

TESAUROS E TERMINOLOGIA

Regina Helena van der Laan, CRB-10/514*
Glória Isabel Sattamini Ferreira, CRB-10/176**

RESUMO: Focaliza o controle de termos utilizados no processo de indexação, como fator facilitador ao acesso às informações. Conceitua tesouro, partindo do princípio que a indexação se faz por conceitos e que estes são representados por meio de termos. Estabelece a relação entre o tesouro e a terminologia, pois, ambos tem como objeto de trabalho o termo. Busca nos princípios da Terminologia elementos para a construção de Tesouros com a finalidade de obter uma maior precisão do vocabulário controlado facilitando a recuperação de informações.

Palavras-chave: tesouro; terminologia; recuperação da informação.

INTRODUÇÃO

A proposta deste trabalho surgiu das reflexões que temos realizado ao longo do nosso exercício profissional, principalmente, como docentes na área de Representação Temática e supervisoras do estágio curricular do Curso de Biblioteconomia.

O processo de indexação por assuntos reveste-se de um profundo e cuidadoso trabalho de análise dos textos e a representação dos temas tratados por uma expressão que pode ser alfabética ou numérica.

A complexidade do processo de indexação cada vez mais se amplia, uma vez que os sistemas de recuperação da informação (SRI) estão conectados em rede, não mais atendendo apenas um público específico. Hoje não mais podemos traçar um claro perfil de nosso usuário e de suas necessidades. A automação agilizou o acesso as

* Regina Helena van der Laan - Professora Assistente do Departamento de Ciências da Informação da UFRGS. Mestre em Educação pela PUC/RS. Doutoranda em Letras pela UFRGS. E-mail: vanderla@adufgrs.ufrgs.br

** Glória Isabel Sattamini Ferreira - Professora Assistente do Departamento de Ciências da Informação da UFRGS. Mestre em Educação pela PUC/RS. E-mail: glo@pro.via-rs.com.br

informações, mas se tornou um fator preocupante para o indexador, que necessita um maior controle sobre o vocabulário utilizado para realmente atender as necessidades de informações de um grupo heterogêneo de usuários.

Percebe-se vários problemas de recuperação das informações devido ao uso de vocabulários livres ou motivados pela imprecisão de definição dos termos para os vocabulários controlados. A falta de padronização gera a perda de informações, pois no momento da busca somente serão recuperados os documentos que coincidentemente foram indexados pelo termo usado na busca.

No processo de indexação estamos lidando com os conceitos de uma área específica do conhecimento, estes são representados através de termos. O termo é o objeto de estudo da Terminologia. Partindo destas reflexões pretendemos neste trabalho destacar princípios da Terminologia que possam ser utilizados na construção de Tesauros, visando maior precisão na definição dos descritores.

LINGUAGEM INDEXAÇÃO

Em indexação trabalhamos com os conceitos expressos nos documentos pelos autores. Apesar de indexarmos palavras, estas tem que serem entendidas como rótulos que podem ser expressos de várias formas. O indexador está trabalhando com uma rede conceitual de uma área específica do conhecimento. A compreensão do que é conceito e de suas relações irá facilitar o trabalho de indexação, que não mais será feito por palavras, mas sim por termos representativos daqueles conceitos.

“En Terminología se parte del concepto y se busca su denominación, aunque para identificar y fijar un concepto es indispensable contar co una denominación o con algún otro signo. (Wüster, 1998, p.39)

Para Wüster (1998) conceito são as características comuns de um determinado objeto percebidos pelos seres humanos.

Este autor enfatiza que conceito é um elemento do pensamento. O conjunto das características individuais do objeto é a intenção conceitual. Em oposição temos a extensão conceitual que podemos definir como o conjunto de objetos individuais que possuem as mesmas características.

Os termos que representam estes conceitos constituem o vocabulário utilizado em indexação ou uma linguagem documental.

Segundo Guinchart e Menou (1994, p.133) “A linguagem documental é uma linguagem convencional utilizada por uma unidade de informação para descrever o conteúdo dos documentos. Com o objetivo de armazená-los e recuperar as informações que eles contém.”

Linguagens de indexação ou linguagem documentária, para Van Slype (1991,p.21), é “. . . todo sistema de signos que permita representar el contenido de los documentos con el fin de recuperar los documentos pertinentes en respuesta a consultas que tratan sobre ese contenido.”

As linguagens documentárias podem ser notacionais, como os sistemas de Classificação Decimal de Dewey (CDD) ou Classificação Decimal Universal (CDU), ou alfabéticas, como os cabeçalhos de assunto ou os tesauros. Estes sistemas podem ser ainda classificados por sua abrangência. Os sistemas anteriormente citados, são universais, abrangem todas as áreas do conhecimento humano. Existem, ainda, os sistemas específicos, cuja abrangência refere-se a uma única área do conhecimento.

Qualquer que seja a linguagem utilizada é necessário se estabelecer uma relação entre a linguagem natural usada pelo usuário e a linguagem utilizada pelo sistema. de recuperação da informação.

Para isto, são estruturados instrumentos de controle de termos buscando estabelecer um vocabulário comum, visando facilitar o processo de comunicação entre o usuário e o sistema.

Estes instrumentos são os vocabulários controlados que segundo Barité (1997) podem ser entendidos como: “Conjunto estruturado de conceptos destinado a la representación del contenido de los documentos, y que comprende la organización lógica de dicho conceptos en distribuciones por clases o disciplinas, y las relaciones recíprocas y estables que mantiene entre ellos.” Como por exemplo os Tesauros; os Sistemas de Classificação; os Cabeçalhos de Assunto.

Tesouro é um exemplo de vocabulário controlado. A palavra *Tesouro* tem sua origem etimológica no latim thesaurus, que se originou do grego thesaurós; tinha o significado de tesouro ou armazém/repositório de palavras. Segundo Motta (1987, p.25), Tesouro é um “Sistema de vocabulário baseado em conceitos, incluindo termos preferidos (descritores), termos não preferidos (não descritores) e suas inter-relações, que se aplica a um determinado ramo do conhecimento e que se destina a controlar a terminologia utilizada para a indexação/recuperação de documentos.”

Podemos definir Tesouro, ainda, como: “Linguagem documentária dinâmica que contém termos relacionados semântica e logicamente, cobrindo de modo compreensivo um domínio do conhecimento.” (Gomes, 1990, p.16), ou ainda, como uma “. . . lista estruturada de termos associados, empregados por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, a (sic) nível de entrada, e para permitir aos pesquisadores a recuperação de informação que procura.” (Cavalcanti, 1978, p.27)

O Tesauro tem como função principal o controle terminológico do vocabulário utilizado em uma área específica do conhecimento, indicando as relações entre os conceitos a serem indexados. Em síntese podemos afirmar que tesauro é:

- a) linguagem especializada;
- b) estruturado conforme rede conceptual, apresenta relações hierárquicas (gênero/espécie; todo/parte) e relações associativas;
- c) estabelece preferência entre os termos através das relações de equivalência, determinando o termo preferido, forma de grafia preferida; uso de siglas, etc.

Podemos verificar a estreita relação entre este instrumento de controle de vocabulário e a Terminologia, ambos tratam de conceitos de uma área especializada e buscam uma linguagem comum entre os especialistas, como fator facilitador da comunicação de novos conhecimentos.

Na construção de um Tesauro é estabelecido um controle do vocabulário visando com que cada conceito seja expresso por um único e inequívoco termo ou descritor. Para este fim utiliza-se de várias fontes, tais como outros Tesauros da mesma área ou área afim, dicionários, vocabulários, esquemas de classificação, índices de publicações periódicas, assim como outros documentos da literatura especializada em que se vão controlar os termos.

Deve ser observado, ainda, o fato que os conceitos se organizam em um sistema conceitual e, portanto, estabelecendo uma rede de relações entre eles. Conforme Dahlberg (1978), estas relações podem ser:

Relações Hierárquicas ou Abstrativas

Quando dois conceitos apresentam características idênticas, um deles possui uma característica a mais. relação gênero/espécie

Exemplo: árvore
 árvore frutífera
 macieira

Relações Partitivas

Quando um dos conceitos é parte do outro, ou, relação todo/parte

Exemplo:
 árvore
 raízes, tronco, galhos, folhas, flores, frutos

Nestes casos estrutura-se uma relação de subordinação, pois a espécie está subordinada ao gênero e as partes ao todo.

Relações de Oposição

Quando um conceito é o oposto, o contrário do outro.

Exemplo: preto - branco

As relações hierárquicas e as partitivas referem-se principalmente as relações entre os objetos, já as relações de oposição referem-se mais a relação entre propriedades dos conceitos.

Relações Funcionais

Aplica-se principalmente a conceitos que expressam processos.

Exemplo: Produção - produto - produtor - comprador

Relações Associativas

Estas relações podem ocorrer entre os conceitos, por exemplo, como:

a) Disciplina - objeto estudado

Entomologia/Insetos

b) Processo ou operação - Agente ou instrumento

Aquecimento/Combustíveis

Automação/Computadores

c) Ação - Resultado

Tecelagem/Tecidos

Além das relações entre os conceitos, o indexador deve observar os casos de palavras sinônimas, quase sinônimas ou indicar a preferência de grafia. Nestes casos estabelece-se uma relação de equivalência entre os descritores de um mesmo conceito.

TERMINOLOGIA E TESAURO

A comunicação entre os indivíduos muitas vezes é prejudicada devido à utilização de termos que não possuem o mesmo significado entre os interlocutores.

A comunicação se faz através da linguagem, escrita, falada ou simbólica. Os signos utilizados devem fazer sentido tanto para o emissor como para o receptor da mensagem. Na área científica é fundamental que usemos uma terminologia compreensível aos nossos pares.

Para Currás (1995, p.23) terminologia é um “processo ontológico pelo qual se chega à normalização dos sistemas formados pelas diferentes linguagens especializadas, de onde o conceito fixado pelo

termo está em relação semântica com o resto dos termos desse sistema.” A autora conclui dizendo que: “A finalidade da terminologia se resume no melhor entendimento dos especialistas entre si e destes com o seu meio.” (op. cit., p.23)

Segundo Cabré (1993) o interesse pela terminologia tem sua origem na preocupação dos cientistas dos séculos XVIII e XIX, com a proliferação de termos e suas relações que dificultavam a comunicação.

O grande desenvolvimento científico, que vem ocorrendo, principalmente no final deste século, tem provocado o surgimento de novos conceitos, novas áreas de especialização, proporcionando o aparecimento de novos termos e, como no passado, dificultando a comunicação.

O vocabulário utilizado pelo especialista já não pertence exclusivamente a um grupo seleto de pesquisadores, os meios de comunicação socializam esta terminologia, muitas vezes, mais confundido o leigo do que o auxiliando a compreender os fatos divulgados. Os próprios pesquisadores utilizam-se de vocabulários de acordo com seus grupos de pesquisa, não existindo padronização nem mesmo em uma mesma área do conhecimento.

Segundo Cabré (1993, p.43) “. . . la terminología nace de la necesidad manifestada por los especialistas de ordenar las denominaciones de sus sistemas de conceptos, con la finalidad de conseguir una comunicación profesional más fiable.”

Um dos aspectos postulados pela terminologia é a normalização dos termos, fixando o uso de um termo e descartando a utilização de outros termos para o mesmo conceito. A padronização do vocabulário técnico e científico, melhora a comunicação entre os especialistas área. Este é sem dúvida um dos objetivos da Ciência da Informação,

considerando que é a linguagem o veículo para a indexação e transferência das informações.

A padronização do vocabulário contrapõe-se a dinâmica da língua. As normas são estáticas a língua é dinâmica. O falante está constantemente criando novos termos, alterando os padrões.

Segundo Faulstich (1998, p.2), "A polifuncionalidade da unidade lexical, no discurso científico, no discurso técnico ou no discurso de vulgarização científica pode produzir mais de um registro ou mais de um conceito para o mesmo termo."

A autora continua afirmando que:

"Como método para classificar as variantes terminológicas, criaram-se critérios de classificação, de acordo com as tendências sistemáticas em que os termos se repartem. Surgiram, portanto, dois grupos. Um de variantes terminológicas propriamente lingüísticas e outro de variantes terminológicas de registro." (1998, p.4)

As variações terminológicas lingüísticas classificam-se em:

- a) variante terminológica fonética - o registro pode ser feito de acordo com a forma falada;
- b) variante terminológica morfológica - há alternância de estrutura de ordem morfológica;
- c) variante terminológica sintática - duas construções sintagmáticas alternam-se com função de predicação de uma UTC (Unidade Terminológica Conceitual);
- d) variante terminológica lexical - há o apagamento de algum item da estrutura lexical sem alterar o conceito;

e) variante terminológica gráfica - o termo é registrado sob mais de uma forma.

As variantes terminológicas de registro ocorrem no plano horizontal, no plano vertical e no plano temporal, classificando-se em:

a) variante terminológica geográfica - são expressões para designar um mesmo conceito utilizado por falantes de uma mesma língua em regiões diferentes. Por exemplo, sinaleira utilizado no Rio Grande do Sul, farol no nordeste ou semáforo no sudeste. Ocorrem no plano horizontal.

b) variante terminológica de discurso - são as expressões utilizadas para um mesmo conceito nos diferentes níveis de discurso, nível científico, técnico ou de divulgação científica. Isto ocorre para permitir o entendimento do texto pelos diferentes públicos.

c) variante terminológica temporal - são designações para um mesmo conceito que concorrem durante um período de tempo simultaneamente, firmando-se a preferência por uma em substituição a anterior e deixando-a em desuso.

Como podemos observar alguns princípios norteadores da Terminologia podem auxiliar na construção de Tesauros possibilitando uma maior precisão no momento da determinação dos descritores. A padronização dos termos que representam determinados conceitos em uma área do conhecimento, tem o objetivo de facilitar a comunicação entre o produtor do conhecimento e o usuário de conhecimento. Desta forma acreditamos que a relação das duas áreas, através da associação de métodos e do aporte da lingüística vindo da Terminologia, poderá resultar na estruturação de vocabulários controlados mais adequados às necessidades e compreensão dos usuários.

Concluiríamos afirmando que ambas as áreas possuem um caráter inter e transdisciplinar. Ambas têm o mesmo objeto de trabalho, o conceito. E que uma se encontra a serviço da outra. Como Cabré (1993,

p.112) afirma: “La terminología constituye una de las bases de los tesauros y clasificaciones temáticas, puesto que la expresión de cualquier relación de contenido entre los conceptos se efectúa a través de un término.”

THESAURO AND TERMINOLOGY

ABSTRACT: The aim of this paper is the control of terms used in the indexation process, as provide factor to the access to the information. It defines the thesauro and focuses indexation starting from concepts and these are representing by terms. It set up the relation between tesauro and the Terminology. The term is the work object of these two areas. It look for principles of the Terminology elements for the construction of Tesauros, with the purpose to get more accuracy of the controled vocabulary in the information retrieval.

REFERÊNCIAS BIBLIOGRÁFICAS

BARITÉ, Mario Guido. **Glosario sobre Organización y Representación del Conocimiento**: Clasificación, Indización, Terminología. Montevideo, CSIC, Indice, 1997.

CABRÉ, Maria Teresa. **La Terminología**: teoria, metodologia, aplicaciones. Barcelona, Antártida/Empúries, 1993.

CAVALCANTI, Cordelia R. **Indexação e Tesauro metodologia e técnicas**. Brasília, ABDF, 1978.

CURRÁS, Emília. **Tesauros: linguagens terminológicas**. Brasília, IBICT, 1995.

DAHLBERG, Ingetraut. Teoria do Conceito. **Ciência da Informação**, Rio de Janeiro, v.7, n.2, p.101-107, 1978.

FAULSTICH, Enilde. Termo e Variação: Tendências no Português do Brasil. In: **Socioterminologia**. Brasília, Unb, [1998]. (Excerto, parte II)

GOMES, Hagar Espanha (coord.) **Manual de Elaboração de Tesouros Monolíngües**. Brasília, Programa Nacional de Bibliotecas das Instituições de Ensino Superior, 1990.

GUINCHAT, Claire; MENOU, Michel. **Introdução Geral às Ciências e Técnicas da Informação e documentação**. Brasília, IBICT, 1994.

MOTTA, Dilza Fonseca da. **Método Relacional como Nova Abordagem para a Construção de Tesouros**. Rio de Janeiro, SENAI/DN/DPEA, 1987

VAN SLYPE, Georges. **Los Lenguajes de Indezación**: concepción, construcción y utilización en los sistemas documentales. Madrid, Salamanca; Fundación Germán Sánchez Ruipérez, 1991

WÜSTER, Eugen.. **Introducción a la Teoría General de la Terminología y a la Lexicografía Terminológica**. Barcelona, Institut Univertari de Lingüística Aplicada/Universitat Pompeu Fabra, 1998.