

A REPRESENTAÇÃO DOCUMENTÁRIA NO PROCESSO DE INDEXAÇÃO COM O MODELO DE LEITURA DOCUMENTÁRIA PARA TEXTOS CIENTÍFICOS E LIVROS: uma abordagem cognitiva com protocolo verbal.

RESUMO - A representação documentária na indexação realiza-se na análise de assuntos para a determinação da tematicidade intrínseca e extrínseca e depois na etapa de tradução com uso de uma linguagem documentária. A representação por conceitos realizada para a determinação da tematicidade intrínseca durante a análise de assuntos é demonstrada pelo Modelo de Leitura documentária para a indexação de textos científicos e de livros. Para tanto, realizou-se uma investigação sobre leitura documentária de catalogadores para indexação de livros em bibliotecas universitárias. A aplicação da técnica introspectiva de Protocolo Verbal Individual com catalogadores de uma amostra de nove bibliotecas universitárias propiciou a análise de procedimentos de catalogação de assuntos que permitiram verificar as partes da estrutura textual do livro exploradas bem como a localização dos termos identificados e selecionados. Os resultados obtidos proporcionaram a adaptação do Modelo de Leitura Documentária para indexação de livros.

Palavras-chaves: Modelo de Leitura Documentária; Indexação; Protocolo Verbal

Mariângela Spotti Lopes Fujita
Professora Titular - Universidade
Estadual Paulista – UNESP – *Campus*
de Marília. Departamento de Ciência
da Informação.

fujita@marilia.unesp.br

THE DOCUMENTARY REPRESENTATION IN THE INDEXING PROCESS WITH THE MODEL FOR DOCUMENTARY READING SCIENTIFIC TEXTS AND BOOKS: A COGNITIVE APPROACH TO VERBAL PROTOCOL.

Abstract - The documentary representation in indexing is carried out, at first, in the analysis of subjects to establish the intrinsic and extrinsic aboutness, and then, in the translation stage using the documentary language. The representation by concepts for determining the intrinsic aboutness during the subject analysis is shown by the Documentary Reading Model for indexing of scientific texts and books, Therefore, it was carried out a search on documentary reading for cataloguers for indexing books in academic libraries. A research on catalogers' documentary reading for indexing books was carried out in nine university libraries. The application of the introspective technique of Individual Verbal Protocol with the catalogers allowed an analysis of subject cataloguing procedures, which made it possible to verify the textual structure parts of a book, as well as to locate the terms identified and selected. The results led to an adaptation of the Documentary Reading Model to book indexing.

Keywords: Documentary Reading Model; Indexing; Verbal Protocol

1. INTRODUÇÃO

A representação documentária está presente no processo de indexação em dois momentos distintos: na análise de assunto quando se realiza a identificação e seleção de conceitos mediante representação por conceitos universais tais como a ação, o objeto que sofreu a ação, o agente que praticou a ação e etc, bem como na tradução, etapa na qual os termos serão traduzidos para os termos de uma linguagem documental que compatibilizará tanto os termos indexados quanto os termos que identificam as necessidades dos usuários.

O estudo de leitura documentária para indexação em abordagem cognitiva propiciou a elaboração de um modelo de leitura documentária para textos científicos (Fujita e Rubi, 2006a, 2006b) que ora propõe-se adaptação para a catalogação de assuntos de livros em bibliotecas universitárias e à formação inicial de indexadores aprendizes. Para a adaptação foi desenvolvido estudo metodológico de abordagem cognitiva com uso da técnica introspectiva de protocolo verbal e análise qualitativa dos resultados da transcrição de protocolo verbal de catalogadores na tarefa de catalogação de assuntos de livros em bibliotecas universitárias. Os resultados obtidos da análise qualitativa indicou propostas para a elaboração do Modelo de Leitura Documentária de livros que foi incluído em Manual de Ensino de Leitura Documentária acompanhado de procedimentos de análise de assunto para catalogação de assuntos de livros.

Para demonstrar a representação por conceitos utilizou-se a metodologia do Modelo de Leitura Documentária para a indexação de textos científicos e para a catalogação de livros.

2. A REPRESENTAÇÃO DOCUMENTÁRIA NO PROCESSO DE INDEXAÇÃO: ELUCIDANDO A DIFERENÇA ENTRE A REPRESENTAÇÃO POR CONCEITOS NA ANÁLISE E A REPRESENTAÇÃO POR LINGUAGENS NA TRADUÇÃO

A qualidade da representação documentária é avaliada pela recuperação de conteúdos documentários pertinentes. A esse respeito Mai (2000) considera que “Se os documentos são representados pobremente ou inadequadamente, a qualidade da busca será igualmente deficiente”. Esta relação direta de causa e efeito entre a representação e a recuperação é intermediada pela indexação.

Moreiro (2004, p.48) considera que o processo de indexar consiste em descrever e caracterizar um documento com a ajuda de representações dos conceitos nos conteúdos com a finalidade de permitir uma busca eficaz das informações incluídas na coleção documental.

A representação do conhecimento contido em conteúdos documentários é o objetivo principal da indexação, entendido de modo mais amplo. Em âmbito mais específico e considerando-se as perspectivas teóricas da indexação conceitual, a representação por conceitos assume função preponderante entre o significado do conteúdo documentário e o termo que o representa. Em continuidade, o termo que representa o conteúdo documentário é representado por uma linguagem documentária. Entendemos, dessa forma, que durante o processo de indexação se realiza, naturalmente, tanto a representação por conceitos durante a análise quanto a representação por linguagem durante a tradução. Entretanto, para demonstrar a ambivalência da representação documentária durante o processo de indexação quando se realiza a análise e, depois, a tradução, é preciso considerar que a representação durante a análise pode ser realizada por conceitos.

O processo de indexação na literatura publicada tem, geralmente, 2 e 3 etapas indicadas para indexadores e catalogadores:

- Processo de 2 etapas:

Reconhecimento e extração dos conceitos informativos e tradução desses conceitos na linguagem documental (CHAUMIER, 1986, p.28);

Análise dos documentos e das perguntas para a seleção dos conceitos explícitos ou implícitos e armazenamento dessas palavras chave tal como estão ou sua normalização por meio de um vocabulário controlado (GIL LEIVA, 2008, p.245)

Reconhecimento dos conceitos que contém a informação nos documentos a indexar e a representação desses conceitos pela linguagem documentária (WALLER, 1999, p.149)

Determinação do assunto e tradução do assunto pela linguagem de indexação (MAI, 2000);

Análise conceitual e representação ou formalização (TAMAYO & VALDEZ, 2008, p.38)

- Processo de 3 etapas:

Análise de conteúdo, determinação do assunto e conversão para uma linguagem de indexação (CLEVELAND & CLEVELAND, 1990, p.104);

Exame do documento e determinação de seu conteúdo, identificação e seleção dos conceitos principais de seu conteúdo e seleção dos termos de indexação (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992)

Processo de análise do documento é a análise do documento pelo seu assunto, processo de descrição do assunto é a formulação de uma frase de indexação e o processo de análise de assunto é a tradução da descrição do assunto por uma linguagem de indexação (MAI, 2000, p.277)

Determinação da tematicidade ou assunto do conteúdo de um item, Análise conceitual para decidir quais aspectos de um item deverão ser representados em um registro bibliográfico e Tradução dos conceitos ou aspectos dentro de um vocabulário controlado, criando os pontos de acesso (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 1985 apud OLSON & BOLL, 1991, p.87).

Outros autores, porém, indicam 4 e 5 etapas no processo de indexação como demonstrado por Gil Leiva (2008, p.247):

Quatro etapas	a) Conhecimento do conteúdo conceitual do documento, b) Extração dos conceitos em linguagem natural, c) Tradução desses conceitos com a linguagem documental, e d) Busca de outros conceitos pertinentes não expressados pelo autor	Dijk & Slype (1972, p. 105)
	a) Contato com o documento; b) Identificação dos conceitos explícitos e implícitos do documento, c) Tradução dos conceitos expressados em linguagem natural por descritores, d) Estabelecimento de ligações sintáticas entre os descritores	Slype (1991, p.116)
Cinco ou mais etapas	a) Revisão dos objetivos desta operação, se necessário, b) Conhecimento prévio do documento, c) Determinação de seu tema principal, d) Identificação dos elementos do conteúdo que se devem descrever e a extração dos termos correspondentes, e) Verificação da pertinência dos termos selecionados, f) Tradução dos termos da linguagem natural pelos termos correspondentes da linguagem documental, se procede, g) Verificação da pertinência desta descrição, e h) Formalização da descrição quando o sistema prevê regras especiais de apresentação ou de redação	Guinchat & Menou, 1983, p. 179

Quadro 1: Etapas na indexação
Fonte: GIL LEIVA, 2008, p. 247

Outros autores, na literatura, explicam que a representação do assunto do documento consiste de duas, três ou quatro etapas do processo de indexação. A explicação para a quantidade de 2, 3 e 4 etapas é resumida por Mai (2000, p. 277) da seguinte forma:

O procedimento em duas etapas consiste de uma etapa em que o assunto é determinado e outra etapa em que o assunto é traduzido e expresso na linguagem de indexação. O procedimento de três etapas insere uma etapa na qual o assunto é formulado explicitamente ou implicitamente. No procedimento de quatro etapas, a tradução do assunto para a linguagem de indexação consiste de duas etapas. O indexador primeiro traduz o assunto de seu vocabulário no vocabulário utilizado na linguagem de indexação. Em seguida, o indexador constrói a entrada de assunto na linguagem de indexação na forma de termos de índice, um código de classificação, ou um cabeçalho de assunto.

A partir da explicação do procedimento em três etapas, Fujita (2007b) apresenta resultados de observação da leitura de indexadores durante a realização da primeira fase do processo de indexação e verifica, assim como Chu e O'Brien (1993, p.439), que a análise de assunto identifica e seleciona os principais tópicos do assunto de um documento, precedendo a fase de tradução desses tópicos pela linguagem de indexação. Além disso, confirma que durante a leitura do indexador a análise de assunto subdivide-se

nas etapas de identificação e seleção de conceitos (FUJITA, 2003a), conforme Norma ABNT (12.676).

Na identificação de conceitos é que se realiza a representação por conceitos do que compreendemos por indexação conceitual. Esta é a fase em que ocorre, conforme Fujita (2007b) a determinação da tematicidade intrínseca e extrínseca como parte da análise de assunto. Dessa forma, a análise de assunto, como primeira etapa da indexação, consiste de quatro fases, iniciadas pela determinação da tematicidade intrínseca do documento e finalizadas pela determinação extrínseca: a) determinação da tematicidade intrínseca, b) Identificação de conceitos, c) Seleção de conceitos, e d) determinação da tematicidade extrínseca.

A determinação da tematicidade intrínseca é a representação por conceitos do assunto mais significativo do conteúdo considerando-se o autor, por outro lado, a determinação da tematicidade extrínseca é a representação do assunto mais significativo considerando-se o leitor usuário.

Esta constatação de um desdobramento da etapa de análise de assunto por meio da identificação e seleção de conceitos confirma a visão mais complementar de autores como Cleveland & Cleveland (2001) e Lancaster (2003) na qual os indexadores atribuem termos que descrevem a necessidade do usuário e não somente do conteúdo. Para Lee & Schleyer (2012, p.1747) os indexadores analisam o conteúdo de um documento conceitualmente e decidem não somente do que trata o documento, mas também daquilo que interessa ao usuário ou a um grupo de usuários. Dessa forma, a representação vai além do conteúdo documentário e se complementa com o contexto sociocognitivo do indexador que inclui o conhecimento de vocabulários especializados de áreas de domínios utilizadas por usuários do sistema de recuperação da informação, durante representação de suas necessidades de busca e recuperação.

Neste processo de determinação de tematicidade, tanto intrínseca quanto extrínseca, é necessário o leitor indexador como variável imprescindível à tarefa de indexação. Entretanto, o indexador como variável está sujeito à subjetividade do processo

de determinação de tematicidade tendo em vista os diferentes domínios de assuntos e a diversidade de leitores usuários especialistas ou não.

O indexador e seu contexto sociocognitivo são determinantes para a diminuição da subjetividade no sentido de que quanto mais compreensão tiver do processo e da subjetividade nele envolvido, mais condições ele terá de resolver metodologicamente os problemas.

Indiscutivelmente, a determinação de assunto é a etapa crucial em que o indexador como leitor é a principal variável, ainda que sujeito às suas concepções de análise de assunto e à política de indexação do sistema de informação.

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 por questionamento para a identificação daqueles conceitos que são elementos essenciais na descrição do assunto:

- a) O documento possui em seu contexto um objeto sob efeito de uma atividade?
- b) O assunto contém um conceito ativo (por exemplo, uma ação, uma operação, um processo, etc)?
- c) O objeto é influenciado pela atividade identificada?
- d) O documento possui um agente que praticou esta ação?
- e) Este agente refere-se a modos específicos para realizar a ação (por exemplo, instrumentos especiais, técnicas ou métodos)?
- f) Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?
- g) São identificadas algumas variáveis dependentes ou independentes?
- h) O assunto foi considerado de um ponto de vista, normalmente não associado com o campo de estudo (por exemplo, um estudo sociológico ou religioso)?

(ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p.2)

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. Contudo, nem a literatura e nem a Norma esclarecem se essa abordagem sistemática significa um método de identificação de conceitos e nem se estão definidos os procedimentos para tal.

O que podemos inferir sobre a existência de um procedimento é que as questões são relacionadas à busca de conceitos. A primeira questão, por exemplo, deverá identificar no texto a presença do conceito "objeto"; a segunda, a "ação"; a terceira, se o "objeto" identificado sofre influência da "ação"; a quarta, o "agente" que praticou a "ação" e assim por diante...

Austin (1974), ao idealizar o sistema de indexação PRECIS muito antes dessa Norma, apoiou-se teoricamente na "Gramática de casos" de Fillmore (1977). De acordo com Cintra (1983), Fillmore (1977) define "casos" como um conjunto de conceitos universais que identificam julgamentos a respeito de assuntos como "quem fez isso", "a quem aconteceu isso", "o que foi mudado", etc.

Trabalhando em torno desses "conceitos universais", o sistema PRECIS recomenda, para a etapa de identificação de conceitos da análise de assunto, uma análise conceitual baseada na interrogação do texto:

Questionamento	Conceitos do sistema PRECIS
O QUE ACONTECEU?	(AÇÃO)
A QUE OU A QUEM ISTO ACONTECEU?	(OBJETO DA AÇÃO - SISTEMA CHAVE)
O QUE OU QUEM FEZ ISTO?	(AGENTE DA AÇÃO)
ONDE ACONTECEU?	(LOCAL)

Quadro 2: Análise conceitual do sistema PRECIS

A identificação desses "referenciais" ou "conceitos universais" dentro de um texto podem servir como metodologia de análise para representação de conceitos. Dentro do Manual do Modelo de Leitura Documentária (FUJITA, 2003b, p.247-8) a explicação de como representar por conceitos é demonstrada como segue:

Tomemos como exemplo o conceito *agente* que pode ser definido por *aquela ou algo que realizou a ação*. Isso significa que o conceito agente poderá ser representado por uma palavra no texto, que dependerá do contexto, para identificar-se com a idéia de agente. Assim, asseguramos que esses conceitos poderão ser identificados em qualquer texto, garantindo uma uniformidade de identificação de conceitos e de compreensão global do texto que, de outra forma, não seria possível por não termos parâmetros de compreensão.

Ex.: Destruição de plantações de café pela geadas.

O agente neste caso é a geadas, pois ela “praticou” a ação de destruição de plantações de café.

O resultado da representação por conceitos realizada na fase de identificação de conceitos obterá um conjunto de termos que expressarão a tematicidade intrínseca do documento com garantia literária. Esse conjunto de termos, antes do indexador, deixar de ter o documento “em mãos”, passa pela seleção, fase na qual o indexador revisará cada termo com o objetivo de determinar quais os que representam as necessidades de recuperação de um usuário ou de um grupo de usuários. Essa fase consistirá na determinação da tematicidade extrínseca do documento.

Ao passar para a etapa de tradução, o indexador não terá mais o documento “em mãos” e sim o conjunto de termos identificados e selecionados para representarem o conteúdo significativo daquele documento. Com uso de uma linguagem documentária adotada pelo sistema de recuperação da informação cada termo, então, será procurado e uma vez obtida a confirmação de que é compatível, total ou parcialmente e nesse caso adequado, será considerado um termo autorizado para a representação tanto para a indexação quanto para atendimento às questões de buscas pelo usuário.

Portanto, a representação por conceitos na análise de assunto para identificação e seleção de conceitos é diferente da representação realizada na tradução porque, em primeiro lugar, é realizada durante a análise de conteúdo do documento, ou seja, com o documento na íntegra e em segundo lugar, porque utiliza conceitos com os quais são representados os termos e assim identificada a tematicidade intrínseca do documento. Na tradução, de outro modo, a representação se dá com os termos extraídos do documento analisado, portanto, fora do contexto documentário em que o autor desenvolveu o conteúdo significativo e utiliza uma linguagem documentária que representa o vocabulário de áreas de especialidade com a finalidade de compatibilizar os termos identificados com os termos buscados pelo usuário.

A representação por conceitos na análise de assunto é parte integrante do Modelo de Leitura Documentária que se apresenta tanto para textos científicos quanto para livros

e garantindo que os conceitos a serem representados não se modificam quando o documento apresenta diferentes tipologias ou estruturas textuais.

3. A ANÁLISE CONCEITUAL E A EXPLORAÇÃO DA ESTRUTURA TEXTUAL NA CATALOGAÇÃO DE ASSUNTOS EM LIVROS

Considerando-se, segundo Guinchat e Menou (1994), a existência de uma “cadeia documentária” composta pelas operações de coleta, tratamento e difusão realizada em unidades de informação, o tratamento documentário é uma operação intermediária que cuidará desde o controle e registro material do documento, descrição do conteúdo até a sua armazenagem para que seja possível a difusão. Dias (2001, p.4) contextualiza o tratamento em sistemas de informação e recuperação da informação e o define como “[...] a função de descrever os documentos, tanto do ponto de vista temático (ou de descrição do conteúdo). O tratamento, pelos próprios aspectos de forma e conteúdo do documento, apresenta uma dicotomia de operações de tratamento descritivo do suporte material da informação e do tratamento temático de conteúdo da informação. Esta dicotomia que se apresenta no tratamento documentário é explicada, de um lado, pelo desenvolvimento teórico e metodológico distintos alcançados pelas duas áreas e de outro, pela diferença existente entre os aspectos da informação – o material e o conteúdo, que exigem tratamento diferenciado e comportam dois níveis:

- o da **forma** - análise descritiva ou bibliográfica - o tratamento físico da informação ligado com o suporte;
- e o do **conteúdo** - tratamento temático da informação e destina-se à representação condensada do assunto intrínseco ou extrínseco tratado em um determinado documento.

Contudo, é preciso considerar o tratamento dentro de um contexto de sistema de informação e nessa perspectiva, os tratamentos de forma e conteúdo, embora operacionalmente distintos, são dependentes um do outro, principalmente, se consideramos o contexto de bibliotecas. Na biblioteca o formato descritivo utilizado é o

catalográfico, a maioria em MARC21, que conterà o resultado das operações de tratamento de conteúdo documentário, o número de classificação, obtido pela classificação, os cabeçalhos de assuntos determinados pela catalogação de assunto e, em alguns casos, o resumo derivado da elaboração de resumo. Para análise do contexto de tratamento, consideramos importante associar os dois níveis de tratamento, uma vez que a biblioteca é composta por partes interligadas (inserção de documentos, classificação, catalogação, indexação etc.) com objetivo comum de disponibilizar a informação da melhor maneira possível.

Na literatura, a primeira fase da indexação, a análise ou análise de assunto, é tema de estudos pelo seu caráter procedimental estar sujeito a uma tarefa intelectual que depende do conhecimento prévio do profissional que executa a tarefa. Fujita (2003) realizou pesquisas sobre análise de assunto na perspectiva da leitura documentária realizada por indexadores. A análise de assunto, enquanto operação de tratamento de conteúdo documentário é realizada por indexadores, classificadores, resumidores e catalogadores com diferentes tipos de documentos e em diferentes sistemas de informação.

A catalogação de assunto para Mey e Silveira (2009) é a ferramenta para determinação dos chamados “pontos de acesso” que são responsáveis pela extração das características em comum do documento com o usuário interessado. E para isso são citadas as fontes principais de informação para o levantamento das informações do ponto de acesso. Para livro elas citam:

- Página de rosto
- Verso da página de rosto
- Outras páginas que antecedem a página de rosto
- Capa
- Colofão
- Encartes, bolsos ou pastas, se houver
- Apêndices e anexos se houver
- Glossários, bibliografias e índices, se houver.

E ainda acrescentam que para informações específicas de autor e assunto devem ser consultadas:

- Orelhas (s)
- Prefácio
- Sumário
- Introdução
- Dados biográficos em partes específicas.

Sauperl (2002) realizou observação da catalogação de assuntos por doze catalogadores de três diferentes instituições com aplicação de protocolo verbal. Após as análises, obtém resultados que apresenta para cada uma das etapas da Catalogação de Assunto: onde começam, quais partes do livro são analisadas, e quais informações do livro são utilizadas para a catalogação de assuntos. Conforme relato de pesquisa de Sauperl (2002), as etapas da análise conceitual do livro relacionado-as à exploração da estrutura textual podem ser resumidas no quadro a seguir:

	Etapas da análise conceitual	Partes do livro analisadas
1º	Exame preliminar do livro para identificação de termos representativos	título, autor/responsabilidade intelectual, série ficha catalográfica
2º	Confirmação de termos identificados	sumário
3º	Identificação da intenção do autor, de esclarecimento sobre termos e relações entre termos	resumo, prefácio, algum subtítulo interessante e a introdução
4º	Seleção de termos identificados	

Quadro 3 - Partes do livro analisadas para análise conceitual de livros na catalogação de assunto.

A Norma 6029 da Associação Brasileira de Normas Técnicas divide o livro em parte externa e interna. Na parte externa dois elementos são importantes:

- Quarta capa ou contracapa ou costa – a regra de apresentação diz “devem ser impressos o ISBN, conforme ABNT NBR 10521, e o código de barras. Opcionalmente podem constar o resumo do conteúdo e o endereço da editora.” (2006, p. 5)
- Orelhas – embora seja um elemento opcional, a norma diz que quando utilizado deve conter dados biográficos do autor e comentários sobre a obra, e podem constar indicação do público ao qual a obra se destina e outras informações. (2006, p. 5)

Na parte interna do livro, destacam-se os elementos pré-textuais, textuais e os pós-textuais considerados como boas fontes de informação para a identificação de conceitos:

Pré-textuais:

- Folha de rosto - contém o título e subtítulo da obra e no verso a ficha catalográfica que fornecem pistas sobre o conteúdo do livro;
- Listas de ilustrações e tabelas;
- O sumário apresenta na ordem de apresentação do texto os elementos que compõe o trabalho publicado;
- Prefácio traz um resumo da obra e geralmente comentários de terceiros;

Textuais:

- Introdução, que apresenta o tema da publicação e uma visão geral do contexto da obra;
- Desenvolvimento - é o conteúdo do trabalho, a abordagem do tema, e a abrangência da pesquisa;
- Conclusão - as considerações finais do trabalho, o desfecho da pesquisa;

Pós-textuais:

- Posfácio – as informações destacadas após a conclusão do trabalho;
- Referências – As fontes usadas pelo autor para elaboração do trabalho.

Silveira (2006) realizou pesquisa para avaliar a aplicabilidade do Modelo de leitura documentária para a catalogação de assuntos de livros mediante coleta de dados introspectiva com Protocolo Verbal com um estudante de graduação em Biblioteconomia, uma bibliotecária aposentada e sem experiência em catalogação de assunto, e uma bibliotecária especializada em catalogação de assuntos.

Os resultados obtidos revelaram que os sujeitos se apóiam na estrutura textual durante a leitura documentária para indexação, analisam principalmente capas, orelhas, sumário, nota do autor, introdução, títulos e subtítulos de capítulos, palavras em destaque, índice etc. Observou-se que com os questionamentos sugeridos no Modelo de Leitura Documentária, o catalogador de assunto se sente desafiado a encontrar no texto termos que correspondam a seus conceitos fazendo desse modo uma análise mais profunda do livro. As partes principais da estrutura textual onde foram identificados os termos são: nota do autor, introdução e capas (1.ª e 4.ª capas, cf ABNT NBR 10521).

4. A OBSERVAÇÃO DO PROCESSO DE CATALOGAÇÃO DE ASSUNTOS COM USO DO MODELO DE LEITURA DOCUMENTÁRIA: APLICAÇÃO DO PROTOCOLO VERBAL INDIVIDUAL

Com dados coletados em pesquisas de observação da leitura de indexadores, por meio de Protocolo Verbal (FUJITA, NARDI, FAGUNDES, 2003), concluiu-se que a leitura documentária, com suas características próprias, apresenta dificuldades cujas causas podem ser a falta de procedimentos comuns que assegurem a leitura e identificação de conceitos e também o domínio da área de assunto, próprio do especialista, mas não do indexador. Considerando, ainda, que a exploração da estrutura textual foi uma estratégia observada frequentemente na leitura documentária dos indexadores, realizou-se elaboração de Modelo de Leitura Documentária para indexação de textos científicos com a possibilidade de uso combinado da exploração da estrutura textual com questionamento para identificação de conceitos:

CONCEITO (ANÁLISE CONCEITUAL)	QUESTIONAMENTO (NORMA 12.676)	PARTE ESTRUTURA TEXTUAL DA
Objeto	O documento possui em seu contexto um objeto sob efeito de uma atividade?	Introdução (objetivos)
Ação	O assunto contém um conceito ativo (por exemplo, uma ação, uma operação, um processo etc)?	Introdução (objetivos)
Agente	O documento possui um agente que praticou esta ação?	Introdução (objetivos)
Métodos do agente	Este agente refere-se a modos específicos para realizar a ação (por exemplo, instrumentos especiais, técnicas ou métodos)?	Metodologia
Local ou ambiência	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	Metodologia
Causa e efeito	São identificadas algumas variáveis dependentes ou independentes?	Resultados; discussão de resultados
Ponto de vista do autor; perspectiva	O assunto foi considerado de um ponto de vista, normalmente não associado com o campo de estudo (por exemplo, um estudo sociológico ou religioso)?	Conclusões

Quadro 4: Modelo de Leitura Documentária para textos científicos: identificação de conceitos por questionamento em partes da estrutura textual

Fonte: FUJITA, 2003b

Conforme Fujita (2003b, FUJITA, RUBI, 2006a, 2006b) o modelo de leitura obtido (QUADRO 1) consiste, fundamentalmente, da combinação das sistemáticas de identificação de conceitos análise conceitual (primeira coluna) e abordagem sistemática da Norma 12.676 (segunda coluna) com a localização dos conceitos em parte da estrutura textual (terceira coluna).

A elaboração desse modelo de leitura foi uma proposição de aprimoramento e evolução da metodologia de indexação a partir da metacognição do indexador que poderá ser oferecida na formação do indexador em leitura documentária.

O modelo de leitura de textos científicos de Fujita (2003b) foi proposto para desenvolver a identificação e seleção de conceitos durante a leitura documentária de acordo com as concepções orientadas para o conteúdo e a demanda. A etapa de identificação de conceitos é realizada durante a Análise de assunto por meio de uma concepção orientada para o conteúdo. Na leitura, os aspectos lógicos, lingüísticos e cognitivos envolvidos na indexação representam fatores de interferência, cabendo ao indexador a habilidade necessária para poder realizar a análise conceitual efetiva do documento.

Tendo em vista os resultados da aplicabilidade do Modelo de Leitura como recurso pedagógico com alunos do Curso de Graduação em Biblioteconomia em 11 coletas de dados com a metodologia do Protocolo Verbal Individual e uma em Grupo, além de Entrevista Retrospectiva em Grupo realizada com os sujeitos que participaram das 4 coletas de dados é proposta a adequação do Modelo de Leitura Documentária conforme recomendações indicadas a partir de problemas e dificuldades observadas (FUJITA, 2007a):

MODELO DE LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO DE TEXTOS CIENTÍFICOS (ADAPTADO)

REFERÊNCIA DO TEXTO:

PASSO-A-PASSO:

1° Observação da estrutura textual e localização do conteúdo do artigo científico: localizar no texto científico os elementos que o compõem, tal como, introdução, metodologia, resultados, discussão dos resultados e conclusão. Verifique o conteúdo pertinente a cada parte do texto.

2° Identificação dos conceitos: considera-se importante, observar as partes da estrutura textual e realizar o questionamento, indicados na grade do modelo de leitura, para que a identificação dos conceitos seja correta.

Para explicações detalhadas referentes à utilização do modelo de leitura, volte ao manual explicativo;

Observações importantes:

- os três conceitos objeto, ação e agente são principais, estão interligados e são dependentes um do outro;
- o conceito objeto deve ser o primeiro a ser identificado e a partir dele deverão ser identificados a ação e, em seguida, o agente, nesta ordem;
- os três conceitos principais deverão ser representados **por apenas um termo cada**;
- termos vazios de significado, como “avaliação”, “estudo”, “análise” e etc, não devem representar os conceitos principais e os demais conceitos;
- o conceito ação pode ser representado por termo que denomina uma doença ou um fenômeno, mas, na maioria dos textos é um verbo no substantivo, como “coagulação”;

- nem todos os conceitos serão, necessariamente, identificados por termos.

CONCEITOS	QUESTIONAMENTO PARA IDENTIFICAÇÃO DE CONCEITOS	PARTES DA ESTRUTURA TEXTUAL
OBJETO e PARTE(S) DO OBJETO (algo ou alguém que está sob estudo do pesquisador)	O documento possui em seu contexto um objeto sob efeito desta ação?	INTRODUÇÃO (OBJETIVOS)
AÇÃO (processo sofrido por algo ou alguém)	O assunto contém uma ação (podendo significar uma operação, um processo etc.)?	INTRODUÇÃO (OBJETIVOS)
AGENTE (aquele ou algo que realizou a ação)	O documento possui um agente que praticou esta ação?	INTRODUÇÃO (OBJETIVOS)
MÉTODOS (métodos utilizados para realização da pesquisa)	Para estudo do objeto ou implementação da ação, o documento cita e/ou descreve modos específicos, por exemplo: instrumentos especiais, técnicas, métodos, materiais e equipamentos?	METODOLOGIA
LOCAL OU AMBIÊNCIA (local físico onde foi realizada a pesquisa)	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	METODOLOGIA
CAUSA E EFEITO Causa (ação+objeto)/Efeito	Considerando que a ação e o objeto identificam uma causa, qual é o efeito desta causa?	RESULTADOS; DISCUSSÃO DE RESULTADOS; CONCLUSÕES

Quadro 5: Versão adaptada do Modelo de leitura documentária para indexação de textos científicos
Fonte: FUJITA, 2007a

A versão adaptada do Modelo de Leitura Documentária foi utilizada na adaptação para a catalogação de assuntos de livros em abordagem cognitiva no contexto de bibliotecas universitárias a partir dos resultados obtidos de sua aplicação à tarefa de catalogação de assuntos livros com catalogadores utilizando a técnica introspectiva de coleta de dados com Protocolo Verbal Individual.

O Protocolo Verbal ou “Pensar Alto” é uma técnica introspectiva de coleta de dados que consiste na verbalização dos pensamentos dos sujeitos. Introspecção, segundo Cavalcanti (1989), é um exame de processos mentais que promove uma análise pelo sujeito de seu próprio processo de pensamento. À medida que o sujeito realiza uma

tarefa, verbaliza como resolve os problemas em relação ao vocabulário, procedimentos, dificuldades e a compreensão das idéias principais do texto.

A aplicação do Protocolo Verbal Individual nesta pesquisa teve objetivo de observar a tarefa de catalogação de assuntos de livros realizada por bibliotecários catalogadores durante a leitura documentária para identificar as partes da estrutura textual em que são identificados os conceitos. A aplicação constou de três procedimentos que se subdividem em etapas:

Procedimentos anteriores à coleta de dados:

Definição do universo da pesquisa: catalogadores de nove Bibliotecas universitárias.

Seleção do Texto-Base: - Livro para catalogação original: a estrutura do livro se resume em capa, orelhas, falsa página de rosto, página de rosto, sumário, introdução, desenvolvimento, conclusão, referencias bibliográficas, crédito das imagens e contracapa.

Definição da tarefa: foi solicitado aos catalogadores que fizessem a catalogação de um Livro, tese ou dissertação para catalogação original a partir dos registros bibliográficos em suas variações: registro aproveitável¹ (RA) e o registro idêntico² (Identidade Total – IT) conforme estavam habituados e que externalizassem seus pensamentos enquanto estivessem realizando essa tarefa. Conversa informal com os sujeitos: No dia da aplicação do protocolo foi realizada uma conversa informal com cada um dos participantes com intenção de esclarecer os objetivos da pesquisa, esclarecendo-os da ocultação da identidade dos participantes, de modo a preservá-los e da função dos pesquisadores durante a realização da tarefa, qual seja, operar o gravador e estimular o sujeito quanto ao “*Think Aloud*”.

Familiarização com a tarefa do “*Think Aloud*”: Para a familiarização com a tarefa de “Pensar Alto”, foi utilizado um texto com o propósito de descontrair e ao mesmo tempo

¹ RA - Registro Aproveitável: algumas informações são idênticas em relação ao documento que está sendo catalogado e ao que foi recuperado (no contexto da catalogação cooperativa, abreviatura utilizada pelos profissionais bibliotecários da Rede de Bibliotecas da UNESP).

² IT - Identidade Total: informações idênticas em relação ao documento que está sendo catalogado e ao que foi recuperado (no contexto da catalogação cooperativa, abreviatura utilizada pelos profissionais

esclarecer o objetivo da pesquisa, bem como apresentar procedimentos para auxiliar o sujeito no desempenho da tarefa.

Procedimentos durante a coleta de dados:

Gravação do “Pensar Alto” durante a leitura do texto-base: antes do início da gravação foi entregue aos sujeitos, o Modelo de Leitura documentária, e o livro. Durante a realização da tarefa foi lembrado ao sujeito a necessidade de exteriorização de seus pensamentos.

Procedimentos posteriores à coleta de dados:

- Transcrição literal das gravações das falas dos sujeitos (Protocolo Verbal Individual);
- Leitura detalhada dos dados em busca de fenômenos significativos e recorrentes para construir categorias de análise;
- Construção de categorias de análise tendo em vista o objetivo de identificar as partes da estrutura textual do livro utilizadas para a identificação de conceitos do modelo de leitura documentária.

5. O PROCESSO DE CATALOGAÇÃO DE ASSUNTOS E A IDENTIFICAÇÃO DE CONCEITOS COMBINADA COM A EXPLORAÇÃO DA ESTRUTURA TEXTUAL DE LIVROS: RESULTADOS

A análise dos protocolos verbais individuais dos catalogadores realizada com o objetivo de verificar se durante a catalogação de assuntos de livros houve a identificação de conceitos combinada com a exploração da estrutura textual permitiu constatar que os bibliotecários realizam o exame do documento percorrendo várias partes da estrutura textual do livro, externa e interna. Apesar disso, a identificação de conceitos ocorre em partes onde não existe texto apenas elementos descritivos, tal como a catalogação na fonte ou a última capa. Na exploração da estrutura textual realizada pelos catalogadores, poucos identificaram termos consultando partes relevantes para a indexação (partes que apresentam mais conteúdo textual sobre o assunto do livro), tais como o **sumário** e o **prefácio** (pré-textual) e o **desenvolvimento do livro** (textual).

Os catalogadores, de modo geral, recuperaram mais termos em partes pouco recomendadas para o processo de indexação, como **catalogação na fonte** e **última capa**.

Nenhum dos bibliotecários analisados apresentou estratégias diretamente relacionadas à exploração da estrutura textual dos livros e não demonstraram preocupação com a identificação de conceitos durante a catalogação de assuntos para recuperação dos documentos pelo usuário.

Entretanto, modificações relacionadas à conceitos e questionamentos foram realizadas em função dos resultados das pesquisas de Sauperl (2002) e Silveira (2006) sobre identificação de termos através de conceitos combinados à exploração da estrutura textual externa e interna considerando-se:

- o exame preliminar da parte externa do livro para análise de elementos da contracapa e da orelha e da parte interna do livro analisando os elementos pré-textuais (título, resumo, autor, ficha catalográfica e série) para uma primeira consulta e identificação de termos;
- a análise do sumário na parte interna pré-textual e na textual a introdução, títulos e subtítulos de capítulos, conclusão, índice ao final do livro para confirmação dos termos identificados no exame preliminar e identificação de novos termos;
- é necessário identificar os termos principais de cada capítulo quando o livro apresentar capítulos de diferentes autores;
- os conceitos ação, objeto, agente, método, local ou ambiência poderão ser identificados na parte interna do livro com os elementos pré-textuais título, resumo, ficha catalográfica, porém devem ser confirmados pelo sumário e introdução;
- o conceito de tempo poderá ser incluído para identificação na parte interna, pré e pós-textual do livro acrescido do questionamento: o estudo foi desenvolvido em local e período específicos? É relevante representá-los na Catalogação de Assunto?
- a identificação da intenção do autor, de esclarecimentos sobre termos e relações entre termos tendo em vista a possibilidade de identificação do conceito “perspectivas ou ponto de vista”;

- é necessário identificar os conceitos de causa e efeitos quando o livro apresentar conclusões.

6. CONCLUSÕES

Estes resultados da observação da tarefa de catalogação de assuntos de livros por catalogadores de bibliotecas universitárias revelam a necessidade de uma metodologia sistematizada de catalogação de assuntos que combine as estratégias de identificação de conceitos com a exploração da estrutura textual do livro que é diferente da estrutura textual de textos científicos cujo formato é de artigos científicos, teses, dissertações, relatórios científicos, trabalhos completos de eventos científicos e etc. com a superestrutura de argumentação contida nos itens: Introdução, metodologia, resultados e conclusões. Por isso, o Modelo de leitura documentária para indexação de textos científicos foi adaptado para a versão de catalogação de assuntos de livros com a modificação principal da terceira coluna que se refere a partes da estrutura textual do livro conforme se constata no Modelo de Leitura Documentária para a indexação na catalogação de assuntos de livros em bibliotecas:

MODELO DE LEITURA DOCUMENTÁRIA PARA INDEXAÇÃO NA CATALOGAÇÃO DE ASSUNTOS DE LIVROS EM BIBLIOTECAS

PASSO-A-PASSO:

1º Observação da estrutura textual e localização do conteúdo do livro:

- Verifique o conteúdo pertinente a cada parte do livro através de exame das partes externas e internas da estrutura textual;

- Localize no livro os elementos que o compõem, tal como:

- parte externa: contracapa e orelhas;
- parte interna pré-textual: folha de rosto (título e subtítulo e no verso a série e a ficha catalográfica que fornecem pistas sobre o conteúdo do livro), listas de ilustrações e tabelas, prefácio, resumo e sumário, se houver;
- parte interna textual: introdução, metodologia, desenvolvimento (com resultados se houver) e conclusão;
- parte interna pós-textual.

2º Identificação dos conceitos: considera-se importante, observar as partes da estrutura textual e realizar o questionamento, indicados na grade do modelo de leitura, para que a identificação dos conceitos seja correta.

Para explicações detalhadas referentes à utilização do modelo de leitura, volte ao manual de ensino;

Observações importantes:

- os três conceitos objeto, ação e agente são principais, estão interligados e são dependentes um do outro;
- o conceito objeto deve ser o primeiro a ser identificado e a partir dele deverão ser identificados a ação e, em seguida, o agente, nesta ordem;
- os três conceitos principais deverão ser representados **por apenas um termo cada**;
- termos vazios de significado, como “avaliação”, “estudo”, “análise” e etc, não devem representar os conceitos principais e os demais conceitos;
- o conceito ação pode ser representado por termo que denomina uma doença ou um fenômeno, mas, na maioria dos textos é um verbo no substantivo, como “coagulação”;
- nem todos os conceitos serão, necessariamente, identificados no conteúdo do livro a não ser que o tema os contemple, por exemplo os conceitos de causa e efeito, serão representados por termos desde que o livro contenha conclusões ou que apresente parte específica sobre metodologia para identificar o conceito de método..

CONCEITOS	QUESTIONAMENTO PARA IDENTIFICAÇÃO DE CONCEITOS	PARTES DA ESTRUTURA DO LIVRO
OBJETO e PARTE(S) DO OBJETO (algo ou alguém que está sob estudo do autor)	O documento possui em seu contexto um objeto sob efeito desta ação?	SUMÁRIO E INTRODUÇÃO
AÇÃO (processo sofrido por algo ou alguém)	O assunto contém uma ação (podendo significar uma operação, um processo etc.)?	SUMÁRIO E INTRODUÇÃO
AGENTE (aquele ou algo que realizou a ação)	O documento possui um agente que praticou esta ação?	SUMÁRIO E INTRODUÇÃO
MÉTODO (métodos utilizados para realização da pesquisa)	Para estudo do objeto ou implementação da ação, o documento cita e/ou descreve modos específicos, por exemplo: instrumentos especiais, técnicas, métodos, materiais e equipamentos?	SUMÁRIO INTRODUÇÃO OU METODOLOGIA
TEMPO (ano, período ou época)	O estudo foi desenvolvido em período específico? É relevante representá-los na Catalogação de Assunto?	SUMÁRIO INTRODUÇÃO OU METODOLOGIA
LOCAL OU AMBIÊNCIA (local físico onde foi realizada a pesquisa)	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	SUMÁRIO INTRODUÇÃO OU METODOLOGIA
PONTO DE VISTA DO AUTOR	O assunto foi considerado de um ponto de vista, normalmente não associado com o campo de estudo (por exemplo, um estudo sociológico ou religioso)?	INTRODUÇÃO E PREFÁCIO
CAUSA E EFEITO Causa (ação+objeto)/Efeito	Considerando que a ação e o objeto identificam uma causa, qual é o efeito desta causa?	CONCLUSÕES OU PARTE FINAL

Quadro 6: Versão do Modelo de Leitura Documentária para catalogação de assuntos de livros

REFERENCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. (2006). *NBR 6029*: Informação e documentação — livros e folhetos — apresentação. Rio de Janeiro.

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

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

CAVALCANTI, M. C. (1989). *I-n-t-e-r-aç-ã-o leitor-texto*: aspectos de interpretação pragmática. Campinas: UNICAMP. 271 p.

CHAUMIER, J. (1986). *Analisis y lenguajes documentales*: el tratamiento lingüístico de la información documental. Barcelona: Editorial Mitre. 172p.

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

CINTRA, A. M. C. (1983). Elementos de lingüística para estudos de indexação. *Ciência da Informação*, Brasília, 12:1 5-22.

CLEVELAND, D. B.; CLEVELAND, A. D. (1990). *Introduction to indexing and abstracts*. 2.ed. Englewood: Libraries Unlimited. 329p.

FILLMORE, C.J. (1977) Em favor do caso. In: *A semântica na linguística moderna*: o léxico. Rio de Janeiro: Livraria Francisco Alves. p. 277-365.

FUJITA, M. S. L. (1999). Análise e síntese documentárias para compreensão de leitura de textos didáticos: uma proposta de aplicação do sistema de indexação *Precis*. *Informare*: Cadernos do Programa de Pós-graduação em Ciência da Informação, Rio de Janeiro, Vol.5, No.2.

FUJITA, M.S.L. (2003^a). A identificação de conceitos no processo de análise de assunto para indexação. *Revista Digital de Biblioteconomia e Ciência da Informação* 1: 60-90.

FUJITA, M. S. L. (2003^b). *A leitura documentária do indexador*: aspectos cognitivos e lingüísticos influentes na formação do leitor profissional. 321 f. Tese (Livre-Docência em Análise Documentária e Linguagens Documentárias Alfabéticas) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.

FUJITA, M. S. L. (2007a). La lectura documentaria como disciplina curricular en el curso de Biblioteconomia: contenido y metodologias del abordaje sócio-cognitivo para metacognición del indexador aprendiz. In: FRANCISCO JAVIER GARCIA MARCO. (Org.). IBERSID 2007: avances y perspectivas en sistemas de información y documentación en el entorno digital. Zaragoza: Universidad de Zaragoza, p. 87-92.

FUJITA, M.S.L. (2007b). Leitura em análise documentária: uma contribuição à formação de indexadores. Marília: Faculdade de Filosofia e Ciências da UNESP. (Relatório final de pesquisa, Bolsa PQ/CNPq).

FUJITA, M. S. L.; NARDI, M.I.A.; FAGUNDES, S.A. (2003). Observing documentary reading by verbal protocol. *Information Research*, v. 8, n. 4, jul. Disponível em: <<http://www.informationr.net/ir/8-4/paper155.html>>. Acesso em 3 maio 2007.

FUJITA, M. S. L., RUBI, M. P. (2006a). Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. *DataGramZero*, Rio de Janeiro, Vol.7, No.3, 1-18.

FUJITA, M. S. L., RUBI, M. P. (2006b). Modelo de lectura profesional para la indización. *Scire* (Zaragoza), Vol.12, 47 - 70.

GIL LEIVA, I. (2008). *Manual de indización: teoría y práctica*. Gijón: Ediciones TREA, 2008. 429p.

GUINCHAT, C.; MENO, M. (1994). *Introdução geral às ciências e técnicas da informação e documentação*. 2.ed. rev. aum. Brasília: IBICT.

LANCASTER, E.W. (2003). *Indexing and abstracting in theory and practice*. Champaign, IL: Intellitext Corporation.

LEE, D. H.; SCHLEYER, T. (2012). Social tagging is no substitute for controlled indexing: a comparison of medical subject headings and citeulike tags assigned to 231,388 papers. *Journal of the American Society for Information Science and Technology*, 63(9):1747–1757.

MEY, E. S. A.; SILVEIRA, N. C. (2009). *Catálogo no plural*. Brasília: Briquet de Lemos/Livros.

MAI, J-E. (2004). Analysis in indexing: document and domain centered approaches. *Information and Processing Management* 41: 599-611.

MAI, J-E. (2000). Deconstructing the Indexing Process. *Advances in Librarianship* 23: 269-298.

MAI, J-E. (2004). The role of domains, documents, and decisions in indexing. In *Knowledge Organization and the Global Information Society: Proceedings of the Eight International ISKO Conference*. Advances in Knowledge Organization, pp. 207-213.

MOREIRO GONZÁLEZ, J. A. (2004). *El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural*. Gijón: Alfagrama. 291p.

OLSON, H. A.; BOLL, J. J. (2001) *Subject analysis in online catalogs*. 2.ed. Englewood: Libraries Unlimited. 333 p.

ŠAUPERL, A. (2002). *Subject determination during the cataloging process*. Lanham: Scarecrow Press.

SILVEIRA, Telma Jaqueline Dias. (2006). *Indexação de livros em ambiente de biblioteca: um modelo de leitura*. Trabalho de Conclusão de Curso. (Graduação em Biblioteconomia e Documentação) - Universidade Estadual Paulista Júlio de Mesquita Filho. Orientador: Mariângela Spotti Lopes Fujita.

TAMAYO, A. M. M.; VALDEZ, J. C. (2008). *Indización y clasificación en bibliotecas*. Buenos Aires: Alfagrama. 235p.

WALLER, S.; MASSE, C. (1999) *L'analyse documentaire: une approche methodologique*. Paris: ADBS. 319p.