

TESAURO DE CIÊNCIA DA INFORMAÇÃO: Considerações teóricas e comentários

Heloisa Rios Gusmão

RESUMO

GUSMÃO, Heloisa Rios. *Tesouro de Ciência da informação: considerações teóricas e comentários*. *Trans-in-formação*, 1(3): 43 - 48, set/dez, 1989.

Análise da versão preliminar do Tesouro de Ciência da Informação publicado pelo IBICT tendo por base a literatura sobre o assunto.

Unitermos: *Tesouro - Ciência da informação*

INTRODUÇÃO

Possuir uma estrutura própria, ser controlada, ser padronizada, ser hierarquizada, são características das linguagens documentárias das quais o tesaurus é das mais conhecidas. Esse instrumento, destinado a representar a informação para fins de indexação e recuperação de documentos, apresenta qualidades que o torna mais completo ao ser comparado com outros instrumentos de controle de vocabulário.

Por apresentar as relações de associação além das relações hierárquicas e de equalência, o tesaurus difere dos vocabulários controlados e organizados alfabeticamente, (Aitchison & Gilchrist, 1979).

A principal função do tesaurus é poder, no momento da indexação, representar o(s) assunto(s) de que trata o documento para que no momento da busca haja uma coincidência entre a "pergunta" do usuário e a "resposta" do sistema de recuperação da informação.

Portanto, a estrutura do tesouro é que irá refletir a filosofia da sua concepção permitindo que ele cumpra sua função.

Com a publicação do Tesouro em Ciência da Informação constatou-se a viabilidade de elaborar uma análise fundamentada na literatura sobre o assunto e a possibilidade de tecer comentários cujo objetivos é uma contribuição mais efetiva.

O TESAURO DE CIÊNCIA DA INFORMAÇÃO

A publicação da versão preliminar do "Tesouro de Ciência da Informação" pelo IBICT ocorreu num momento oportuno em que as atenções estão particularmente voltadas para o tema tão sedutor que é a elaboração de tesouros.

Apresenta, o tesouro do IBICT, dois tipos de saída: saída alfabética estruturada e saída alfabética por categoria, deixando de fora uma parte importantíssima que é a definição dos termos. Em vista desta falha, não foi possível compreender certos relacionamentos como por exemplo a falta de ligação entre os descritores ARQUIVO PÚBLICO e ARQUIVO ESTADUAL, MUNICIPAL e NACIONAL.

Quanto à origem dos fundos, os arquivos podem ser públicos e privados; os públicos podem ser: estadual, municipal e nacional - o que caracteriza estes tipos de arquivos é a sua entidade mantenedora. Outra forma que pode ser compreendida é de acordo com o tipo de acesso, isto é, público quando é destinado ao povo, à coletividade. De qualquer maneira deve haver uma relação entre esses descritores.

Segundo DAHLBERG (1978) é necessário que os termos sejam definidos para que a cada um deles se atribua somente um conceito. No exemplo ADMINISTRAÇÃO DE UNIDADE DE INFORMAÇÃO verifica-se a existência de dois conceitos, não podendo, portanto, ser caracterizado como descritos uma vez que descritores nomeiam um conceito e não assunto (reunião de idéias).

Um tesouro é uma lista estruturada de termos agrupados através de relações verticais (hierárquicas - TG/TE) e relações horizontais (associativas - TA). O relacionamento, isto é, a ligação entre os termos, que representam os conceitos, é que compõe a estrutura do tesouro. Portanto, nenhum termo poderá aparecer sem estar ligado a um outro.

No "Tesouro de Ciência da Informação", as disciplinas são apresentadas sem algum tipo de relação com outros termos, figurando apenas um qualificador, como no exemplo: CATALOGAÇÃO (DISCIPLINA).

De acordo com LANCASTER (1985) termos podem ter mais de um significado no contexto de um mesmo tesauro e pode-se usar um qualificador entre parênteses a fim de distingüir os significados. No caso do exemplo citado, o qualificador (disciplina) só permitirá fazer a distinção entre o termo CATALOGAÇÃO-operação e CATALOGAÇÃO-disciplina. Um qualificador jamais poderá substituir a ligação daquele termo com outros.

Não havendo necessidade de apresentar a estrutura das disciplinas arroladas no tesauro, seria mais adequado que seu controle fosse feito em lista à parte - como identificadores.

Outro descritor que aparece isolado é BILBIOTECA DE EMPRÉSTIMO. Sem definição torna-se difícil entender o descritor uma vez que apresenta dois conceitos: biblioteca uma unidade de informação, e empréstimo-atividade que pode ser realizada na biblioteca e deveria estar subordinado ao descritor biblioteca como um termo específico (TE).

A função do tesauro é poder representar os assuntos dos documentos, feita no momento da indexação, e das solicitações de busca, feita no momento em que o usuário busca uma informação no sistema. Para que haja uma coincidência entre a pergunta do usuário e a resposta do sistema, um termo não pode ora ser um processo, ora ser documento, como se apresenta o descritor ÍNDICE BIBLIOGRÁFICO, nas duas categorias.

O mesmo acontece com o termo TESAURO: aparece na categoria E - Processo, e Categoria B - Documento. Tesauro não pode ser um processo; basta verificar sua definição e concluir que é um instrumento.

A divisão da área estudada em categorias é o primeiro estágio para se agrupar os termos em classes e sub-classes e conseqüentemente montar sua estrutura hierárquica.

AITCHISON e GILCHRIST (1979) sugerem que a escolha das categorias ou facetas seja feita de acordo com as necessidades do usuário e do objetivo do sistema de informação. As categorias identificadas no "Tesauro de Ciência da Informação" não mostram nas classes e sub-classes uma vez que os termos saíram em listagens alfabéticas. A parte alfabética estruturada mostra apenas as hierarquias em ordem alfabética mas não sistemática.

Na categoria G - Profissão aparece o termo Associação de Bibliotecas. Na lista alfabética estruturada este termo está em ligação a outros termos que associada à falta de definição torna impossível sua identificação: seria o movimento associativo ou nome da organização?

Nessa categoria foram também identificados termos que pertencem a outras categorias, como:

Bibliotecário - agente

Currículo-entidade

Ensino - operação

Em relação às categorias verifica-se que os termos não foram categorizados com base nas suas definições, o que gerou a impossibilidade de divisão em classes e sub-classes e, portanto, falhas na estrutura.

Parece que o princípio para escolha das categorias não foram aqueles estabelecidos por RANGANATHAN(1987) e/ou pelo "CLASSIFICATION RESEARCH GROUP" (1970). Neste caso, sugere-se que nas próximas versões sejam incluídos os princípios que regeram aquela escolha.

Ao apresentar a relação hierárquica, o tesouro do IBICT refere-se ao termo específico (TE) como sendo uma parte ou tipo; ao mesmo tempo refere-se à relação todo-parte como relação associativa. Ela não é nem hierárquica nem associativa. O termo específico, isto é, um conceito (TE) com todas as características de outro conceito (TG).

A relação partitiva se dá quando é possível reconhecer no objeto suas partes; ela não se dá entre conceitos e é representada pelos códigos TGP/TEP, conforme preconizam as Diretrizes do IBICT sobre o assunto (1985).

Não existe pesquisa suficiente para determinar as bases teóricas da relação de associação.

MILHAILOV & GILTAREVSKI (1971) citam como relações conceituais mais importantes as do tipo espécie/gênero (gênero/espécie), subordinação (coordenação), semelhança funcional, causa/efeito (efeito/causa), parte/todo (todo/parte).

Para SOERGEL (1974) as relações de associação complementam as relações hierárquicas; elas existem entre todos os conceitos que têm uma característica em comum. Segundo AITCHISON e GILCHRIST (1979) a relação de associação ocorre quando um termo está relacionado a outro conceitualmente e não hierarquicamente.

Por não estarem definidas de forma objetiva pelos autores, as relações de associação são representadas de forma confusa - seu código ora aparece com TR ora como TA. A melhor maneira de representá-las seria TA (termo associado) uma vez que TR (termo relacionado) não sugere esse tipo

de relação. Num tesouro todos os termos estão relacionados a outro ou outros, isto é, possuem um tipo de relação, seja lógica ou ontológica logo, TR não indica qual o tipo de relação. Acredita-se que o embasamento teórico ainda é suficiente para determinar o estabelecimento das relações associativas, portanto seria necessário que o tesouro do IBICT incluísse na metodologia qual o procedimento adotado para identificação dessas relações.

CONCLUSÃO

É de grande importância para a área a elaboração de um instrumento como o Tesouro de Ciência da Informação. As considerações sobre a obra bem como sugestões podem ser usadas como auxílio às edições futuras uma vez que, sendo o tesouro um dos instrumentos de controle do vocabulário para fins de indexação/recuperação da informação, necessita apresentar suas peculiaridades de maneira bem explícita. É o que difere dos outros instrumentos destinados a controlar os vocabulários usados em sistemas de informação.

SUMMARY

GUSMÃO, Heloisa R. *Thesaurus of science information: theoretical considerations and commentaries*. *Trans-in-formação*, 1(3) 43 - 48, sept/dec., 1989.

An overview of the structure of the Tesouro de Ciência da Informação (IBICT) and proceeds to a discussion based on literature.

Key words: *Tesouro - Science information*

REFERÊNCIAS BIBLIOGRÁFICAS

- AIRCHISON, J., GILCHRIST, A. (1979). **Manual para construção de tesouros**. Rio de Janeiro: BNG/BRASILART.
- CLASSIFICATION RESEARCH GROUP (1970). **Faceted classification** London: Library Association.
- DAHLBERG, I. (1978). Teoria do conceito. **Ciência da Informação**, Brasília, v.7, n.2,p. 101-107.
- IBICT. **Diretrizes para elaboração de tesouros monolíngues**. Brasília: 1984.
- IBICT. **Tesouro sobre leitura**. Brasília, 1985.
- KANDELAKI, T.L. (1981') Les sens des terms et le systèmes de sens des terminologies scientifiques. In: _____. **Textes choisis de terminologia**. Quebec: Univ. Laval, v.1,p.173.
- LANCASTER, F.W. (1985). **Vocabulary control for information retrieval**. 2. ed. Arlington: Information Resources Press.
- MIHAILOV, A.I., GILTAREVSKI, R.S. (1971). **An introductory course on Informatics/.Documentation**. The Hague: FID.
- RANGANATHAN, S.R. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.
- SOERGEL, D. (1974). **Indexing languages and thesauri: construction and maintenance**. Los Angeles: Melville Publ.
- VICKERY, B.C. (1975). **Classification and indexing in Science** 3rd. ed. London: Butterworths.