

Metodologia de elaboração de tesauro conceitual: a categorização como princípio norteador

■ **Maria Luiza Almeida Campos**

Doutora em Ciência da Informação, professora do Departamento de Ciência da Informação e da Pós Graduação em Ciência da Informação UFF/IBICT

■ **Hagar Espanha Gomes**

Livre docente

Historicamente, o tesauro documentário surgiu como uma relação estruturada de termos constituídos, quase exclusivamente, de uma única palavra. Posteriormente evoluiu para algumas palavras compostas sem, no entanto, estabelecer bases para isso. A elaboração de um tesauro implica bases classificatórias e as categorias fundamentais permitem correto posicionamento dos conceitos nas classes e a organização de domínio; e a definição de cada conceito é o elemento que vai comprovar objetivamente tanto as relações hierárquicas como as partitivas e associativas com outros conceitos.

Palavras chave: Tesauro conceitual; Metodologia; Categorização

Recebido em 01.07.2006

Aceito em 19.09.2006

1 Introdução

Tesouro tem sido definido como linguagem documentária. Um olhar mais atento sobre o objeto/referente, evidencia que se trata de uma relação de termos de um domínio, relacionados entre si, com objetivo de indexação/recuperação em um sistema de recuperação de informação. A palavra *linguagem* está sendo tomada em sentido figurado, pois não lida com discursos, apenas deles extrai um vocabulário especializado, visando à garantia literária. O tesouro em ação num serviço de recuperação de informação requer instruções que poderiam ser, também, numa linguagem figurada, uma gramática; esse conjunto – *vocabulário* e *gramática* se constituiriam, então, numa linguagem documentária. Este texto se limita ao *vocabulário*: sua natureza e estrutura.

De fato, o vocabulário do tesouro não é constituído de palavras da linguagem natural – como ocorreu nos primeiros tempos de seu surgimento – mas de uma lista de termos¹, entendendo-se aqui como termo o signo verbal que designa um referente, ou, num nível maior de abstração, que denota o conceito. A análise de tal referente se dá a partir de um determinado domínio. Assim, o contexto do tesouro é o domínio e não o discurso. Após análise do objeto/referente a que aquelas palavras ou expressões se referem, tem-se o termo, com sua forma final estabelecida, bem como as relações entre ele e outros conceitos, revelados na análise. Por sua natureza verbal, a denominação do objeto/referente mascara, por vezes, sua verdadeira natureza, dificultando sua posição na estrutura do tesouro e é a análise conceitual que irá revelá-la. Tesouro conceitual seria, então, um tesouro com base em conceitos: seu nome indica que cada termo denota um conceito, ou seja, uma unidade de conhecimento.

Na construção de um tesouro conceitual são necessários princípios para:

- Estabelecimento do termo/conceito, e
- Estabelecimento das relações entre eles.

As bases teórico-metodológicas aqui apresentadas se propõem a fornecer elementos para a elaboração de tesouros, sem descurar, no entanto, a dinâmica que ocorre no conhecimento, de cujos conceitos os termos decorrem – quando se requer *neonímia* –, dinâmica que afeta ou a própria denominação ou o conteúdo dos termos. São clássicos os exemplos de *átomo* e *hipnose*.

A teoria do conceito, voltada para o referente, e a teoria da classificação facetada contribuem para a elaboração de tesouros conceituais porque estabelecem bases para identificação dos conceitos, dos termos e das relações entre eles, e, ainda, para sua ordenação sistemática.

Desde os anos oitenta temos associado a teoria do conceito à teoria da classificação para a elaboração de tesouros, (GOMES, 1990). De fato, Campos (CAMPOS, 1986) já havia alertado para o sistema de classificação subjacente ao tesouro. Mas, ao contrário de tesouros que consideram a parte sistemática como uma parte alfabética dos termos gerais com sua respectiva hierarquia, preconizamos o uso das categorias Ranganathianas para ordenar as classes gerais (facetas) no interior de cada uma delas, construindo um verdadeiro sistema e revelando, assim, a estrutura conceitual de um dado domínio. A definição construída para cada termo/conceito torna explícito o

¹ A maioria dos tesouros atualmente se refere a termos, mas as decisões adotadas para o tratamento de grupos de palavras – impropriamente chamados de termos compostos – mostra que a abordagem ainda é lingüística e não conceitual.

conteúdo do conceito e fornece os elementos para o relacionamento entre os termos/conceitos.

O conceito é, de fato, o ponto de partida para estabelecer as relações conceituais e determinar a forma verbal mais adequada para representá-lo. O conceito é entendido como *unidade de conhecimento*, como propõe Dahlberg (1978), incluindo a definição como elemento essencial para a fixação do conceito.

Este novo tesouro, ou melhor, esta nova metodologia para a elaboração de tesouros, está, assim, fundamentada nas questões que envolvem o conceito e as categorias. Como toma por base o termo, é denominado, por alguns estudiosos, tesouro terminológico. Embora Currás (1995) em seu livro *Tesouros*, inclua o subtítulo linguagens terminológicas, ela não integra a terminologia na construção do tesouro, limitando-se a apresentar o tema num capítulo isolado. Considera-se que a expressão tesouro terminológico não é apropriada, pois o que se quer evidenciar é o conteúdo conceitual de uma etiqueta lingüística e não a palavra ou a expressão verbal. Assim, as bases para a elaboração do tesouro conceitual se encontram fundamentalmente nos princípios que enfatizam o processo de conceituação e em sua ordenação sistemática.

2 A elaboração de tesouros: um histórico

A palavra *thesaurus* etimologicamente vem do grego e do latim e significa *tesouro* tendo sido usado durante muitos séculos para designar léxico, ou *tesouro* de palavras. Esta palavra popularizou-se a partir da publicação do *Thesaurus of English Words and Phrases*, de Peter Mark Roget, em Londres, 1852. O subtítulo de seu dicionário expressa bem o objetivo: *classified and arranged so as to facilitate the expression of ideas and to assist in literary composition* (ROGET, 1925)

Em 1950, Hans Peter Luhn, do Research Center da IBM nos Estados Unidos, foi o primeiro a utilizar o termo *Thesaurus* para nomear seu sistema de palavras autorizadas com uma estrutura de referências cruzadas (FOSKETT, 1985, p. 270). Ao organizar seu sistema ele percebeu que uma simples listagem alfabética não solucionaria o problema de localizar a palavra/ idéia mais adequada à recuperação. Alguma relação entre estas palavras deveria ser estabelecida, para que a lista pudesse apresentar uma estrutura mais sólida de referências cruzadas; era necessário definir, de alguma forma, as famílias de noções entre as palavras, isto é, evidenciar que idéias afins estavam ligando uma palavra à outra. Ao nomear esta nova lista de palavras autorizadas deu o nome de *thesaurus*, influenciado pelo trabalho de Roget que, em seu dicionário analógico, expõe esta questão da seguinte forma:

a revisão de um catálogo de palavras de significado análogo vai sugerir, com freqüência, por associação, outras sucessões de pensamento. A apresentação dos assuntos sob aspectos novos e variados pode expandir grandemente a esfera de nossa visão mental (ROGET, 1925, p. XVIII).

Deste modo, um novo tipo de linguagem documentária está nomeado - o tesouro de recuperação de informação - que veio se contrapor às listas de cabeçalhos de assunto e servir como instrumento de auxílio aos sistemas que utilizavam um único termo (unitermo). Outras listas de

termos que apresentavam alguma relação entre eles passaram a chamar-se, também, tesouro.

Vickery, nos anos 60, demonstra quatro significados diferentes usados na literatura de ciência da informação para o termo *tesouro*, sendo a interpretação mais comum a de uma lista alfabética de palavras, onde cada palavra é seguida de outras que a ela se relacionam (VICKERY, *apud* FOSKETT, 1985, p. 270).

No início da década de 70, através do programa UNISIST (UNESCO, 1973, p. 6), a Unesco, define *tesouro* para a área de ciência da informação, sob dois aspectos:

a) Segundo a estrutura:

É um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento.

b) Segundo a função:

É um dispositivo de controle terminológico usado na tradução da linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema (linguagem de documentação, linguagem de informação) mais restrita.

Percebe-se a preocupação da Unesco em apresentar definições que pudessem atender tanto à área de elaboração de tesouro (definição *a*), quanto à área de organização e recuperação de informação (definição *b*). Estas definições, de alguma forma, vêm sendo utilizadas na literatura até os dias atuais.

A evolução histórica do *tesouro de recuperação* pode ser traçada a partir de duas vertentes: uma, que toma nitidamente como base o Unitermo uma única palavra -, e outra, influenciada pela teoria da classificação facetada.

Lancaster apresenta um diagrama – FIG. 1 – em que evidencia a evolução histórica do tesouro. De um lado, a influência da América do Norte, caracterizada pela abordagem alfabética, e, de outro, a abordagem sistemática, de influência inglesa principalmente.

2.1 A vertente norte-americana

Os tesouros elaborados na América do Norte, especificamente nos Estados Unidos, foram fruto do desenvolvimento que ocorreu a partir do cabeçalho de assuntos para o unitermo; talvez fosse mais adequado fazer referência a uma ruptura, porque se introduziu um novo modelo, estranho ao anterior: de um sistema pré-coordenado para sistemas pós-coordenados.

Durante a década de 60, os tesouros foram sistematicamente aperfeiçoados. Em 1960, o Centro de Informação do Ministério da Defesa dos Estados Unidos (anteriormente *Armed Services Technical Information Agency* - ASTIA) produz seu primeiro tesouro. Em 1961, o *American Institute of Chemical Engineers* – AICHE - publica o *Chemical Engineering Thesaurus*, que era um derivado direto do trabalho da Dupont, primeiro tesouro a ser colocado à venda (LANCASTER, 1986, p. 31). Baseado no AICHE, o *Thesaurus of Engineering Terms* é publicado, em 1964, pelo

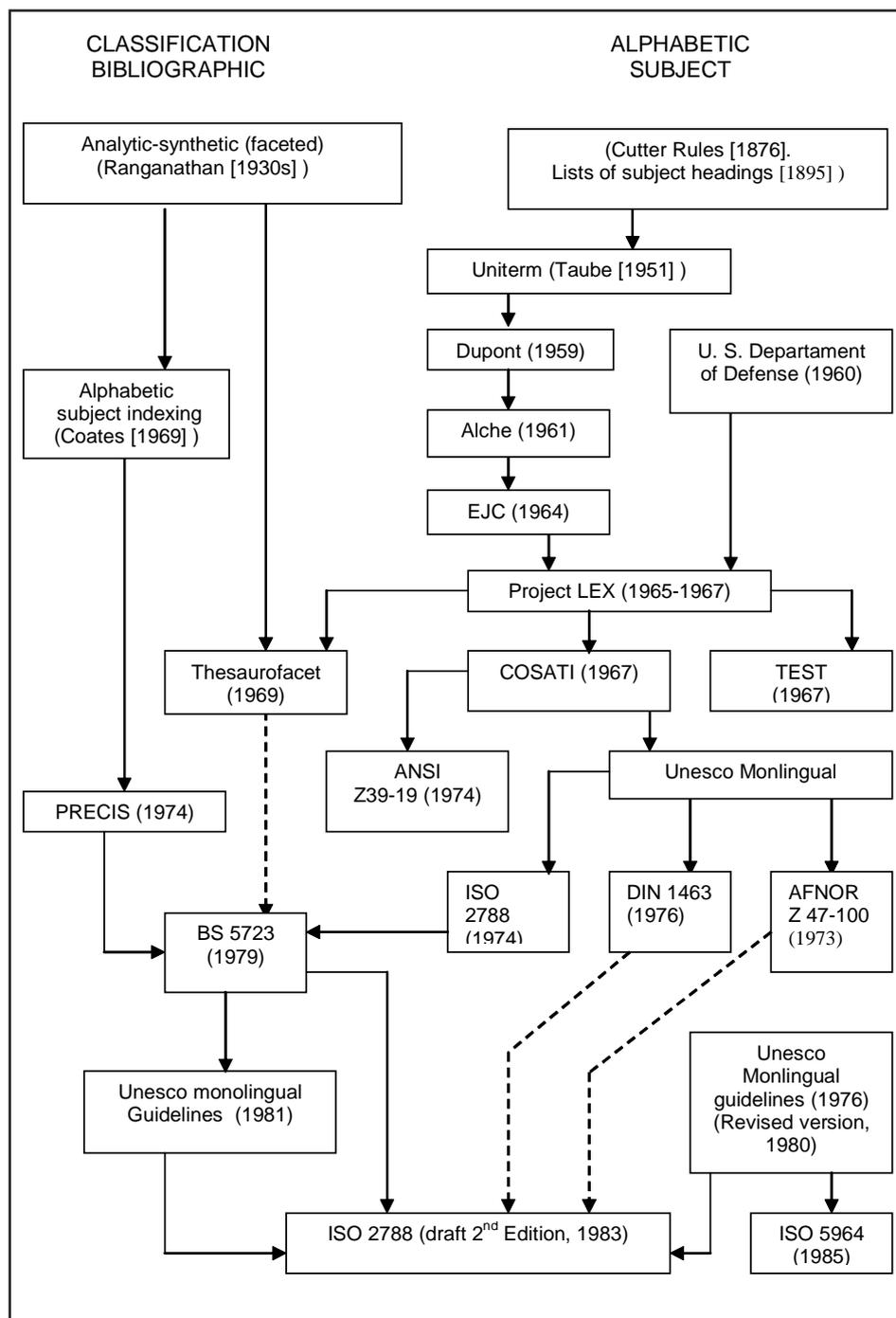


Figura 1 – Evolução histórica

Fonte: LANCASTER, 1986

Engineers Joint Council - EJC - , com a finalidade de cobrir toda a área de engenharia. No ano seguinte, o *Engineers Joint Council* e o *Department of Defense* - DoD - dos Estados Unidos estabeleceram um acordo visando a reunir ambos os tesouros, que abarcavam temas ligados à engenharia, com considerável participação da indústria. Através do projeto *Lex*, cuja meta era estabelecer princípios comuns de construção e uso e, também, criar um único instrumento para ambas as instituições, prepara-se um manual de construção de tesouros e publica-se o *Thesaurus of Engineering and Scientific Terms* - TEST - , em 1967. Endossado e publicado pelo *Committee on Science and Technical Information* - COSATI - , órgão oficial do *Federal Council for Science and*

Technology dos Estados Unidos, o manual foi recomendado como fonte para a construção de tesouros, tendo servido de base para as diretrizes e normas produzidas, posteriormente, pela ANSI - *American National Standardization Institute* (ANSI, 1981) e pela Unesco (1973). Por sua vez, as diretrizes da Unesco deram origem às normas nacionais e internacionais. (LANCASTER, 1986).

2.2 A vertente europeia

Até o final dos anos 50 e início dos anos 60, os tesouros possuíam somente arranjo alfabético, tipo de arranjo incapaz de representar bem as relações úteis entre os termos (RIVIER, 1992, p. 72). Suas limitações levaram ao emprego, mais tarde, de princípios de classificação.

Esta vertente, mostrada no diagrama de Lancaster FIG.1, produziu vários tesouros e contribuiu para o desenvolvimento teórico dos tesouros. De especial relevância é a criação do *Classification Research Group* - CRG -, em Londres, voltado para desenvolver estudos e pesquisas voltadas para a classificação facetada, de Ranganathan, do qual fizeram parte expoentes como Vickery, Foskett, Aitchison, Farradane, Austin, Mills e tantos outros, que influenciaram toda uma geração de profissionais interessados em representação e recuperação da informação.

As bases classificatórias do tesouro foram expostas por Jean Aitchison, a partir de sua experiência no desenvolvimento do *Thesaurifacet*, dedicado à engenharia elétrica. A partir da teoria da classificação de Ranganathan ela constrói o esquema de classificação, com orientação para a organização de uma sintaxe (uso da notação para construir a classificação de um documento) tendo um índice alfabético, sob forma de tesouro, decorrente do próprio sistema de classificação.

O pressuposto das categorias, presente na abordagem analítico-sintética de Ranganathan também teve impacto nos sistemas pré-coordenados, ao oferecer uma sintaxe mais adequada à organização e recuperação da informação. São representativos os princípios da indexação alfabética de assunto, desenvolvidos por Coates para o catálogo de assuntos do *British Technology Index*, e o PRECIS - *Preserved Context Indexing System* -, de Austin, desenvolvido para o índice impresso da *British National Bibliography*.

Apesar dessa vertente avançar quanto aos aspectos relacionados a uma apresentação sistemática, muitos tesouros apresentam um comportamento semelhante ao dos tesouros americanos, que adotam base lingüística na construção do vocabulário. Esta afirmativa pode ser comprovada em normas nacionais (AFNOR Z 47-100, BSI 5723, ANSI/NISO Z 39-2005) e internacionais (UNESCO e ISO 2788), entre outras.

2.3 O tesouro conceitual: o conceito e as categorias

Como apresentado por Lancaster (1986), o tesouro surgiu como uma ruptura em relação ao cabeçalho de assunto, tomando por unidade a palavra, em geral uma palavra técnica, ou *term* em inglês. No entanto, logo se percebeu que algumas palavras sozinhas eram insuficientes para

designar um conteúdo de informação. A solução foi considerar a possibilidade de que, em alguns casos, a unidade seria uma expressão verbal; as bases teóricas para seu estabelecimento, no entanto, não foram desenvolvidas de forma satisfatória (GOMES & CAMPOS, 2004). Diretrizes e normas continuam a se referir a pré-coordenação e fatoração sintática, abordagens lingüísticas que privilegiam a forma lexical e, portanto, variam de tesouro para tesouro, de norma para norma e de língua para língua. E interferem no relacionamento entre os termos. De fato, a maioria dos tesouros produzidos nos Estados Unidos, dentro da linha tradicional americana que privilegia a organização alfabética dos termos, não apresentam uma base teórica explícita. Mas seus autores avançam em relação ao modelo anterior, o cabeçalho de assunto: a unidade de trabalho passa a ser o termo e não o assunto; os diferentes tipos de relação, que nos cabeçalhos vinham sob forma de referência cruzada, se apresentam de forma mais estruturada. No entanto, as instruções para a seleção da unidade de trabalho ainda apresentam fortes bases lingüísticas; aproveitá-las em outras línguas é sempre problemático, porque não há uma correspondência lexical nas diferentes línguas para a denominação de um mesmo conceito. A norma para elaboração de vocabulários controlados (ANSI/NISO Z39.19-2005) sugere, inclusive, o uso de adjetivos, em algumas circunstâncias, evidenciando seu divórcio do conceito como unidade de representação.

Este aspecto, o da unidade de representação, não resolvido pelas duas grandes vertentes - a americana e a dos classificacionistas - parece encontrar solução a partir da década de 70 com Dahlberg em sua *Teoria Analítica do Conceito voltada para o Referente* (aqui referenciada apenas como teoria do conceito) (DAHLBERG, 1978).

Dahlberg afirma que o conceito se estabelece quando três elementos estão presentes: o referente/objeto, suas propriedades e um termo que o sintetize, ou seja, que carregue com ele todas as propriedades do referente. Diz-se, então, que o termo denota o conceito. Então, é por meio do conceito que se estabelece a relação do termo com o objeto/referente. Influenciada pelos analíticos, ela desenvolve sua teoria que fornece bases seguras para estruturar as hierarquias (relações lógicas). Posteriormente, tendo tomado conhecimento do método analítico-sintético de Ranganathan, incorpora o princípio de categorização como uma solução para a organização dos conceitos num sistema de conceitos, não importando a finalidade de aplicação e propõe princípios para o estabelecimento de relacionamentos associativos. (DAHLBERG, 1978a). Seu método leva à elaboração da definição analítica ou conceitual. De fato, segundo Dahlberg, existe uma relação de equivalência entre o termo e sua definição. Como os tesouros se inserem num domínio, é este que vai delimitar os aspectos, as propriedades - ou seja, as características dos conceitos - que figurarão na definição. Dahlberg, como Ranganathan, parte da *lógica*, para o estabelecimento das relações hierárquicas; mas avança fornecendo os elementos para outros relacionamentos, desnecessários num esquema de classificação.

Em nosso ponto de vista, o tesouro conceitual reúne dois princípios fundamentais para sua elaboração: o conceito, como unidade de representação, e o uso de categorias, como base para organização de sistemas de conceitos. No Brasil, as primeiras tentativas têm início na década de 80. (GOMES, 1990).

3 O método de categorização e sua importância para a elaboração de tesouros conceituais

Categorização, simplesmente, é o processo cognitivo de dividir o mundo da experiência humana em grupos gerais ou categorias amplas compreendendo certos componentes que compartilham similaridade imediata em termos de atributos num dado contexto (BINWAL, 2001).

O uso das categorias para a organização de conceitos em uma determinada área de interesse foi introduzido por Ranganathan no âmbito da documentação, a partir de sua teoria da classificação facetada, na qual utiliza a noção de categoria para a análise dos assuntos contidos nos documentos e para a organização dos componentes desses assuntos (isolados) em um esquema de classificação. Dahlberg utiliza, também, a noção de categoria sob dois aspectos: como um recurso para o entendimento da natureza do conceito e para a formação de estruturas conceituais. Entretanto, Dahlberg não considera estes dois aspectos de forma excludente, muito pelo contrário, ela os vê como complementares.

As categorias possuem a propriedade de possibilitar a sistematização de todo o conhecimento da realidade e podem ser identificadas no momento da determinação do conceito, ao serem inferidas predicções verdadeiras e finais a respeito de um item de referência desta realidade observada. (DAHLBERG, 1978a, p. 21). Mas fornecem, ainda, princípios para estruturar todas as classes de conceitos de um domínio.

Na elaboração de tesouros, esta metodologia fornece os princípios para agrupar conceitos de mesma natureza em classes gerais ou facetadas para construir cadeias e renques, séries verticais e horizontais de conceitos, respectivamente. O entendimento das classes que compõem um dado domínio é de suma importância para a elaboração de tesouros, pois permite uma maior compreensão do conceito e da organização das relações entre os conceitos.

Dahlberg enfatiza a importância fundamental da categoria na estruturação do conceito e do sistema de conceitos:

Podemos ver que as categorias têm uma capacidade de estrutura: não apenas estruturam, de fato, todos os nossos elementos de conhecimento e unidades do conhecimento; elas fornecem, ao mesmo tempo, por este meio, o esqueleto, os ossos e tendões para estruturar todo o nosso conhecimento. Com seu uso consciencioso, então, o corpo do nosso conhecimento pode se manter unido, pode se mover, pode se manter flexível - e pode crescer organicamente (DAHLBERG, 1978a. p. 34).

Inicialmente, Ranganathan concebeu cinco categorias fundamentais: personalidade, matéria, energia, espaço e tempo.

Ele inicia definindo a categoria Tempo, que apresenta a menor dificuldade em sua identificação. Esta categoria está de acordo com o que geralmente entendemos por aquele tempo. A categoria Espaço segue a categoria Tempo em sua dificuldade de identificação. E também '*está de acordo com o que geralmente entendemos por aquele termo*'. A categoria Energia é um pouco mais difícil do que *espaço e tempo*. Falando geralmente, sua

manifestação é ação de uma espécie ou outra. A identificação da categoria fundamental Matéria é mais difícil ainda do que Energia. Suas manifestações são de duas espécies – material e propriedade. Aqui ele explica que madeira tanto é uma propriedade de mesa de madeira, como material de que é feita. (Mas não vale aprofundar esta questão aqui). Finalmente, sobre a categoria Personalidade ele não apresenta uma definição; diz que é a que apresenta a maior dificuldade. *'It is too elusive. It is ineffable* ². E explica: Se uma certa manifestação é facilmente determinada como não sendo *tempo, espaço, ou energia, ou matéria*, pode-se dizer que é uma manifestação da categoria fundamental Personalidade. Este é o método dos resíduos. Qualquer entidade que não seja uma manifestação de tempo ou espaço, energia ou matéria, deve ser uma manifestação de Personalidade. (RANGANATHAN, 1967, p. 399-401). Em relação a esta última categoria fundamental Ranganathan parece usar um princípio do hinduísmo: “Não é isso, não é isso.”

Na verdade, aplicar a categorização é analisar o domínio a partir de recortes conceituais que permitem determinar a identidade dos conceitos (categorias) que fazem parte deste domínio. Por exemplo:

- na engenharia mecânica: máquinas operatrizes (personalidade); seus componentes (partes); esmerilhamento, polimento, frezamento (processos); os profissionais envolvidos (agentes); e assim por diante.
- na economia: teorias, política econômica (instrumentos/agentes); planejamento (operações); e assim por diante.
- na construção civil: edificações, como casas, edifícios, teatros, praças (personalidade); técnicas de edificação (técnicas); cimento, pedra, cerâmica, vidro (materiais); os profissionais envolvidos (agentes); e assim por diante.
- na biblioteconomia: instituições e organizações, bibliotecas, documentos, usuários, suporte documental (personalidade); tratamento documentário, recuperação de informação, aquisição (processos); linguagens documentárias (agentes dos processos); profissionais envolvidos (agentes dos processos); e assim por diante. (CAMPOS; GOMES; MOTTA, 2004).

A etapa de categorização auxilia no processo de pensar o domínio, não gerando qualquer registro e serve para orientar os profissionais no levantamento dos termos. Ela consiste em identificar as possíveis classes gerais (categorias) de conceitos que a área do conhecimento, objeto do tesouro, comporta. O exercício de categorização torna claro o domínio temático do tesouro e, como consequência, estabelece as bases para seleção dos termos, nas fontes de onde eles serão retirados.

Posteriormente, o Classification Research Group - CRG - na Inglaterra, desdobrou estas categorias para a elaboração de classificações facetadas. Mas elas podem ser aplicáveis a qualquer temática de tesouros. São, na verdade desdobramentos das categorias fundamentais:

Coisas, substâncias, entidades
que ocorrem naturalmente
produtos
instrumentos
constructos mentais

² Preferimos manter o texto em inglês pela dificuldade de uma tradução exata.

Suas partes
constituintes
órgãos

Sistemas de coisas

Atributos de coisas
qualidades, propriedades, incluindo
estrutura
medidas
processo, comportamento

Objeto da ação (paciente)

Relações entre coisas, interações
efeitos
reações

Operações sobre coisas
experimentos, ensaios
operações mentais

Propriedades de atributos, relações e operações

Lugar, condição

Tempo

Tal lista facilita a análise de qualquer área de assunto ou de qualquer documento. Como se pode observar, os desdobramentos seguem a ordem PMEST.

Como exemplo, citamos algumas facetas - e suas manifestações - que integrariam o sistema de conceitos relativos a Proteção Sanitária:

| Faceta | Manifestações |
|-------------------------|--|
| Coisas | Resíduos, efluentes, lodo, substâncias perigosas, substâncias tóxicas |
| Operações sobre coisas | Tratamento de águas residuais, tratamento de resíduos, tratamento de lodo, disposição final, processos de decomposição da matéria orgânica, processos de estabilização da matéria orgânica |
| Agentes do processo | Agentes de biodegradação/decomposição, poluentes, instrumentos utilizados no tratamento de resíduos |
| Equipamento/Instalações | Unidades de tratamento de resíduos |

O exemplo acima permite perceber, em extensão, a realidade empírica que está sendo objeto de representação conceitual. Esse é um dos aspectos mais interessantes do método de faceta. Quando se consegue fazer este exercício, compreende-se melhor quais as facetas relevantes para uma área.

Como se pode verificar, o método é dedutivo: ele fornece as bases para estabelecer as grandes classes, no interior das quais são incluídas as classes de conceitos, com suas cadeias e renques.

Exemplo recente desta abordagem é a elaboração do *Tesouro de folclore e cultura popular brasileira* (IPHAN, 2005). Graças à tecnologia disponível, é possível apresentar o tesouro em suas formas alfabética e sistemática. Nesta última, a ordem dos conjuntos se baseia nas categorias fundamentais (PMEST). As definições e as relações entre os termos estão presentes na parte alfabética.

4 Considerações finais ou uma abertura para questões teóricas que fundamentam a metodologia de tesouros conceituais

O método de categorias preconizado por Ranganathan trouxe para o campo da documentação o princípio norteador para a organização de conceitos em um dado domínio. O tesouro, entendido em finais da década de 70 como um sistema de conceitos, agrega este princípio através da elaboração do *Thesaurfacet de Aitchison* (AITCHISON, 1972). Na década de 80 Dahlberg traz para o fazer *tesaurográfico* a teoria do conceito (1978), agregando princípios terminológicos relacionados ao conteúdo conceitual e sua definição, estabelecendo assim uma teoria consolidada para a determinação do que se entenderia por menor unidade em um tesouro, ou seja, o conceito representado por um termo.

Dahlberg agrega princípios Ranganathianos relacionados à categorização de um domínio como o método que permite estruturar todo o domínio.

Pode-se afirmar que o método de categorização e o exercício de entendimento do conteúdo conceitual do termo, ou seja, sua definição, são considerados os elementos norteadores para a elaboração de tesouros conceituais. Entretanto, esta afirmativa nos leva a uma pergunta clássica: Define-se para *classificar* (categorizar) ou *classifica-se* (categoriza-se) para definir? Fica a questão.

Methodology for construction of conceptual thesaurus : categorization as a theoretical principle.

Historically, thesaurus documentation consisted of a structured list of terms constituted almost of one-worded terms. Later, it evolved to include some compound words but no guidance was offered as to when to use them or not. A thesaurus is viewed as a classifying system and the Fundamental Categories provide the basis for a correct positioning of the concepts in their classes and organization of the domain. Definition of each concept is the element that will objectively prove the different relationships with other concepts, be it hierarchical, partitive or associative .

Key-words: *Conceptual thesaurus; Methodology; Categorization*

6 Referências

- AFNOR Z 47-100:1981. Règles d' établissement des thésaurus monolingues. Paris, AFNOR, 1981. 20 p.
- AITCHISON, Jean. Thesaurifacets: a new concept in subject retrieval schemes. In: WELLISCH, H. (Ed.). *Subject retrieval in the seventies*. Connecticut: Greenwood, 1972. p. 72-98.
- ANSI Z39.19:1980. *American National Standard Guidelines for Thesaurus Structure, Construction, and Use*. Bethesda, Md, NISO, 1980. 20p.
- ANSI/NISO Z 39.19: 2005 — *Guidelines for the construction, format and management of monolingual controlled vocabularies*. Bethesda, Md, NISO, 2005. 176 p.
- BINWAL, J. C.; LALHMACHHUANA, C. Knowledge representation: concept, techniques and analytico-synthetic paradigm. *Knowledge organization*, v. 28, n. 1, p. 5-16, 2001.
- BSI BS5723:1979. *Guidelines for the establishment and development of monolingual thesauri*. London, BSI, 1979. 36 p.
- CAMPOS, Astério T. Linguagens documentárias. *Revista de Biblioteconomia de Brasília*, v. 14, n. 1 p. 85-88, jan./jun. 1986.
- CAMPOS, Maria Luiza de Almeida. *Linguagem documentária: teorias que fundamentam sua elaboração*. Niterói, RJ: Eduff, 2001a.
- CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha; MOTTA, Dilza Fonseca. *Elaboração de tesouro documentário - Tutorial*. Disponível em: < <http://www.conexaorio.com/bit/tesouro/index.htm> > Acesso em 29 ago. 2006.
- DAHLBERG, I. A Referent-oriented analytical concept theory of interconcept. *International Classification*. v. 5, n. 3, p. 142-150, 1978.
- DAHLBERG, I. *Ontical structures and universal classification*. Bangalore: Sarada Ranganthan Endowment, 1978a. 64p.
- FOSKETT, D. Thesaurus. In: *Subject and information analysis*. New York: M. Dekker, 1985. p. 270-316.
- GOMES, Hagar Espanha. *et alii. Manual de Elaboração de Tesouros Monolingues*. Brasília: CNPq/PNBU, 1990. 78p.
- GOMES, Hagar Espanha; CAMPOS, Maria Luiza de Almeida. Tesouro e normalização terminológica: o termo como base para o intercâmbio de informações. *DatagramaZero*, dez. 2004. Disponível em: <http://www.datagramazero.org.br> Acesso em 29 ago. 2006.
- GOMES, Hagar Espanha; MOTTA, Dilza Fonseca da; CAMPOS, Maria Luiza de Almeida. *Revisitando Ranganathan*. Rio de Janeiro, ago. 2006. Disponível em: < <http://www.conexaorio.com/bit/revisitando/revisitando.htm> > Acesso em 29 ago. 2006.
- IPHAN. Centro Nacional de Folclore e Cultura Popular. *Tesouro de Folclore e Cultura Popular Brasileira*. Rio de Janeiro, 2005. Disponível em <http://www.museudofolclore.com.br/tesouro>. Acesso em 28 ago. 2006.
- LANCASTER, F. W. *Vocabulary control for information retrieval*. 2nd ed. Arlington: Information Resources, 1986. 270p.
- RANGANATHAN, S.R. *Prolegomena to library classification*. Bombay: Asia Publishing House, 1967. 640p.
- RIVIER, A. Construção de Linguagens de Indexação: aspectos teóricos. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 21, n. 1, p. 56-99, jan./jun. 1992.
- ROGET, P. M. *Thesaurus of English words and phrases*. New York: Longmans, 1925. 691p.
- UNESCO. *Guidelines for the establishment and development of monolingual thesauri*. [s.n.t]. 37p.
- WUESTER, E. L'Étude scientifique générale de la Terminologie, zone frontalière entre la Linguistique, la Logique, l' Ontologie, l' Informatique et les Sciences des Choses. In: RONDEAU, G.; FELBER, F. (Org.). *Textes Choisis de Terminologie. I. Fondements théoriques de la terminologie*. Québec: Girsterm, 1981. p. 57-114.