

# FUNDAMENTOS TEÓRICOS DA REPRESENTAÇÃO DO CONHECIMENTO E DA INFORMAÇÃO: UMA DISCUSSÃO SOBRE O INDEXADOR<sup>1</sup>

Email:  
dilzabastos@gmail.com  
annalima27@gmail.com  
lourdes.maria@gmail.com

Dilza Ramos Bastos, Ana Paula Lima dos Santos, Lourdes Maria dos Santos

## *Resumo*

Discorre sobre a representação do Conhecimento e da Informação e atesta a importância do indexador no processo de busca/recuperação da informação. Apresenta um breve histórico sobre a indexação, diferenciando a indexação pré-coordenada e pós-coordenada. Como metodologia na elaboração do artigo foi empregada a Revisão de Literatura. O trabalho tem como objetivo refletir sobre o profissional indexador e suas atividades no âmbito da recuperação da informação e sua importância para a sociedade da informação contemporânea. Conclui ressaltando a importância do permanente estudo da indexação por parte dos profissionais atuantes nesse processo, visando à melhoria da recuperação da informação e produção do conhecimento.

Palavras-chave: Indexação pré-coordenada. Indexação pós-coordenada. Indexador. Recuperação da informação.

## *Abstract*

This work discusses the Representation of Knowledge and Information and attests the importance of the indexer in the information search/retrieval process. It presents a brief history about indexing, differentiating pre-coordinate and post-coordinate indexing. The Literature Review methodology was used in the elaboration of the paper. The work aims to reflect on the professional indexer and its activities in the scope of information retrieval and its importance to the contemporary information society. It concludes by highlighting the importance of the permanent study of indexation by the professionals working in this process, aiming at improving information retrieval and knowledge production.

Keywords: Pre-coordinate indexing. Post-coordinate indexing. Indexer. Information retrieval.

<sup>1</sup> O trabalho é uma comunicação da pesquisa intitulada *Aplicação de metodologia de indexação temática a textos jurídicos*, que se encontra em andamento, fazendo parte do Programa de Incentivo à Produção do Conhecimento Técnico Científico na área da Cultura da Fundação Casa de Rui Barbosa.

## INTRODUÇÃO

A Representação do Conhecimento e da Informação nunca foi tão relevante e pertinente. Logo, aprofundar os estudos sobre essa temática é essencial, uma vez que, a sociedade da informação tem buscado e produzido uma grande quantidade, tanto de informações, quanto de conhecimentos. Dessa forma, o profissional da informação deve se adequar para oferecer melhores serviços de informação, bem como aperfeiçoar as técnicas de processamento da informação procurando sempre o melhor caminho para criar e aperfeiçoar ferramentas adequadas e funcionais visando atender a um público cada vez mais exigente e ao mesmo tempo perdido, diante de tantas informações. A internet apesar de ser um “mundo” onde se encontra tudo, não dispensa um mediador ou um especialista capaz de filtrar as informações, dentro dos aspectos, de: autoridade, conteúdo, confiabilidade necessários para que se produzam conhecimentos íntegros. Isto é: mesmo estando interagindo de forma independente na busca por informações, o usuário conta com ferramentas/operadores dos sistemas disponíveis, se bem que em muitas vezes se torna necessária ou uma melhor oportunidade a mediação de profissionais da informação, para que se chegue a resultados positivos. Podemos ressaltar assim que esses sistemas previamente são construídos com dispositivos e estratégias para que eles próprios atuem como mediadores entre a informação e o usuário.

Nesse sentido, cabe ressaltar a importância do profissional indexador nesse processo, uma vez que será ele quem evidenciará o que trata o documento. Quanto mais capacitado for esse profissional, maior será a qualidade da indexação feita por ele e, conseqüentemente, maior será a qualidade da recuperação da informação.

Este artigo tem como objetivo refletir sobre o profissional indexador e suas atividades no âmbito da recuperação da informação, bem como sobre sua importância para a sociedade da informação contemporânea, uma vez que os acessos online e os atendimentos virtuais têm aumentado. Quanto à coordenação de termos que representam o assunto das obras, nos sistemas informatizados, o uso de descritores ou de cabeçalhos de assunto irá depender de uma política institucional e cada instituição irá usar a que for melhor para os serviços que oferecem ao público específico que atende. Além do mais, concordamos com Brasil (2002, p. 4), que ao discorrer sobre a questão do uso de tecnologias nos catálogos informatizados, afirma que

atualmente, o uso de instrumentos de Controle Terminológico para o apoio à indexação e recuperação da informação torna-se imperativo diante de sistemas informatizados, acesso on-line a base de dados na Internet, organização de bibliotecas virtuais e outros instrumentos para a organização e disseminação da informação, que exigem melhor controle da terminologia para eficiente recuperação e filtragem de informações.

Por isso, o profissional indexador deve perseguir o estudo permanente para executar o melhor serviço, ou seja, a indexação “perfeita”.

Nesse contexto se torna fundamental o domínio não apenas das técnicas, mas dos conceitos que envolvem as técnicas para executar a indexação.

## *2 INDEXADOR: DESCRITORES OU CABEÇALHOS DE ASSUNTO?*

De acordo com Naves (2001, p. 190) o indexador é o profissional responsável “por todo o processo de análise de assunto, tendo a sua figura ocupado um papel de destaque neste trabalho, pois a ele é creditado, em grande parte, o sucesso ou insucesso de um sistema de recuperação da informação”. Recai sobre este profissional uma responsabilidade maior no que se refere à ação de indexar e assim dar maior visibilidade ao conteúdo, considerando-se também que pode ampliar os pontos de acesso à obra em relação à tematicidade.

Ainda de acordo com Naves (2001, p. 191) é preciso se ter cuidado quando se fala em indexador e se atentar a precisão conceitual deste termo. Uma vez que “nas literaturas inglesa e americana, o termo indexador é aplicado tanto àquela pessoa que elabora índices de textos ou livros quanto àquela que faz a indexação acadêmica”. O termo é adotado para se referir a todos os profissionais que fazem o tratamento de assunto que abrange a análise do documento, a seleção de conceitos e a tradução desses conceitos em uma linguagem própria do sistema de recuperação da informação.

Diante de tantas indagações quanto ao futuro do profissional indexador e quanto às tecnologias emergentes, frente ao tratamento de bibliotecas virtuais e acervos digitais, Naves (2001) questiona também se haveria lugar para o indexador humano, porém ela mesma fornece a resposta afirmando que, pelo menos, até hoje não se transferiu para máquina a tarefa que se faz presente nas atribuições do indexador humano, como a abstração, a percepção, a interpretação, dentre outras que são inerentes à mente humana.

Para Araujo Junior (2007, p. 20 e 24) a indexação é a tradução de um documento em termos documentários, ou seja, em descritores, cabeçalhos de assunto, palavras chaves, que têm como objetivo “expressar o conteúdo do documento ou como o processo de atribuir termos ou códigos de indexação a um registro de documentos, termos ou códigos esses que serão úteis posteriormente na recuperação da informação”. Desse modo, a indexação pode ser manual ou automática, sendo que na primeira sem o auxílio da atribuição automática de termos ou extração, ou seja, a indexação manual de termos é a indexação feita pelo homem. Já a indexação automática é qualquer procedimento que permita identificar e selecionar os termos que representam o conteúdo dos documentos sem a intervenção direta do documentalista.

Na perspectiva de Silva e Fujita (2004), o conceito de indexação apareceu a partir da elaboração de índices, porém hoje está mais atrelado ao conceito de análise de assunto. Com a necessidade de uma recuperação da informação mais rápida e precisa por parte das instituições que trabalham com a informação, naturalmente houve uma evolução da prática da indexação, com uma nova roupagem metodológica e instrumentos mais diversificados e voltados mais para o contexto específico do documento.

Silva e Fujita (2004, p. 136) ressaltam ainda que “a partir da evidência da Documentação como área científica na década de sessenta e do surgimento dos serviços de informação em áreas especializadas”, a indexação e a elaboração de resumos, que são utilizados na prestação de serviços bibliográficos para a recuperação de artigos de periódicos científicos, ganharam notoriedade e espaços até hoje reconhecidos. O termo “indexação” tornou-se ainda mais abrangente ao ser destacado por Bradford (1961), para análise de documentos.

No campo da Análise Documentária existem vertentes diferentes. Na linha teórica de Gardin (1981 apud SILVA; FUJITA, 2004, p. 136), por exemplo, a indexação é entendida

como “uma operação de representação documentária com a finalidade pragmática de Recuperação da Informação”. No entanto, sob a ótica de outros teóricos, “principalmente ingleses e norte-americanos, a indexação é a própria Análise Documentária, compostas das mesmas etapas operacionais com o objetivo de representação do conteúdo informacional” para que índices sejam elaborados.

A partir dessa evolução que determinou a importância do contexto do documento para uma efetiva recuperação da informação, Silva e Fujita (2004, p. 136) afirmam que “a área da indexação passa a incorporar os estudos dirigidos à compreensão do conteúdo dos textos a serem analisados” e que esses estudos estão de forma muito clara, inseridos “em correntes teóricas” sendo fácil “confundir na literatura, a função da indexação perante a necessidade de análise de conteúdo”, observando dessa forma na literatura duas vertentes teóricas: a francesa e a inglesa.

A corrente francesa assume a expressão Análise Documentária, introduzida por Gardin, enquanto que na corrente inglesa a análise documentária e a indexação compreendem os mesmos processos, incluindo a análise de assunto como a etapa inicial da indexação (SILVA; FUJITA, 2004).

De acordo com Cesarino e Pinto (1978, p. 273) “todas as linguagens de indexação exercem a mesma função nos sistemas de recuperação da informação”. São elas:

- representar o assunto de uma forma consistente;
- permitir a coincidência entre a linguagem do indexador e a do pesquisador;
- possibilitar ao indexador alternar o nível de pesquisa, do específico para o mais geral ou o contrário, de acordo com a necessidade do usuário.

Ainda de acordo com esses autores podemos encontrar na literatura duas vertentes para as linguagens de indexação, partindo de diferentes critérios. A separação mais conhecida divide as linguagens em sistemas alfabéticos e sistemas classificados. Os sistemas alfabéticos usam termos da própria linguagem natural, já os sistemas classificados têm por base as classificações arbitrárias “do conhecimento humano, dando normalmente uma notação simbólica para as classes, e determinam uma ordenação com base lógica, de acordo com os símbolos usados”. (CESARINO; PINTO, 1978, p. 273).

O conceito mais utilizado na literatura sobre cabeçalhos de assunto o define como “palavra ou grupo de palavras que expressam o conteúdo de um documento”. (CESARINO; PINTO, 1978, p. 273). Logo, a primeira forma de organização iniciou-se com as bibliografias que listavam as obras por autor. Diante disso começou a se impor a necessidade das listagens por assunto. Essas listas vinham no final das listas de autor e com a padronização das apresentações de assuntos essas ordenações começaram a ser organizadas de forma alfabética ou classificadas. Alguns dos fatores que foram determinantes para o surgimento dos cabeçalhos de assunto foram:

- os títulos das obras não representavam de forma adequada o assunto tratado;
- problemas associados as subdivisões de assuntos;
- obras com mais de um assunto;
- livros com assuntos relacionados;
- obras que relacionavam os assuntos a épocas e lugares diversificados.

Essas primeiras “regras” foram elaboradas por Charles Ammi Cutter (1837-1903), em 1876, antes disso os cabeçalhos eram atribuídos de acordo com o catalogador. Para Cutter, se a indexação não fosse construída com regras pré-estabelecidas de uma forma precisa, não haveria como o usuário encontrar a entrada correta para determinado assunto. Assim, este autor desenvolveu três princípios na elaboração de um catálogo alfabético de assunto:

- a) O princípio da especificidade – onde o assunto deveria dar acesso pelo termo mais específico e não pela classe a que está subordinado.
- b) O princípio do uso – o princípio da conveniência de acordo com as necessidades dos usuários.
- c) O princípio sindético – se baseia no alfabeto dos cabeçalhos de assunto, fazem aproximações de assuntos e ao mesmo tempo, dividem assuntos relacionados e que hoje conhecemos pelas remissivas “ver” e “ver também”.

Brasil et al (2002) concluem que a representação de assuntos presentes em documentos se relaciona ao tipo de sistema adotado pela instituição, que por sua vez depende da tecnologia utilizada e, conseqüentemente, sistemas pré ou pós coordenados têm a mesma dependência. Sendo assim, observam que a maioria das bibliotecas brasileiras se utiliza dos cabeçalhos de assunto para o tratamento dos seus acervos, porém para o tratamento de obras especializadas não é o mais indicado, por conta da especificidade temática, logo para coleções especializadas os descritores são mais eficientes.

Na verdade, a mudança para a utilização da pós-coordenação se deu por ter sido a pré-coordenação suplantada em razão do advento das novas tecnologias da informação, pois os sistemas informáticos passaram a ser capazes de realizar operações booleanas, combinando assim termos e “cruzamentos” de forma a associar ou excluir conceitos durante as operações de busca. Portanto, o cabeçalho de assunto, constituído pela prévia combinação de conceitos não faria mais sentido, já que posteriormente as estratégias de busca podem realizar a associação desses conceitos. Por exemplo, no cabeçalho de assunto “Educação – Trabalhador rural” encontram-se dois conceitos distintos, que estão assim pré-coordenados representando uma relação entre o conceito Educação e o conceito Trabalhador rural, provavelmente representando o assunto de uma obra que trata da educação de trabalhadores rurais. Na pós-coordenação os dois conceitos estarão representados pelos descritores “Educação” e “Trabalhador rural”, podendo ser associados na estratégia de busca para esse caso, mas também associados para outras necessidades informacionais. Desse modo, a pós-coordenação permite acesso múltiplo à informação.

O assunto da obra é assim apresentado pelo descritor ou pelo conjunto de descritores representando os conceitos constantes nessa obra, portanto, salienta Vickerey (1986) que é necessário decidir como o conhecimento pode ser representado de forma que as representações possam ser manipuladas.

### *3 CONSIDERAÇÕES FINAIS*

Normalmente, o potencial informativo de um documento é analisado em dois aspectos: os formais e os temáticos. Segundo o aspecto temático, o indexador deve prezar pelo resgate de

todos os pontos de acessos possíveis, ou seja, o conteúdo do documento deve ser analisado e desdobrado por meio do assunto/conteúdo, apresentando os aspectos específicos que visam possibilitar sua estratificação. Já no aspecto formal, o indexador deverá analisar a informação com o objetivo de investigar tanto o objeto de análise quanto o processo de produção de um item informativo; observar a definição contextual; a intenção do documento levando em conta a coleção na qual estão inseridos, os objetivos organizacionais do serviço de recuperação da informação e as necessidades de informação dos usuários.

Nesse sentido, o especialista em indexação não deve apenas dominar as técnicas e rotinas de serviço, mas, sim priorizar a contextualização do documento, bem como as teorias que a norteiam, sempre priorizando o usuário e se certificando se o que ele está indexando realmente retratando o documento.

Vale ressaltar que para que esse processo se dê com qualidade e eficácia, que seja feita uma avaliação contínua dos serviços informacionais oferecidos aos usuários. Desse modo, se certificar se os serviços oferecidos realmente atendem às necessidades dos usuários, caso contrário esses serviços devem ser ajustados, não perdendo de vista que a informação é dinâmica e o que anteriormente representava um conceito, pode hoje não representar. Por isso acompanhar essa dinâmica se torna imperativo para que a representação da informação esteja sempre em processo evolutivo.

Todavia, não há como garantir que o conteúdo seja plenamente percebido e representado pelo indexador. Tal questão é abordada por Mai (2001), ao afirmar que para o indexador, por mais experiência que tenha, é quase impossível que ele consiga precisar todos os fatos, as ideias e os significados que estejam presentes ou associados a um documento, haja vista que cada pessoa tem uma percepção diferente sobre o mesmo assunto, que pode divergir dependendo do momento, do lugar em que estão inseridos.

Dessa forma, salienta Bates (1998) que antecipar os termos pelos quais os pesquisadores procurariam as informações e atendendo assim às suas necessidades, é o desafio para o indexador. Segundo Sayão (1996), os pesquisadores se encontrariam em um estado transitório caracterizado por um alto grau de indefinição em relação ao assunto em pauta, estando ainda nebulosas suas próprias interrogações. A representação do conteúdo do documento é de grande importância para atender à demanda informacional, pois o pesquisador procura referências que permitam reconstruir seu conhecimento e orientar seu trabalho.

Alguns teóricos consideram que o texto possui diferentes significados de acordo com o uso particular que uma pessoa pode fazer do conteúdo em dado momento, porém um documento pode ser relevante para uma necessidade de informação sem tratar diretamente dela. Por essa razão, o que se almeja em um sistema documentário é a captação do conteúdo informativo do documento de forma a traduzi-lo em uma linguagem que seja intermediária entre o usuário e o documento.

## *REFERÊNCIAS*

ARAÚJO JUNIOR, Rogério Henrique de. **Precisão no processo de busca e recuperação da informação**. Brasília: Thesaurus, 2007.

BATES, Marcia J. Indexing and access for digital libraries and the Internet: human, database, and domain factors. **Journal of the American Society for Information Science**, vol. 49, n. 13, p. 1185-1205, nov. 1998.

BRADFORD, S. C. **Documentação**. Rio de Janeiro: Fundo de Cultura, 1961.

BRASIL, Maria Irene et al . Vocabulário sistematizado: a experiência da Fundação Casa de Rui Barbosa. In: CONGRESSO INTERNACIONAL DE ARQUIVOS, BIBLIOTECAS, CENTROS DE DOCUMENTAÇÃO E MUSEUS (1. : 2002 : São Paulo), 2002, São Paulo. Integrar: 1º CONGRESSO INTERNACIONAL DE ARQUIVOS, BIBLIOTECAS, CENTROS DE DOCUMENTAÇÃO E MUSEUS: textos. São Paulo: Imprensa Oficial do Estado, 2002. v. 1. p. 81-93.

CESARINO, Maria Augusta da Nóbrega; PINTO, Maria Cristina Mello Ferreira. Cabeçalho de assunto como linguagem de indexação. **Revista da Escola de Biblioteconomia da UFMG**. Belo Horizonte, v.7, n. 2, p. 268-288, 1978.

MAI, J.E. Semiotics and indexing: an analysis of the subject indexing process. **Journal of Documentation**, London, v. 57, n. 5, p. 591-522, sep. 2001.

NAVES, Madalena Martins Lopes. Estudos de fatores interferentes no processo de análise de assunto. **Perspect. ciênc. inf.**, Belo Horizonte, v. 6, n. 2, p. 189-203, jul./dez. 2001.

SAYÃO, Fernando. Bases de dados: a metáfora da memória científica. **CI. Inf.**, Brasília, v. 25., n. 3, p. 314-318, set./dez. 1996.

SILVA, Maria R; FUJITA, Mariângela S. L. A prática da indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, Campinas, v. 16, n. 2, p.133-161, maio/ago, 2004.

VICKERY, B. C. Knowledge representation: a brief review. **Journal of Documentation**, vol.42, no 3, p. 145-159, 1986.