

Ensino em organização e representação do conhecimento: linguagens e normas



Cibele Araújo Marques dos Santos
Universidade de São Paulo
cibeleac@usp.br



Eduardo de Abreu de Jesus
Universidade de São Paulo
eduardo.jesus@usp.br



João Ricardo de Luca
Universidade de São
Paulo
joao.luca@usp.br

1 Introdução

As ferramentas utilizadas para organização e representação da informação contida em documentos impressos ou digitais compõem os sistemas de organização do conhecimento, constituídos pelas linguagens documentárias como vocabulários controlados, listas de cabeçalhos de assuntos, sistemas de classificação e tesouros, bem como, os outros tipos de esquemas: as taxonomias e ontologias (CARLAN e BRÄSCHER, 2011). Tais sistemas, especialmente as linguagens documentárias de

acordo com Campos (2006), podem ser definidos pelo controle do vocabulário e pela dinâmica estabelecida de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.

Estas linguagens abrangem aportes teóricos e metodológicos que incluem a linguística, a lógica, a indexação, a terminologia e a construção de tesouros e vocabulários controlados, bem como o uso de sistemas de classificação e listas de cabeçalhos de assunto. A internet e a web semântica permitiram o surgimento de formatos e padrões como esquemas XML, o SKOS (*Simple Knowledge Organization System*) e as ontologias.

Assim, a interdisciplinaridade e complexidade destas ferramentas torna imperativo estudar estes recursos e repensar o ensino das linguagens documentárias e demais sistemas de organização do conhecimento para os profissionais da informação.

Em um projeto desenvolvido para identificar as práticas de ensino e estudar as tendências dos formatos, padrões e normas destes sistemas foi possível observar que a literatura da organização e representação do conhecimento apresenta uma lacuna na pesquisa sobre metodologias de ensino-aprendizagem das linguagens e esquemas de organização da informação.

2 Metodologia

Foi realizada a pesquisa bibliográfica nas bases de dados internacionais da área de Ciência da Informação LISA e ISTA. Foram pesquisadas também as bases multidisciplinares Web of Science e SCOPUS, mas nestas não foram encontrados artigos. Nas bases nacionais foram pesquisadas a BRAPCI, onde foram encontrados apenas 3 artigos sobre ensino de indexação e na SCIELO, que não trouxe resultados. A estratégia de busca utilizou palavras-chaves como “linguagens documentárias”, “classificações bibliográficas”, “tesauro” e “vocabulário controlado” cruzando com “ensino” em português e inglês, nas palavras do título, texto e resumo e na base LISA pesquisou-se utilizando o tesauro com estratégia semelhante. Alguns artigos foram selecionados, principalmente da base LISA para análise.

3 Características dos Sistemas de Organização do Conhecimento

O ensino de ferramentas de organização do conhecimento como as linguagens documentárias demanda contribuições interdisciplinares teóricas e práticas que apresentam certa complexidade, considerando as características dos sistemas e esquemas existentes.

Para Leise (2008), que apresenta artigo bem didático, um vocabulário controlado se diferencia de uma lista de indexação por diversos fatores: tempo, forma e finalidade de elaboração, e as características essenciais de um vocabulário controlado são as relações entre seus termos, como o controle de sinonímia, a estrutura hierárquica e a relação associativa. O autor adota a terminologia utilizada pela norma ANSI/NISO Z39.19 (Vocabulário Controlado) para se referir às linguagens documentárias e como a norma inclui as taxonomias neste escopo (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2010). Também alerta para a diversidade terminológica quanto às linguagens e a dificuldade de definição e conceituação dos termos da área. O artigo define e explica utilizando quadros, além do vocabulário controlado, o anel de sinônimos, o controle de autoridade, a taxonomia e os tesouros. Traz também o conceito de faceta e sua importância para a categorização e estruturação de um vocabulário controlado.

Hedden (2008) também define de forma didática os conceitos de Vocabulário Controlado, como uma lista de termos para a elaboração da indexação e da categorização, sendo controlado porque define o termo a ser utilizado na indexação e sua importância é visível, mais facilmente, quando tratamos da recuperação de informações. Para a autora, o tesouro contempla além do conceito de vocabulário controlado, uma estrutura hierárquica que fornece informações sobre cada termo e seus relacionamentos. Considera a taxonomia como um vocabulário controlado com estrutura hierárquica, mas isenta de relações de equivalência e sinonímia e outros requisitos do tesouro tradicional.

A taxonomia de navegação para Leise (2008) corresponde a um vocabulário controlado para organizar websites ou outros conteúdos digitais, criando-se etiquetas, para categorias de navegação e diferentes níveis de hierarquização, que devem ser

designadas para refletir um contexto particular, como já acontece com os vocabulários, e devem ser usadas em conjunto com sistemas de gestão de conteúdo de websites.

Enquanto o design e a construção de vocabulários controlados têm sido apresentados em nível técnico detalhadamente ao longo das últimas décadas, o nível metodológico tem sido negligenciado, de forma que pesquisadores têm argumentado que a elaboração dessas linguagens precisa se basear nas análises e entendimentos dos contextos em que se inserem (MAI, 2008).

Informações e orientações sobre gestão dos sistemas de organização do conhecimento são importantes, pois a literatura sobre manutenção e revisão de linguagens documentárias e vocabulários controlados é escassa e dispersa, e também são poucos os trabalhos sobre o ensino desta atividade.

A manutenção de um vocabulário controlado, como enfatiza Leise (2008), deve ser constante, considerando a adição/supressão/ revisão de termos, reorganização das hierarquias, modificação, adição e deleção da etiqueta da faceta, atualização, controle de sinônimos, mudanças no processo de controle e definição de propriedade do vocabulário.

Além disso, Leise (2008) destaca que embora existam muitos conceitos e ideias novas a serem aprendidos na área, muitos indexadores consideram a elaboração de um vocabulário controlado como uma extensão de sua atividade de indexação.

A construção de vocabulários controlados está relacionada à indexação, segundo Hedden (2008), no que diz respeito a seus objetivos. Para a autora, as habilidades necessárias para a indexação são as mesmas habilidades necessárias para o desenvolvimento das linguagens e vocabulários, e incluem decidir sobre a melhor palavra para um conceito, criação de termo geral e específico em uma hierarquia, definição de entrada principal e os termos preferidos/não preferidos, referência e relação entre os termos.

A questão que se coloca aqui é: como desenvolver estas habilidades para que os estudantes de graduação e os profissionais graduados possam aprender as competências necessárias para criação e gestão de sistemas de organização do conhecimento. Também existe a necessidade de aprender a selecionar linguagens documentárias quando necessário, compatibilizar linguagens, desenvolver e ou utilizar

softwares e formatos, aprender as normas internacionais para vocabulários controlados e os esquemas para interoperabilidade.

4 Ensino e normas

A literatura pesquisada trouxe artigos sobre a tipologia das linguagens documentárias, suas definições e estrutura, usos, aplicações e normas, mas foram poucos os trabalhos sobre o ensino dessas linguagens.

Para Caro-Castro (2006), nas linguagens documentárias o que importa é o que está codificado lexicamente, não gramaticalmente, razão pela qual se eliminam as formas flexionadas dos termos substantivos escolhidos para descritores, assim estes são unidades das linguagens com função de meta-representação e que não mais apresentam ambiguidade, por terem sido normalizadas. Isto significa que foram controladas as relações entre conceito e descritor. Para isso, são utilizados métodos que estabelecem relações de equivalência e especifica-se o significado dos descritores mediante códigos e notas de alcance. A autora lembra também que a organização de um campo de conhecimento em um tesouro, por exemplo, favorece sua compreensão ao indicar não apenas conceitos representados, mas as inter-relações que os unem.

Esteban-Navarro (1995) reforça, em trabalho apresentado em evento sobre o ensino de linguagens documentárias, que bibliotecários devem ter conhecimentos de princípios para manutenção destas linguagens nas unidades de informação e desenvolvem estas competências em curso com apresentação da metodologia de aplicação dos princípios básicos de terminologia para gestão de tesouros, incluindo fundamentos científicos e exercícios práticos. Considera como aportes teóricos necessários: a terminologia, estudos lexicográficos e terminológicos em área específica, fontes de trabalho terminográfico de localização, identificação e seguimento de novos termos e conceitos mediante busca, seleção e recuperação de documentação técnico-científica representativa, atual e explícita, atribuição de códigos de classificação aos conceitos dos registros terminológicos, gestão de sistemas conceituais pelo uso e gestão de linguagens documentárias, terminologias normalizadas mediante implantação de sistemas de informação científicos, publicação e difusão de documentos secundários e controle da difusão de termos através de estudos bibliométricos.

A presença das linguagens documentárias nos estudos de graduação, segundo Espelt (1995), que apresentou trabalho no mesmo evento, não se limita à definição, características básicas e aplicação prática das linguagens mais utilizadas, mas também devem focar na terminologia e na análise de facetas. Para a autora, é importante que os profissionais estejam familiarizados com as teorias, princípios e diretrizes terminológicas. O aluno deve aprender a ser consciente que sua função se baseia na existência de vários níveis de linguagem e seu trabalho consiste em mover-se entre a terminologia dos documentos, a terminologia do sistema de informação e a terminologia dos usuários.

A normalização dos termos facilita o tratamento dos documentos, lembra Esteban-Navarro (1995), e favorece a progressiva homogeneidade do vocabulário, a recuperação e a difusão da informação, permitindo que autores, profissionais da informação e usuários utilizem os mesmos termos para designar os mesmos conceitos. Também a análise de como as definições dos conceitos e o controle dos sistemas terminológicos e conceituais podem ser aplicados à gestão de linguagens documentárias, já que fornece vocabulários baseados nos princípios de estruturação e relação conceitual, que comunicam a situação precisa de um conceito dentro de um campo temático. A linguagem documentária aparece como normalizadora nos níveis morfológico, sintático e semântico para garantia da estabilidade e eficiência na organização da informação.

O autor expõe a importância do ensino da definição de conceitos terminológicos e de suas regras de elaboração baseadas na norma ISO 704-1987. A aprendizagem das técnicas de construção de definições permite adquirir um recurso fundamental para manter e construir essas linguagens, já que o estabelecimento de relações entre os termos deve ser precedido por fixação restrita da extensão dos conceitos. Para obter univocidade são utilizadas diversas técnicas de desambiguação ou eliminação da sinonímia ou polissemia que de acordo com Espelt (1995), incluem as operações lógicas entre conceitos e as operações aplicadas na definição são instrumentos para determinar sinonímia e polissemia entre termos, para estabelecer especificidade da linguagem e aplicação correta.

Também no controle da relação entre os conceitos, o método de construção dos sistemas de conceitos é a melhor estratégia no entender de Esteban-Navarro (1995) para conhecer o sentido exato do conceito, já que estes se apresentam em relação aos outros

dentro do campo conceitual, e como fundamento básico das estruturas lógico-semânticas das linguagens documentárias que refletem a sistematização do conhecimento existente em um dado campo, segundo seu estado de desenvolvimento e segundo a tendência da taxonomia dominante. Apresenta a análise dos principais critérios de agrupamento estrutural ou de construção das taxonomias para mostrar como as relações genéricas e as ontológicas são as hierarquias, indicando também os diversos sistemas de representação gráfica que podem ser utilizados.

É importante o ensino da aplicação destes tipos de relações em casos concretos, para mostrar os três tipos de sistemas de conceitos documentários: os hierárquicos, os partitivos e os mistos, e critérios de agrupamento de conceitos que utilizam temas, disciplinas ou facetas, bem como as regras para apresentação dos termos nos tesouros de acordo com a normativa internacional (ESTEBAN-NAVARRO, 1995).

Para Espelt (1995), deve-se também considerar que o conhecimento das características dos textos científicos e técnicos faz parte do preparo para análise do conteúdo dos documentos. Além disso, considera conveniente o conhecimento dos sistemas de conceitos e os métodos aplicados na definição de termos para a correta aplicação das linguagens documentárias na indexação e recuperação da informação. A autora lembra que a equipe de elaboração de um tesouro deve ser composta por especialistas da área temática, terminólogos e documentalistas. Do mesmo modo, destaca a importância de desenvolver exercícios com os alunos como análise de textos científico-técnicos, definições de termos com vistas a se tornarem descritores, listas de termos com relações e elaboração de sistemas de conceitos. Igualmente, as classificações facetadas são importantes para a análise do conteúdo dos documentos e também para a qualidade da indexação, e reforça ainda que a literatura nesta área de ensino é pouca e os manuais de indexação são muito gerais sobre isso, de forma que o ensino de linguagens documentárias e de indexação devem ter pautas específicas.

Moreiro-Gonzalez (1995), salienta que, ao mesmo tempo, existe uma quantidade significativa de alunos com pouco preparo em gramática, ou até mesmo com dificuldades para a interpretação do texto que analisam. Por isso, propôs que a aprendizagem ocorra em fases, em um primeiro momento com a introdução de estudos linguísticos preparatórios, o ensino dos sistemas de classificação, das técnicas de indexação e

resumo tendo nesta fase um componente aplicativo predominante. Em uma segunda fase deve ocorrer o aprofundamento do tema, com a inclusão de uma variedade de elementos e conhecimento que permitam transformações representativas no processo.

Clarke e Zeng (2012) mostram que os princípios para a construção de um tesouro mudaram radicalmente da norma ISO 2788 para a ISO 25964, publicada em 2011, o que consideramos que se justifica pelas mudanças ocorridas no mundo com a tecnologia da informação, com a necessidade da interoperabilidade para a Web Semântica, onde a distinção entre termo e conceito é fundamental, contribuição importante implementada na nova norma. Assim, o SKOS (*Simple Knowledge Organization System*), recomendação do W3Consortium foi desenvolvido para a publicação de vocabulários e tesouros na Web.

Em comparação com outras soluções o SKOS oferece uma alternativa cuja aplicação é simples e rápida (PASTOR-SÁNCHEZ; MARTÍNEZ-MÉNDEZ; RODRÍGUEZ-MUÑOZ, 2012). Os autores informam que o SKOS se define formalmente como uma ontologia que permite representar qualquer tipo de sistema de organização do conhecimento mediante RDF. Seu âmbito de aplicação se estende praticamente a totalidade dos vocabulários controlados: classificações, tesouros, cabeçalhos de assunto, taxonomias, glossários, etc. No SKOS, os elementos de um vocabulário se representam mediante conceitos entre os quais se estabelecem relações semânticas, hierárquicas, associativas, de idioma, de sinonímia, de alternativas de acesso a um conceito através de etiquetas.

5 Considerações finais

Esta pesquisa permitiu identificar uma lacuna que existe na literatura, principalmente no Brasil, sobre metodologia e práticas de ensino de linguagens documentárias, a importância do tema para professores e profissionais, bem como a especificidade dos conhecimentos necessários para o desempenho na área, tanto para uso como para elaboração de linguagens documentárias.

O levantamento bibliográfico realizado trouxe trabalhos sobre definição, métodos de elaboração e gestão de linguagens documentárias, e aportes teóricos e práticos para

ensino, formatação e uso dos sistemas de organização do conhecimento. Foram mostradas questões relacionadas às competências que precisam ser ensinadas, a interdisciplinaridade e tendências para desenvolvimento de sistemas e interoperabilidade que precisam ser incluídas nos cursos.

Devido à relevância das questões ligadas a ensino na área, consideramos importante a continuidade deste trabalho, analisando a literatura e experiências realizadas para propor formas de melhoria dos processos de ensino-aprendizagem e desenvolvimento de competências.

Referências

CAMPOS, M. L. A., GOMES, H. E. Metodologia de elaboração de tesauro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, v.11, n.3, 348-359, 2006.

CARLAN, E.; BRÄSCHER, M. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **RICI: Revista Ibero-americana de Ciência da Informação**, v. 4, n. 2, p. 53-73, ago./dez. 2011.

CARO-CASTRO, Carmen. Léxico y documentación: del lenguaje natural al lenguaje documental. In: ROMAN, Maria Teresa Fuentes; REY, Jesus Torres del. *Nuestras palabras: entre el léxico y la traducción*. Frankfurt: Veuver, 2006. p. 127-145.

CLARKE, Stella G. Dextre; ZENG, Marcia Lei. From ISO 2788 to ISO 25964: the evolution of thesaurus standarts towards interoperability and data modeling. **ISQ** n. 1, v. 24, p. 20-26, 2012.

ESPELT, C. Bases teóricas em la enseñanza de lenguajes documentales. I Conferencia ISKO- Espanha, Madrid, 4-5 Novembre, 1993. **Proceedings**. Saragosa, 1995.

ESTEBAN-NAVARRO, M. A. Aplicaciones de la Terminologia para la docência de la gestión de lenguajes documentales. I Conferencia ISKO- Espanha, Madrid, 4-5 Novembre, 1993. **Proceedings**. Saragosa, 1995.

HEDDEN, H. Controlled vocabularies, thesauri and taxonomies. **The Indexer**, v. 26, n. 1, p. 33-34, 2008.

LEISE, F. Controlled vocabularies: an introduction. **The Indexer**, v. 26, n. 3, p. 121-126, 2008.

MAI, Jens-Erik. Actors, domains, and constraints in the design and construction of controlled vocabularies. **Knowledge organization**. v. 35, n. 1, p. 16-29, 2008.

MOREIRO-GONZÁLEZ, José Antonio. La enseñanza de los lenguajes documentales en las diplomaturas de Biblioteconomía y Documentación: problemática y tendencias. I Conferencia ISKO- Espanha, Madrid, 4-5 Noviembre, 1993. **Proceedings**. Saragosa, 1995.

MORGAN, J.; BAWDEN, D. Teaching knowledge organization: educator, employer and professional association perspectives. **Journal of Information Science**, v. 32 p. 108, 2006.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **Guidelines for the construction, format, and management of monolingual controlled vocabularies**, 2010. (ANSI/NISO Z39.19-2005 R2010.)

PASTOR-SÁNCHEZ, J.; MARTÍNEZ-MÉNDEZ, F.; RODRÍGUEZ-MUÑOZ, J. Aplicación de SKOS para la interoperabilidad de vocabulários controlados en el entorno de linked open data. . **El profesional de la información**, v. 21, n. 3, p. 245-253, mayo-junio 2012.

Apoio: Universidade de São Paulo. Programa Ensinar com Pesquisa.