is it? **American Documentation**, v. 19, n. 1, p. 3-5, 1968.

CERVO, A. L.;BERVIAN, P. A.; DA SILVA, R. **Metodologiacientífica**. 6. ed. São Paulo: Pearson Prentice Hall, 2007.

CRUZ, M. C. A. **Linguagens de indexação em bibliotecas universitárias**: estudo analítico em território nacional. 2019. 82 p. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2019.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v. 2, n. 7, p. 101-107, 1978. Disponível: <http://revista.ibict.br/ciinf/article/view/115/115> Acesso em: 18 ago. 2020.

DODEBEI, V. L. **Tesouro**: linguagem de representação de memória documentária. Niterói: Intertexto, 2002.

FERNEDA, Edberto. **Introdução aos Modelos Computacionais de Recuperação de Informação**. Rio de Janeiro: Ciência Moderna, 2012. v. 1. 184 p.

FUJITA, M. S. L.; TOLARE, J. B. Vocabulários controlados na representação e recuperação da informação em repositórios brasileiros. **Informação & Informação**, v. 24, n. 2, p. 93-125, 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/125804> Acesso em: 10 set. 2021.

GOLUB, K., SCHMIEDE, R. TUDHOPE, D. Recent applications of Knowledge Organization Systems: introduction to a special issue. **International Journal on Digital Libraries**. 20, p. 205-207, 2019. Disponível em: <https://link.springer.com/article/10.1007/s00799-018-0264-8> Acesso em: 10 set. 2021.

HJØRLAND, B. Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science. **Information**, v.3, n. 135, 2021. Disponível em: <https://www.isko.org/cyclo/ir> Acesso em 23 jul. 2021.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. Information and documentation - Thesauri and interoperability with other vocabularies - Part 1: Thesauri for information retrieval. **ISO 25964-1**. Geneva: International Organization for Standardization, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. Information and documentation - Thesauri and interoperability with other vocabularies - Part 2: Interoperability with other vocabularies. **ISO 25964-2**. Geneva: International Organization for Standardization, 2013.

LLANES PADRÓN, D.; FUJITA, M. S. L.; BASTOS, F. M. Os registros de autoridade em sistemas de informação: uma perspectiva biblioteconômica e arquivística. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. **Anais...** Belo Horizonte: UFMG, 2014. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/2606> Acesso em 21 jul. 2021.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22> Acesso em 5 maio, 2020.

SARACEVIC, T. Information Science. **Journal of the American Society for Information Science**, v. 50, n.12, p. 1051-1063, 1999.

ZENG, M. L. Knowledge organization systems (KOS). **Knowledge Organization**, v.35, n.2/3, p.160-82, 2008. Disponível em: https://www.academia.edu/26672820/Knowledge_Organization_Systems_KOS Acesso em: 06 maio 2019.

ZHANG, Y.: OGLETREE, A.; GREENBERG, J. Controlled Vocabularies for Scientific Data: Users and Desired Functionalities. **Proceedings of the Association for Information Science and Technology**, v. 52, n. 1, p. 1-8, Jan. 2015.