

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

GT-2 – Organização e Representação do Conhecimento

A CONSTRUÇÃO DE TESAUROS NA PERSPECTIVA DOS MANUAIS DE INDEXAÇÃO

Mariângela Spotti Lopes Fujita - Universidade Estadual Paulista (Unesp)

Maria Carolina Andrade e Cruz - Universidade Estadual Paulista (Unesp)

Bruna Otreira Muniz Patrício - Universidade Estadual Paulista (Unesp)

THE CONSTRUCTION OF THESAURI IN THE PERSPECTIVE OF THE INDEXING MANUALS

Modalidade da Apresentação: Comunicação Oral

Resumo: Os sistemas de organização do conhecimento abrangem as linguagens de indexação e, por conseguinte, o tesauro. O tesauro é utilizado como controle de vocabulário para representação documentária, na indexação e na recuperação de informações. Os manuais de indexação são documentos que oficializam e auxiliam o trabalho do indexador. Atuam como ferramentas de organização do conhecimento, buscando padronizar a representação temática dos documentos, através das linguagens de indexação. A literatura normativa sobre gestão e manutenção de tesauros é realizada por organismos internacionais e nacionais, com avanços no controle de vocabulário em contexto digital para a gestão e compartilhamento de conceitos e termos. A norma mais atualizada é a norma ISO 25964, publicada em 2011. Com o objetivo de investigar a construção de tesauros, esta pesquisa efetuou análise de manuais de indexação com critérios fundamentados nas diretrizes da norma ISO 25964, precisamente na seção 13. Para isso, foi feito o exame de quatro manuais de indexação de redes de bibliotecas brasileiras, cujos resultados revelaram que todos atendem à maioria dos critérios. Contudo, as versões desses manuais não trazem informações sobre o uso de *softwares* de gestão, a indicação do uso de normas para construção de tesauros e a atualização de seus respectivos manuais.

Palavras-Chave: Manuais de Indexação; Gestão de Tesauros; Linguagem de Indexação.

Abstract: The Knowledge Organization Systems comprise the subject indexing languages and, therefore, the thesaurus. Thesaurus is used as vocabulary control for the documentary representation in subject indexing and in the information retrieval. The indexing manuals are documents that make official and offer aid to the indexer work. It acts as a tool of knowledge organization, seeking the standardization of the document's subject representation through the indexing languages. The standards guideline literature on thesaurus management and maintenance is build up by national and international organisms with the advances on vocabulary control in digital context for concepts and terms management and sharing. The most up to date guideline is the ISO 25964, published in 2011. Aiming to investigate the thesaurus construction, this research examined indexing manuals using

criteria grounded on the ISO 25964 standards guidelines, precisely on its section 13. Analysis of four indexing manuals of Brazilian libraries' networks where then conducted, whose results revealed that all of them comply with most of the criteria. However, the analyzed versions of those manuals do not offer information on the use of thesaurus management software, the indication of guidelines used for the construction of thesauri and the update of its respective manuals.

Keywords: Indexing Manuals; Thesauri; Thesauri Management; Indexing Language.

1 INTRODUÇÃO

A representação documentária diz respeito aos processos e aos sistemas de organização do conhecimento. Nos processos de organização do conhecimento, a representação é o resultado, de forma que, a partir da indexação de um texto ou de uma imagem, por exemplo, obtém-se uma representação por conceitos. A finalidade desse resultado é servir de referente do conteúdo indexado. Por outro lado, nos sistemas de organização do conhecimento, a representação é a finalidade principal, porque são ferramentas, tais como tesouros, sistemas de classificação, taxionomias, entre outras, utilizadas para realizar a tradução dos conceitos por termos autorizados, que fazem o controle de vocabulário para futuras estratégias de busca em metabuscadores, bases de dados bibliográficas, catálogos *online*, entre outros sistemas de recuperação da informação.

Na perspectiva do ambiente institucional e da concepção do conhecimento registrado, Hjørland (2008) distingue a Organização do Conhecimento em sentido específico, como a organização cognitiva ou intelectual do conhecimento e, em sentido amplo, como a organização social do conhecimento.

No sentido mais específico, a Organização do Conhecimento compreende as atividades de “[...] descrição documentária, indexação e classificação realizadas em bibliotecas, bases de dados bibliográficas, arquivos e outros tipos de ‘instituições da memória’ por bibliotecários, arquivistas, especialistas da informação, especialistas de assunto” (HJØRLAND, 2008, p.86).

Como área de estudo, a Organização do conhecimento está preocupada, segundo Hjørland (2008, p.86), “[...] com a natureza e qualidade dos processos de organização do conhecimento (KOP) bem como dos sistemas de organização do conhecimento (KOS) usados para organizar documentos, representações de documentos, atividades e conceitos.”

No contexto da Organização do Conhecimento, Zeng e Chan (2004, p.377) consideram “[...] sistemas de organização do conhecimento um termo geral que se refere às ferramentas que apresentam a interpretação das estruturas de organização do conhecimento”. Segundo

as autoras, em tempos atuais, coexistem diferentes sistemas de recuperação da informação em ambiente digital: bases de dados bibliográficas, catálogos *online*, bibliotecas digitais, repositórios institucionais, grandes diretórios de assuntos derivados da *web*, entre outros. Existem, portanto, muitas tipologias de sistemas de organização do conhecimento que, conforme sua complexidade se classifica em três grupos (ZENG; CHAN, 2004, p. 377-8):

- Listas de termos: glossários, arquivos de autoridades, dicionários e dicionários geográficos;
- Esquemas de classificação e categorização: esquemas de classificação de bibliotecas, taxonomias e esquemas de categorias;
- Vocabulários relacionais: listas de cabeçalhos de assuntos, tesouros, redes semânticas e ontologias.

Tais sistemas de organização do conhecimento são representações de estruturas de organização do conhecimento em nível macro, cujos conteúdos organizados em vocabulários controlados de termos representam conceitos.

O tesouro é um sistema de organização do conhecimento adotado como ferramenta de representação documentária na indexação e na recuperação de informações cujo uso é difundido mundialmente.

A literatura normativa sobre gestão de tesouros desde sua construção é importante e tem evolução significativa, realizada por organismos internacionais e nacionais, os quais implementam seus marcos regulatórios por meio de normas que demonstram avanços em relação às necessidades de controle de vocabulário em contexto digital, com o uso de tecnologias avançadas para a gestão e compartilhamento de conceitos e termos. A norma mais atualizada é a norma ISO 25964, “Information and documentation: part 1: thesauri and interoperability with other vocabularies”, publicada em 2011.

Com o objetivo de investigar a construção de tesouros por serviços de informação que mantêm base de dados bibliográficos, com indexação de recursos de informação, esta pesquisa analisou manuais de indexação, com critérios fundamentados na seção 13 das diretrizes da norma ISO 25964 (2011) sobre gestão de tesouros, construção e manutenção.

2 LINGUAGEM DE INDEXAÇÃO

Para que possamos discutir a construção de tesouros e sua relação com os manuais de linguagem de indexação, faz-se necessária a compreensão do tema linguagem de indexação.

Vogel (2009) analisa e resgata alguns avanços e mudanças na terminologia utilizada pela área da Ciência da Informação, com ênfase em grupos de autores brasileiros e na linha inglesa, quanto ao conceito de linguagem de indexação e sua aplicabilidade. É possível identificar, em seu trabalho, essa pluralidade de terminologias empregadas, como “[...] linguagem de indexação, indexação de informações, recuperação de informações, controle do vocabulário, tradução de conceitos, uso de tesouros e índices, léxico reduzido, regra de uso” (VOGEL, 2009, p.88).

No mesmo estudo, o autor traz algumas definições da área para linguagem de indexação. Vogel (2009, p.85) define linguagem de indexação como uma linguagem artificial de estrutura hierárquica, com fins de recuperação da informação que atende como sistemas de classificação, linguagens de estruturas alfabéticas ou combinatórias, tesouros e índices. O objetivo da linguagem de indexação é proporcionar o controle do vocabulário, de modo a facilitar a recuperação de informação pelo conteúdo do documento. O controle de vocabulário é feito por meio de padronização de termos, durante a representação e busca da informação no sistema de informação.

A linguagem de indexação é a responsável por realizar a mediação entre o documento e o usuário, capaz de adequar a linguagem usada pelo autor a uma linguagem do sistema, com base na forma como os usuários buscam determinado tipo de informação. Essa importância é sinalizada por Autor A (2010, p. 3):

Nesse contexto, a linguagem de indexação tem grande importância dada a sua função de compatibilização da linguagem utilizada por uma comunidade de usuários e entre várias instituições de modo a servir de instrumento de representação tanto na indexação, por indexadores durante o tratamento temático da informação, quanto na recuperação por usuários durante a estratégia de busca.

É importante diferenciar linguagem natural de linguagem de indexação, tendo em vista que ambas estão constantemente inseridas nos processos de representação. A linguagem natural está presente nas diversas formas da comunicação humana, seja na fala, seja na escrita. Tem como principais características a polissemia e a ambiguidade, fatores indesejados na organização e representação da informação e do conhecimento, porque acarretam erro de recuperação e diminuem as possibilidades de disseminação com precisão e especificidade.

O profissional que executa o trabalho de mediar a informação, aproximando documento e usuário, objetiva executar essa tarefa de forma rápida e precisa, evitando

intercorrências capazes de atrapalhar a localização da informação desejada, quer por desistência, devido à quantidade de esforço exigida, quer por ausência de resultados obtidos na busca. Para tal, ao longo de anos de atuação e estudo, a área desenvolveu práticas e teorias que auxiliam esse trabalho, com o principal intuito de eliminar possíveis dificuldades na recuperação da informação.

Foi verificada por Autor A (2016, p. 15) a necessidade de ampliar os estudos na área acerca da representação de documentos, visando à recuperação por assunto, tendo em vista a evolução tecnológica e o aumento de demandas por assuntos com altos níveis de especificidade. A fim de que isso ocorra, é necessário abranger as especificidades que o grupo de usuários da unidade necessita. Não se trata de uma representação dos acervos, porém, de uma forma de aproximar o documento do usuário, trazendo as palavras do autor de maneira padronizada aos usuários.

Segundo a norma norte americana ANSI/NISO Z39.19, a linguagem de indexação é

[...] um vocabulário controlado ou um sistema de classificação com regras para sua aplicação [...] é utilizado para a representação dos conceitos tratados nos documentos e para a recuperação de tais documentos de um armazenamento de informações e de sistema de recuperação (AMERICAN NATIONAL STANDARDS INSTITUTE/NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005, p. 11, tradução nossa).

A linguagem de indexação é responsável por garantir que os termos escolhidos pelos indexadores sejam precisos, de tal sorte que levem em si não somente as ideias que o documento apresenta, mas também a filosofia e a padronização da instituição à qual a linguagem está associada, de modo a permitir que os usuários se satisfaçam e que seja possível o intercâmbio de informações entre bases de dados.

A linguagem de indexação possui algumas tipologias em níveis de complexibilidade e tipos de estrutura distintos, as quais são formas de organizar as informações estudadas no processo de representação de assuntos nas bibliotecas. Essas ferramentas sintetizam as tomadas de decisão já planejadas ou, ainda, atuam como um registro do trabalho executado, objetivando melhor aproveitamento de tempo na descrição. Caso a definição de assunto fosse somente pensada e iniciada no momento em que se recebe o documento para processamento, teríamos uma espera muito grande para disponibilização do documento para uso, além de falhas em padronização, correndo-se o risco de cada processamento ocorrer de uma maneira diferente.

De acordo com Autor A (2016), é importante que as instituições, ao escolherem uma linguagem de indexação, analisem a estrutura a partir das relações semânticas e sintáticas, de modo a garantir a qualidade dos instrumentos usados na indexação, além de permitir que a instituição tenha a visão de qual linguagem é mais apropriada à sua realidade.

Quanto à estrutura, a LI apresenta vocabulário e sintaxe, sendo estes responsáveis pela relação entre termos, a fim de identificar o conteúdo temático dos documentos, e pelas regras de combinação entre termos, com a finalidade de representar o conteúdo temático do documento (PINTO, 1985).

Já as relações semânticas constituem a forma de apresentação dos conceitos, podendo ser divididas em três tipos, cada qual possuindo um método diferente de associar os termos entre si: hierárquicas (relação de termos mais amplos ou mais específicos dentro de determinada classe), associativas (termos de indexação que podem ampliar as possibilidades de busca) e de equivalência (relação de termos substitutivos de indexação e busca). (PINTO, 1985, p.173).

Há também as relações sintáticas que, segundo Pinto (1985, p. 177), são originadas a partir da necessidade de recuperação de dois ou mais conceitos, os quais, juntos, correspondem a um assunto composto. Ademais, para a autora, essa relação sintática, por ser dependente do documento que está sendo analisado, é temporária, diferentemente das relações semânticas, que tendem a ser perpétuas.

Como já frisado anteriormente, a linguagem de indexação é uma linguagem controlada, com fins de representação e recuperação da informação. De acordo com a norma ANSI/NISO Z39.19 (AMERICAN NATIONAL STANDARDS INSTITUTE/NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005), os vocabulários controlados são listas, anéis de sinônimos, taxonomias e tesouros, este último o mais complexo dos citados.

No entanto, não é a intenção deste estudo se estender mais sobre essas tipologias: tem-se apenas um panorama, a fim de que se possa compreender o uso e a construção dos tesouros nas instituições, já que se trata de uma linguagem de indexação.

2.1 Tesouros

Uma ferramenta com o intuito de facilitar processos de recuperação da informação, o tesouro, passou a ter seu uso difundido desde avanços científicos tecnológicos, em meados do século XX. Juntamente com sua difusão, tornou possíveis estudos para seu

aperfeiçoamento, ampliando suas aplicações e métodos de elaboração, com aprofundamento de pesquisas em diversas áreas do saber. Entretanto, os tesouros foram aprimorados para suprir as exigências dos antigos e dos novos campos e subcampos da ciência, já que estes necessitavam de ferramentas de organização da informação mais elaboradas, específicas e flexíveis.

Moreira e Moura (2006, p.2) contextualizam o advento do tesouro como uma “[...] resposta a essa necessidade de recuperação de informação”, porque este tem como principal característica estabelecer relações entre os termos que o compõem. Os autores ainda apontam que os tesouros evoluíram tanto em estrutura quanto nas referências cruzadas, “[...] dando lugar às relações hierárquicas (verticais) e associativas (horizontais). São as relações existentes entre os termos que conferem a um tesouro uma multiplicidade de usos, desde o auxílio na própria função de indexação até o auxílio na efetiva recuperação dos documentos.” (MOREIRA, MOURA, 2006, p.3).

O tesouro é uma linguagem de indexação de vocabulário controlado e pós-coordenado, o que oferece a essa ferramenta maior flexibilidade para atender às necessidades dos usuários, ao desenvolverem suas buscas. Pinto (1985, p.185) aponta como características do tesouro, além das já citadas, a ênfase no uso do sistema pós-coordenado com apenas o mínimo de pré-coordenação para padronização e não comprometimento da significação dos descritores.

Além disso, o tesouro proporciona a possibilidade de estabelecer relações entre os termos empregados, sejam eles os escolhidos para a representação (preferidos), sejam outros termos que também possam compreender o tema (não preferidos). Isso ajuda o usuário a conhecer melhor o sistema de busca que utiliza, tendo mais possibilidades de recuperação pertinente à sua necessidade, podendo dar luz à instituição sobre a terminologia que os usuários estão usando.

Cervantes (2009, p.35) acredita que o tesouro permite uma visão sistêmica da área, sendo, portanto, um objeto de representação. A autora salienta que o fato de permitir ao usuário combinações de termos, isto é, além de já trazer uma estrutura pré-coordenada dos termos, possibilita que o usuário pós-coordene sua busca de acordo com suas necessidades, dando maior precisão à recuperação e mais dinamicidade aos processos de busca.

Convém ressaltar a constituição de um tesouro e as atenções necessárias ao desenvolvê-lo, como, por exemplo, uma estrutura lexical hierarquizada, onde os termos são

dispostos segundo uma lógica hierárquica. Conforme Cervantes (2004, p.36), contribuem para o exercício da representação, além de ter como apoio o conhecimento categorizado em assuntos e em um *corpus* descritivo (de onde são compilados os termos significativos), para fins de uma boa elaboração do tesauro.

Cintra et al. (1994), ao abordar sobre a estrutura da linguagem de indexação, no caso o tesauro, caracterizam seus três tipos de relações entre termos: hierárquica, associativa e de equivalência:

[...] são estruturadas de maneira lógico-semântica. O conjunto nocional básico é apresentado em hierarquias (na vertical), em torno das quais se agregam as unidades informacionais que se relacionam horizontalmente, relações não-hierárquicas convencionalmente denominadas associativas. Nenhuma unidade pode figurar numa LD sem que esteja ligada a outra. [...] [Já as] relações de equivalência estabelecem as remissivas, definindo o conjunto dos não-termos ou não-descritores e dos termos ou descritores. A finalidade dessas remissivas é encaminhar o usuário para os termos preferidos pelo sistema. Constitui-se, desse modo, a chave de acesso ao sistema (CINTRA et al., 1994, p. 31-33).

Na estrutura hierárquica, portanto, os termos são dispostos em classes, sendo possível aprofundar-se para termos mais específicos ou mais gerais, como, por exemplo, numa classe de “Animais”, é possível ter “Aves”, assim como “Tucano” também é um tipo de ave e estará como termo específico deste. Já a relação associativa figura-se no sentido de termos que se localizam no mesmo sentido, ou seja, são subordinados e superordenados em uma mesma classe, mas não representam a mesma ideia; é o caso, por exemplo, de “Tucano” e “Urubu”, ambos “Aves” ocupando o mesmo nível hierárquico, porém, diferentes entre si.

Além dos exemplos acima, há a relação de equivalência, na qual os termos se relacionam no mesmo nível hierárquico, contudo, desta vez, podem passar a mesma ideia. Esse tipo de relação propicia ao usuário ter ciência dos termos que pertencem à mesma temática e identificar qual foi o utilizado na indexação, isto é, o termo preferido, e qual pode ter o mesmo sentido, mas não foi usado na representação, termo não preferido. Nesses casos, o tesauro apresenta o “use”, para que o usuário saiba qual termo empregar para localizar determinada temática.

Cintra et al. (2002, p.46) destacam as relações de um tesauro, a partir do conceito de Gomes (1990), para quem a estrutura do tesauro é pensada como um conjunto de relações que são essenciais para o funcionamento da ferramenta, pois possibilita “[...] ao usuário, (indexador ou consultante) encontrar o(s) termo(s) mais adequado(s), mesmo sem saber, de

início, o nome específico para representar a ideia ou o conceito que ele procura.” Ademais, partindo de termos já conhecidos, graças à estrutura do tesouro, é possível ter ciência de diversos outros que circundam a temática do primeiro, muitas vezes até mais apropriados como termos de busca. As autoras ainda sinalizam a possibilidade de combinação de termos com o uso de conjuntos de operadores sintáticos (operadores booleanos, no caso de tesouros), para a construção de novas relações e outras possibilidades de recuperação.

Portanto, podemos compreender que o tesouro é um instrumento de relações hierárquicas, associativas e equivalentes, organizado com base na categorização do conhecimento e levando em conta diversos pontos para a escolha dos termos a compor o tesouro, como a filosofia da instituição, os usuários (garantia do usuário) e o que o documento apresenta em sua composição (garantia literária).

Com a conceituação básica sobre essa ferramenta em mente, é possível discutirmos as normas que regem sua elaboração e manutenção, assim como, no decorrer deste trabalho, examinar os manuais que abordam esses processos.

2.2 Construção e manutenção de tesouros regidos por normativas

O tesouro é uma das ferramentas mais adequadas para as áreas especializadas, viabilizando a representação e a recuperação da informação, pois emprega controle dos termos que compõem o sistema após a indexação de assuntos (AUTOR A, 2012).

Segundo Moreira e Moura (2006), a elaboração de um tesouro demanda flexibilidade na incorporação de mudanças necessárias para o desenvolvimento da linguagem. Tendo em vista o uso predominante da ferramenta dos tesouros, por sua capacidade de ajustar-se às necessidades de gerenciamento e organização nos sistemas de informação, desenvolveram-se, com o decorrer de evoluções no campo em Ciência da Informação, normas e diretrizes para a viabilização de construção e manutenção de tesouros.

Autores A e B (2012, p. 55) abordam os estudos de Lancaster (1986, 1987, 1995), observando-se, ao longo da evolução das normas sobre tesouros, duas abordagens: abordagem alfabética de assuntos da América do Norte e abordagem norteadada pela classificação bibliográfica da Europa. De acordo com Campos e Gomes (2006), a abordagem norte-americana partiu de um sistema de cabeçalhos de assuntos para o “unitermo”, ou seja, de sistemas pré-coordenados para sistemas pós-coordenados. Já a influência europeia contribuiu em abordagens mais sistêmicas, utilizando os princípios da classificação, como as

promovidas pelos estudos de Ranganathan. Desde 1970, percebe-se que ambas as influências passam a convergir e a buscar, uma na outra, aprimoramentos no que se refere à representação dos documentos (CAMPOS; GOMES, 2006; CERVANTES; FUJITA, 2012).

Em 1970, a UNESCO pública *Guidelines for the Establishment and Development of Monolingual Scientific and Technical Thesauri*. A norma americana ANSI Z39.19, *American National Standards Institute*, contempla diretrizes para a construção de tesouros, no ano de 1974.

De acordo com Autores A e B (2012, p. 124), as diretrizes da UNESCO serviram como base para a primeira norma internacional de construção de tesouros, a *International Organization for Standardization (ISO 2788)*, datada em 1974, “[...] assim como para as normas francesas da *Association Française de Normalization (AFNOR Z47-100)*, em 1973, e a alemã *Deutsches Institut für Normung (DIN 1463)*, de 1976.”

Nas normas internacionais atuais, encontram-se: Z39:10 - *Guidelines for the construction, format, and management of monolingual controlled vocabularies*, elaborada pelo *American National Standards Institute (ANSI)*, da *National Information Standards Organization (NISO)*, publicada em 2005, e a norma ISO 25964, que possui duas partes: “*Information and documentation: part 1: thesauri and interoperability with other vocabularies*”, publicada em 2011, e a “*Information and documentation – thesauri and interoperability with other vocabularies - part 2: Interoperability with other vocabularies*”, de 2013.

O propósito deste trabalho concentra-se em manuais de indexação que abordam a construção de linguagens de indexação, especificamente tesouros. Assim, a primeira parte da norma, a ISO 25964-1 (2011), reúne elementos em desenvolvimento e manutenção de tesouros na seção 13 de sua composição. O título da seção é “Managing thesaurus construction and maintenance” (ISO 25964-1, 2011) e exhibe os seguintes tópicos:

Planning a thesaurus

Na etapa inicial, são apresentadas as seguintes diretrizes: determinar objetivos; determinar as características do tesouro; fazer balanço geral dos recursos; estabelecer responsabilidades e escolher o *software* de gestão do tesouro.

O planejamento é fundamental, uma vez que os interesses e objetivos sejam reunidos e aplicados na elaboração do tesouro. A norma elenca os seguintes pontos a serem considerados no planejamento: “a) O tesouro deve ser usado para quê, e por quem; b) Ele

será limitado pelas restrições do software existente com o qual deverá ser usado; c) Quão conhecedores serão os usuários da área de assunto do tesouro e de uso do tesouro.” (ISO 25964-1, 2011).

Considera-se que, na condição desses aspectos evidenciados pela norma, é necessário reconhecê-los por meio de estudo prévio e defini-los. Também é apontado, na norma, que as pessoas envolvidas no projeto de elaboração do tesouro serão guiadas por esses objetivos, tornando-se a tarefa mais bem definida a todos, para que o trabalho seja realizado de modo mais eficaz.

Ressalta-se que essas orientações são elaboradas para se seguir linearmente. Assim, o próximo passo é a determinação das características do tesouro. Para isso, a norma estabelece as seguintes características a serem delimitadas:

- a) Se o tesouro será usado em formato impresso ou eletrônico, ou ambos;
- b) O modelo de apresentação requerido;
- c) Se formatos especiais são necessários para integração com outros sistemas, tais como sistemas de indexação ou busca;
- d) Que formatos são previstos para atualizações e quão frequentemente serão requeridas (ISO 25964-1, 2011).

Essas definições são essenciais, tendo-se em vista o formato de apresentação do tesouro que melhor se ajuste aos objetivos iniciais propostos. Incluem-se as decisões sobre as características opcionais, como armazenamento de informações sobre definições de termos, relações customizadas, presença dos rótulos de nó. Outras características a serem mais detalhadas, por exemplo, seriam o formato dos termos utilizados no plural ou singular, grafia, maiusculização, extensão máxima do termo, caracteres especiais, se necessários, e se haverá número limite de níveis da hierarquia (ISO 25964-1, 2011).

Para o balanço de recursos do tesouro, a norma indica que “[...] os recursos chave, tais como as pessoas, o financiamento, ferramentas de software e recursos de vocabulário devem ser determinadas.” (ISO 25964-1, 2011, p. 89). Devem-se descrever as finanças voltadas aos recursos humanos e de vocabulário.

No estabelecimento das responsabilidades, o responsável pelo projeto tem a função de delegar as responsabilidades do grupo participante do projeto, delimitando as atividades a serem realizadas por cada pessoa.

A escolha de *software* de gestão de tesouro é uma decisão importante, pois é através do suporte oferecido pelo *software* que a eficiência do serviço se manifestará. Por isso, é

necessária uma avaliação dos *softwares* presentes no mercado, a fim de verificar qual se enquadra melhor nos objetivos e orçamento do projeto.

Early stages of compilation

Nos estágios iniciais da compilação, são reunidas as questões: por onde e como se deve começar, a coleta dos termos e conceitos, e a análise dos termos.

Para responder à questão de como e onde começar a utilizar o tesauro, a norma relata que o ideal é que o tesauro esteja concluído antes de a compilação da base de dados começarem. No entanto, não é fácil saber os termos a serem incluídos no tesauro, sem a base de dados estar construída. Por isso, recomenda que uma parte substancial da base de dados tenha sido elaborada para que os termos sejam inseridos no tesauro, para evitar uma reindexação (ISO 25964-1, 2011, p. 89).

Nas orientações a respeito da coleta de termos e conceitos, a norma informa que, no estudo dos termos, identificam-se as indicações feitas pelos usuários, em que podem ser encontradas nas listas de referências registradas no serviço de referência. A norma enfatiza que, nesse estágio inicial, a tarefa primordial é coletar os termos, observar a fonte de cada um e a frequência de ocorrência (ISO 25964-1, 2011, p. 90).

Na etapa de análise, os termos coletados devem ser organizados sistematicamente, a fim de estudá-los para que posteriormente sejam inseridos no tesauro. É necessário verificar se o *software* de gestão de tesouros tem as funcionalidades próprias para a inserção dos termos (identificando variantes e sinônimos, por exemplo); caso não tenha essas características, os termos podem ser dispostos em uma planilha, organizados em áreas de assuntos ou facetas, e que sejam incluídas sua frequência e fonte.

Construction

Para a construção do tesauro, os termos devem estar organizados em grupos de assuntos/facetos ou hierarquias (ISO 25964-1, 2011). Devem ser estabelecidos os grupos de sinônimos e quase sinônimos. É necessário cuidado com a duplicação de termos, sobreposições ou omissões, reparando-se também no grau de especificidade estipulado. Se forem requeridas, as notas de escopo podem contemplar as fontes e as definições dos termos.

Na sequência da construção, a norma aponta que é preciso definir as estratégias quanto à estrutura hierárquica e ao nível de especificidade conjuntamente, para que sejam evitadas as anomalias na estrutura. A norma retrata a eficiência dos termos agrupados hierarquicamente, em conjunto com as relações hierárquicas e de equivalência, porém,

acrescenta recomendações de que as relações associativas devem ser incluídas no estágio final. Enfatiza-se: “Seja qual for a sequência seguida, o controle de qualidade deve ser empreendido para cobrir aspectos tais como completude, ausência de duplicações, precisão ortográfica, escolha correta de relações etc.” (ISO 25964-1, 2011, p. 91, tradução nossa).

Introduction to the thesaurus

Ressalta-se a necessidade e a importância de se registrar as seguintes informações:

- a) A(s) área(s) de assunto(s) coberta(s), com identificação de áreas central e marginais;
- b) Os idiomas nos quais o tesouro é apresentado;
- c) Confirmação de qual norma nacional e/ou internacional foi seguida;
- d) O significado de todas as convenções, abreviaturas e sinais de pontuação usados de formas não normatizadas;
- e) O número total de termos, com totais separados de termos preferidos e não-preferidos;
- f) Quaisquer regras que regem a seleção de conceitos, incluindo conceitos complexos;
- g) As regras adotadas na seleção de formas dos termos preferidos, incluindo uma referência a qualquer manual de estilo que seja seguido e as regras para decidir suas inter-relações;
- h) As regras de registro empregadas, seguindo (e citando) uma norma nacional ou internacional apropriada quando possível;
- i) Uma declaração sobre a atualização da política, incluindo a frequência, datas e procedimentos, e o nome e endereço da agência responsável a quem os comentários e sugestões devem ser enviados;
- j) Referências às, e reconhecimento de, fontes usadas na compilação e revisão do tesouro (ISO 25964-1, 2011, p. 93).

A norma indica a introdução do tesouro para contemplar tais informações.

Dissemination

A etapa de disseminação aborda: a) integração com um sistema eletrônico, na qual é preciso considerar as necessidades de uso do tesouro, indexação ou pesquisa. A norma recomenda que o tesouro “[...] esteja totalmente integrado com as funções de indexação, navegação e pesquisa” (ISO 25964-1, 2011, p. 93). b) A publicação eletrônica engloba as melhores maneiras de apresentar o tesouro em formato digital. c) No que diz respeito à distribuição limitada de cópia impressa, a norma indica ponderar tal circunstância. d) A publicação convencional possui a vantagem de alcançar um público que não possui acesso a computadores. e) A norma aponta que é preciso depositar uma cópia ou edição futura em

uma biblioteca nacional depositária apropriada. f) Diretórios do *website*, como a Iniciativa Dublin Core de Metadados (Dublin Core Metadada Initiative - DCMI).

A manutenção do tesouro deve ser feita no decorrer do seu período de vida, desde a sua publicação. É importante ter consciência de que, ao inserir termos novos em uma situação de emergência, eles terão impacto na recuperação, pois são disponibilizados para pesquisa. A norma recomenda os dois procedimentos a seguir: os procedimentos de sugestão que implicam um mecanismo de propor mudanças ao tesouro, tanto para usuários quanto para indexadores, e os procedimentos de revisão, os quais devem examinar as sugestões e termos recebidos regularmente.

3 METODOLOGIA

Com o objetivo de verificar comparativamente os aspectos gerenciais da construção de tesouros por serviços de informação que mantêm sistemas de busca e recuperação da informação, com o uso de linguagens de indexação, serão analisados seus manuais de indexação.

Conforme Lancaster (2004), a indexação é um procedimento subjetivo e não objetivo, porque é um processo sociocognitivo, pelo qual, no momento de representação do documento, cada indexador pode realizar tal tarefa de uma maneira diferente. Para isso, procura-se atingir uma coerência nas atividades executadas na indexação. Entre os fatores que influem no aspecto de coerência da indexação estão os instrumentos de auxílio com que conta o indexador (LANCASTER, 2004).

Os manuais de indexação são documentos que oficializam a tomada de decisões dos indexadores. O manual de indexação deve ser elaborado em ordem lógica, contendo as etapas da indexação a serem seguidas, desde a análise de assuntos até a tradução dos conceitos por meio da linguagem de indexação, contemplando regras, diretrizes e elementos da política de indexação (AUTOR A, 2009; RUBI, 2012).

Segundo Narukawa e Sales (2012), é através dos manuais que os novos indexadores são capacitados. Um manual de indexação deve ser uma condensação de três tipos:

1. Manual de operação ou procedimentos: pois ele deve descrever a atividade de indexação, dar instruções sobre a realização dessa tarefa e proporcionar métodos que possibilitem sua execução de maneira uniforme;

2. Manual de política: o manual deve apontar diretrizes sobre as políticas a serem seguidas pelos indexadores no momento da indexação, coordenando, assim, esforços de todas as bibliotecas cooperantes;

3. Manual de organização: deve servir como um repositório das experiências acumuladas dos indexadores mais antigos, a serem aproveitadas para facilitar a aprendizagem pelos mais novos, podendo, com isso, constituir-se num manual de consultas (RUBI, 2012, p. 174).

Os manuais de indexação, de modo geral, são textos descritivos organizados conforme uma lógica operacional, norteados por condutas que obedecem à normalização vigente sobre indexação e análise de assunto, política de indexação para a recuperação da informação e a linguagem de indexação adotada como ferramenta de representação documentária. A finalidade maior de um manual de indexação é atingir metas de gestão documental por uma equipe de profissionais que realizam a indexação e utilizam a linguagem de indexação como ferramenta de representação documentária entre o sistema e o usuário. Desse modo, os manuais de indexação são importantes ferramentas de gestão para a sistematização de condutas operacionais e metodológicas que visam à uniformização de tarefas, em uma equipe de profissionais de um ou mais serviços de informação.

Enfatiza Rubi (2003, p.70):

O manual de indexação de um sistema de informação constitui também sua documentação oficial, pelo fato de estar descrito em ordem lógica de etapas a serem seguidas para a análise de assuntos, por fornecer as regras, diretrizes e procedimentos para o trabalho do indexador e por conter os elementos constituintes da política de indexação adotada pelo sistema. Dessa maneira, é por meio dos manuais, principalmente, que a política de indexação do sistema poderá ser observada.

Da mesma forma, podemos nos referir à observação da linguagem de indexação, cujas orientações de gestão e uso estão inseridas nos manuais de indexação, porque são empregadas no processo de indexação, na etapa de tradução dos conceitos identificados e selecionados na primeira etapa de análise.

Os manuais selecionados para a análise são quatro: “Diretrizes para a construção do tesouro da Rede Virtual de Bibliotecas - Congresso Nacional – RVBI” (2007), “Guia de atualização dos Descritores em Ciência da Saúde – DeCS (2007), “Manual de indexação de assuntos com o uso de vocabulário controlado da USP: versão preliminar” (2006) e o “Manual de política de indexação para bibliotecas universitárias da UNESP” (2017).

Foram escolhidos a partir de uma busca na *web* por manuais de indexação. Inicialmente, foram coletados sete manuais, todavia, com base em uma análise preliminar,

manteve-se a prioridade aos manuais de indexação mais completos em relação aos elementos da norma ISO 25964-1 (2011), como será visto abaixo.

Todos os manuais apresentam seus respectivos contextos históricos de ocorrências que culminaram na elaboração de suas linguagens de indexação. As “Diretrizes para a construção do tesouro da Rede Virtual de Bibliotecas - Congresso Nacional – RVBI” (2007) foram desenvolvidas pelo “Grupo de estudo do tesouro da RVBI”, buscando reunir os critérios adotados na construção do tesouro da RVBI. O “Guia de atualização do DeCs” (2007) foi criado pelo Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde, também conhecido por BIREME, construindo o vocabulário estruturado e trilingue DeCS, a partir do MeSH - *Medical Subject Headings* da *U.S. National Library of Medicine*. Já o “Manual de indexação de assuntos com o uso de vocabulário controlado da USP: versão preliminar” (2006) foi elaborado pelo Grupo de Gerenciamento do Vocabulário Controlado da USP, procurando atender ao Sistema Integrado de Bibliotecas da Universidade de São Paulo (SIBi/USP). Por fim, o “Manual de política de indexação para bibliotecas universitárias da UNESP” (2017) foi preparado pelo Grupo de Trabalho em Política de Indexação da Rede de Bibliotecas UNESP e, além de conter os elementos de política de indexação da rede, o documento também aborda a construção da Linguagem UNESP.

A metodologia de análise para observação da gestão e uso de linguagens de indexação em manuais de indexação segue a estrutura de aspectos selecionados da seção 13 da ISO 25964-1 (2011), “Managing thesaurus construction and maintenance”:

- Planejamento do tesouro: objetivos, características, recursos, responsabilidades, *software* de gestão;
- Estágio inicial de compilação e análise dos termos;
- Construção: estratégia de elaboração das estruturas hierárquicas e determinação de nível de especificidade, relações de equivalência e associativa entre termos;
- Introdução do tesouro: informações sobre a configuração do tesouro;
- Disseminação: integração com a indexação, navegação e pesquisa, formato digital e impresso, diretórios do *website*;
- Atualização: procedimentos de sugestão e de revisão.

4 ANÁLISE DOS MANUAIS DE INDEXAÇÃO

Os manuais de indexação serão analisados na perspectiva dos aspectos apontados na norma ISO 25964-1 (2011), no quadro abaixo, sendo, em seguida, detalhadamente descritos, em ordem de acordo com cada manual.

Quadro 1: Planejamento do tesouro.

Aspectos	Rvbi	DeCs	UNESP	VocaUSP
Planejamento do tesouro				
Objetivos	X	X	X	
Características	X	X	X	
Recursos				
Responsabilidades	X	X	X	X
<i>Softwares</i> de Gestão				
ESTÁGIO INICIAL DE COMPILAÇÃO				
Coleta dos termos	X	X	X	
Análise dos termos	X	X	X	
Construção				
Estratégia de elaboração das estruturas hierárquicas	X	X		X
Nível de especificidade	X	X	X	X
Relações de equivalência	X	X		X
Relações associativas	X	X		X
Introdução do tesouro				
Configuração do tesouro	X	X	X	
Disseminação				
Integração com a indexação	X	X	X	X
Navegação e pesquisa		X		X
Formato digital	X	X	X	X
Formato impresso				
Diretórios do <i>website</i>		X		
Atualização				
Procedimentos de sugestão	X	X	X	X

Procedimentos de revisão	X	X	X	
--------------------------	---	---	---	--

Fonte: Elaborado pelas autoras - 2017.

O manual de “Diretrizes para a construção do tesouro da RVBI” (2007) apontou como *objetivo* reformular e atualizar o vocabulário controlado de base da RVBI, com a aplicação de normas para a elaboração de um tesouro multidisciplinar monolíngue. O manual indica a natureza do tesouro, nos objetivos, incluindo duas áreas de abrangência do tesouro: áreas de núcleos e periféricas.¹ Verificam-se as *características* referentes às padronizações dos termos, como o que deve ser usado e o que não deve. As *responsabilidades* são do grupo de estudo formado pelos bibliotecários da rede, os quais se reúnem periodicamente.

Em estágio inicial de *compilação*, a Rede Virtual de Bibliotecas utilizou o Vocabulário Controlado de Base (VCB), que antes era a linguagem de indexação empregada para a representação de assuntos, para começar a construir o tesouro. As *análises de termos* feitas pelo grupo de estudos se concentraram nos termos do antigo vocabulário, examinando sua frequência a partir das buscas efetuadas na base, além de estudos dos termos sugeridos pelos indexadores.

Quanto à *construção* do tesouro, um dos problemas identificados no VCB era a falta de hierarquização do vocabulário. A *estrutura hierárquica* partiu da instalação de áreas de núcleos e áreas periféricas, a fim de que pudesse haver relações de subordinação entre conceitos. Para alcançar o nível de *especificidade*, é usada uma combinação de descritores, os quais são chamados de especificadores. O *termo equivalente* não pode ser utilizado para a indexação, por isso, é igualmente denominado não-descritor e indicado em USE/NÃO USE. As *relações associativas* relacionaram os termos de categorias diferentes, sem estabelecer relações hierárquicas entre eles.

No manual, estão presentes os elementos de *introdução* ao tesouro, na qual são indicados as áreas de assuntos e o idioma, língua portuguesa, normalizado com o uso do Dicionário Houaiss da Língua Portuguesa e, também, são estipuladas questões de abreviaturas e sinais de pontuação. Nenhuma norma é apontada para a construção da linguagem, assim como não há indicação atual do número total de termos e nem a definição de uma data para atualizações.

¹ Na norma ISO 25964-1 (2011, p. 93, tradução nossa), são chamadas de áreas centrais e marginais.

É notável, no manual de indexação, a *integração com a indexação*, ao observar a preocupação com as relações entre os termos e as indicações do que não deve ser usado na indexação de assuntos. Não há informações a respeito de *navegação e pesquisa*. Entende-se que o *formato* é apresentado na versão *digital*. Há uma *versão impressa*, porém, possui uma nota de que ainda está sendo arquivada.

Os aspectos de *atualização e manutenção* serão responsabilidades da gerência da RVBI, onde serão verificadas as sugestões das bibliotecas participantes em relação à complementação e atualização das áreas trabalhadas.

No “Guia de atualização dos DeCS” (2007), o *objetivo* do manual é descrever os conceitos, regras, responsabilidades e métodos envolvidos na manutenção do DeCS como vocabulário-padrão da Biblioteca Virtual em Saúde. Nas *características* do tesouro, são calculados o número de descritores e as categorias organizadas em hierarquias. As *responsabilidades* também são designadas a um comitê técnico da BIREME.

Como frisado anteriormente, o DeCS foi construído a partir do MeSH, de sorte que os termos originais foram *coletados* de lá. De acordo com as atualizações de coleta e *análise dos termos*, é necessário apresentar justificativa para inserção do novo termo. É *responsabilidade* do comitê técnico analisar os novos termos sugeridos, preservando a integridade da estrutura da linguagem. As revisões anuais que ocorrem no MeSH se refletem nas mudanças do DeCS. Aqui, já podemos verificar os aspectos de atualização ocorrentes na linguagem.

Na *elaboração das hierarquias*, apresentou-se um fluxo chamado de pré-projeto, onde são feitas perguntas para identificar como construir essa hierarquia, como, por exemplo: está dentro dos domínios da saúde? A terminologia está hierarquizada? Está suficientemente representada no DeCS? Com base nas respostas obtidas, será decidida qual a solução mais pertinente: revisão terminológica, ampliação do vocabulário e atualização da estrutura. Consta no manual que a criação de um *termo específico* será válida, se atender ao critério de frequência.

As *relações de equivalência* podem ser utilizadas na recuperação da literatura. Já nas *relações não hierárquicas ou associativas*, os conceitos existentes em categorias ou subcategorias que tenham afinidade entre si podem ser relacionados no DeCS, através do campo TR (termo relacionado).

Nos aspectos de *configuração*, ligados à *introdução do tesouro*, apresentam-se as áreas de assunto cobertas, como Ciência e Saúde (218), Homeopatia (1.950), Saúde Pública (3.486)

e Vigilância Sanitária (830). É uma linguagem trilingue (inglês, espanhol e português), capaz de realizar a busca nesses idiomas distintos. A criação de termos dependerá de sua amplitude, frequência de manifestação, pertinência, repetição de ramos hierárquicos já existentes no DeCS e pós-coordenação. Não consta, no manual, a norma adotada para a elaboração da linguagem. É indicado que a atualização seja realizada anualmente.

Constata-se a integração com a indexação, quando é inicialmente mencionado no manual que o DeCS foi elaborado para fins de indexação de artigos. A respeito da *disseminação*, o tesouro está disponível para consulta no *site* para qualquer pessoa ou instituição que desejar consultá-lo, e seu endereço eletrônico aparece com indicação nas referências bibliográficas do manual.

Para sugerir retirada de termos, acréscimos e modificações, é necessário preencher um documento com a proposição dos descritores, acompanhado do registro bibliográfico completo e a estrutura hierárquica.

O “Manual de Política de Indexação para as Bibliotecas Universitárias da UNESP”, 1ª edição revisada e ampliada de 2017, abrange as três grandes áreas do conhecimento, com suas respectivas subdivisões. O público-alvo é a comunidade universitária, incluindo a equipe técnica das bibliotecas, a qual executa os procedimentos de organização e disseminação da informação. O manual e todo o sistema da biblioteca são disponibilizados em meio eletrônico, disponíveis na página da Internet da Biblioteca.

A determinação de gerência e a presença de profissionais da informação conhecedores da área de assunto são itens referentes aos *recursos* identificados, enquanto o envolvimento oficial de usuário no projeto e a atuação de suporte de profissionais de TI não foram citados no manual analisado. Já os *recursos do vocabulário* advêm totalmente da Linguagem UNESP.

O manual não menciona pretensão de ser multilíngue, assim como não designa editor, apenas informando as unidades a que estão filiados os envolvidos, assim como os responsáveis pelas etapas de pesquisa e gerenciamento do projeto.

Não há identificação de qual *software* é usado para o gerenciamento do tesouro e as restrições que este ofereceria, mas é informada a pretensão de elaboração ou adoção de um *software* específico para o sistema UNESP, apesar da ausência dos requisitos para a escolha desse *software* de gestão.

A UNESP também não elaborou seu tesouro antes da preparação de uma base de dados; dessa forma, a coleta de termos ocorreu mediante importação e modelização de

registros de autoridade das linguagens *Library of Congress Subject Headings* (LCSH), Terminologia da Biblioteca Nacional (TBN) e *Medical Subject Headings* (MeSH), para consulta quanto às revisões e atualizações dos termos. Segundo o manual, a Linguagem UNESP é multidisciplinar e essencialmente fundamentada na LCSH, com incorporação de novos termos e modificações de registros de autoridades de termos importados da Lista de Cabeçalhos de Assunto da Rede Bibliodata (LCARB), em consulta às linguagens citadas.

Na estrutura dos termos a UNESP, por meio de sua linguagem, adotaram-se as hierarquias, destinando-se um capítulo aos processos de decisão quanto à construção e análise dos termos a serem adotados. Os registros de autoridade são feitos em formato MARC21, conforme determinação da Linguagem UNESP, como regra para todos os procedimentos de processamento da informação nas unidades da universidade e a criação do tesouro. O nível de conhecimento dos usuários da área de assunto do tesouro e do próprio tesouro não foi identificado no manual, porém, estes seguem o foco temático de cada biblioteca, que é relacionado aos cursos e atividades de cada unidade.

Quanto à *introdução* ao tesouro, as conformidades com a norma estão presentes nos itens de especificação das áreas cobertas pelo tesouro e suas subdivisões, bem como das referências e reconhecimentos aos recursos necessários para elaboração do tesouro e/ou linguagem. A totalidade de termos constituintes e suas especificações de preferidos e não-preferidos não foram identificadas, apenas constando os dados de 2014 sobre os registros de termos-tópico, da mesma maneira que houve ausência de informações sobre o idioma disponível; confirmação de normas utilizadas para a construção do tesouro ou para regras de registro; uma reunião das convenções, abreviaturas, siglas e sinais de pontuação usados, além de não prever a periodicidade da atualização da política, a qual recebe uma especificação de como deve ser executada, nos casos de sugestões.

O manual de indexação da UNESP e suas diretrizes quanto à linguagem de indexação UNESP está em versão digital disponível para *download*. Não há informações quanto à disponibilização de cópias impressas de nenhuma espécie ou por motivação especial, não sendo igualmente citadas as publicações em diretórios de *website* ou depósito em Centro de Informação.

A sugestão de atualização se dá por meio de formulários e procedimentos descritos no manual, possibilitando a interação de todos os interessados. Já quanto à revisão, não há informações a respeito que permitam a análise.

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

O “Manual de indexação de assuntos com uso do vocabulário controlado USP: versão preliminar” é um guia de uso para os indexadores do Sistema, auxiliando no processo de representação e atendendo aos usuários do Banco de Dados da USP – DEDALUS (bibliotecários de referência e usuários finais dos serviços da biblioteca).

Quanto ao item *planejamento*, podemos localizar a designação dos executores e objetivo do manual, no entanto, não ficam claras outras questões, como a carga de conhecimento sobre tesouro e dos assuntos ali representados pelos usuários do sistema. Da mesma maneira, não há informação sobre o *software* de gestão para o uso do tesouro e quais as implicações que ele acarretaria; também não há lista de requerimentos necessários para a escolha de *software*, mas é sabido que atualmente a USP adota o TemaTres 2.1. para tal finalidade.

É possível identificar todos os envolvidos na elaboração do manual e dos trabalhos relacionados ao tesouro e vocabulário controlado, porém, trata-se de uma lista de responsabilidades, sem definições claras sobre atividades e especialidades, assim como não há especificação de envolvimento oficial de usuários e profissionais de TI. Não foi mencionada a intenção de desenvolver um tesouro multilíngue.

O manual evidencia tratar-se de um trabalho continuado, o qual teve seu início no estudo de um vocabulário controlado próprio, Vocabulário Controlado da USP, apresentado como base de dados denominada SIBIX 650. Sua elaboração visa a atender à demanda de indexação das 40 Bibliotecas que integram o SIBi/USP, no Banco de Dados DEDALUS, abrangendo todas as áreas em que a USP se envolve academicamente. Não são claras as formas de coleta e análise dos termos, todavia, é possível entender que são possivelmente exportados do vocabulário controlado, o qual faz uso dos descritores, qualificadores e remissivas, por meio de tabelas.

Os termos foram manuseados em grupos hierárquicos, a fim de construir a estrutura de entrada do tesouro, com macroestrutura que traz as áreas abrangidas pela ferramenta. A estrutura propicia utilizar outras formas de combinação para estratégia de busca, como os operadores booleanos e suas combinações, formando equações de busca. Um capítulo é dedicado a exemplificações de como aplicar os operadores booleanos no manual, com ênfase nas aplicações do sistema USP.

Os itens sobre a introdução do tesouro contêm informações, como o número total de termos com especificação de que os 42.702 termos são divididos nas categorias descritores,

qualificadores, remissivas e elos falsos, dados de 2006. As áreas de assunto cobertas são listadas com informações sobre as grandes áreas do conhecimento, sem informações sobre idiomas cobertos, normas seguidas ou regras, tanto para reger a seleção de conceitos complexos quanto para seleção de forma dos termos preferidos; ademais, não se tem nenhuma tabela que trouxesse as siglas, sinais e abreviaturas usadas, no manual ou no tesouro.

Sobre as disposições que envolvem a disseminação, foi localizada para análise apenas o quesito de integralização, em que o manual sinaliza que vocabulário e tesouro estão integrados. Não é abordada, de modo direto e claro, a publicação do tesouro e seu formato, sabendo-se que estão acessíveis no *site* oficial da SIBi/USP. Não há menção de distribuição impressa em nenhuma forma, incluindo cópia depositada em Centro de Informação ou em diretório de *website*, como prevê a Norma ISO 25964-1.

É ressaltado, no manual, como se dará a atualização, incluindo os bibliotecários e os usuários da rede. A sugestão de atualização ocorre por formulário, e é necessária justificativa fundamentada para a sua análise, entretanto, não são definidos os procedimentos para a revisão ou sua periodicidade.

5 CONSIDERAÇÕES FINAIS

O grande avanço na gestão de tesouros é o controle de vocabulário em contexto digital, com o uso da interoperabilidade entre vários vocabulários de sistemas de informação, redes de bibliotecas, bibliotecas digitais e repositórios. A Norma ISO 25964 (2011) incorpora, em suas diretrizes, aspectos evolutivos: os aspectos de mono e multilinguismo, interoperabilidade linguística, interoperabilidade entre vocabulários controlados, aplicação da análise de facetas na construção do tesouro, definição de protocolos e formatos necessários para o intercâmbio de dados e a proposta de modelização de dados.

Para isso, a seção 13 da Norma ISO 25964 (2011) sobre gestão de tesouros indica tópicos importantes que orientam a construção e a manutenção. Tais tópicos, em princípio, devem constar dos manuais de indexação, para que seja possível a compreensão, em diferentes períodos, da evolução do controle de vocabulário. Mais do que isso, é importante que haja uma orientação normativa para a gestão de tesouros.

Na análise realizada com os quatro manuais de indexação, observou-se que suas versões não são atualizadas, não há informação sobre o uso de *softwares* de gestão de

tesauros e nem de uso de norma para construção de tesauros. Recomenda-se a necessidade de atualização de suas diretrizes e de orientação normativa, a qual poderia ser fundamentada na Norma ISO 25964 (2011). Dessa forma, o aprimoramento e a atualização na construção de tesauros poderiam ser obtidos pela complementação do que é indicado pela Norma.

Por outro lado, é preciso lembrar e frisar que a função mais importante do tesouro é servir de mediador entre o vocabulário do conteúdo documentário, durante a representação pelo indexador, e o do usuário, ao longo da representação de sua necessidade de busca.

Em nota explicativa, a Norma ISO 25964-1 (2011) caracteriza o aspecto funcional do tesouro, ao assinalar que seu propósito é orientar tanto o indexador quanto o pesquisador para a seleção do mesmo termo preferido. Contudo, essa função somente é possível com a disponibilização do tesouro para consulta ao usuário, durante a estratégia de busca.

Embora esse não tenha sido um dos critérios de análise, ressalta-se que apenas um dos manuais, o do DeCs, informa que disponibiliza seu tesouro para consulta. Registra-se que o Sistema de Bibliotecas da USP (SIBI/USP) recentemente disponibilizou seu Vocabulário USP (VocaUSP), mediante interface de busca com o *software* TemaTres, versão 2.1. Essa funcionalidade dupla do tesouro é uma grande vantagem de visibilidade e, principalmente, de interação do usuário com o sistema de informação, para conhecimento do vocabulário e de sua estrutura de relacionamentos entre termos, que, certamente, ampliará e dará precisão à sua busca.

REFERÊNCIAS

AMERICAN NATIONAL STANDARD/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.

Campos, M. L. A.; Gomes, H. E. Metodologia de elaboração de tesouro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 3, p. 348-358, 2006. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/273>>. Acesso em: 6 jul. 2017.

CERVANTES, B. M. N. **A construção de tesauros com a integração de procedimentos terminográficos**. 2009. 198 f. Tese (Doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências - Marília, 2009.

CERVANTES, B. M. N. **Contribuição para a Terminologia do Processo de Inteligência Competitiva**: estudo teórico e metodológico. Marília, 2004. 183f. Dissertação (Mestrado) – Universidade Estadual Paulista, Faculdade de Filosofia e Ciências - Marília, 2004.

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

Cervantes, B. M. N.; Fujita, M. S. L. Uma perspectiva diacrônica de parâmetros e modelos de construção de tesouros. In: Cervantes, B. M. N. (Org.). **Horizontes da organização da informação e do conhecimento**. Londrina: EDUEL, 2012.

CINTRA, A. M. *et al.* **Para entender as linguagens documentárias**. São Paulo: Polis; APB, 1994. 72p.

CINTRA, A. M. *et al.* **Para entender as linguagens documentárias**. 2.ed. São Paulo: Polis; APB, 2002. 96p.

AUTORES A, B, C. Linguagens de indexação em bibliotecas universitárias brasileiras: diagnóstico preliminar das regiões sul e sudeste. In: ENCONTRO DA ASSOCIAÇÃO DE EDUCAÇÃO E PESQUISA EM CIÊNCIA DA INFORMAÇÃO DA IBEROAMÉRICA E CARIBE - EDICIC, 10., 2016, Belo Horizonte. No prelo.

DIRETRIZES PARA CONSTRUÇÃO DO TESAURO DA REDE VIRTUAL DE BIBLIOTECAS - CONGRESSO NACIONAL – RVBI. Desenvolvido pelo Grupo de Estudo do Tesouro da RVBI. Brasília: Senado Federal, Secretaria de Biblioteca, 2007. 76 p. Disponível em: <<http://www2.senado.leg.br/bdsf/item/id/81838>>. Acesso em: 06 jul. 2017.

(Org.) **A indexação de livros**: a percepção de catalogadores e usuários de bibliotecas universitárias. Um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Cultura Acadêmica, 2009.

As linguagens de indexação em bibliotecas nacionais, arquivos e sistemas de informação na América-latina. In: SEMINÁRIO NACIONAL DE BIBLIOTECAS UNIVERSITÁRIAS. 16., 2010, São Conrado, 2010. II SEMINÁRIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS. **Anais...** São Conrado, Rio de Janeiro, 2010.

Política de indexação para biblioteca: elaboração, avaliação e implantação. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2016.

Guia de Atualização do DeCS. BIREME / OPAS / OMS. São Paulo: BIREME / OPAS / OMS. Disponível em <<http://bvsmodelo.bvsalud.org/download/decs/DeCS-GuiaAtualizacao.pdf>>. Acesso em: 6 jul. 2017.

Hjørland, B. What is Knowledge Organization (KO)? **Knowledge Organization**., v.35, n.2/3, p.86-101, 2008.

ISO 25964-1. Information and documentation: part 1: thesauri and interoperability with other vocabularies. Geneva: International Organization for Standardization, 2011.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. Brasília: Binquet de Lemos, 2004.

MANUAL DE INDEXAÇÃO DE ASSUNTOS COM O USO DE VOCABULÁRIO CONTROLADO DA USP: versão preliminar. Universidade de São Paulo Sistema Integrado de Bibliotecas, 2006. Disponível em: <https://edisciplinas.usp.br/pluginfile.php/378251/mod_resource/content/1/Vocabulario_manualdeindexa%C3%A7%C3%A3o.pdf>. Acesso em: 6 jul. 2017.

MANUAL DE POLÍTICA DE INDEXAÇÃO PARA BIBLIOTECAS UNIVERSITÁRIAS DA UNESP. FUJITA, M. S. L. (Coord.). 1. ed. rev. e ampl. – São Paulo: Unesp, 2017. Disponível em <<http://www.biblioteca.unesp.br/portal/arquivos/manual-politica-indexacao>>. Acesso em: 6 jul. 2017.

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

MOREIRA, M. P., MOURA, M. P., Construindo tesouros a partir de tesouros existentes: a experiência do TCI Tesouro em Ciência da Informação. **DataGramaZero**, v.7, n.4, ago. 2006.

NARUKAWA, C. M.; SALES, R. A relevância e influência de linguagens documentárias na política de tratamento da informação. In: GIL LEIVA, I.; FUJITA, M.S.L. (Eds.). **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012.

OLIVEIRA, D. A., ARAUJO, R. F. de. Construção de linguagens documentárias em sistemas de recuperação da informação: a importância da garantia do usuário. **Encontros Bibli**, Florianópolis, v.17, n. 34, p.17-30, maio/ago., 2012.

PINTO, M. C. M. F. Análise e representação de assuntos em sistemas de recuperação da informação; linguagens de indexação. **Revista da Escola de Biblioteconomia da UFMG**, v. 14, n. 2, p. 169-186, 1985.

RUBI, M. P. A. Política de indexação. In: GIL LEIVA, I.; FUJITA, M. S. L. **Política de indexação**. São Paulo: Cultura Acadêmica; Marília: Oficina Universitária, 2012.

RUBI, M. P. Os princípios da política de indexação na análise de assunto para catalogação: especificidade, exaustividade, revocação e precisão na perspectiva dos catalogadores e usuários. In: FUJITA, M. S. L. (Org.). **A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias**. Um estudo de observação do contexto sociocognitivo com protocolos verbais. São Paulo: Cultura Acadêmica, 2009.

RUBI, M. P. Elementos de política de indexação em manuais de indexação de sistemas de informação especializados. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 8, n.8, p. 57-66, 2003. Disponível em:

<http://www.brapci.ufpr.br/brapci/index.php/article/download/12889>

VOGEL, M. J. M. A influência da Jean-Claude Gardin e a linha francesa na evolução do conceito de linguagem documentária. **Perspectivas em Ciência da Informação**, São Paulo, v. 14, p. 80-92, 2009

ZENG, M. L.; CHAN, L.M. Trends and issues in establishing interoperability among knowledge organization system. **Journal of the American Society for Information Science and Technology**; v.55, n.5, p. 377-395, 2004. Disponível em:

<http://onlinelibrary.wiley.com/doi/10.1002/asi.10387/epdf?r3_referer=wol&tracking_action=preview_click&show_checkout=1&purchase_referrer=onlinelibrary.wiley.com&purchase_site_license=LICENSE_DENIED>. Acesso em: 6 jul. 2017.