

Relações semânticas em tesouros: contribuições da abordagem pragmática

Semantic relationships in thesauri: pragmatic approach contributions

Leila Cristina Weiss

Doutoranda do Programa de Pós-graduação em Ciência da Informação da Universidade Federal de Santa Catarina – UFSC.
Bibliotecária coordenadora da Biblioteca Setorial do Colégio de Aplicação da Universidade Federal de Santa Catarina – UFSC.
E-mail: leilacw@gmail.com

Marisa Bräscher

Doutorado em Ciência da Informação pela Universidade de Brasília – UnB.
Professora do Departamento de Ciência da Informação da Universidade Federal de Santa Catarina – UFSC.
E-mail: marisa.brascher@gmail.com

Resumo

Analisa como a abordagem semântica pragmática pode ser adotada para o estabelecimento de relações semânticas em tesouros. A partir da caracterização da abordagem pragmática nos estudos de OC e a verificação de como essa abordagem se encontra nas normas de elaboração de tesouros, descreve contribuições da abordagem pragmática às recomendações das normas ANSI/NISO Z39.19(2005) e ISO2594-1(2011). Para as relações de equivalência, essa contribuição diz respeito à percepção de que a distinção do que foi considerado igual ou equivalente para fins de revocação na Recuperação da Informação é uma medida que minimiza as implicações éticas de se escolher uma forma de expressão como termo preferido e proporciona maior flexibilidade ao tesouro. Para as relações hierárquicas, verifica-se a importância de se considerar as diferenças conceituais, pois se as hierarquias são estabelecidas apenas a partir de premissas universalistas podem se tornar demasiadamente rígidas. Para as relações associativas verifica-se a importância da bibliografia para a identificação das relações que podem guiar o usuário a localizar informações úteis ao alcance de seus objetivos.

Palavras-chave: Relações semânticas. Pragmatismo. Tesouros. Organização do Conhecimento.

Abstract

It analyzes how the pragmatic semantic approach can be adopted for establishing semantic relationships in thesauri. In the context of Knowledge Organization, it describes the pragmatic approach contributions to the recommendations of ANSI/NISO Z39.19 (2005) and ISO2594-1 (2011) norms. For equivalence relationships, this contribution relates to the perception that the distinction of what was considered equal or equivalent for recall purposes in Information Retrieval is a measure that minimizes ethical implications of choosing an expression form as a preferred term and it provides greater flexibility to thesaurus. For hierarchical relationships, there is the importance of considering the conceptual differences, because if hierarchies are established only from universalist assumptions, they may become too rigid. For associative relationships, there is the importance of bibliography for identifying relationships that can guide the user to find useful information to reach his or her goals.

Key words: Semantic relationships. Pragmatism. Thesauri. Knowledge Organization.

Introdução

Um importante componente dos tesouros são as relações semânticas. Khoo e Na (2006) apontam a existência de um interesse crescente no estudo de tais relações. Para Hjørland (2003, p. 87, tradução nossa), “a unidade básica na organização do conhecimento é a relação semântica entre dois conceitos, e essas relações estão envoltas em teorias”. Teorias podem representar pontos de vistas distintos e, assim, apresentar diferentes relações semânticas entre os mesmos conceitos. As dificuldades que se apresentam para representar essas diferenças levam a certo consenso entre pesquisadores da área em considerar o estabelecimento de relações semânticas em tesouros um processo arbitrário. “A arbitrariedade da seleção dos agrupamentos e relacionamentos não pode ser vista, no entanto, como algo totalmente aleatório. Ela se deve aos aspectos que desejamos destacar numa determinada representação do conhecimento” (CAFÉ; BRÄSCHER, 2011, p. 26). Deve, portanto, haver correspondência entre o sistema e a realidade que este representa.

Essas e outras questões são estudadas na área de Organização do Conhecimento (OC), que é interdisciplinar e possui ligação com a área de semântica (estudo do significado). De acordo com Hjørland (2007b) as abordagens diferentes para OC implicam diferentes pontos de vista sobre semântica. Peregrin (2004) destaca dois paradigmas dominantes em semântica, o pragmático e o positivista. Na abordagem positivista a linguagem é retratada como um meio de representar o mundo e a realidade. Na abordagem pragmática, a linguagem é vista como um meio de interação e o significado de uma expressão não é a entidade representada por ela, mas sim a sua função na interação. A abordagem positivista baseia-se no pressuposto de que a linguagem é essencialmente um sistema de nomes para representações de coisas, enquanto que na abordagem pragmática assume-se que a linguagem é uma coleção de meios para fazer coisas. (PEREGRIN, 2004).

Observa-se que a área de Ciência da Informação (CI) também sofre influência das abordagens positivista e pragmática. Vega-Almeida, Fernández-Molina e Linares (2009) apontam que os pressupostos teóricos pragmáticos também estão presentes no paradigma social em CI. Tanto a abordagem pragmática quanto a social em CI enfocam aspectos sócio-culturais e contextuais envolvidos nos processos de comunicação do conhecimento.

Na OC o tipo de abordagem influencia as características e funções de um Sistema de Organização do Conhecimento (SOC). Diante de uma perspectiva em que a linguagem é vista

como um meio de interação, e o significado, como a função que uma expressão exerce na interação, os SOC devem também ser instrumentos de interação que levem em conta os aspectos contextuais em que se inserem.

Os tesouros são tipos de SOC e possuem normas internacionais que orientam sua elaboração e que apresentam algumas práticas já consolidadas na OC. As recomendações das normas, de modo geral, são baseadas em processos utilizados para o desenvolvimento de tesouros já existentes e considerados eficientes. Essas recomendações, ainda que não indiquem de forma explícita seus pressupostos teóricos, podem indicar em qual abordagem de OC se inserem. O que torna possível identificar recomendações cuja abordagem pragmática possa trazer contribuições.

Assim, diante da influência que a área de semântica exerce na OC, da importância das relações semânticas para a estrutura conceitual dos tesouros, bem como demais Sistemas de Organização do Conhecimento (SOC), nesta pesquisa analisamos como a abordagem pragmática pode ser adotada para o estabelecimento de relações semânticas em tesouros e apresentamos as contribuições da abordagem pragmática na OC às recomendações das normas ANSI/NISO Z39.19 (2005) e ISO2594-1 (2011) para o estabelecimento de relações semânticas em tesouros.

A investigação das relações semânticas se deve ao importante papel que desempenham nos tesouros e demais SOC. Mas, apesar disso, não se apresentam tanto quanto outros aspectos desses sistemas em pesquisas da área de CI e no campo de OC. Os estudos “têm-se centrado mais em conceitos e termos, mas o foco irá cada vez mais se deslocar para a identificação, processamento e gestão das relações, para alcançar uma maior eficácia e refinamento das técnicas em ciência da informação.” (KHOO; NA, 2006, p. 158, tradução nossa).

A investigação da abordagem semântica pragmática justifica-se por ser um paradigma recente em pesquisas no campo de OC, principalmente nas de natureza empírica. Neste trabalho procura-se aproximar a base teórica que sustenta a abordagem semântica pragmática da aplicação prática, refletida nas normas de elaboração de tesouros.

A base teórica que sustenta a abordagem pragmática na OC é descrita em Weiss e Bräscher (2014) e Weiss (2014). A pesquisa que deu origem aos trabalhos apontados empregou técnicas de análise de conteúdo por categorias, conforme definidas por Bardin (1979). Nesses trabalhos analisou-se a literatura da área de CI que trata sobre pragmatismo e pragmática. Por

meio de tal análise identificaram-se ideias compartilhadas e/ou complementares que foram organizadas em enunciados para identificar a compatibilidade entre as características da abordagem pragmática e as recomendações das normas para elaboração de tesouros.

Na análise da literatura que aborda o pragmatismo e pragmática na CI o critério empregado para a formação das categorias foi o semântico¹. Os temas foram identificados na medida em que os trabalhos do corpus de pesquisa foram analisados. Ou seja, as categorias ou classes não foram previamente definidas. O processo empregado então foi o “por milha”, como denominado por Bardin (1979).

Para identificar a compatibilidade entre as características da abordagem pragmática e as recomendações das normas para elaboração de tesouros o critério empregado para a análise categorial foi novamente o semântico. Mas nesse caso as categorias, compatibilidade ou incompatibilidade com os apontamentos da abordagem pragmática, foram definidas antes da análise das normas. O processo empregado então foi o “por caixas”, como denominado por Bardin (1979).

Com a aplicação da análise de conteúdo, foi possível caracterizar a abordagem pragmática nos estudos de OC e verificar como essa abordagem se encontra nas normas de elaboração de tesouros. Tais resultados foram essenciais para o alcance do objetivo final da pesquisa, que são apresentados neste trabalho. Abordamos algumas possíveis contribuições do arcabouço teórico, aqui denominado abordagem pragmática na OC, às recomendações para o estabelecimento de relações semânticas em tesouros das normas ANSI/NISO Z39.19 (2005) e ISO2594-1 (2011).

Na abordagem pragmática na OC a informação é vista como um fenômeno que é construído. Essa construção se dá em circunstâncias diversas e envolve diferentes atores que podem ter pontos de vista e necessidades distintas, devido aos diversos aspectos contextuais envolvidos na produção e uso da informação e do conhecimento. Assim, a avaliação da qualidade da informação deve levar em conta esses pontos de vista distintos. Tanto a informação quanto a sua qualidade não são vistos como fenômenos constantes e avaliar uma fonte de informação a partir de premissas de verdade ou falsidade é visto como perigoso, pois se deve avaliá-la em relação à forma como aborda as controvérsias e diferentes pontos de vista.

¹ De acordo com Bardin (1979) os critérios da análise de conteúdo podem ser semânticos (categorias temáticas) e sintáticos (os verbos, os adjetivos), dentre outros.

Assim como as enciclopédias, e outras fontes de informação do gênero, os SOC também devem apresentar os diferentes pontos de vista encontrados na literatura de uma área do conhecimento.

Considera-se mais importante possibilitar uma comunicação efetiva pela compreensão dos diferentes pontos de vista do que a escolha de um como sendo o melhor para representar uma realidade que é, por natureza, heterogênea. As relações semânticas poderiam desempenhar papel crucial nessa comunicação, ao guiar o usuário a partir do seu ponto de vista inicial até aqueles defendidos pela frente de pesquisa sobre o tema.

Relações semânticas em tesouros

As normas ANSI/NISO Z39.19 (2005) e ISO2594-1 (2011) apresentam recomendações para estabelecer relações semânticas em tesouros. Os três tipos principais de relações semânticas, de equivalência, hierárquica e associativa, coincidem em ambas as normas. No quadro 1 descrevem-se algumas possíveis variações dos diferentes tipos de relações e seus exemplos.

Quadro 1: tipos de relações semânticas em tesouros

Tipo de relação	Exemplo
Equivalência	
Sinônimos	ONU / Organização das Nações Unidas
Variante lexicais	estoque / stock
Quase sinônimos	água do mar / água salgada
Hierárquica	
Genérica	aves / papagaios
Instância	mar/ mar Mediterrâneo
Todo / Parte	encéfalo / tronco encefálico
Associativa	
Causa / Efeito	acidente / lesão
Processo / Agente	medir temperatura / termômetro
Processo / Contra-agente	fogo / anti-chamas

Fonte: ANSI/NISO Z39.19 (2005, p. 42, tradução nossa)

A seguir apresentamos, com base nas recomendações das normas, cada tipo de relação de forma mais detalhada e as contribuições que a abordagem pragmática pode trazer para o estabelecimento dessas relações em tesouros e Sistemas de Recuperação da Informação.

Relações de equivalência

Além dos três tipos de relações de equivalência apresentadas no quadro 1, a norma ANSI/NISO Z39.19 (2005) descreve mais dois: postagem genérica, que é recomendada apenas para os termos de áreas periféricas a do tesouro, ou quando o número de documentos a ser indexado com uma determinada subclasse for muito pequeno, a ponto de não justificar a sua adição ao tesouro; e referência cruzada para os elementos de um termo composto.

A norma ISO25964-1 (2011) apresenta duas subdivisões para as relações de equivalência que coincidem com as da ANSI/NISO Z39.19 (2005): a) sinônimos e; b) quase-sinônimos.

A relação de equivalência que na ANSI/NISO Z39.19 (2005) é denominada postagem genérica, na ISO25964-1 (2011, p. 48) se fala em “termos específicos agrupados em um conceito mais amplo”.

Além disso, por se tratar de uma norma destinada também à tesouros multilíngues a ISO25964 (2011) apresenta recomendações para o estabelecimento de relações de equivalência entre linguagens. Recomenda-se que

Em um tesouro multilíngue, todas as linguagens têm *status* igual e, se possível, cada conceito deve ser representado em todas as linguagens do tesouro. Uma coleção que tem sido indexada usando uma das linguagens pode ser pesquisada de forma igualmente efetiva usando uma das outras linguagens. (ISO25964-1, 2011, p. 50, tradução nossa)

Também se descreve que é possível tratar diferentes dialetos ou sublinguagens como sendo linguagens separadas. Por exemplo, inglês americano, inglês britânico e inglês indiano. De modo semelhante, a terminologia preferida pelos cientistas pode ser apresentada como uma sublinguagem diferente. (ISO25964-1, 2011).

A partir das recomendações das normas e da análise da abordagem pragmática, podemos dizer que as relações de equivalência dependem da escolha do termo preferido e podem contribuir significativamente com a revocação. De maneira geral, os sistemas de busca estão

preparados para efetuar a pesquisa apenas com o termo preferido, ou seja, ao buscar por um sinônimo o usuário é orientado a realizar a busca com o termo preferido. Dessa maneira não há distinção entre o termo não preferido e o termo preferido na busca, no entanto, em determinados contextos o significado dos dois termos pode não ser o mesmo e o usuário pode se interessar exatamente por aquilo que diferencia um termo do outro, naquele contexto de uso.

O possível interesse em distinguir os termos que foram considerados equivalentes fica evidente quando pensamos em quase sinônimos e postagem genérica. Mas é importante lembrar que sinônimos perfeitos são praticamente inexistentes e que podem existir implicações éticas em se escolher uma forma de expressão e não outra. Assim, informar o usuário sobre o que foi considerado equivalente poderia minimizar o viés que a escolha de um termo como o termo preferido pode gerar. Conforme Hjørland (2008a) recomenda, se não podemos eliminar o viés, devemos ao menos ser mais conscientes e responsáveis com relação a ele, determinando explicitamente a perspectiva representada no sistema de informação.

Isso é importante tanto pela dificuldade em se distinguir todos os possíveis contextos de uso já existentes, como também pela dificuldade em se prever os contextos que poderão ocorrer. O significado de um termo não está apenas no presente e no passado, está também no futuro, ou seja, pode mudar de acordo com o que as pessoas farão com ele, como ele será utilizado (PEIRCE² apud HJØRLAND, 1998a) e isso não pode ser previsto. Assim, dois termos que no momento da elaboração do tesouro são realmente sinônimos, e a distinção entre eles naquele momento realmente não teria utilidade, com o passar do tempo podem deixar de ser, e a distinção vir a ser útil.

A forma exata como isso poderá ser feito ainda é uma questão a ser resolvida, uma possibilidade seria a formação de *clusters* dos resultados para cada um dos termos possíveis, como já ocorrem em algumas bases de dados que agrupam os resultados por autor, ou por ano da publicação, por exemplo. No caso dos sinônimos, o critério de agrupamento seriam os diferentes termos que na busca inicial foram assim considerados.

Por exemplo, quando consultamos o Descritores em Ciências da Saúde (DeCS)³, verificamos que o termo “Esclerose Múltipla” possui três sinônimos “Esclerose Disseminada”,

² PEIRCE, C. S. What pragmatism is. **The Monist**, v. 15, p. 161-181. 1905. Publication manual of the American Psychological Association. 4. ed. Washington, DC: APA, 1994.

³ O DeCS é um vocabulário estruturado que foi desenvolvido com base no Medical Subject Headings (MeSH) da U.S National of Medicine (NLM). Integra a metodologia LILACS, que é um componente da Biblioteca Virtual em Saúde em contínuo desenvolvimento, constituído de normas, manuais, guias e aplicativos, destinados à coleta,

“Esclerose Múltipla Aguda Fulminante” e “MS (Esclerose Múltipla)”. Ao fazer uma busca com qualquer um desses termos na base de dados MEDLINE, que utiliza o DeCS para a indexação, recupera-se exatamente o mesmo resultado, que é de 38656 referências. Dentre os trabalhos recuperados, talvez nem todos utilizem o termo “Esclerose Múltipla”. Assim, o que se propõe é disponibilizar aos usuários a possibilidade de refinar a busca com termos considerados sinônimos, por meio de agrupamentos e tornando possível ao usuário identificar o que foi considerado sinônimo e escolher apenas os resultados da busca com determinado(s) termo(s), se considerar mais útil para sua pesquisa.

A título de ilustração, digamos que das 38656 referências recuperadas talvez 30000 poderiam ser agrupadas em “Esclerose Múltipla”, pois esse é o termo utilizado no documento indexado; 8000 agrupadas sob o termo “Esclerose Disseminada”, por ser o termo pelo qual a doença era denominada no momento em que estes trabalhos foram escritos; e 600 seriam agrupadas em “Esclerose Múltipla Aguda Fulminante”, pois este é o termo que os autores utilizaram por tratar de uma forma específica da doença, e assim sucessivamente.

Pode-se pensar que isso não seria necessário por que o usuário pode escolher o texto completo como campo de busca, opção presente em muitos SRI, e assim fazer a pesquisa pelo termo que considera mais adequado, o que recupera apenas os documentos que contem tal termo no texto. Mas isso, em muitos casos, poderia resultar em baixa precisão, recuperando muitos documentos que apenas citam tal termo, tratando apenas brevemente o assunto.

Os profissionais da informação podem encontrar motivos para considerar que o que se propõe não seja realmente necessário, pois talvez já existam ferramentas que, devidamente combinadas poderiam oferecer o mesmo resultado. Descrever todas essas possibilidades ultrapassam o objetivo dessa pesquisa. Mas é importante lembrarmos que dominar o uso de tais ferramentas também pode ser bastante complexo, fazendo com que os usuários subutilizem os SRI.

Essa sugestão de como pode ser feita a distinção dos termos que foram considerados equivalentes e a formação de *clusters* é um exemplo de alternativas que levam em conta aspectos da abordagem pragmática, na qual se evidencia a necessidade dos SOC serem mais flexíveis e dinâmicos, pois considera-se que a informação é construída, não apenas inventada

seleção, descrição, indexação de documentos e geração de bases de dados. Que tem como objetivo de permitir o uso de terminologia comum para pesquisa em três idiomas – inglês, português e espanhol – proporcionando um meio consistente e único para a recuperação da informação independentemente do idioma.

ou descoberta, e depende do contexto social. A dependência do contexto e da ação humana para a construção da informação faz com que sua qualidade não seja constante. As práticas e o uso da informação e da linguagem aparecem como elementos essenciais para a avaliação do valor potencial da informação e de seus possíveis significados em determinado contexto. Nesse sentido destacamos os trabalhos de Frohmann (2004, 2008a, 2008b), Redón Rojas (1996), Hjørland (1998a, 2003, 2007a, 2007b, 2008a, 2008b, 2009), Thellefsen e Thellefsen (2004), Alvarenga (1998), Almeida (2010, 2012a, 2012b, 2012c), Novelino (1998), dentre outros.

Relações hierárquicas

As normas ISO25964-1 (2011) e ANSI/NISO Z39.19 (2005) apontam a existência de três tipos de relações hierárquicas, gênero/espécie, todo/parte e de instância. Além desses, ambas as normas descrevem que um conceito pode ser logicamente designado como membro de mais de uma classe, o que se denomina relacionamento polihierárquico. A ISO25964-1(2011) também traz recomendações para a solução de problemas em tesouros multilíngues, que ocorrem quando os mesmos não possuem uma estrutura simétrica. A ANSI/NISO Z39.19 (2005) descreve, ainda, a possibilidade de se usar indicador de faceta para mostrar os princípios da divisão entre um conjunto de termos irmãos, aqueles que fazem parte de uma mesma classe.

Para o estabelecimento de relações hierárquicas do tipo gênero/espécie ambas as normas recomendam que se faça o teste “todos – alguns”. A ANSI/NISO Z39.19 (2005) apresenta como exemplo a aplicação do teste aos conceitos de “plantas suculentas” e “cactos”, caso em que alguns membros da classe plantas suculentas são conhecidos como cactos e que todos os cactos, por definição e, independentemente do contexto, são plantas suculentas, e assim a relação hierárquica pode ser implementada no tesouro. Já no caso de se aplicar o teste aos conceitos de “plantas do deserto” e “cactos” tal relação não deve ser implementada, pois algumas plantas do deserto são conhecidas como cactos e que alguns, mas não todos, os cactos são plantas do deserto. Assim, a esses termos não deve ser atribuída uma relação hierárquica no tesouro. (ANSI/NISO Z39.19, 2005).

A ISO25964-1(2011) apresenta o mesmo esquema para o teste mudando apenas os conceitos utilizados como exemplos, que no caso foram aves, papagaios e animais de estimação.

Papagaios seriam hierarquicamente subordinados a aves, mas não a animais de estimação. Mas descreve também que

Esse argumento não se aplica no contexto de um tesouro especialista dedicado a animais domésticos, em que os únicos papagaios no quadro de referência são animais de estimação. Nesse caso, “papagaios” pode ser subordinado a “animais de estimação” na mesma hierarquia. Tais aproximações devem ser aplicadas com cautela, especialmente em ambientes de rede em que os registros de um sistema podem ser misturados com os de outro. Quando é necessária a interoperabilidade, as relações estabelecidas devem ser universalmente aceitas. (ISO25964-1, 2011, p.59, tradução nossa)

Quanto a relação hierárquica todo-parte a norma ANSI/NISO Z39.19 (2005, p. 49, tradução nossa) descreve que

Esta relação abrange situações em que um conceito é inerentemente incluído em outro, independentemente do contexto, para que os termos possam ser organizados em hierarquias lógicas, com o todo tratado como um termo mais amplo. Essa relação pode ser aplicada a alguns tipos de termos.

Apresenta três exemplos em que esse tipo de relação pode ocorrer: sistemas e órgãos do corpo, locais geográficos e estrutura hierárquica organizacional, empresarial, social ou política, e afirma que tais exemplos não são exaustivos.

Porém na ISO25964-1 (2011, p. 60, tradução nossa) afirma-se que “relações todo-parte cobrem uma gama limitada de situações em que a parte de uma entidade ou sistema pertence exclusivamente a um todo particular”, essas situações ocorreriam apenas entre sistemas e órgãos do corpo, locais geográficos, disciplinas ou campos do discurso, estrutura social hierárquica.

Ambas as normas recomendam que quando um termo que representa a parte pertencer a mais de um todo, como no caso do termo rodas, que poderia ter relação todo-parte com carros e bicicletas, não se deve implementar uma relação hierárquica todo-parte. Caso se considere apropriado, o tipo de relação que poderia ser estabelecida seria uma relação associativa.

A relação hierárquica de instância identifica a ligação entre uma categoria geral de coisas ou eventos, expresso por um substantivo comum, e uma instância individual dessa categoria, muitas vezes, um nome próprio. O exemplo apresentado para esse tipo de relação em ambas as normas é “Alpes” e “Himalaia” como termos subordinados a “regiões montanhosas”, que não são tipos nem partes de “regiões montanhosas”, mas na verdade representam uma instância particular e são nomes próprios.

Quando um conceito pode ser logicamente designado como membro de mais de uma classe ou categoria um relacionamento polihierárquico também pode ser implementado. O termo “pianos”, por exemplo, pode ser subordinado a “instrumentos de cordas” e “instrumentos de percussão.

Além das recomendações referentes a essas possibilidades de relações hierárquicas, a norma ANSI/NISO Z39.19 (2005) descreve que quando os termos são organizados em hierarquias indicadores de faceta podem ser utilizados para mostrar os princípios da divisão entre um conjunto de termos irmãos (termos que compartilham o mesmo termo mais amplo). Alerta que, embora a sua função seja semelhante a de termos genéricos, indicadores de faceta não são termos e não devem ser utilizados como termos de indexação.

Indicador de faceta é utilizado na apresentação de um tesouro quando se emprega a análise facetada. De acordo com a ANSI/NISO Z39.19 (2005, p. 14) a análise facetada pode ser útil em tesouros que possuem milhares de termos, no qual uma estrutura puramente hierárquica pode se tornar confusa e a organização a partir da análise facetada pode facilitar seu manuseio.

A ISO25964-2 (2011) aponta que a análise facetada é útil em gerar hierarquias conforme as regras para relações hierárquicas válidas apenas para conceitos que pertencem à mesma categoria geral. A escolha das facetas pode variar dependendo do domínio, mas de maneira geral se usam as categorias fundamentais como objeto, matéria, agente, ação, espaço, tempo, etc. Caso se considere necessário pode-se subdividir essas facetas.

Diante dessas recomendações para o estabelecimento de relações hierárquicas apresentadas pelas normas, podemos dizer que o teste todos – alguns representa um exemplo claro de premissa universalista para o estabelecimento de relações semânticas. O que é criticado na abordagem pragmática, na qual se considera que as relações semânticas “devem ser entendidas principalmente quanto ao domínio específico, reveladas pelas (e construídas pelas) disciplinas científicas” (HJØRLAND, 2003, p. 107). O teste “todos – alguns” é um teste de lógica referente à semântica formal, ou abordagem positivista em semântica, que não deve ser ignorado, mas deve ser entendido assim como a semântica formal “como uma forma de destacar e materializar cada expressão contribuindo para as inferências de acordo com o contexto onde ocorrem” (PEREGRIN, 2004, p. 12).

Nas recomendações das normas descreve-se que o teste todos – alguns não precisam ser aplicados em um tesouro de uma área específica, mas isso é descrito como uma espécie de exceção à regra, quando na verdade deveria ser o contrário. Pois a maioria dos tesouros se destina a domínios específicos. Assim, a resposta do teste todos – alguns poderia ter como base a análise da bibliografia do domínio ao qual o tesouro se destina. Ou seja, como base na literatura da área identifica-se que relações hierárquicas são pertinentes e devem, portanto, ser estabelecidas no tesouro. Essa prática pode garantir maior consistência no estabelecimento de relações hierárquicas e deve ser compreendida como um critério que é aplicado em contextos específicos, e não como um critério para relações hierárquicas universais.

Os outros dois tipos de relações hierárquicas, todo-parte e de instância, conforme descrevemos, destinam-se apenas a determinadas situações, que são devidamente explicadas e justificadas nas normas e a respeito das quais, com base na abordagem pragmática na OC, não temos maiores contribuições a fazer.

A análise facetada é aplicada cada vez mais na OC, e no desenvolvimento de tesouros. Ela é mais flexível e adaptável do que as demais regras aplicadas para o estabelecimento de relações hierárquicas e pode ser usada como um item complementar para o desenvolvimento da estrutura hierárquica do tesouro.

Em tesouros que contêm milhares de termos, a estrutura puramente hierárquica pode se tornar confusa em alguns casos e dificultar o uso por parte dos usuários. Nesse sentido, a análise facetada e o uso de indicadores de faceta vão ao encontro do que é defendido na abordagem pragmática na OC, que os tesouros e demais SOC devem servir como um mapa, um guia para o usuário encontrar a informação que necessita para o alcance de seus objetivos.

Relações associativas

Relações associativas cobrem ligações entre pares de conceitos que não são hierarquicamente relacionados, mas são semanticamente e conceitualmente associados de tal forma que a relação entre eles deve ser explicitada no tesouro, pois pode sugerir termos adicionais ou alternativos para uso na indexação ou RI. Como orientação geral para se estabelecer esse tipo de relação estaria o fato de sempre que um termo é usado, o outro deve estar implícito no quadro de referência compartilhado pelos usuários do tesouro. Além disso, um dos termos é frequentemente um componente necessário de qualquer explicação ou definição do outro. O termo células, por exemplo, é necessário para a definição de citologia, bem como o termo aves é necessário para definição de ornitologia. (ANSI/NISO Z39.19, 2005; ISO25964-1, 2011).

A norma ANSI/NISO Z39.19 (2005) define dois grupos principais de relações associativas, entre termos da mesma hierarquia e entre termos que pertencem a hierarquias diferentes. Entre os termos que pertencem à mesma hierarquia recomenda-se que a relação associativa não seja estabelecida entre termos irmãos mutuamente excludentes, por exemplo rosas e narcisos que pertencem à categoria flores. Mas pode ser implementada entre conceitos ligados por uma relação familiar ou derivacional (quando um conceito é derivado do outro). Conforme o exemplo apresentado na figura 1.

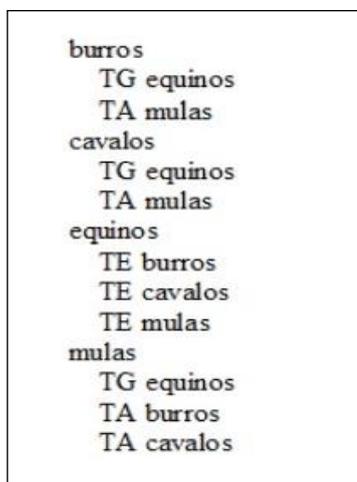


Figura 1: Relação Derivacional
Fonte: ANSI/NISO Z39.19 (2005, p. 52)

Nos casos de termos pertencentes a uma mesma hierarquia, ou termos irmãos, a norma ANSI/NISO Z39.19 (2005) também recomenda que a relação associativa seja estabelecida entre aqueles que possuem significados sobrepostos. Por exemplo, barcos e navios, que podem ser

definidos distintamente com precisão, sem formar um par de termos equivalentes, mas muitas vezes podem ser usados como sinônimos, assim os usuários devem ser alertados ou lembrados, por meio da relação associativa entre eles.

A norma ISO25964-1 (2011) também recomenda o estabelecimento de relação associativa entre termos com significados sobrepostos, sendo este o critério de distinção dos dois grupos principais de relações associativas que essa norma apresenta, relação associativa entre termos com significados sobrepostos (que podem ou não pertencer a uma mesma categoria) e outros casos de ligações associativas.

Esse último grupo descrito na ISO25964-1 (2011), no item ‘outros casos de ligações associativas’, é similar ao grupo de relações associativas entre termos de diferentes hierarquias que a norma ANSI/NISO Z39.19 (2005) descreve, no qual as relações associativas podem ser estabelecidas de acordo com variada gama de critérios, como por exemplo causa/efeito (morte/luto), objeto/propriedade (veneno/toxidade), ação/produto (tecelagem/tecidos) e processo/agente (controle de temperatura/termostato), dentre outros.

O que nos chama a atenção nessa recomendação é a definição de relações associativas como aquelas que podem sugerir termos adicionais ou alternativos para uso na indexação ou RI. Essa descrição inicial vai ao encontro de uma importante discussão presente na abordagem pragmática, sobre o usuário ser informado da existência de pontos de vistas diferentes dos seus.

No entanto, a partir dessa definição, as normas descrevem como orientação geral para se estabelecer esse tipo de relação o fato de que sempre que um termo for usado, o outro deve estar implícito no quadro de referência compartilhado pelos usuários do tesouro. Com base na abordagem pragmática, podemos questionar como esse quadro de referência será identificado, pois se for identificado prioritariamente a partir das opiniões dos usuários sobre o assunto, podem ocorrer alguns problemas. Por exemplo, qual a finalidade de o tesouro trazer o termo ômega 3 relacionado ao termo peixes de águas profundas se a coleção da unidade de informação não contem nenhum documento sobre ômega 3? Ao encontrar o termo no tesouro, os usuários podem criar a expectativa de encontrar material sobre o assunto, mas o resultado da busca será o silêncio.

Esse exemplo diz respeito à identificação de conceitos e de termos, mas ilustra de forma clara o que é defendido na abordagem pragmática, que os componentes dos SOC devem ser identificados a partir da bibliografia da área. Podemos complementar essa ilustração

descrevendo, também como exemplo, que os usuários poderiam considerar que o ômega 3 é um ácido graxo presente na gordura de animais bovinos e suínos, mas na literatura indexada pelo tesouro, o ômega 3 é considerado um ácido graxo encontrado em peixes de águas profundas. Se uma relação associativa for implementada entre “ômega 3” e “animais bovinos” um usuário interessado em informações adicionais sobre o ômega 3, ao verificar tal relação pode decidir pesquisar também “animais bovinos” e não encontrará o que necessitava, pois na bibliografia o ômega 3 é considerado proveniente de peixes de águas profundas e a literatura indexada sobre animais bovinos não conterá informações sobre ômega 3.

O que descrevemos é simplesmente a título de ilustração, sabemos que isso não acontece na realidade, os usuários possuem um certo conhecimento sobre a literatura da área. No entanto, essa literatura é muitas vezes bastante extensa e possui pontos de vista que não são compartilhados por todos os usuários de forma homogênea. Desse modo, e conforme a abordagem pragmática na OC, as relações semânticas e demais componentes dos SOC devem ser identificados a partir da análise da bibliografia da área.

E como a função das relações associativas, conforme é apontado inclusive nas normas para elaboração de tesouros, é sugerir termos adicionais para a indexação e RI e os termos representam conceitos que são definidos a partir de diferentes pontos de vista, a abordagem semântica pragmática se mostra especialmente útil para orientar o estabelecimento desse tipo de relação semântica. Pois, conforme se enfatiza nessa abordagem, um SOC deve servir como um guia para o usuário encontrar a informação que necessita. Essa informação pode ou não ser aquilo que o usuário desejava ou esperava encontrar quando iniciou sua busca, mas deve ser potencialmente útil para que ele alcance seu objetivo. Nesse sentido destacamos os trabalhos de Blair (2003), Hjørland (1998a, 2007b), Hjørland e Christensen (2002), Frohmann (1990, 2009), Gracioso (2010).

Considerações finais

As relações semânticas são elementos essenciais em tesouros e tem recebido menor atenção do que termos e conceitos nas pesquisas da área de CI. Configuram-se, portanto, em um objeto de pesquisa relevante. O estabelecimento de relações semânticas em tesouros pode ser feito de maneiras e de acordo com pressupostos teóricos diferentes, que nem sempre são apresentados de forma explícita, mas isso não significa que não existam. Assim, nessa pesquisa analisamos como a abordagem semântica pragmática pode ser adotada para o estabelecimento de relações semânticas em tesouros.

Para as relações associativas destaca-se a importância da bibliografia para a identificação das relações que podem guiar o usuário a localizar informações úteis ao alcance de seus objetivos. A importância dada à bibliografia talvez não se aplique de forma tão efetiva no caso das relações de equivalência, nas quais a identificação da variedade de termos que podem representar um conceito é essencial e para isso os usuários, e não apenas a bibliografia, podem ser consultados. No caso das relações de equivalência percebemos que os aspectos contextuais envolvidos na construção do conhecimento, que são múltiplos e heterogêneos e acarretam implicações éticas para a OC, podem ser úteis para o desenvolvimento de tesouros e demais SOC que sejam mais orientados à comunicação do que à representação do conhecimento. Da mesma maneira, para a função de comunicação, destaca-se a importância da prática para a formação dos significados.

Um aspecto enfatizado na abordagem pragmática diz respeito à dificuldade em se identificar efetivamente os significados dos termos. Diferentes significados podem ser identificados por diferentes pessoas e comunidades ou, ainda, por uma mesma pessoa ou comunidade que, com o tempo, podem atribuir significados distintos aos termos. A relação de equivalência é importante para unir os diferentes termos em um único ponto de acesso de assunto de uma busca por determinado conceito. Possibilitar a distinção do que foi considerado igual ou equivalente para fins de revocação na RI é uma medida que minimiza as implicações éticas de se escolher uma forma de expressão como termo preferido e proporciona maior flexibilidade ao SOC e ao SRI. Flexibilidade que é uma característica da linguagem e tão importante para a comunicação, mas que os SOC, por vezes, têm como prioridade eliminar, o que, conforme descrevemos, é oposto ao que defende a abordagem pragmática na OC.

Nas relações hierárquicas, assim como nas de equivalência, entender que os significados variam de acordo com o contexto é um aspecto importante, pois se as hierarquias são estabelecidas apenas a partir de premissas universalistas podem se tornar demasiadamente rígidas. O teste todos – alguns, um exemplo clássico da lógica formal e do paradigma positivista em semântica, poderia ser utilizado como critério para a consistência no estabelecimento de relações hierárquicas, mesmo numa abordagem pragmática, contanto que se aplique ao que é aceito como verdade na bibliografia do domínio ao qual o tesouro se destina. Dessa maneira, é um critério aplicado em contextos específicos, e não apenas como um critério geral para identificar relações hierárquicas universais.

Na abordagem pragmática encontramos contribuições para o estabelecimento dos três tipos de relações semânticas dos tesouros, conforme descrevemos. Entre essas contribuições destacam-se aquelas que se referem às relações associativas, pois essas são consideradas as mais difíceis de definir e sobre as quais ainda não existe pesquisa suficiente para determinar suas bases teóricas, conforme apontam Marroni (2006), Austin (1993), Lancaster (1986) e Gomes (1984), dentre outros. Nessa abordagem, as relações associativas devem ser estabelecidas principalmente a partir da análise da literatura, de modo a levar em conta a existência de diferentes pontos de vista ali presentes, e servir como um guia para o usuário localizar informações possivelmente inesperadas e adicionais, mas ainda assim úteis para que alcance seu objetivo.

Referências

ALMEIDA, C. C. Conceito como signo: elemento semiótico para análise e mediação da informação. **Scire**, Zaragoza, v. 18, p. 49-56, 2012a.

_____. Mediação como processo semiótico: em busca de bases conceituais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 5, 2012b.

_____. Pragmatismo e semiótica de Peirce na organização da informação. **Scire**, Zaragoza, v. 16, p. 15-20, 2010.

_____. The methodological influence of Peirce's pragmatism on knowledge organization. **Knowledge Organization**, v. 39, p. 204-215, 2012c.

ALVARENGA, L. Bibliometria e arqueologia do saber de Michel Foucault: traços de identidade teórico-metodológica. **Ciência da Informação**, Brasília, v. 27, n. 3, p. 253-261, 1998.

ANSI/NISO Z39.19:2005. **Guidelines for the construction, format and management of monolingual controlled vocabularies**. Bethesda: NISO, 2005.

AUSTIN, D. **Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngues**. Brasília: IBICT/SENAI. 1993. 83p.

BARDIN, L. **Análise de conteúdo**. Lisboa: 70, 1979. 225 p.

BLAIR, D. C. Information Retrieval and the Philosophy of language. **ARIST**, v. 37, p. 2-50, 2003.

CAFÉ, L.; BRASCHER, M. Organização do Conhecimento: teorias semânticas como base para estudo e representação de conceitos. **Informação e Informação**, Londrina, v. 16, n. 3, p. 25-51, jan./jun. 2011.

FROHMANN, B. Documentary ethics, ontology, and politic. **Arch Sci**, v. 8, 2008a. Disponível em: <<http://dx.doi.org/10.1007/s10502-008-9073-y>>. Acesso em:

_____. Documentation Redux: Prolegomenon to (Another) Philosophy of Information. **Library Trends**, v.52, n. 3, 2004.

_____. Revisiting "what is a document?". **Journal of Documentation**, v. 65, n. 2, 2009.

_____. Rules of indexing: a critique of mentalism in information retrieval theory. **Journal of Documentation**, v. 46, n. 2, 1990. Disponível em: <<http://dx.doi.org/10.1108/eb026855>>. Acesso em:

_____. Subjectivity and information ethics. **JASIST**, v. 59, n. 2, 2008b. Disponível em: <<http://dx.doi.org/10.1002/asi.20742>>. Acesso em:

GOMES, H. E. **Diretrizes para elaboração de tesouros monolíngues**. Brasília: IBICT, 1984. 70p.

GRACIOSO, L. S. Parâmetros teóricos para elaboração de instrumentos pragmáticos de representação e organização da informação na Web: considerações preliminares sobre uma possível proposta metodológica. **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 1, n.1, p. 138-158, 2010.

HJØRLAND, B. Concept theory. **JASIST**, v. 60, n. 8, p. 1519-1536, 2009.
<http://dx.doi.org/10.1002/asi.21082>

_____. Deliberate bias in knowledge organization? **Advances in Knowledge Organization**, v. 11, p. 256-260, 2008a.

_____. Fundamentals of Knowledge Organization. **Knowledge Organization**, v.30, n.2, p.87-111, 2003.

_____. Information: Objective or Subjective/Situational? **JASIST**, v. 58, n. 10, 2007a.
<http://dx.doi.org/10.1002/asi.20620>

_____. Information retrieval, text composition, and semantics. **Knowledge Organization**. v. 25, n. 1-2, 1998a.

_____. Semantic and knowledge organization. **ARIST**, v. 41, p. 367-405, 2007b.
<http://dx.doi.org/10.1002/aris.2007.1440410115>

_____. What is knowledge organization (KO)? **Knowledge Organization**, v. 35, n. 2, p. 86-101, 2008b.

_____.; CHRISTENSEN, F. S. Work tasks and socio-cognitive relevance: a specific example. **JASIST**, v. 53, n. 11, p. 960-965, 2002.

ISO 25964-1:2011. **Thesauri and interoperability with other vocabularies**. Part 1: thesauri for information retrieval. Geneve: International Standard Organization, 2011.

KHOO, C.; NA, J. C. Semantic relations in information science. **ARIST**, v. 40, p. 157-228, 2006.

LANCASTER, F. **Vocabulary control for information retrieval**. Washington, D.C.: Information Resources, 1986. 233p.

MARRONI, G. N. **Identificação e delimitação de relações associativas em tesouros: um estudo de caso na área do direito do trabalho**. 2006. 127 f. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2006.

NOVELINO, M. S. F. A linguagem como meio de representação ou de comunicação da informação. **Perspect. Cienc. Inf.**, Belo Horizonte, v. 3, n. 2, p. 137 - 146, jul./dez.1998.

PEREGRIN, J. Pragmatism und semantik [Pragmatism and semantics]. In: FUHRMANN, A.; OLSSON, E. J. **Pragmatisch denken** [Thinking pragmatically]. p.

89-108. Frankfurt am Main, Germany: Ontos. 2004.

RENDÓN ROJAS, M. A. Hacia um nuevo paradigma em bibliotecología. **Transformação**, Campinas, v. 8, n. 3, p. 17-31, set./dez. 1996.

THELLEFSEN, T.; THELLEFSEN, M. Pragmatic semiotics and knowledge organization. **Knowledge Organization**, v. 31, n. 3, p. 177-187, 2004.

VEGA-ALMEIDA, R. L.; FERNÁNDEZ-MOLINA, J. C.; LINARES, R. Coordinadas paradigmáticas, históricas y epistemológicas de la Ciencia de la Información: una sistematización. **Information Research**, v. 14, n. 2, p.1-20, June 2009. Disponível em: <<http://informationr.net/ir/14-2/paper399.html>>. Acesso em: 14 ago. 2014.

WEISS, L. C. **Relações semânticas em tesouros**: um estudo da abordagem pragmática. Florianópolis, 2014. Dissertação (Mestrado) – Programa de Pós-graduação em Ciência da Informação, Universidade Federal de Santa Catarina, Florianópolis, 2014.

WEISS, L. C.; BRÄSCHER, M. Pragmática na organização do conhecimento. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014: além das nuvens, expandindo as fronteiras da Ciência da Informação. **Anais...** Belo Horizonte: Associação Nacional de Pesquisa em Ciência da Informação, 2014. p. 681-701. Disponível em: <<http://enancib2014.eci.ufmg.br/programacao/anais-do-xv-enancib>>. Acesso em: 26 nov. 2014.

Artigo submetido em: 22 jun. 2015

Artigo aceito em: 21 jun. 2016