

## A IDENTIFICAÇÃO DE CONCEITOS NO PROCESSO DE ANÁLISE DE ASSUNTO PARA INDEXAÇÃO

*Mariângela Spotti Lopes Fujita*

### **Resumo**

A indexação compreende a análise de assunto como uma das etapas mais importantes do trabalho do indexador. A análise de assunto tem como objetivo identificar e selecionar os conceitos que representam a essência de um documento. A identificação de conceitos, realizada durante a leitura documentária, envolve esforço adicional de compreensão de texto e as dificuldades apresentadas possibilitaram a motivação para investigar, por meio de revisão de literatura, a identificação de conceitos a partir da leitura documentária, da tematicidade e das concepções de análise de assunto. Pelas análises obtidas considera-se que as concepções de análise orientadas para o conteúdo e para a demanda são decisivas para o bom desempenho do indexador durante a leitura para a identificação de conceitos na medida em que vincula-se à busca pela tematicidade do texto.

### **Palavras-chave**

Identificação de conceitos; Análise de assunto; Indexação

## IDENTIFICATION OF CONCEPTS IN THE PROCESS OF SUBJECT ANALYSIS FOR INDEXING

### **Abstract**

The indexing comprehends the subject analysis as one of the most important work phases of the indexer. The subject analysis aims to identify and to select the concepts that represent the essence of the document. The identification of concepts, made during the documentary reading, comprehends additional effort of text comprehension and the difficulties presented made the motivation to investigate, by literature review, identification of concepts from the documentary reading, aboutness and conceptions of subject analysis. By obtained analysis is considered that the conceptions of analysis oriented to the content and to the demand are decisive to the good performance of the indexer during the reading to the identification of concepts as far as it is related to the search for the text aboutness.

### **Keywords**

Identification of concepts; Subject analysis; Indexing

## 1 INTRODUÇÃO

A indexação como ato de construir índices é prática bastante antiga no tratamento de documentos. Basta sabermos que em "bibliotecas" da Antigüidade já existiam listas dos documentos ali armazenados. Entretanto, a partir do momento que a ordenação dessas listas necessitou de uma organização por assunto foram estabelecidas profundas mudanças na abordagem do ato mecânico de construir índices, ou seja, introduziu-se um processo de análise do conteúdo dos documentos.

A indexação como processo de análise documentária é realizada mais intensamente desde o aumento de publicações periódicas e da literatura técnico científica de modo geral, que impulsionaram a necessidade de criação de mecanismos de controle bibliográfico em centros de documentação especializados. Assim, as bibliografias como mecanismos de controle bibliográfico surgiram fora do âmbito das bibliotecas tradicionais e apresentavam uma evolução nas técnicas de tratamento

da informação, dando impulso teórico-prático, naquela ocasião, a uma nova área: a Documentação.

Dentro desse bojo de evolução de técnicas de tratamento da informação, está ligada a análise documentária como extensão do tratamento temático que comporta a geração de resumos e a indexação. A indexação, vista por Chaumier (1980, p. 42) como a "[...] parte mais importante da análise documentária" é uma combinação metodológica altamente estratégica entre o tratamento do conteúdo de documentos e sua recuperação por um usuário. Além de estratégica, demonstra uma relação estreita entre o processo e a finalidade da indexação.

Segundo o World Information System for Science and Technology <sup>1</sup>(1981, p. 84) há que se considerar a indexação sob dois pontos de vistas distintos: enquanto processo que consiste em descrever e identificar um documento com ajuda de representações dos conceitos contidos em um documento e quanto à sua finalidade

---

<sup>1</sup> Sistema internacional vinculado à UNESCO e conhecido pela sigla UNISIST

permitindo busca e acesso à informação armazenada.

A indexação em análise documentária, sob o ponto de vista dos sistemas de informação, também é reconhecida como a parte mais importante porque condiciona os resultados de uma estratégia de busca. O bom ou mau desempenho da indexação reflete-se na recuperação da informação feita através de índices.

Isso nos leva a considerar que a recuperação do documento mais pertinente à questão de busca é aquele cuja indexação proporcionou a identificação de conceitos mais pertinentes ao seu conteúdo, produzindo uma correspondência precisa com o assunto pesquisado em índices.

Na *identificação de conceitos*, o indexador, após o exame do texto, passa a abordá-lo de uma forma mais lógica a fim de selecionar os conceitos que melhor representem seu conteúdo.

Entretanto, a identificação de conceitos analisada sob os aspectos da leitura, da busca pela tematicidade e de concepções

de leitura, envolve um considerável grau de complexidade que, certamente, acarreta dificuldades ao indexador como verificado em estudo de observação da leitura de quatro indexadores do Centro Coordenador Nacional do Sistema Especializado na Área de Odontologia, antiga Sub-Rede Nacional de Informação em Ciências da Saúde Oral do convênio BIREME/KELLOGG/USP (FUJITA, 1998), visando analisar a importância da identificação de conceitos como estratégia de leitura.

Considerando-se que a identificação de conceitos, realizada durante a leitura documentária, é o objetivo da análise de assunto na indexação, as dificuldades observadas possibilitaram a motivação para investigar, por meio de revisão de literatura, a identificação de conceitos a partir da leitura documentária, da tematicidade e das concepções de análise de assunto.

## 2 A IDENTIFICAÇÃO DE CONCEITOS NA INDEXAÇÃO: CONCEITUAÇÃO E PROCESSO

De acordo com os "Princípios de indexação", formulados pela World Information System for Science and Technology (1981, p. 84), a indexação é definida como "[...] o ato de descrever ou identificar um documento em termos de seu conteúdo", sendo que, para a norma 5693 da International Standardization for Organization (1985, p. 2), a indexação é vista como

[...] a representação do conteúdo dos documentos por meio de símbolos especiais, quer retirados do texto original, quer escolhidos numa linguagem de informação ou de indexação.

A indexação, como operação de representação documentária, desenvolve-se de acordo com um processo composto de operações básicas. De acordo com os "Princípios de indexação" do World Information System for Science and Technology (1981, p. 84),

Durante a indexação, os conceitos são extraídos do documento através de um processo de análise, e então traduzidos para os termos de instrumentos de indexação (tais como tesouros, listas de cabeçalhos de assunto, esquemas de classificação, etc.)

O processo de indexação, portanto, compreende dois estágios: o analítico, em

que é realizada a compreensão do texto como um todo, a identificação e a seleção de conceitos válidos para a indexação e o estágio de tradução, que consiste na representação de conceitos por termos de uma linguagem de indexação:

- Determinação do assunto: estabelecimento dos conceitos tratados num documento;
- Representação de conceitos por termos de uma linguagem de indexação: a tradução dos conceitos nos termos da linguagem de indexação.

Segundo Vickery (1980), o processo de indexação comporta ainda mais um estágio, o de sumarização, entre o analítico e o de tradução. Essa diferença de etapas ou estágios é explicada pelo fato de que Vickery desdobrou o estágio analítico em dois, o de análise e o de sumarização, que pode ser entendido como síntese. Nos "Princípios de indexação" o estágio de determinação do assunto, ou análise de assunto, engloba a operação de síntese, como se vê a seguir.

O primeiro estágio, a análise de assunto, razão de nosso estudo por abrigar o processo de leitura, é subdividido em outros três estágios:

- compreensão do conteúdo do documento;
- identificação dos conceitos que representam este conteúdo;
- e seleção dos conceitos válidos para recuperação.

Em observação à parte, o texto de "Princípios" chama a atenção para o fato de que "[...] na prática esses três estágios se superpõem." (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981, p. 86), porém não explicita em que momento. Mais à frente observaremos que, de fato, esses três sub-estágios são superpostos durante a leitura do documento.

Para a *compreensão do conteúdo documento*, o texto refere-se a documentos gráficos e não-gráficos. Para os gráficos (livros, monografias, jornais, periódicos, relatórios, teses) aponta a impraticabilidade de uma leitura extensiva do texto, embora a considere ideal.

Mesmo assim, para que o indexador não negligencie nenhuma informação relevante, aponta um roteiro de partes importantes do texto que merecem especial atenção, durante sua leitura: título, introdução e as primeiras frases de capítulos e parágrafos; ilustrações, tabelas, diagrama e suas explicações; conclusão; palavras ou grupos de palavras sublinhadas ou impressas com tipo diferente. Porém, adverte que os primeiros itens do texto apresentam, geralmente, as intenções do autor, enquanto que as partes finais comunicam o alcance dessas intenções. Por isso, não recomenda a indexação somente pelo título e/ou pelo resumo.

Na *identificação de conceitos* (segundo estágio do estabelecimento de conceitos), o indexador, após o exame do texto, passa a abordá-lo de uma forma mais lógica a fim de selecionar os conceitos que melhor representem seu conteúdo. Para isso, recomenda que a identificação de conceitos seja feita obedecendo a um esquema de categorias existente na área coberta pelo documento, como por ex.: o fenômeno, o processo, as propriedades, as operações, o material, o equipamento, etc.

Embora o texto de “Princípios de indexação” não se refira à leitura durante os estágios de identificação e seleção de conceitos, é possível observar que está subtendida na frase citada e grifada anteriormente “[...] após o exame do texto, passa a abordá-lo de uma forma mais lógica [...]”.

E a *seleção de conceitos* é necessária tendo em vista os objetivos para os quais as informações são indexadas. Assim, nem todos os conceitos identificados serão necessariamente selecionados.

A publicação dos "Princípios de indexação" pelo World Information System for Science and Technology (1981), gerou a primeira norma para análise, identificação de assuntos e seleção de termos de indexação publicada pela ISO, sob número 5693 em 1985, com o título "Documentation - methods for examining documents, determining their subjects, and selecting indexing terms". Em 1992 a ABNT publicou a tradução dessa mesma norma, sob número 12.676, intitulada "Métodos para análise de documentos - determinação de

seus assuntos e seleção de termos de indexação".

A Norma 12.676, da Associação Brasileira de Normas Técnicas (1992, p. 2), indicou para a indexação três fases ou estágios:

“a) exame do documento e estabelecimento do assunto de seu conteúdo;

b) identificação dos conceitos presentes no assunto;

c) tradução desses conceitos nos termos de uma linguagem de indexação”

No item “*exame do documento*” a Norma 12.676, ao mesmo tempo em que considera ideal a leitura total do documento aponta sua impraticabilidade operacional, oferecendo ao indexador a possibilidade do texto ser analisado através do exame cuidadoso de partes como o título, resumo, sumário e etc.

Ao recomendar o exame do documento através de uma leitura de suas partes, não especifica que tipo de documento possui essas partes, ou mesmo, se todos os documentos, indistintamente, possuem. Adverte em nota de rodapé que “não se recomenda indexar apenas por qualquer

um destes elementos isoladamente [...]” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p. 2).

Após o exame do documento, a Norma 12.676 indica o estágio de *Identificação de Conceitos* como:

[...] uma abordagem sistemática para identificar aqueles conceitos que são os elementos essenciais na descrição do assunto. (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p. 2, grifo nosso).

A abordagem sistemática da norma para identificação de conceitos, portanto, vai além do "esquema de categorias existente na área coberta pelo documento" proposto em "Princípios de indexação" (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981, p. 87) porque recomenda um questionamento do texto através de questões preparadas para identificar determinados conceitos essenciais.

A identificação de conceitos, segundo a norma (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p. 2), é realizada após o exame do documento quando, então, o indexador deverá seguir uma abordagem sistemática para a identificação daqueles conceitos que são elementos essenciais na descrição do

assunto. Observe-se que a referida norma não se refere a uma interrupção da leitura, mas sim ao exame que corresponde à exploração de partes do texto. A leitura, aqui, parece estar implícita porque a identificação de conceitos deve ser feita por um questionamento.

Pode-se supor, portanto, que, a abordagem sistemática é um questionamento que o indexador realiza para melhor extrair conceitos enquanto estiver fazendo a leitura das partes do texto, ainda que, a Norma 12.676 da Associação Brasileira de Normas Técnicas não explicita quais questões seriam mais indicadas para cada parte do texto. Além disso, não faz nenhuma menção a respeito de a leitura ser melhor realizada pela “abordagem sistemática” do questionamento e de essa ser considerada como estratégia de leitura, recomendando que "após examinar o documento [...]" o indexador deve abordá-lo sistematicamente. Isso poderia significar que a identificação de conceitos independe da leitura? Supondo que a Norma 12.676 da Associação Brasileira de Normas Técnicas considerasse a identificação de conceitos independente da leitura, ela certamente teria indicado,

no estágio de “exame do documento”, a necessidade de extrair, pelo menos, um enunciado de assunto com o qual seria realizada a identificação de conceitos. Assim, presume-se que a identificação e seleção de conceitos deva ser realizada durante a leitura.

A seleção de conceitos, diferente dos "Princípios de indexação" (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981), é incluída pela Norma 12.676 dentro do item "Identificação de conceitos", recomendando que o indexador não precisa representar como termos de indexação todos os conceitos identificados, mas selecioná-los conforme objetivos de uso dos termos.

Mais adiante, a Norma inclui o item "Seleção de termos de indexação" relativa à tradução dos conceitos em termos de indexação e aqui, então, fica estabelecida a diferença entre a seleção dos conceitos identificados durante a leitura e a seleção dos termos de indexação em linguagens documentárias para a representação dos conceitos selecionados.

Na etapa de “Tradução” desses conceitos na linguagem de indexação do sistema, a norma recomenda procedimentos para a verificação de descritores controlados e a preparação de uma lista daqueles termos para os quais não houve, no tesouro, uma exata representação de assuntos encontrados no documento.

A mesma norma, com o item "controle de qualidade", aponta a necessidade de qualidade e consistência em indexação e ainda relaciona os fatores que a garantem: a imparcialidade do indexador, o conhecimento do indexador sobre o campo coberto pelos documentos a serem indexados, as vantagens do contato direto com o usuário e a receptividade das linguagens de indexação para novos termos.

Considerando o esclarecimento sobre a função de cada uma das etapas da indexação, é possível afirmar que uma das etapas consideradas mais importantes do trabalho do indexador é a Análise de assunto, que tem como objetivo identificar e selecionar os conceitos que representam a essência de um documento. Trata-se de um processo complexo, pois exige esforços do profissional (indexador)

para seguir uma metodologia adequada a fim de obter resultados satisfatórios. A eficácia desse trabalho pode ser avaliada pelos resultados obtidos pelo usuário no momento da recuperação da informação.

### **3 A ANÁLISE DE ASSUNTO PARA IDENTIFICAÇÃO DE CONCEITOS: UMA ANÁLISE DE CONCEPÇÕES**

Verifica-se, na literatura, que a palavra assunto tem várias interpretações. Em vista disso, o processo também pode ser denominado Análise temática, Análise documentária, Análise conceitual, ou mesmo Análise de conteúdo.

Para Langridge (1977), citado por Albrechtsen (1993), o termo Análise de assunto envolve o conhecimento do conteúdo dos documentos e a determinação de suas características significantes.

Chu e O'Brien (1993, p. 439) consideram a análise de assunto como a fase inicial do processo de indexação, a qual decidirá sobre os principais tópicos do assunto de um documento, precedendo a fase de tradução desses tópicos de acordo com a linguagem documentária adotada pelo sistema.

O termo análise de assunto é o mais utilizado, mas grande parte dos autores o considera como a etapa de tradução dos conceitos extraídos dos documentos para um vocabulário controlado ou o processo de indexação como um todo.

De acordo com Vickery (1968, p. 356), análise de assunto é vista pelo seu produto, ou seja:

Análise da informação significa derivar de um documento o conjunto de palavras que servem como uma representação condensada desse documento. Esta representação pode ser usada para identificar o documento, para prover pontos de acesso na busca, para indicar seu conteúdo, ou como substituto para o documento.

Segundo Naves (2000), a análise de assunto envolve grande complexidade, pois, além do problema da terminologia, existe a influência direta das pessoas que a executam, conhecida como subjetividade, pela qual o indexador interpõe seus próprios valores em sua atuação de intermediário entre autores e usuários, ainda que a tarefa do indexador seja determinar, de forma precisa, o conteúdo do documento.

Para Cesarino e Pinto (1980), existem duas situações nas quais os profissionais fazem a análise de assunto. A primeira seria ao receber um documento e inseri-lo num sistema de informação. Fazem uma análise para determinar seu conteúdo informativo, visando ao objetivo do sistema e à necessidade do usuário. A segunda acontece quando, ao receberem um pedido de informação fazem uma análise deste com o objetivo de compreender a necessidade de informação requerida pelo usuário, identificam os conceitos existentes no pedido e traduzem os mesmos para a linguagem do sistema. Estas duas situações têm por objetivo identificar a necessidade informacional do usuário.

Destaca-se, portanto, que o processo de análise de assunto reveste-se de uma subjetividade característica, dadas as circunstâncias e elementos envolvidos, pois, a partir da leitura do documento pelo indexador, é realizado um processo de comunicação interativo entre três variáveis: leitor, texto e contexto. Cada uma dessas variáveis estará sujeita a diferentes condições, mas é o indexador como leitor a variável mais influente nessa interação para análise de assunto,

porque precisa realizar a compreensão da leitura mediante sua cognição.

Destacamos que, como o indexador tem o objetivo de tornar o assunto conhecido pelos usuários interessados, a função desse profissional é “[...] aumentar a visão do que os outros podem ler em um texto” (HUTCHINS, 1977, p. 19).

Para alguns autores, como veremos adiante, a análise de assunto implica determinar a tematicidade do documento mediante a identificação e seleção dos conceitos que compõem o assunto ou tema principal e secundários.

Sendo o assunto a informação relevante abordada no texto, é preciso ressaltar, ainda, que a seleção do assunto ou informação relevante sofre a influência da política de indexação do sistema de informação ao qual se insere o indexador. A instituição decidirá se o tema extraído do documento será o mais específico, ou se considerará um nível mais genérico, baseando-se no perfil do usuário que estabeleceu atender.

Conforme Albrechtsen (1993, p. 221), dependendo dos objetivos institucionais,

percebe-se qual a concepção de análise de assunto que o sistema de informação segue e, conseqüentemente, o indexador levará esse aspecto em questão. Consideram-se, assim, diferentes concepções de análise que, certamente afetam o desempenho do indexador enquanto leitor. A esse respeito, Albrechtsen (1993, p. 220) classifica os diferentes pontos de vista em três tipos de concepções:

- concepção simplista: considera os assuntos como entidades objetivas absolutas, que podem derivar de uma abstração lingüística do documento ou de somas usando métodos estatísticos de indexação. De acordo com essa concepção a indexação pode ser totalmente automatizada;
- concepção orientada para o conteúdo: envolve uma interpretação do conteúdo do documento que vai além dos limites da estrutura superficial léxica e gramatical. A análise de assunto do conteúdo de documentos envolve identificação de tópicos ou assuntos que não estão explicitamente colocados na estrutura textual superficial do documento, mas que são facilmente percebidos por um indexador humano. Envolve,

portanto, uma abstração indireta do documento.

- concepção orientada pela demanda: considera o assunto como instrumento para transferência de conhecimento, portanto, direcionado para uma finalidade pragmática de informação e conhecimento. Conforme esta concepção, documentos são criados para comunicação do conhecimento, e assuntos devem, portanto, ser ajustados para funcionar como instrumentos de mediação e transmissão desse conhecimento para qualquer pessoa interessada. Dessa forma, quando o indexador analisa um documento não concentra-se em representar ou resumir a informação explícita ou implícita, mas questiona-se: como eu poderia tornar esse conteúdo ou parte dele, visível para o usuário potencial? Que termos deverei utilizar para levar esse conhecimento até o leitor interessado?

Verifica-se que as três concepções possuem vantagens e desvantagens. A principal vantagem de adotar uma concepção simplista se refere ao barateamento de computadores e de softwares, o que significa uma indexação

automática a baixos custos. Entretanto, seu principal inconveniente se refere à impossibilidade de transferência do conhecimento do ponto de vista social.

A concepção orientada para o conteúdo, ainda que seja estabelecida para o aprimoramento das técnicas e o trabalho do indexador, pode ser simples ao focalizar apenas a representação dos documentos, sem considerar seus possíveis usos.

A concepção orientada para a demanda tem a vantagem de permitir a transferência e a disseminação do conhecimento, mas envolve um alto grau de responsabilidade ao distinguir a prioridade de determinados assuntos a usuários potenciais para assegurar sua utilização. Neste caso, torna-se necessário enfatizar o exame de como o texto é analisado para a definição de um assunto, levando-se em conta o contexto onde ele está inserido.

Embora, de acordo com Naves (1996), as duas últimas concepções – indexação orientada pelo conteúdo e orientada pela demanda – sejam complementares, julgamo-las mais do que complementares:

são intrínsecas porque, no momento em que o indexador está lendo e procurando identificar e selecionar conceitos para a determinação do assunto do documento, está objetivando encontrar o assunto que lhe é familiar devido à sua prática de indexação e também definir o que pode interessar ao usuário do sistema de informação.

A questão da identificação do assunto do documento nos faz resgatar que a atividade está vinculada à leitura, esclarecendo que o indexador realiza as duas operações, identificação e seleção de conceitos, *durante* a mesma e que a tradução dos termos que representam conceitos em descritores da linguagem do sistema só deve ser feita após seu término para que a análise seja conceitual e voltada para a demanda. A preservação do conteúdo do documento é uma garantia da relevância de recuperação, objetivo da boa indexação de conteúdo.

Isso nos remete aos princípios de indexação do sistema PRECIS que, no início da década de 70, preconizou a "preservação do contexto" baseado principalmente em uma análise cuja proposição era a identificação de

conceitos a partir da investigação da estrutura profunda do texto, usando um princípio de organização dos assuntos em facetas de acordo com uma ordem de citação estruturada em uma entrada de duas linhas e três posições que revelava a estrutura superficial (FUJITA, 1989). Esse princípio distinguia muito bem duas partes para funcionamento do sistema PRECIS: a parte sintática, formada pela estrutura de entradas e gramática composta de operadores de função atribuídos aos conceitos durante a análise conceitual, e a parte semântica, em que os termos identificados pelos operadores eram traduzidos por termos de um vocabulário controlado. A análise conceitual do PRECIS era baseada, então, na idéia do autor e utilizava a terminologia do texto, sem nenhum tipo de associação com instrumentos de controle de vocabulário.

Se o princípio de "preservação do contexto", baseado principalmente na análise conceitual, tinha como proposição a identificação de conceitos a partir da investigação da estrutura profunda do texto, a contribuição da lingüística textual, fundada na gramática transformacional, torna-se decisiva para a

transcrição difícil entre forma e conteúdo (PINTO MOLINA, 1994). Então, como a base teórica do PRECIS é a gramática transformacional, o uso de sua análise conceitual dá garantia de que o conteúdo será representado na identificação de conceitos.

Neste momento, convém esclarecer a dupla função da seleção de conceitos que ocorre em dois momentos diferentes da análise de assunto: durante a identificação de conceitos para a determinação do assunto e, após a identificação de conceitos, durante a tradução dos termos que representam os conceitos para os termos da linguagem documentária adotada pelo serviço de análise. Esse esclarecimento dará mais respaldo e deverá iluminar a continuidade de novos estudos em análise documentária, porque confirma que as concepções de leitura orientada para o conteúdo e para a demanda devem ser intrínsecas, caracterizando a leitura documentária. Nessa, a concepção de leitura orientada para o conteúdo deve orientar a identificação de conceitos e a concepção orientada para a demanda, a seleção de conceitos.

Resgatando o processo de análise de documentos pela Norma 12.676 da Associação Brasileira de Normas Técnicas, já mencionado anteriormente, verificamos que a identificação e a seleção de conceitos são etapas da análise de assunto. Ao recomendar uma abordagem sistemática para a identificação de conceitos, a Norma 12.676 revela um forte indicativo dos princípios da concepção orientada para o conteúdo, porque, conforme Albrechtsen (1993, p. 220):

[...] envolve identificação de tópicos ou assuntos que não estão explicitamente colocados na estrutura textual superficial do documento [...].

Pode-se, também, observar que a identificação de conceitos feita pela abordagem sistemática é uma análise orientada para o conteúdo.

Portanto, os "Princípios do UNISIST" e a Norma 12.676 contêm as duas concepções, a orientada para o conteúdo, porque propõem a abordagem sistemática para a identificação de conceitos e a orientada para a demanda, porque orientam a seleção dos conceitos conforme objetivos de uso dos termos, confirmando o aspecto intrínseco das

duas concepções de análise de assunto durante a leitura documentária.

#### **4 IDENTIFICAÇÃO DE CONCEITOS E A BUSCA PELA TEMATICIDADE**

Com a evidência da identificação de conceitos na análise de assunto orientada pelo conteúdo e demanda, há necessidade de orientação sobre esse processo. Ainda que já nos tenhamos referido à sistemática de identificação de conceitos por questionamento, indicada pela Norma 12.676 e a tenhamos comparado à análise conceitual do PRECIS, considerando esta uma boa orientação, não questionamos a ordem e a relevância dos conceitos conforme a tematicidade intrínseca (relevância autor) e a tematicidade extrínseca (saliência leitor) do conteúdo do documento.

O sistema PRECIS foi, também, influenciado pela "análise em facetas" idealizada por Ranganathan (1960), um bibliotecário hindu que compreendia o conteúdo de um documento como um conjunto de assuntos específicos relacionados entre si dentro de uma perspectiva particular. Dentro dessa perspectiva facetada, Derek Austin, como

idealizador do PRECIS, propôs a análise conceitual do PRECIS com elementos da análise em facetas.

Ao demonstrar a concepção do sistema de indexação PRECIS, Fujita (1989, p. 5) relata uma síntese evolutiva dos cem anos de estudos realizados em torno da indexação alfabética de assunto desde a publicação, em 1876, da obra básica de Charles Ammi Cutter "Rules for dictionary catalog" até a idealização do sistema de indexação PRECIS por Derek Austin em 1974.

Por esta síntese evolutiva, é possível observar que, embora a preocupação principal dos estudos seja o produto final, ou seja, a geração do índice, a análise que envolve a transformação do conteúdo em índice está expressa tanto pela proposição de categorias quanto pelos sistemas para indexação, como se observa na evolução dos principais estudos teóricos:

- *J. Kaiser* (1911) - com a publicação do trabalho "Systematic Indexing" propõe a análise de assuntos compostos pela combinação de três categorias: um

"concreto", um "processo" e "lugar";

- *S. R. Ranganathan* (1933) - desenvolveu um esquema de classificação baseado na análise de facetas e o uso de cinco categorias fundamentais: Personalidade, Matéria, Energia, Espaço e Tempo;
- *E. J. Coates* (1960) - em seu livro "Subject Catalogues", apresenta a formulação de cabeçalhos de assunto específicos por categorias: coisa - parte - matéria - ação;
- *J. W. Metcalfe* (1959) - admite que a entrada deve ser direta e discute o propósito da catalogação de assunto como sendo o de indicar somente a classe de assunto em que está inserido;
- *M. F. Lynch* (1973)- criou e desenvolveu os índices articulados de assunto num estudo de índices para o Chemical Abstracts;
- *J. E. L. Farradane* (1977) idealizou um sistema de indexação que adota nove

operadores relacionais, para indicar as relações entre termos em etapas de discriminação no tempo e no espaço;

- *POPSI* (Postulated-based Permuted Subject Indexing Language), idealizado por Neelameghan e Gopinath (1975), é um sistema inteiramente baseado em princípios classificatórios e que utiliza cabeçalhos de classificação como termos de entrada na produção dos índices cuja padronização é derivada das categorias da classificação de dois pontos de Ranganathan;
- *T. C. Craven* (1978) - idealizou inicialmente o sistema NEPHIS (Nested Phrase Indexing System) e depois, em consequência de uma evolução experimental, o sistema LIPHIS (Linked Phrase Indexing System). Ambos são sistemas de indexação automática;
- *D. Austin* (1974) - idealizou para a British National

Bibliography (BNB) o PRECIS, cujo funcionamento se fundamenta em estruturas semântica e sintática e em esquema de operadores de função.

A noção de índice sempre esteve muito ligada ao processo de indexação. Os índices outrora existentes em sistemas de recuperação da informação, tais como os antigos catálogos de fichas de biblioteca, foram considerados dentro de uma perspectiva classificatória, porque os chamados cabeçalhos de assunto eram compostos sob influência da terminologia classificatória e não do texto e seu conteúdo.

O grande elemento transformador dentro da indexação alfabética a marcar os estudos teóricos foi a análise em facetas proposta por Kaiser, Ranganathan e seus seguidores, demarcando a possibilidade de maior especificidade e uniformidade com o uso dos conceitos essenciais: espaço, tempo, processo, concreto, coisa, ação etc.

Após Ranganathan, o Classification Research Group desenvolveu a aplicação

dos princípios da análise em facetas assumindo a influência da classificação facetada e passou a utilizar e desenvolver uma metodologia facetada (PIEDADE, 1983, p. 79-80). Vickery (1975, p. 181-189), citado por Esteban Navarro, (1999, p. 74), por exemplo, ampliou a quantidade de facetas propostas por Ranganathan: Personalidade, Matéria, Energia, Espaço e Tempo (PMEST) para Tipo, Estrutura, Constituintes, Propriedades, Processos, Operações, Técnicas, Generalidades.

Segundo Esteban Navarro (1999, p. 73) :

[...] a faceta permite descobrir as relações que mantêm entre si os conceitos mediante a formulação de uma série de perguntas peculiares para o domínio disciplinar em que se situa o assunto do documento [...].

Nesse sentido, as facetas relacionadas ao assunto “materiais dentários”, por exemplo, seriam reveladas a partir dos seguintes conceitos:

- *tipo* de materiais dentários: Materiais dentários metálicos e Materiais Dentários não-metálicos;
- *constituintes*: Ouro, alumínio, porcelana, prata;

- *propriedades*: Resistência à fratura, fotoelasticidade, rigidez;
- *processos*: Amalgamação, polimerização;
- *operações*: Vibração;
- *técnicas de laboratório*: Fase Gama; etc.

Na visão de Esteban Navarro (1999, p. 79) a identificação de conceitos na indexação deve utilizar questões construídas a partir da "análise das facetas que caracterizam um conjunto de assuntos relacionados entre si". Se retomarmos a recomendação dos "Princípios de indexação" do World Information System for Science and Technology e da Norma 12.676 para identificação de conceitos<sup>2</sup>, veremos que a base é a mesma, ou seja, a influência da análise em facetas permeia o processo de análise em indexação. Os autores citados mencionam a identificação do tema referindo-se a conceitos, categorias e facetas, que podemos considerar como a mesma coisa,

<sup>2</sup> "A escolha dos conceitos pode obedecer a um esquema de categorias reconhecidas como importantes no campo coberto pelo documento, ex.: o fenômeno, o processo, as propriedades, as operações, o material, o equipamento, etc." (WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY, 1981, p. 87)

porque o tema é constituído da presença de conceitos.

Na metodologia proposta por Tálamo (1987), o processo de indexação consiste em identificar o tema de um documento por meio de um mecanismo de perguntas e respostas agrupadas por generalidades e que respondem a cada uma das seguintes questões fundamentais: Quem? (ser), O que? (tema), Como? (modo), Onde? (lugar) e Quando? (tempo). Conforme a autora, identificando essa estrutura temática encontra-se o objetivo principal do texto, isto é, as informações relevantes separando-as das acessórias.

Kobashi (1994) baseou-se no mecanismo de perguntas conceituais de Lasswell (1971), utilizadas por Garcia Gutierrez e Lucas (1987), um método analítico para fins de indexação que abrange: Who, What, Whem, Where, Why. Destaca que a categoria “Quem?” não foi identificada em textos técnicos científicos, enquanto que a categoria “O que?” é essencial por ser o *elemento nuclear da estrutura temática*. Segundo a autora, as categorias Quando?, Onde? e Como? são categorias acessórias da principal “O que?”, podendo assim aparecer, ou não, no texto,

independente da “*ordem de precedência de entre elas*”.

O tema, portanto, possui uma estrutura temática composta por conceitos ou categorias ou facetas cuja identificação decorrerá da análise conceitual do documento. A composição das categorias identificadas formularão o tema do documento em questão.

A respeito de “onde” localizar conceitos depende da identificação da estrutura temática, contudo, conforme a legibilidade e a estrutura textual do documento, o tema poderá estar formulado de forma clara no “objetivo” do trabalho e, quando isso não acontecer, será preciso a identificação dos conceitos dentro da estrutura textual do documento.

A questão de determinação do tema, é preciso esclarecer, refere-se à determinação do assunto, de forma idêntica. Apenas consideramos como tema, e não assunto porque as citações acima, assim, mencionam.

Para alguns pesquisadores da área, é relevante nos referirmos à tematicidade (*aboutness*) do documento quando se

busca pesquisar sobre a problemática da identificação do tema. A tematicidade é pertinente à análise de assunto porque estamos tratando de seu objetivo principal que é a identificação do assunto ou tema mediante análise conceitual composta de identificação e seleção de conceitos. Podemos dizer que o assunto ou tematicidade do documento é o cerne principal e mais carente de esclarecimentos dentro dos estudos em análise documentária.

Conforme esclarece Albrechtsen (1993), o conceito de *aboutness* passou a ser pesquisado em substituição ao conceito de *subject* que, mais recentemente, retornou em outras pesquisas.

Todd (1992, p. 101) afirma que o termo *aboutness* é usado em literatura mais recente como sinônimo de assunto e relata, desde Cutter até Borko e Bernier, as noções acerca do termo assunto:

- assunto, do ponto de vista de Cutter, é tema ou tópico de pesquisa, estando ou não determinado no título;
- Kaiser descreve assunto como coisas em geral, reais ou imaginárias, e as

condições que estão ligadas a elas; noções estabelecidas pelas categorias concreto e processo;

- Ranganathan fala do conteúdo de um documento como um termo assumido, um isolado;
- Coates identifica o assunto como uma abstração da idéia geral que está contida no contexto de uma unidade da literatura;
- Vickery fala do tema dos quais os livros, partes de livros, partes de artigo estão escritos; como um agregado complexo de aspectos específicos; composto de termos elementares;
- Borko e Bernier definem assunto como “the foci of work”, o tema central para o qual a atenção e os esforços do autor foram direcionados. São aqueles aspectos do trabalho que contêm as idéias do assunto, explicações, ou interpretações, as quais podem ser indexadas.

Elucidando a definição acima, acreditamos que a tematicidade sempre será o conteúdo relevante do documento, ainda que, algumas variáveis influenciem na determinação desse conteúdo como os interesses informacionais dos usuários do sistema de informação, entre outras. Portanto, a escolha do tema de um documento sempre estará relacionada com os interesses de tais usuários, independente da quantidade de informações referentes ao tema selecionado.

Todd (1992, p. 102) diz que, segundo a opinião de Wilson (1985), podemos entender também que o grau de relação entre tematicidade e significado é variável porque depende do

[...] uso que a pessoa pode encontrar da tematicidade do documento numa certa época, e o mesmo documento pode vir a ter diferentes significados para o mesmo leitor em diferentes épocas, entretanto o documento possui uma atinência fundamental.

O termo *aboutness*, originário da língua inglesa e usado por Fairthorne (1969) e outros, pode significar ‘do que trata um texto’, em português. Em língua portuguesa há divergências entre os

pesquisadores para se traduzir *aboutness*: para alguns, pode ser “tematicidade”<sup>3</sup>, por se considerar como um substantivo ligado ao termo *temático*, enquanto outros adotam “atinência”<sup>4</sup>.

O conceito de assunto (*subject*), segundo Albrechtsen (1993, p. 219), voltou atualmente a se constituir como área central de estudo, visto que as abordagens realizadas em torno do conceito de *aboutness* ou *tematicidade* tendiam a manejar documentos como fontes isoladas de conhecimento, à exceção de Hutchins (1977) e Beghtol (1986). A re-inserção do conceito de assunto por Blair (1990), Hjörland (1992), Weinberg (1988) e Soergel (1985) enfatizou a função primária da indexação em servir à busca pelo conhecimento e recomendar que o indexador não deve direcionar a análise exclusivamente sobre o conteúdo, mas antecipar o impacto e o valor de um determinado documento para uso potencial.

<sup>3</sup> Preferimos usar tematicidade por considerar que esse termo está mais relacionado com a noção de tema do documento.

<sup>4</sup> O termo atinência foi empregado na tradução do livro “Indexação e resumos: teoria e prática” de Lancaster e é utilizado por Naves em seus trabalhos (1996, 2000)

Sob essa influência, Begthol (1986) fez distinção entre *aboutness* e *meanings*; *aboutness* é o conteúdo intrínseco do documento, que independe do uso temporal que um indivíduo possa fazer do mesmo em análise e que o faz possuir uma tematicidade relativamente permanente e um número variável de *meanings* (significados), podendo ser medido de acordo com o uso particular do documento tendo em vista os usuários. Em suma, por *aboutness* deve-se entender o conteúdo relativamente permanente do documento e por *meanings*, o significado compreendido pelo usuário.

Elucidando a definição de Begthol (1986), acreditamos que a tematicidade sempre será o conteúdo relevante do documento, porém, algumas variáveis irão influenciar na determinação desse conteúdo, como os interesses informacionais dos usuários do sistema de informação, entre outras. Portanto, a escolha de um assunto de um documento sempre estará relacionada com os interesses de tais usuários, independente da quantidade de informações referentes ao assunto selecionado.

Outra forma de se referir a essa questão é apresentada por Cavalcanti (1989), pesquisadora da área de leitura em Linguística Aplicada. Para a autora, a tematicidade intrínseca é tema importante para o autor e tematicidade extrínseca é o tema importante do ponto de vista do leitor. A autora denomina a tematicidade intrínseca de “saliência autor” e a extrínseca de “relevância leitor”.

Hutchins (1977, p. 33) destaca que o indexador, durante a busca pela compreensão do texto, procura identificar assuntos familiares contrariamente ao leitor comum que, durante a leitura normal, visa a encontrar informações novas para ampliar o seu conhecimento sobre o assunto. Na opinião da autora, o indexador, no momento da indexação, procura identificar primeiramente, até de maneira automática, temas familiares ao conhecimento prévio que adquiriu sobre a área de assunto na qual trabalha, mas é também preciso atentar para os temas novos que possam interessar aos usuários do documento ao mesmo tempo em que possibilitam ao indexador ampliar seu vocabulário sobre a terminologia da área.

O ideal é que haja uma equivalência entre a relevância do assunto do documento tanto para o indexador como para o usuário. Aquele atingirá esse objetivo se elaborar informações documentárias (índices e resumos) consistentes.

Considerando, portanto, que a determinação da tematicidade intrínseca e extrínseca faz parte da análise de assunto, entendemos que o indexador, durante a leitura documentária, realiza a análise de assunto na fase inicial do processo de indexação.

Entenda-se, até aqui, que a identificação e a seleção de conceitos são operações características de uma análise de assunto cuja concepção é orientada para o conteúdo e para a demanda e, portanto, voltada para a preservação do contexto do documento antes de acontecer a operação de tradução de termos para uma linguagem de indexação. Ocorre que a identificação e seleção de conceitos visam à composição de um enunciado temático que identifica o assunto do texto, sendo esse uma condição necessária à compreensão global do texto.

## 5 A LEITURA NO PROCESSO DE IDENTIFICAÇÃO DE CONCEITOS

Considerando o fato de que o conteúdo do documento estará melhor representado se a identificação e seleção de conceitos forem realizadas dentro da concepção orientada para o conteúdo e para a demanda, isso nos leva à necessidade de compreensão de leitura pelo indexador porque a análise orientada para o conteúdo pressupõe a explicitação do significado do texto, uma situação que não se resolve sem que haja compreensão de leitura.

Para analisarmos a importância da leitura para a análise de assunto em indexação, consideramos que a concepção orientada para o conteúdo está, de fato, comprometida com uma compreensão de leitura para identificação e seleção de conceitos, uma vez que:

[...] envolve uma interpretação do conteúdo do documento que vai além dos limites da estrutura superficial léxica e gramatical (ALBRECHTSEN 1993, p. 220).

Em análise da literatura, vemos que Foskett (1973), antes dos “Princípios de Indexação” do World Information System for Science and Technology (1981) e da

Norma 12.676, acreditava que a determinação do assunto de um documento deveria ser feita pela leitura, porém, como considerava que não havia tempo disponível para leitura do documento na íntegra, sugeriu partes do texto a serem lidas, alertando que o título é, muitas vezes, escolhido para chamar a atenção e não para indicar o assunto abordado.

Da mesma forma, Chaumier (1980, p. 43) considerou que, para fins de indexação, o conhecimento do conteúdo do documento "Faz-se através de uma leitura rápida, ou 'leitura em diagonal' do documento" enumerando, em seguida, as partes do texto para uma leitura mais precisa. Quando aborda a escolha de conceitos, não faz nenhuma referência à leitura, afirmando que a escolha depende de uma "verdadeira análise conceitual do documento".

Van Slype (1991, p. 116), contudo, ao estruturar o processo de indexação humana em quatro etapas (exame do conteúdo do documento, seleção dos conceitos, tradução dos conceitos em descritores e estabelecimento de relações sintáticas entre os descritores), recomenda

ao documentalista, na primeira fase, uma leitura rápida "em diagonal" de partes do documento para somente ver do que trata; na seleção de conceitos, indica claramente que "à medida que realiza a leitura, o documentalista identifica os conceitos [...]" e, mais à frente, esclarece quais noções devem ser retidas ou selecionadas.

A leitura para indexação, dentro da Norma 12.676, dos "Princípios" do World Information System for Science and Technology e dos teóricos da área, excetuando-se Van Slype, parece completamente desvinculada do processo de identificação e seleção de conceitos, sendo considerada como mero exame de partes do documento.

O que está evidente, dentro das abordagens de Foskett (1973), Chaumier (1980) e da Norma 12.676, é o aspecto mais técnico do exame do documento, sem se questionar ou identificar o indexador como leitor.

Com relação à compreensão, podemos considerá-la como condição à leitura, ou seja, não existe leitura sem compreensão. Então, quando falamos em leitura para

indexação, podemos dizer que o indexador necessita compreender o texto para identificar e selecionar conceitos, pois somente o fará a contento se houver compreensão. Também na visão de Farrow (1995, p. 243):

[...] o processo de indexação consiste da compreensão do documento a ser indexado, seguido pela produção de um conjunto de termos de indexação.

Pensando em termos de uma leitura documentária, Farrow (1991, p. 151) considera "[...] razoável assumir que indexadores compreendem o texto essencialmente do mesmo modo que os leitores fluentes [...]", porém, sob a influência de condições de tempo, objetivo definido, modelo a ser produzido e áreas de assunto específicas com estrutura textual padronizada dos documentos induzindo-os à um processo repetitivo e automático.

Lancaster (1993, p. 20-21) aborda a questão da leitura em indexação quando examina "A prática da indexação" em que se preocupa com as restrições de tempo e de quantidade de documentos do serviço de indexação ao ponderar que "Ao indexador raramente é dado o luxo de poder ler um documento do começo ao

fim". Mais à frente, também, refere-se às partes mais interessantes e oportunas à leitura do indexador, salientando o exame de partes do documento recomendado pela Norma 12.676. Porém, de maneira muito apropriada, considera como pressuposto o fato "[...] de que é possível ler o documento a ser indexado" e que o "[...] motivo é a decisão sobre o que incluir na indexação", baseando-se nos interesses da comunidade a que serve.

O mais importante, ainda, é a advertência que Lancaster (1993, p. 22) faz aos indexadores que realizam a "análise conceitual" influenciados pelo vocabulário do sistema, pois julga que isso compromete a representação do conteúdo do documento. Por isso discorda da Norma 12.676 (1992, p. 2) quanto à afirmação de que "tanto a análise quanto a transcrição devem ser realizadas com o auxílio de instrumentos de indexação [...]".

Da mesma forma acreditamos que o uso do vocabulário controlado durante a leitura para a identificação e seleção de conceitos poderá impedir a compreensão do contexto do documento e sua adequada representação, pois, o fato de

termos não associados ao vocabulário não serem considerados significativos estabelece que o indexador realize uma análise bastante superficial do conteúdo do documento, ou seja, não está representando realmente as idéias do autor, apenas ajustando-as a palavras, de uma forma muito simplista.

É preciso esclarecer ainda que, quando a Norma 12.676 (1992, p. 2), no item “Exame do documento”, aborda a questão de uma leitura completa do documento como “impraticável e nem sempre necessária” para a indexação, possivelmente está distinguindo, neste momento, a leitura documentária do processo global de leitura. Embora a norma não comente os motivos pelos quais admite ser impraticável a leitura completa do documento, é razoável supormos que isso se deve ao fato de que o trabalho de um indexador não se restringe a poucos documentos se considerarmos a totalidade do acervo de uma biblioteca.

A atividade de indexação inicia-se com a leitura do documento a ser analisado. Essa leitura, a documentária, difere da leitura comum porque exige outros

procedimentos, ainda que os conhecimentos necessários para um bom entendimento de um texto sejam comuns a ambas.

Na leitura documentária o leitor é caracterizado como indexador, podendo ser um bibliotecário ou indivíduo com formação superior na área de assunto com a qual trabalha (leitor-especialista). Esse leitor-indexador tem objetivo definido: identificação e seleção de conceitos que representem o conteúdo do texto e que coincidam com as necessidades informacionais da comunidade usuária do sistema de informação, não sendo possível a ele realizar uma “crítica da ciência e a avaliação das contribuições veiculadas pelos textos” (GINEZ DE LARA, 1993, p. 50), considerando que, na maioria das vezes, não apresenta conhecimento especializado sobre o assunto que indexa. Porém, isso não significa que o indexador bibliotecário não possa vir a ser um especialista por não possuir formação superior na área do assunto em que trabalha; pelo contrário, pode tornar-se um especialista com os anos de prática na atividade de indexação e cursos especializados.

A leitura realizada pelo indexador é também caracterizada como uma leitura dinâmica, não havendo necessidade de se fazer a leitura do item documental por completo, exceto na indexação para a elaboração de índices internos ou alfabético-remissivos – aqueles que acompanham um livro ao seu final, porque, como bem afirma Collinson ([1971], p. 18), “[...] em primeiro lugar o livro deverá ser lido inteiramente, duas ou três vezes, como um todo”, demonstrando que, nesse tipo de indexação, exige-se do indexador uma leitura completa da obra.

Como visto anteriormente, o indexador, por realizar uma leitura com objetivos profissionais, sofrerá a pressão da falta de tempo devido à grande quantidade de material que necessita ler para indexar. Na leitura para fins de indexação, portanto, o leitor-indexador deverá utilizar estratégias próprias de leitura documentária que lhe facilitem atingir o seu objetivo.

## **6 CONSIDERAÇÕES FINAIS**

A análise de assunto é a etapa mais importante do trabalho do indexador. Tem como objetivo identificar e selecionar os conceitos que representam a

essência de um documento. O processo de identificação de conceitos envolve certo grau de complexidade por exigir do indexador o uso de metodologia adequada para garantir bons resultados na recuperação, o que pressupõe o conhecimento de abordagens sistematizadas ao texto. Além disso, pela análise de literatura, a identificação de conceitos depende da tematicidade do texto e está atrelada à leitura do indexador e às suas concepções de análise de assunto adquiridas pela sua formação, objetivos e políticas de indexação.

A identificação e a seleção de conceitos são operações características de uma análise de assunto cuja concepção é orientada para o conteúdo e para a demanda. Portanto, voltada para a preservação do contexto do documento antes da tradução.

Considerando o fato de que o conteúdo do documento estará melhor representado se a identificação e seleção de conceitos for realizada dentro da concepção orientada para o conteúdo e para a demanda, isso nos leva à necessidade de compreensão de leitura pelo indexador porque a análise orientada para o

conteúdo pressupõe a explicitação do significado do texto, uma situação que não se resolve sem que haja compreensão de leitura.

Os aspectos teóricos que fundamentam os estudos sobre leitura documentária, indicam a importância da identificação de conceitos na análise de assunto. Por isso, deu-se especial enfoque aos estudos sobre tematicidade que envolve a determinação do assunto dentro da análise de assunto e às propostas metodológicas da indexação indicadas pela normalização e pelos estudos teóricos. Os estudos sobre tematicidade revelaram a necessidade de distinção entre tematicidade intrínseca (*aboutness* - inerente ao conteúdo documento) e extrínseca (*meanings* - significado para o usuário do sistema) e comprova a concepção de análise de assunto orientada para a demanda.

A seleção de conceitos é parte integrante da identificação de conceitos realizada durante a análise de assunto e existe para o indexador prever a adequação dos conceitos representados à recuperação conforme demanda do usuário.

Considera-se que a concepção de análise está diretamente vinculada com sua formação educacional (concepção orientada pelo conteúdo) e com a postura do sistema de informação (concepção orientada pela demanda) e não pelo fato de ele ser um leitor menos ou mais habilitado. Por isso, recomenda-se que a formação do indexador seja direcionada para a importância da identificação e seleção de conceitos feita durante a análise de assunto conforme uso de metodologia adequada.

## REFERÊNCIAS

ALBRETCHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. *The Indexer*, London, v.18, n. 4, p. 219-24, 1993.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 12676: Métodos para análise de documentos - determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro, 1992. 4 p.

AUSTIN, D. *PRECIS: a manual of concept analysis and subject indexing*. London: Council of the British National Bibliography, 1974. 551 p.

BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of Documentation*, London, v. 42, n. 2, p. 84-113, 1986.

BLAIR, D. C. *Language and representation in information retrieval*. Amsterdam: Elsevier Science Publisher, 1990.

CAVALCANTI, M. C. *Interação leitor-texto: aspectos de interpretação pragmática*. Campinas: UNICAMP, 1989. 271 p.

CESARINO, M. A. N.; PINTO, M. C. M. F. Análise de assunto. *Revista de Biblioteconomia de Brasília*, Brasília, v. 8, n. 1, p. 32-43, jan./jun. 1980.

CHAUMIER, J. *Travail et methodes du/de la documentaliste: connaissance du problème*. Paris : ESF/Libraries Techniques, 1980. Exposé 3, Chap.3: L'indexation, p.42-7.

CHU, C. M.; O'BRIEN, A. Subject analysis: the critical first stage in indexing. *Journal of Information Science*, London, v. 19, n. 6, p. 439-54, 1993.

COLLINSON, R. L. *Índices e indexação: guia para indexação de livros, e coleções de livros, periódicos, e coleções de livros, periódicos, partituras musicais, com uma seção de referência e sugestões para leitura adicional*. Tradução de Antônio Agenor Brinquet de Lemos. São Paulo: Polígono, [1971]. 223 p.

CRAVEN, T. C. Linked phrase indexing. *Information Processing and Management*, New York, v. 14, p. 469, 1978.

ESTEBAN NAVARRO, M. A. E. Elementos, actividades y critérios para la identificación, comprensión y selección de conceptos en la indización analítica. In: GARCIA MARCO, F. J. G. M. *Organización del conocimiento en sistemas de información y documentación*. Zaragoza: Capítulo Español de la ISKO, Universidad Carlos III de Madrid, 1999. v. 3, p. 69-93.

FAIRTHORNE, R. A. Content analysis, specification, and control. *Annual Review of Information Science and Technology*, Medford, NJ, v. 4, p. 73-109, 1969.

FARRADANE, J. A. A comparison of some computer produced permuted alphabetical subject indexes. *International Classification*, Munich, v. 4, n. 2, p. 94-101, 1977.

FARROW, J. All in the mind: concept analysis in indexing. *The Indexer*, v. 19, n.4, p.243-7, 1995.

- FARROW, J. A cognitive process model of document indexing. *Journal of Documentation*, London, v. 47, n. 2, p. 149-166, 1991.
- FOSKET, A. C. *A abordagem temática da informação*. Tradução de Agenor de Briquet de Lemos. São Paulo: Polígono, 1973. Tradução de: Subject approach to subject information.
- FUJITA, M. S. L. *PRECIS na língua portuguesa: teoria e prática de indexação*. Brasília: UnB/ABDF, 1989.
- \_\_\_\_\_. *A leitura em análise documentária*. 1998. 184 f. Relatório final (Projeto Integrado de Pesquisa) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista; CNPq, Marília.
- GARCIA GUTIERREZ, A.; LUCAS, R. *Documentación automatizada de los médios informativos*. Madrid: Paraninfo, 1987.
- GINEZ DE LARA, M. L. *A representação documentária: em jogo a significação*. São Paulo, 1993. 133 f. Dissertação (Mestrado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo.
- HJÖRLAND, B. The concept of subject in information science. *J. Doc.*, London, v. 48, n. 2, p.172-200, 1992.
- HUTCHINS, W. K. On the problem of aboutness in document analysis. *Journal of Informatics*, v. 1, p. 17-35, 1977.
- INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *Documentation - methods for examining documents, determining their subjects, and selecting indexing terms*. Suíça: ISO, 1985. 5p. (ISO 5963-1985 (E))
- KAISER, J. O. *Systematic indexing*. London: Pitman, 1911.
- KOBASHI, N. Y. *A elaboração de informações documentárias: em busca de uma metodologia*. 1994. 195 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo.
- LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Trad. de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos/Livros, 1993.
- LANGRIDGE, D. *Classificação: abordagem para estudantes de biblioteconomia*. Tradução de Rosali P. Fernandes. Rio de Janeiro: Interciência, 1977. 120 p.
- LASSWELL, H. D. A estrutura e a função da comunicação na sociedade. In: COHN, G. *Comunicação e indústria cultural*. São Paulo: Nacional/EDUSP, 1971.
- LYNCH, M. F.; PETRIE, J. H. A program suite for the production of articulated subject indexes. *Computer Journal*, Oxford, v. 16, p. 46-51, 1973.
- METCALFE, J. *Subject classifying and indexing of libraries and literature*. New York: Scarecrow, 1959.
- NAVES, M. M. L. *Fatores interferentes no processo de análise de assunto: estudo de caso de indexadores*. 2000. 257 f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte.

- NAVES, M. M. L. Análise de assunto: concepções. *Revista de Biblioteconomia de Brasília*, Brasília, v. 20, n. 2, p. 215-226, jul./dez, 1996.
- NEELAMEGHAM, A.; GOPINATH, M. A. Postulated-based permuted subject indexing (POPSI). *Library Science with a slant to documentation*, v. 12, n. 3, p. 79-87, 1975.
- PIEDADE, M. A. R. *Introdução à teoria da classificação*. 2. ed. rev. aum. Rio de Janeiro: Interciência, 1983. 221 p.
- PINTO MOLINA, M. Interdisciplinary approaches to the concept and practice of written text documentary content analysis (WTDCA). *Jornal of Documentation*, London, v. 50, n. 2, p. 111-133, jun. 1994.
- RANGANATHAN, S. R. *Colon Classification*. E. Goldston: London, 1933.
- RANGANATHAN, S. R. *Colon Classification*. 6. ed. Bombay: Asia Publishing House, 1960.
- SOERGEL, D. *Organizing information – principles of database and retrieval systems*. New York: Academic Press, 1985.
- TÁLAMO, M. F. G. M. *Elaboração de resumos*. Escola de Comunicação e Artes, 1987. 14 f. Datilografado.
- TODD, R. T. Academic indexing: what's it all about? *The Indexer*, London, v. 18, n. 2, p. 101-104, apr. 1992.
- VAN SLYPE, G. *Los lenguajes de indización: concepción, construcción y utilización en los sistemas documentales*. Trad. Pedro Hípola e Félix de Moya. Madrid: Fundación Germán Sánchez Ruipérez; Pirámide, 1991. 200p. Tradução de: Les langages d'indexation: conception, construction et utilisation dans les systèmes documentaires.
- VICKERY, B. C. Analysis of information. In: KENT, A., LANCOUR, H. (Ed.). *Encyclopedia of library and information science*. New York: Decker, 1968. v. 1, p. 355-384.
- \_\_\_\_\_. *Classification and indexing in science*. Londres: Butterworths Scientific Publications, 1975
- \_\_\_\_\_. *Classificação e indexação nas ciências*. Tradução de Maria Christina Girão Pirolla. Rio de Janeiro: BNG/Brasilart, 1980. 274 p.
- WEINBERG, B. H. Why indexing fails the researcher. *The Indexer*, London, v. 16, n. 1, p. 3-6, 1988.
- WILSON, P. Subject and the sense of position. In: CHAN, C. et al. *Theory of subject analysis: a manual*. Littleton, Colorado: Libraries Unlimited, 1985. p. 306-23.
- WORLD INFORMATION SYSTEM FOR SCIENCE AND TECHNOLOGY. Princípios de indexação. *R. Esc. Bibliotecon. UFMG*, v. 10, n. 1, p. 83-94, 1981.

**Mariângela Spotti Lopes Fujita**

Profa. Assistente Dra. Do Departamento de  
Ciência da Informação da UNESP – Campus  
de Marília, SP  
Pesquisadora pelo CNPG

[goldstar@flash.tv.br](mailto:goldstar@flash.tv.br)

**Agradecimentos:**

Ao CNPQ pelo auxílio, confiança e  
valorização dispensados à pesquisa  
durante dez anos de colaboração mútua.

Artigo aceito para publicação em:  
30 maio 2003