

EMANUEL GUEDSON FERREIRA GUEDES

**O CONCEITO ABOUTNESS NA ORGANIZAÇÃO E
REPRESENTAÇÃO DO CONHECIMENTO**

Marília/SP
2009

EMANUEL GUEDSON FERREIRA GUEDES

O CONCEITO *ABOUTNESS* NA ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Faculdade de Filosofia e Ciências da Universidade Estadual Paulista – UNESP, como requisito para obtenção do título de mestre em Ciência da Informação.

Orientador: Dr. João Batista Ernesto de Moraes

Linha de pesquisa: Organização da Informação.

Marília/SP
2009

Guedes, Emanuel Guedson Ferreira.
G924o O conceito *aboutness* na Organização e Representação do Conhecimento/
Emanuel Guedson Ferreira Guedes. – 2009.
90 f. ; 30 cm.

Dissertação (Mestrado em Ciência da Informação) – Faculdade de
Filosofia e Ciências, Universidade Estadual Paulista, 2009.
Orientador: João Batista Ernesto de Moraes

1. *Aboutness*. 2. Tematicidade. 3. Atinência. 4. Organização e
Representação do Conhecimento. I. Autor. II. Título.

CDD 025.4
CDU 025.4:17

EMANUEL GUEDSON FERREIRA GUEDES

**O CONCEITO *ABOUTNESS* NA ORGANIZAÇÃO E REPRESENTAÇÃO DO
CONHECIMENTO.**

Dissertação apresentada para obtenção de
título de Mestre ao Programa de Pós-
Graduação em Ciência da Informação da
Faculdade de Filosofia e Ciências,
Universidade Estadual Paulista “Júlio de
Mesquita Filho”.

BANCA EXAMINADORA

Dr. João Batista Ernesto de Moraes (Orientador)
Faculdade de Filosofia e Ciências da UNESP
Marília/SP

Dra. Mariângela Spotti Lopes Fujita
Faculdade de Filosofia e Ciências da UNESP
Marília/SP

Dra. Maria de Fátima Gonçalves Moreira Tálamo
Escola de Comunicações e Artes da USP
São Paulo/SP

Marília, 23 de setembro de 2009.

À minha família.

Agradecimentos

À minha família por todo apoio e compreensão durante o desenvolvimento desse trabalho;

À Noemi, minha companheira (dessa e de outras labutas), pelo companheirismo, amor e amizade;

Aos colegas da turma de Mestrado e demais amigos;

Aos professores e funcionários do PPGCI;

À Banca Examinadora: Profas. Mariângela S. L. Fujita e Maria de Fátima G. M. Tálamo;

Ao Prof. João Batista pela orientação do trabalho;

À equipe de trabalho (equipe LOST) na Biblioteca Mário de Andrade (em especial à Izabel por me apresentar o “Lúcio”);

A todos que contribuíram direta ou indiretamente para a conclusão desse trabalho;

***“Ao expandirmos o campo do
conhecimento apenas aumentamos o
horizonte da ignorância”.***

Henry Miller.

RESUMO

A identificação de assuntos, ocorrida na etapa analítica de um documento submetido ao tratamento temático, constitui um campo de estudo carente de metodologias que auxiliem o leitor profissional na seleção de conceitos que propiciem os subsídios necessários à elaboração dos produtos documentários decorrentes desse processo. Dessa forma, os estudos em torno do conceito *Aboutness* têm adquirido um papel evidente nos estudos do campo documental, tendo em vista, que o processo de atribuição de termos e conceitos a partir da leitura documental deve sintetizar satisfatoriamente o conteúdo informativo do documento tornando esse procedimento um momento chave para todas as fases subsequentes do processo. Desse modo, a Ciência da Informação tem buscado também em outros campos, subsídios que colaborem para a abordagem temática, de forma a auxiliar a sua compreensão de acordo com a finalidade de sua representação em um sistema de informação. Realizamos assim a abordagem da literatura da Ciência da Informação em torno do conceito *aboutness* com vistas à melhor compreensão dessa abordagem e para tanto, são explorados além da literatura relacionada, quatro dos principais periódicos nacionais da área, selecionando artigos que tenham apresentado ou discutido essa questão. Verificou-se que o termo *aboutness* foi traduzido em geral por *atinência*, *tematicidade* e *sobrecidade*, sendo que não houve um consenso quanto a um dos três termos, no entanto, enquanto *sobrecidade* caiu em desuso, *atinência* e *tematicidade* permanecem em utilização. Além disso, apesar de pouco explorado em língua portuguesa, esse conceito traz implicações conceituais para a área, que podem influenciar o desenvolvimento teórico dos processos de identificação de assunto ao buscar uma abordagem que visa não apenas a identificação do *aboutness* inerente ao documento, mas também do *aboutness* atribuído pelo indexador ou leitor, também tratado por *meaning* (significado), e dessa forma admite que mesmo havendo a possibilidade de atribuição de diferentes sentidos e interpretações a um documento, esse possui um tema inerente e fundamental que não se altera.

Palavras-chave: *Aboutness*, *Atinência*, *Tematicidade*, *Organização* e *Representação* do Conhecimento.

ABSTRACT

The subject determination that occurred in the analytical stage of a document submitted to the subject analysis, is a field of study in need of methodologies that can help the professional readers in the selection of concepts that provide the subsidies to elaborate documentaries products. Thus, studies on the concept Aboutness have acquired a clear role in the documentary field studies in order that the process of assigning terms and concepts from the documentary reading should adequately summarize the information content of the document by making the procedure a key moment for all subsequent stages of the process. Thus, information science has sought in other fields, subsidies to help in the subject analysis in order to assist in your understanding in accordance with the purpose of their representation in an information system. We approach the literature of information science around the concept of aboutness with a purpose to better understanding this conception of subject analysis and both are explored in addition to the related literature, four major national journals in the field, selecting items that have submitted or discussed this concept. It was found that the term aboutness has been translated by *atênência*, *tematicidade* and *sobrecidade*, and there was no consensus on one of three terms, however, while *sobrecidade* has fallen into disuse, *tematicidade* and *atênência* remain in use. Furthermore, although little explored in portuguese, this concept has conceptual implications for the area, which can influence the theoretical development of processes to identify the subject by seeking an approach that aims not only to identify the aboutness inherent in the document, but also aboutness assigned by the indexer or user, also handled by meaning (*significado*), and thus admits that even with the possibility of assigning different meanings and interpretations of a document that has an inherent and fundamental issue that does not change.

Keywords: *Aboutness*, Subject Analysis, Knowledge Organization and Representation.

LISTA DE ILUSTRAÇÕES

Figura 1: Esquema de progressão temática linear	60
Figura 2: Esquema de progressão temática paralela.....	60
Figura 3: Esquema de progressão temática combinada.....	61
Figura 4: Relação dos termos identificados nos artigos na RCI	73
Figura 5: Relação dos termos identificados nos artigos na DGZ.....	74
Figura 6: Relação dos termos identificados nos artigos na PCI.	74
Figura 7: Relação dos termos identificados nos artigos	76
Gráfico 1: Total de artigos selecionados por percentagem de revistas	71
Gráfico 2: Percentual de artigos identificados a partir de cada periódico	72
Quadro 1: Comparativo de citações sobre a utilização do termo <i>aboutness</i>	76
Quadro 2: Comparativo de citações sobre a utilização dos termos <i>aboutness</i> e <i>atinência</i>	77
Quadro 3: Comparativo de citações da utilização do termo <i>atinência</i>	78
Quadro 4: Comparativo de citações da utilização dos termos <i>aboutness</i> e <i>tematicidade</i>	78

LISTA DE TABELA

Tabela 1: Periódicos nacionais analisados	70
--	----

LISTA DE ABREVIATURAS

AD	Análise Documental
CI	Ciência da Informação
RCI	Revista Ciência da Informação
DGZ	Revista Datagramazero
ORC	Organização e Representação do Conhecimento
PCI	Revista Perspectivas em Ciência da Informação
RI	Recuperação da Informação
SEER	Sistema Eletrônico de Editoração de Revistas
TR	Revista Transinformação
TTI	Tratamento temático da informação

SUMÁRIO

	INTRODUÇÃO	14
I	O TRATAMENTO TEMÁTICO DA INFORMAÇÃO	21
II	O CONCEITO <i>ABOUTNESS</i>.....	31
	<i>O termo aboutness</i>	<i>34</i>
	<i>O aboutness na Ciência da Informação.....</i>	<i>36</i>
	<i>Aboutness e assunto</i>	<i>49</i>
	<i>A interdisciplinaridade no estudo do aboutness.....</i>	<i>53</i>
	<i>A Linguística documental</i>	<i>55</i>
	<i>A Linguística textual.....</i>	<i>56</i>
	<i>Tema, rema e progressão temática</i>	<i>57</i>
	<i>A abordagem dos aspectos lógicos e cognitivos</i>	<i>61</i>
III	PROCEDIMENTOS METODOLÓGICOS	64
	<i>Corpus da pesquisa.....</i>	<i>64</i>
	<i>Revista Ciência da Informação</i>	<i>65</i>
	<i>Revista Perspectivas em Ciência da Informação</i>	<i>65</i>
	<i>Revista Transinformação</i>	<i>66</i>
	<i>Revista Datagramazero</i>	<i>66</i>
	<i>Período cronológico abordado.....</i>	<i>67</i>
	<i>Termos definidos para busca.....</i>	<i>67</i>
	<i>Busca dos termos e procedimentos de seleção dos artigos....</i>	<i>68</i>
	<i>Análise dos resultados.....</i>	<i>69</i>
IV	CONSIDERAÇÕES	79
	REFERÊNCIAS	85

INTRODUÇÃO

O livro "O último leitor" do mexicano David Toscana apresenta um bibliotecário bastante peculiar: Lúcio, que vive em uma pequena aldeia no interior do México chamada Icamole, onde se não bastasse ser o único bibliotecário da única biblioteca, é também o único leitor da cidade. Dessa forma, Lúcio acaba desenvolvendo um método bastante peculiar para selecionar o seu acervo e o divide em dois tipos: os livros bons e os livros que vão para o inferno, embora um pouco longa, a citação a seguir apresenta um pouco do *modus operandi* e do pensamento de Lúcio, que discutimos logo após (TOSCANA, 2005):

Em Icamole, Lúcio confia no seu cérebro a ponto de descartar tudo o que lhe ensinaram em Monterrey. Controle de empréstimos? Eu não empresto nada. Conservação? Meus livros vão durar pouco tempo; quando eu morrer, eles poderão fazê-lo. E, acima de tudo, desprezou os métodos de catalogação. Um especialista explicou o modo de ordenar os livros conforme o assunto, a data de publicação, a nacionalidade do autor e outras variáveis, atribuindo-lhes números e letras. Jamais falou de separar os livros bons dos ruins, em compensação garantiu que a classificação principal se baseava em ficção e não-ficção. Lúcio se decepcionou totalmente ao ouvir o discurso do especialista. Não podia acreditar que tal classificação tivesse sido feita por gente de livros, de letras; não era possível que ficassem sem palavras a ponto de chamar uma coisa pelo que ela não é. Além disso, onde ficava a fronteira entre uma coisa e outra? Onde se encaixavam as memórias de um presidente? Um romance histórico? As vidas dos santos? De que lado está *Testemunho de um soldado*? Se existem contradições entre dois livros de história ou dois livros sagrados, quem decide qual deles deve ser ficção? Lúcio fechou sua caderneta e não ouviu mais aquele tagarela. Suas idéias estavam claras. Um livro de história falava de coisas que aconteceram, enquanto um romance fala de coisas que acontecem, e, assim, o tempo da história contrasta com o do romance, que Lúcio chama de presente permanente, um tempo imediato, tangível e autêntico.

Essa angústia enfrentada por Lúcio, colocada em uma linguagem bem menos "científica" que as pretendidas nos trabalhos acadêmicos, sintetiza o problema abordado nessa pesquisa: como podemos analisar o conteúdo de uma obra e determinar sobre o que ela trata? Sabemos que a magnitude das questões

envolvidas extrapola o âmbito do controle bibliográfico e atinge um nível relacionado ao próprio entendimento da cognição humana e está também relacionada aos juízos de valor e ideologias do próprio leitor entre outras variáveis, e dessa forma, como questionou o próprio Lúcio, onde vamos encontrar a fronteira entre uma coisa e outra? Ou em uma linguagem mais especializada: qual o assunto pode ser determinado em um documento para definir a sua entrada no catálogo?

Ressalvadas as especificidades de cada tipo de produção (científica, literária, jornalística, quadrinhos, poesia, etc.) os problemas de análise e compreensão do conteúdo são bastante similares, além de muitas vezes o próprio enquadramento de uma obra em um gênero específico, não ser uma tarefa das mais simples.

Essa pesquisa surgiu em decorrência do trabalho de conclusão de curso de Biblioteconomia “Análise documentária e semântica discursiva: uma proposta de análise temática em folhetos de Cordel”, ocasião na qual já havíamos abordado esse problema. Naquele momento, buscou-se a semântica discursiva como recorte teórico, para investigar uma forma de identificação de temas da literatura de cordel.

Nesse primeiro trabalho partiu-se do problema que a Ciência da Informação e a Biblioteconomia, ainda não estabeleceram métodos precisos de análise temática e, portanto, havíamos iniciado a abordagem dessa questão, com uma aplicação prática de um método específico de análise, que poderia auxiliar os leitores profissionais na obtenção do “*assunto*” ou temas de um documento.

Dessa forma tomou-se contato com o termo “*aboutness*” como o conceito que representaria a noção de que existe em todo documento uma idéia principal, um tema fundamental, que não se altera em relação ao uso que lhe é atribuído.

No entanto, percebeu-se que a discussão em torno do que é o “*aboutness*” – “caro à crença nas essências” (LARA, 2006) - continua em aberto e que há uma carência em nossa área, da fundamentação teórica desse conceito e que o mesmo foi pouco discutido em língua portuguesa.

A importância dessa pesquisa se viu no fato que essa discussão está diretamente relacionada com a noção existente na área, do que seja o “assunto” contido em um documento passível de ser representado e incorporado em um sistema de informação. E por sua vez, isso se relaciona diretamente com a

qualidade dos instrumentos de recuperação que serão criados.

Dessa forma, observou-se que esse é um problema que tem permeado todas as propostas de organização e tratamento da informação. A definição do que seja o assunto, ou de como encontrá-lo e defini-lo em um documento é um problema que vem nos acompanhando desde o início da informação registrada e da tentativa de organizar os repertórios documentais.

Assim, essa questão está na base da Classificação, da Catalogação de assuntos, da Indexação, da Análise documental, enfim de todos os processos que envolvam a análise temática.

Os instrumentos de busca e recuperação da informação, catálogos, índices ou cabeçalhos, sempre foram ferramentas fundamentais na tarefa de organização dos acervos, pois a partir deles se tornou possível a organização, a localização e o acesso, por ordem alfabética de títulos ou autores, e a mais importante de todas, devido a sua demanda: a ordenação pelo agrupamento de afinidade temática.

Dessa forma, a partir dos pequenos acervos, razoavelmente simples de serem gerenciados, pela pouca quantidade de materiais e usuários, aos imensos acervos atuais, com um público heterogêneo, um longo percurso histórico precisou ser percorrido.

No entanto, muito antes que possamos considerar que estejamos em um ponto de sofisticação que nos separa definitivamente de nossos ancestrais, na verdade, preservados os contextos, estamos ainda em situações bastante similares.

Enquanto os primeiros “bibliotecários” da Idade Antiga, que desconheciam as nossas questões conceituais ou metodológicas – mas certamente tinham as suas próprias, que desconhecemos – buscavam a formação de um acervo bastante peculiar, formado por tabletes de argila, papiros etc., os profissionais correspondentes atualmente, também precisam atender à formação de acervos constituídos por diversos tipos de materiais que não apenas livros, mas sim, revistas, jornais, materiais áudio-visuais e num estágio mais contemporâneo, uma parcela cada vez maior de profissionais também precisa lidar com uma verdadeira avalanche invisível e *automultiplicadora* de informações e documentos em formato eletrônico, que não podem ser “manuseados” ou mensurados.

Não é possível saber quantos arquivos eletrônicos poderiam caber numa estante ou como organizá-los lado a lado em uma sequência fixa, ordenada e previsível. Dessa forma, o profissional atual, pode se sentir como um “ser” ao redor do qual orbitam os mais diversos elementos (documentos e informações), com sua própria dinâmica, se movendo com uma fluidez de velocidade inalcançável para permitir a sua manipulação.

Mas ainda assim, os profissionais envolvidos nesse contexto precisam dar respostas às demandas que se impõem, e assim criar novas ferramentas, suportadas pelas tecnologias atuais, mas fundamentadas por um longo trajeto de acertos, erros e questionamentos que carregam em si a experiência humana na tentativa de controlar o que é gerado a partir da sua própria experiência.

E essa produção não será interrompida pela nossa incapacidade de gerenciá-la e preservá-la. Mas enfim, em que tudo isto torna os profissionais atuais próximos aos da Antiguidade?

A formação dos acervos físicos ou eletrônicos necessita para sua própria sobrevivência, de um mínimo de coerência. Se alguém se dirige a uma biblioteca ou serviço de informação qualquer (incluam-se aqui os serviços eletrônicos) em busca de materiais relacionados à agricultura, e é obrigado a vasculhar documentos de culinária, saúde e apicultura até chegar ao que deseja, podemos tirar algumas conclusões.

A primeira: o indivíduo foi persistente o suficiente para continuar a sua busca sem se direcionar a outro serviço, mesmo que não soubesse formular exatamente qual era a sua necessidade; a segunda: o estoque de documentos provavelmente não está organizado o suficiente, ou o sistema não oferece as ferramentas adequadas para atendê-lo e assim, a credibilidade desse sistema foi abalada.

Atualmente, a primeira situação não é a mais comum de ocorrer, no entanto a segunda é infelizmente, bastante frequente.

Obviamente, os recursos de informação atuais trouxeram imensos avanços a esse contexto. Mas algumas questões basilares são persistentes.

Quando se busca a organização de um acervo é necessário ter em mente, que o fundamental é o acesso à informação nele contida, e dessa forma,

independente do seu suporte ou arranjo, o primordial é que ele possa prover a informação desejada. Assim, mesmo que tenhamos ferramentas mais elaboradas, a compreensão integral do percurso entre a análise do conteúdo de um documento até a sua síntese, tradução e incorporação em um sistema, continua uma incógnita.

Neste trabalho, realizamos a abordagem desse ponto fundamental, mas ainda pouco explorado: o processo de reconhecimento e representação do conteúdo informativo do documento.

Esse processo tem sido denominado e definido de diferentes formas, mas refletem fundamentalmente um mesmo problema: como se definir o tema, ou os temas de um documento para a sua representação em um sistema de informação? Como ressaltam Dias e Naves (2007), é necessário alertar para a variação terminológica existente na área em torno dessa questão. Prosseguem os autores:

É interessante a observação de que o termo *análise de assunto* apareça na literatura com diversos sentidos. Há uma confusão com relação ao real significado do termo, verificando-se que essa situação acarreta dificuldades generalizadas, tanto para os indexadores, quanto para os professores que ministram disciplinas na área, no momento de transmitir aos alunos conhecimentos sobre esse tema. (DIAS; NAVES, 2007, p.10).

A análise de assunto, conforme Fujita (2003) é o primeiro estágio da indexação, e é subdividida em outros três estágios: a compreensão do conteúdo do documento; a identificação dos conceitos que representam este conteúdo e a seleção dos conceitos válidos para recuperação.

De outro lado, a Análise documental¹ (também conhecida por Análise documentária): “Trata-se de uma modalidade específica de análise de textos, em que esses últimos são desestruturados, sintetizados e transformados em novos textos (resumos documentários e termos de indexação), com base na distinção entre informação essencial e acessória.” (DIAS; NAVES, 2007, p.12).

Em outras áreas como a Análise do Discurso, a Linguística Textual, a Semântica Discursiva, propõe-se uma análise do texto como um todo, observando sua totalidade, e assim, realizar uma análise tão profunda quanto possível.

¹ Ver trabalho de Moraes e Guimarães (2006) onde o termo correto proposto de acordo com a gramática da língua portuguesa, é Análise documental e não Análise documentária, expressão mais comumente utilizada.

Dessa forma, a incorporação do auxílio dessas disciplinas traz inovações bastante pertinentes, principalmente ao demonstrar as variáveis relacionadas à análise do conteúdo que envolve não apenas os aspectos psicológicos do profissional (FUJITA, 2007), mas também suas posições morais e ideológicas (LUCAS, 2000).

Embora tenham passado a existir recortes dessas áreas dentro da Ciência da Informação e da Biblioteconomia, essas últimas possuem objetivos pragmáticos que raramente podem se ater integralmente às propostas de análise sugeridas por essas outras disciplinas.

No entanto, estamos atrelados a um fim bastante ingrato, que é *grosso modo*, uma espécie de “reduzir para ampliar”, ou seja, ao realizar a representação de um documento em um sistema de informação, necessitamos reduzir o conteúdo desse documento a uma ou poucas expressões, que possam ampliar a visão do público interessado, sem que ele tenha qualquer contato com o documento em si, dessa forma garantindo sua maior circulação e possibilidade de acesso.

Essa tarefa incorpora em si mesma, responsabilidades de cunho pragmático: como construir essas ferramentas? Responsabilidades de cunho social: como construí-las de tal modo que a informação neles contida possa ser encontrada, de forma a não “negar” a informação? Responsabilidades de cunho ético: como as “coisas” e as pessoas, serão nomeadas e agrupadas no sistema?

Embora a aproximação inicial a isso possa ser desoladora, frente à magnitude dessas questões, os desafios envolvidos são instigantes e dão razão a busca pelo desenvolvimento de nossa área.

Assim, o problema principal de nossa pesquisa (enquanto sequência da anterior), foi a compreensão de uma das etapas do processo de tratamento temático, a determinação do *aboutness*, tendo em vista seu papel de relevância para a fundamentação teórica em Organização e Representação do Conhecimento.

Dessa forma, propomos compreender o termo *aboutness* e se de fato essa abordagem apresentou avanços significativos no entendimento do processo de definição de assuntos, pois se observou que essa discussão permaneceu em evidência basicamente entre o período da década de 1970 e meados de 1980, sendo que após isso, poucos trabalhos o têm retomado.

Nosso objetivo, portanto, foi a revisão bibliográfica da discussão e conceituação do *aboutness* e sua inserção na literatura brasileira e assim, elaborar uma discussão em torno da teoria do “*aboutness*”, pouco abordada em língua portuguesa com essa nomenclatura, mas de certa forma, essencialmente presente nas discussões que circundam o tratamento temático.

Para essa finalidade, realizamos o levantamento da bibliografia fundamental, na qual incluímos também a discussão do termo assunto, dada a estreita relação entre *assunto* e *aboutness*.

Além disso, para a verificação da inserção na discussão nacional, realizamos ainda uma análise da literatura brasileira em Ciência da Informação, através da abordagem de quatro títulos periódicos reconhecidos pela comunidade. Assim, procedemos a pesquisa em todo o período cronológico destas publicações, do termo *aboutness* e suas traduções: atinência, tematicidade, sobrecidade.

Na primeira seção, abordamos as propostas de tratamento temático da informação, e após sua apresentação concluiu-se que todas elas apresentam a mesma problemática quanto à determinação temática, ou seja, o *aboutness*. Nessa etapa, abordamos também os estudos brasileiros em tratamento temático, através dos trabalhos de Eduardo Dias e Madalena Naves, docentes da Universidade Federal de Minas Gerais e dos trabalhos desenvolvidos pelo Grupo TEMMA, fundado na década de 1980 e do qual participam pesquisadores das universidades paulistas USP e UNESP.

Na segunda seção, apresentaram-se as definições do termo *aboutness*, seu decorrer histórico, sua adaptação em língua portuguesa, e algumas das discussões existentes na análise para determinação do *aboutness*.

Na terceira seção, detalhamos os procedimentos metodológicos e realizamos a apresentação dos periódicos nacionais que foram selecionados para constituírem nosso *corpus* de coleta e análise. Nesse momento, destacamos as formas de levantamento e os termos buscados, além da análise dos dados e a apresentação dos resultados alcançados.

Por último apresentamos as conclusões da pesquisa, e nossas considerações gerais.

I O TRATAMENTO TEMÁTICO DA INFORMAÇÃO

O âmbito do tratamento temático da informação é a área que se encarrega da identificação de assuntos e conceitos presentes em um documento e como estes podem ser representados. Conforme Pinto-Molina (1993) é o processo responsável em gerar subprodutos de um documento a partir das suas representações temáticas, que serão utilizados também como instrumentos de busca.

Conforme relembra Foskett (1973, p.3) em sua obra clássica:

“[...] a quantidade de novas informações produzidas é de tal ordem, que nenhum indivíduo pode alimentar a esperança de estar ao corrente delas, mesmo que seja de uma ínfima parcela. E o problema que temos de enfrentar é o de possibilitar às pessoas que precisam de informações a sua obtenção com o mínimo de gastos (de tempo e de dinheiro) e sem que sejam assoberbadas por grandes quantidades de material irrelevante.

Para atingir esse objetivo os processos de tratamento temático que serão realizados pela biblioteca ou demais centros de informação, valem-se de operações que não se resumem a técnicas ou procedimentos e modelos definidos que possam vir a ser estabelecidos e aplicados como fórmulas padronizadas.

No entanto, o processo de representação temática da informação envolve dois passos principais (NOVELINO, 1996, p.38): “[...]1) análise de assunto de um documento e a colocação do resultado desta análise numa expressão linguística. 2) atribuição de conceitos ao documento analisado”.

Embora se almeje uma sistematização desses processos que possa proporcionar a sua melhoria, ainda encontram-se sérias implicações, e entre as principais, estão as concernentes à análise e definição de assuntos, um dos pontos de maior divergência no tratamento temático.

Derek Langridge (1989 citado por Castañon Moreno, 1992) adverte que, a análise temática tem a ver mais com os documentos do que com os sistemas; desafortunadamente, em não poucos casos, o exame e análise de textos para determinar um tema se converte na busca de idéias ou termos que primeiro e melhor se acomodem as estruturas e símbolos previstos no sistema em que será

incorporado.

Dias (2004) afirma que o tratamento temático da informação é o processo de análise para determinar o modo como o assunto do documento será representado a partir de simulacros para ser inserido na coleção de um sistema de informação ou de recuperação. Sua atividade é de vital importância para o fluxo informacional em virtude da necessidade da busca por assunto ser a mais utilizada pelos usuários.

Dessa forma, o profissional documentalista necessita adquirir conhecimentos auxiliares e aprofundar sua reflexão sobre esse momento da atividade do tratamento da informação, tendo em vista ser determinante na probabilidade de que o documento seja acessado quando necessário.

A questão apontada apresenta um enorme desafio aos profissionais da área dado o grau de complexidade das questões linguísticas, culturais e ideológicas envolvidas, e mesmo com relação ao conhecimento do indivíduo sobre a área à qual o documento pertence.

É importante também observarmos que essa tarefa na realidade foi efetuada de forma empírica desde o primórdio das atividades documentais.

Dessa maneira, encontra-se o profissional da informação, obrigado a repensar como esse procedimento é realizado e as consequências da sua execução, tendo em vista os resultados de uma representação falha devida à má compreensão do conteúdo ou escolhas imprecisas quanto à definição de conceitos e assuntos e suas representações.

O tratamento temático da informação demonstra-se como um processo em que são conjugadas atividades de cunho teórico e prático, que necessitam cada vez mais para atender satisfatoriamente às demandas de informação, de embasamento conceitual e aprofundamento da discussão em torno da formulação de seus procedimentos.

Assim o foco deste trabalho encontra-se na abordagem da etapa em que ocorre a identificação dos conceitos que levarão aos assuntos definidos por este tipo específico de leitor (profissional indexador, etc.). Portanto, este ponto inicial do trabalho documental é:

[...] uma de suas etapas mais complexas, seja por envolver aspectos subjetivos do analista seja, ainda, pelo fato de a literatura da Ciência da Informação vir se dedicando, mais tradicionalmente ao estudo da etapa que a sucede – a representação documental – com ênfase nas linguagens de indexação. (MORAES; GUIMARÃES; 2006).

Portanto, a responsabilidade do profissional envolvido neste momento do tratamento documental, não é pequena, pois em geral, é ele próprio o indexador do sistema em que atua e desta forma, ele assume a responsabilidade por todo o complexo processo de análise de assunto, “[...] é pois a ele creditado, em grande parte, o sucesso ou o insucesso de um sistema de recuperação da informação”. (NAVES, 2001, p.190).

Em vista disto, fica evidente que a leitura do documento se constitui como a fase mais delicada rumo a uma representação eficaz do seu conteúdo e para tanto, de um eficaz estabelecimento temático. Para Fujita (2004) embora essa tarefa possa ser classificada como simples e natural, o processo de leitura para análise de assunto é altamente complexo “[...] porque depende do processamento humano de informações e da cognição de quem lê, de um texto elaborado por um autor e do contexto de ambos, o que determina os objetivos da leitura”.

Porém, como já afirmado, desde seu surgimento as atividades documentais de identificação e representação de conteúdo são realizadas de forma essencialmente prática e segundo Hutchins (1977, p.17, tradução nossa)²:

Há uma atitude certamente muito comum entre os cientistas da informação, que nós não precisamos saber como os indexadores chegaram a uma descrição particular do conteúdo de um documento, tudo que importa é se permite aos usuários encontrar o documento quando requerido.

No entanto, ainda de acordo com o autor (HUTCHINS, 1977, p.17, tradução nossa)³: “Nós poderíamos estar certamente felizes com esta visão, se não estivéssemos todos dolorosamente conscientes da inadequação dos resumos e

² Trecho original: “There is indeed a very common attitude among information scientists that we do not need to know how indexers arrive at a particular description of the contents of a document; all that matters is whether it enables users to find the document when required.”

³ Trecho original: “We could probably feel happier with this view if we were not all painfully aware of the inadequacies of the abstracts and indexes which are actually produced.”

índices produzidos atualmente”.

Levando em consideração o tempo de existência destas atividades, e o pouco referencial teórico existente a respeito, ou ao menos, de referenciais que ofereçam definições e respostas mais eficazes, fica nítido que esta discussão tem sido relegada pelos documentalistas. Conforme afirma Kobashi (1994, f.26):

A abordagem teórica do complexo percurso que vai da análise de textos à sua representação, para fins documentários, é fato bastante recente, se consideramos que a documentação, enquanto atividade prática, tem uma história que remonta ao segundo Milênio A.C.

No entanto, tal atitude nas últimas décadas passou a se alterar, uma série de pesquisadores tem se dedicado à investigação da leitura e análise de assuntos no âmbito do tratamento temático da informação, demonstrando que, se por um lado ainda não foram alcançadas definições e respostas mais concretas, por outro, o assunto tem se tornado alvo de maior atenção.

Com vistas a estes problemas, a Ciência da Informação tem ido buscar em outras áreas, e entre elas a Linguística, suportes que possam oferecer auxílio na solução destas questões, no entanto, é necessário que ela adapte tais ferramentas ao seu próprio espectro e forma de atuação, como aponta Baranow (1983, p.26):

O recorte a ser feito no universo da Linguística para a formação do cientista da informação deverá pautar-se pelos aspectos aplicativos exigidos em sua futura área profissional. Quanto ao enfoque didático, sugerimos partir-se da problematização de fenômenos linguísticos relevantes para a Ciência da Informação. É desaconselhável partir-se de soluções prontas, de teorias estabelecidas como se fossem dogmas [...]

Baranow (1983, p.28) ressalta ainda que: “[...] considerando que a Ciência da Informação continua a lidar principalmente com textos em linguagem natural, seria conveniente acompanhar os recentes progressos na Análise do Discurso, na Linguística do Texto e em campos de pesquisa afins.”

Outros pesquisadores como Van Der Laan (2000) propõem que a aproximação da análise de assunto com a Linguística Textual “[...] pode tornar-se uma importante ferramenta no trabalho do bibliotecário, capacitando-o a compreender melhor os aspectos linguísticos do texto e das questões de busca”. A incorporação de metodologias da Análise do Discurso é por sua vez recomendada

por Lucas (2000, p.52, grifo do autor), ao afirmar que “[...] a *representação* não significa a simples substituição de uma palavra por outra. Esta representação é feita a partir da interpretação e configura a dispersão do sujeito e suas diferentes posições”.

Uma das linhas de tratamento temático, a Análise Documental, independente de suas diferentes concepções: inglesa, espanhola ou francesa (FUJITA, 2003a), e norte-americana ou européia (GUIMARÃES, 2003), carrega em si o objetivo final da representação da informação de forma satisfatória à sua recuperação. Dessa forma, a Análise Documental:

[...] no âmbito da Ciência da Informação e, mais especificamente no que tange à Biblioteconomia, relaciona-se aos processos de condensação e de representação por meio de linguagens documentárias (por meio da classificação e da indexação), com o objetivo específico de produzir resumos e índices. (GUIMARÃES; MORAES; NASCIMENTO; 2005, p.135, grifo nosso)

Conforme relata Guimarães (2003, p. 104), a Análise Documental conhece distintos momentos ao longo da História: “Um resgate histórico da Análise Documentária sugere a identificação de três momentos, representativos de concepções distintas da área: a arte, a técnica e a busca por metodologias que a sustentem e justifiquem cientificamente.”

Ainda conforme Guimarães (2003, p.104-105), a Análise Documental passa por um primeiro momento em que era considerada:

[...] fruto de um talento especial, uma verdadeira habilidade artística em que o emprego do bom senso se aplica a um processo altamente intuitivo: a determinação do conteúdo do documento e sua conseqüente nomeação.

Em um segundo momento, sofrendo os reflexos da Revolução Industrial e a conseqüente explosão da produção documentária “[...] assume uma dimensão mais pragmática e passa a ser vista enquanto produto da aplicação de técnicas, receitas ou regras previamente estabelecidas”.

Porém, a discussão em torno de suas metodologias só começa a ocorrer em um terceiro momento, a partir da década de 1950, onde passam a ser

considerados o tratamento documental em larga escala e as tentativas de que seja realizado de forma automatizada, surgindo assim duas vertentes de pensamento: “[...] a européia (sobretudo na Inglaterra, na França e na Espanha), mais voltada para a questão do processo em si, e a norte-americana primordialmente centrada na consistência dos produtos.”(GUIMARÃES, 2003, p.105).

Formalmente conceituada como “Análise Documental” a partir principalmente de Gardin (1981, p. 29 *citado por* FUJITA, 2003,f.05), conhece diferentes concepções na Europa (FUJITA, 2003): a linha espanhola entende a Análise Documental enquanto dividida em dois processos: a análise da forma e a análise do conteúdo, ou seja, a primeira enquanto **representação descritiva** (descrição física ou bibliográfica) do documento, visando seu acesso físico e a segunda, sua **representação temática** (resumo, classificação, indexação) visando seu acesso temático. Enquanto a linha francesa a entende como tratamento apenas do conteúdo, ou seja, refere-se apenas ao tratamento temático, no qual a indexação é uma de suas etapas, sendo então o resultado deste tratamento. E por fim, a vertente inglesa representada nas pesquisas em indexação principalmente por Lancaster (2002) que entende a indexação como processo idêntico à Análise Documental “[...] incluindo-se a análise de assuntos como etapa inicial da indexação”.

As etapas que constituem o processo da Análise Documental são, para Kobashi (1994) a **análise**, a **síntese** e a **representação**. Através das quais o conteúdo do documento será representado em diferentes níveis, uns mais compactos, como no caso da indexação, que se limita a termos descritores ou expressões, ou níveis menos sintéticos tais como os resumos, que praticamente se caracterizam uma obra nova em novo formato, visando abarcar os pontos principais da obra original.

A primeira etapa no processo de análise temática tem por finalidade “[...] identificar a organização metodológica do discurso do autor/produtor através da segmentação do texto e em seguida isolar conceitos/palavras-chave tradutoras do conteúdo desses segmentos” (CUNHA, 1989, p.57). Essa etapa é considerada a fase mais importante desse processo, pois tem por finalidade extrair e revelar o assunto, tema, conceitos que melhor representam o núcleo do documento.

A realização dessa análise ocorre através da leitura documental, portanto

“[...] a leitura constitui atividade fundamental da Análise Documentária, pois resulta na seleção dos conceitos que, posteriormente, serão representados por termos que serão recuperados o documento pelo usuário.” (FUJITA, 2003, fl.8).

Para Fujita (2004, p.4) “a atitude do leitor frente ao texto, anteriormente vista como recepção passiva de mensagens, passou a considerar o processamento mental de compreensão da informação e evoluiu para uma perspectiva de interação entre o leitor e o texto”. Essa concepção da compreensão na leitura envolvendo o leitor se expandiu significativamente nas últimas décadas.

É grande a responsabilidade do leitor indexador frente à análise de assunto, a ele compete a compreensão do assunto do documento, de modo a extrair sua essência e principais aspectos. Sua leitura deve ser respaldada na neutralidade, de modo que questões subjetivas tais como valores e opiniões pessoais, não interfiram na leitura, identificação e representação do assunto. Quanto a essa neutralidade Cintra (1989, p.31) afirma a sua impossibilidade:

[...] tendo em vista qualquer objetivo de leitura, mesmo que seja uma análise documentária. Sendo a linguagem intrinsecamente comprometida com o cultural e o ideológico, tanto o processo de produção do texto, quanto o de recepção não se isentam destes componentes. E é esse comprometimento que garante o caráter interativo da leitura.

Em relação a esse caráter subjetivo da leitura, a observação de Fujita (2004) se faz também pertinente:

[...] supõe-se que o indexador não tenha plena consciência desse processo que envolve seu conhecimento sobre o texto e seu contexto de trabalho, que fixa os objetivos e delimita seu desempenho no ato de indexação. Em consequência, supõe-se também que o indexador não consiga atentar para a obviedade de que a fase inicial da indexação, a análise de assunto, realiza-se pela leitura e que condições específicas para o desenvolvimento desse processo, como o conhecimento prévio e estratégias construídas durante sua vida, têm importância fundamentais.

Esse processo de síntese a partir da leitura consiste na tradução dos elementos principais contidos no documento, que no primeiro momento foram selecionados e depois traduzidos em linguagem documental (CUNHA, 1989, p.60). É nessa etapa que está inserida a complexidade das tarefas de seleção e de

condensação, “[...] na medida em que implica atribuir valor às informações contidas no texto”. (KOBASHI, 1994, p.24) a partir dos instrumentos de linguagens documentais. Por último ocorre a etapa de representação, que conforme Fujita (1998, p.22):

[...] possui duas naturezas distintas: na primeira, a representação é construída através de um processo de condensação intensiva do texto original, gerando os diferentes tipos de resumo; na segunda, a representação é realizada através do uso de uma linguagem documentária, que tem como função a normalização das unidades significantes ou conceituais presentes no texto original: indexação.

No Brasil, destacamos o trabalho realizado pelo Grupo TEMMA (SMIT, 1987; CUNHA, 1989), do qual participam pesquisadores da USP e da UNESP. O grupo desenvolve suas atividades desde a década de 1980, em torno do desenvolvimento teórico e experimentações quanto ao tratamento temático, trazendo propostas de formulação de um quadro conceitual em Análise Documental, aplicação em diferentes tipologias discursivas e textuais, a realização de análise em textos literários, contos, textos jornalísticos entre outras abordagens e explorações, além de uma série de trabalhos individuais de seus membros, o grupo trabalha com as propostas de benefícios advindos dos conceitos a serem incorporados, oriundos de outras áreas do conhecimento como a Linguística Textual, Análise do Discurso e a Lógica, que estabelecem relações interdisciplinares com a Ciência da Informação.

Com base no que foi apresentado, a proposta desse trabalho justifica-se pela relevância que os estudos temáticos possuem nas áreas da Ciência da Informação e a necessidade de avanços pragmáticos nesse sentido, embasados por estudos teóricos que possam sustentá-los. Frisamos desde o início a falta de referencial teórico nesse campo, conforme Fujita (2003, p.78, grifo nosso):

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.

Desta forma, entendemos que o estudo da tematicidade apresenta-se enquanto ponto essencial ao progresso teórico e conceitual dos processos englobados pelo tratamento temático da informação e em um âmbito ainda maior

pela organização e representação do conhecimento, e assim, é fundamental que tenhamos em vista que deste aprofundamento deverão surgir os avanços necessários aos procedimentos seguintes de tratamento e representação.

Portanto, consideramos de total relevância a preocupação com a qualidade na geração dos produtos documentários, conforme afirma Moreiro (2002, p.53, tradução nossa): “Os resultados do tratamento documental devem garantir a qualidade e eficácia dos produtos que oferecem, se se quer justificar e assegurar a própria existência dos repertórios documentais.”⁴

Entendemos assim que:

“A qualidade das tarefas correspondentes a análise de conteúdo documental (especificadas na indexação e resumo) resulta, pois, fundamental para permitir uma satisfatória recuperação da informação e uma adequada explicação dos conteúdos aos usuários.” Moreiro (2002, p.54, tradução nossa)⁵.

Afirmamos que como um processo onde as etapas não podem ser encaradas como meros procedimentos separados uns dos outros, mas sim como um ciclo em que cada etapa está diretamente ligada à sua anterior dela dependendo, e as primeiras por sua vez podem perder o sentido quando as últimas não atendem ao mesmo nível de qualidade. Consideramos ainda que além da observação do produto documental resultante do tratamento temático, devemos nos preocupar também com a observação dos modelos de leitura, com os conceitos e fundamentos que tornam estas questões em problemas científicos a serem solucionados pela Ciência da Informação, tendo em vista seu desenvolvimento e consolidação, onde consideramos pertinente a observação realizada por Lucas (1997):

[...] não devemos substituir as questões, difíceis, interessantes e promissoras, [...] por "objetivos operacionais", que se atingem a curto prazo. Nem devemos somente disciplinar o exercício dos recursos da inteligência humana com relação ao arquivo textual, mediante dispositivos de classificação, indexação, que derivam mais da gestão administrativa e do sonho logicista de língua ideal, que da pesquisa científica fundamental.

⁴ Trecho original: “Los resultados del tratamiento documental deben garantizar la calidad y eficacia de los productos que ofrecen si se quiere justificar y asegurar la propia existencia de los repertorios documentales.”

⁵ Trecho original: “La calidad de las tareas correspondientes al análisis de contenido documental (especificadas en la indización y el resumen) resulta, pues, fundamental para permitir una satisfactoria recuperación de información y una adecuada explicación de los contenidos a los usuarios.”

Discutimos em seguida, a teoria em torno do conceito “*aboutness*”, que é em essência, o cerne principal dos trabalhos acima abordados.

II O ABOUTNESS

O conceito *aboutness* alcançou maior propagação na Ciência da Informação nos últimos anos, notadamente a partir da década de 1970 (BEGHTOL, 1986), no entanto, a área não chegou a um consenso quanto à sua definição, pois sua estreita relação com o termo assunto e a operação análise de assunto faz com que diversas vezes sejam considerados meros sinônimos ou variações.

No entanto, a dificuldade de definição e a própria relação de proximidade com outros termos (assunto, tema, etc.) tem causado inclusive o questionamento da validade ou utilidade (HJØRLAND, 2007) de uma definição conceitual do termo.

Obviamente, se essa questão levanta discordâncias mesmo entre os teóricos de outros países, no Brasil, enfrentamos além das questões comuns e fundamentais do processo de análise temática, a própria adaptação da expressão ao idioma.

Nesta seção, buscou-se apresentar o histórico e as definições de *aboutness*, além das discussões encontradas na literatura. Embora a busca pela clara distinção entre assunto e *aboutness* tenha sido pretendida, percebeu-se que a relação intrínseca dos termos denota a quase impossibilidade de tratá-los de forma isolada, pois conforme Joudrey (2005) os conceitos foram intrinsecamente entrelaçados desde que a Biblioteconomia e Ciência da Informação tenderam a tratar os termos enquanto sinônimos. Dedicamos uma subseção à discussão do termo assunto.

De qualquer maneira, mesmo que a dificuldade com a identificação de assuntos possua raízes mais abrangentes, destacamos que essa preocupação, começou a ganhar corpo na área em termos de discussão teórica, após o trabalho de Charles Ammi Cutter (1876), *Rules for a Dictionary Catalog*, na qual buscou conforme relata Foskett (1973) estabelecer diretrizes para a criação de catálogos de assuntos, e dessa forma, apresentou suas considerações quanto ao problema e mesmo que em alguns casos, suas soluções raiassem a sublimidade da inocência (FOSKETT, 1973), a obra de Cutter trouxe colaborações influentes até os dias de hoje, marcando inclusive como afirma Hickey (1976, p. 273) o início de uma

dicotomia entre as regras e procedimentos de identificação descritiva e a evolução dos princípios e práticas da análise temática.

Os objetivos de Cutter (JOUDREY, 2005) têm guiado os princípios da Organização da Informação há mais de cento e vinte e cinco anos e suas idéias ainda norteiam muitas práticas atuais. Cutter (1876) estabeleceu três princípios ao redor do catálogo de assuntos, nos quais destacava a importância de que o catálogo pudesse prover o acesso aos documentos quando o assunto buscado fosse conhecido, mas que também a biblioteca pudesse mostrar os documentos de um determinado assunto existentes em seu acervo.

Nesse primeiro momento em que a Biblioteconomia passou a ter preocupações de ordem mais conceituais e teóricas, o termo *aboutness*, não era utilizado, e seu lugar era de fato preenchido pelos termos assunto, tema, tópico.

No entanto, esses termos tiveram uma larga utilização que acarretou em uma espécie de generalização, perdendo dessa forma, a sua delimitação mais precisa dentro do processo de tratamento documental, e a abordagem a partir do conceito *aboutness*, buscou claramente ser uma tentativa de distinção da natureza temática dos documentos, enquanto uma “propriedade” da obra independente de fatores externos.

Mas de qualquer forma, uma pergunta que subjaz ao nosso tema, é qual a necessidade de uma definição conceitual a respeito de um processo tão comum quanto dizer “do que trata” um livro, um artigo, ou uma obra de qualquer natureza? Ou como questiona Hjørland (2007), como essa definição pode ser útil para estabelecer diretrizes de tratamento de assunto?

Lancaster (2003), com uma visão bastante objetiva questiona se realmente precisamos de uma compreensão teórica de atenção para realizar uma indexação mais eficiente: “Não bastará que sejamos capazes de reconhecer que um documento tem interesse para determinada comunidade pelo fato de contribuir para nossa compreensão dos tópicos X, Y e Z?

O trabalho de Lancaster ainda traz um apontamento quanto às preocupações de Hjørland (2001), pelo fato deste ter se dado “[...] ao trabalho de tentar diferenciar termos como ‘assunto’, ‘tópico’, ‘tema’, ‘domínio’, ‘campo’ e ‘conteúdo’[...]”.

Para Lancaster (2003, p.15), enquanto essa discussão pode ter algum interesse acadêmico:

“[...] não tem qualquer importância prática para o indexador, que fará bem se ignorar essas diferenças semânticas e simplesmente atribuir ao item os rótulos que o tornarão utilmente recuperável pelos membros de uma comunidade-alvo.”

Esbarrando em uma corroboração da afirmação de Lancaster, sabemos que no dia a dia comum isso não é alvo de maiores questionamentos, ou seja, podemos dizer sem grandes preocupações “do que trata” um livro, um filme etc., no entanto o tratamento de assuntos em um sistema de informação, obviamente acarreta várias implicações que pouco tem a ver com as existentes no cotidiano e certamente a mais basilar é a que a interpretação dada a alguma obra em uma conversa comum, não possui o propósito de ser registrada ou ser a chave de acesso a própria obra que foi discutida. Por outro lado no trabalho documental, esses são justamente os seus propósitos naturais.

Assim, além de se buscar a construção de uma ferramenta de auxílio ao público geral ou a uma comunidade específica, a dificuldade de revisão de possíveis erros é uma realidade de grande parte das instituições, ou seja, sistemas com atribuições de assuntos equivocadas, provavelmente permanecerão nessa situação, dessa forma, essa é uma preocupação que deve estar na gênese de qualquer sistema.

No entanto, conhecemos as dificuldades em obter um documento com um conteúdo específico e a agilidade necessária, pois não basta saber exatamente sobre o que é a sua requisição, mas também como formular a sua busca de forma que possa ser identificada com o vocabulário do sistema, ou ao menos quando o sistema possui um vocabulário próprio.

Dessa forma, sabemos que as tecnologias de informação têm apresentado grandiosos avanços e conquistas nesse sentido, inclusive ao permitir a busca em linguagem natural, em todo o conteúdo dos documentos. Mas ainda não tem sido capazes de responder, ao menos sozinhas, as questões quanto à compreensão do conteúdo do documento ou a sua relevância para a necessidade do usuário.

Assim, pesquisas que investiguem os fundamentos teóricos desses problemas, são de considerável auxílio ao avanço dos próprios sistemas de informação, sejam eles eletrônicos ou tradicionais.

O termo aboutness

Inicialmente, o termo *aboutness* possui um problema relativo à sua apresentação em nosso idioma, pois mesmo já sendo utilizado na literatura nacional, não há consenso sobre sua tradução ou adaptação gramatical em língua portuguesa.

Desse modo, é possível encontrar o termo traduzido como “atência” (DIAS; NAVES, 2007; ALVARENGA, 2001) e “tematicidade” (GUIMARÃES; MORAES, 2006; MEDEIROS, 1986; FUJITA, 2003; 2004), ou determinado simplesmente como um sinônimo para “assunto”, seguindo a opinião de Hjørland (1997) - embora o termo assunto⁶ também possua divergências teóricas quanto à sua conceituação em Ciência da Informação.

Para Baranow (1983) em uma das primeiras obras em português que observam o problema da tradução do termo:

[...] o primeiro passo na indexação é determinar do que trata um determinado documento. É o que poderíamos chamar de ‘concernência’ (do verbo **concernir**, adj. **concernente**), termo mais adequado ao vernáculo do que o esdrúxulo anglicismo ‘sobrecidade’ (**aboutness**).

Para Medeiros (1986, p.140) em seu estudo sobre terminologias em CI, (o termo) deve ser traduzido por “tematicidade” sendo essa a palavra mais apropriada, pois apesar de ser um neologismo “[...] foi criado observando-se os padrões gramaticais da língua portuguesa (derivação do adjetivo **temático** com o sufixo — (d)ade)”.

⁶ A discussão do termo assunto é realizada em subseção própria na página 50.

Dias e Naves (2007), como já mencionamos, preferem a utilização do termo “*atinência*”, pois optaram por seguir a tradução em português (realizada por Briquet de Lemos⁷) para o livro *Indexação e Resumos: teoria e prática*, do autor inglês Lancaster (2003), enquanto Fujita (2003, p. 79), considera que tematicidade “[...] está mais relacionado com a noção de tema do documento” e pode significar ‘*do que trata um texto*’.

De qualquer forma, enquanto o termo *atinência* é registrado em dicionários brasileiros (FERREIRA, 1975; CALDAS AULETE, 2009), como a qualidade ou condição do que é atinente (Atinente: 1. Que se refere, diz respeito, concerne a, que tem a ver com); o termo tematicidade não está presente, sendo o mais próximo o termo “*temática*”.

Dessa forma, caso venha a se respeitar o vocabulário da língua portuguesa, o termo correto seria *atinência*, ou caso contrário, a adoção de tematicidade se torna de fato um neologismo, conforme observado anteriormente por Medeiros (1986), o que obviamente não invalida o uso da expressão desde que possa atender ao propósito.

Quanto à etimologia da expressão *aboutness*, ‘*about*’ possui vários sinônimos em inglês (CHRIST, 1975): *near, round, around, encircling, through, over, touching, respecting, concerning*. O termo *ness*, por sua vez é um sufixo (MICHAELIS, 2008) utilizado para a formação de substantivos abstratos, por exemplo: *good* (bom) *goodness* (bondade).

Apesar da intenção de não utilizar termos estrangeiros que possuam correspondência em português, não adotamos para esse trabalho nenhuma das traduções de *aboutness*. Devido a isso, optou-se por manter o termo original ao longo do trabalho, e preservar as traduções utilizadas pelos autores nacionais, quando estes estiverem em foco. Daí a possível utilização dos termos de forma aparentemente desordenada.

Além da Ciência da Informação, outras áreas também exploram o assunto, como a Psicologia, a Informática (mais especificamente as pesquisas em

⁷ O tradutor, Briquet de Lemos, em nota de rodapé sobre a tradução do termo, diz o seguinte: “O autor emprega os termos ingleses *about* e *aboutness*. O primeiro traduzimos por ‘trata de’ e o segundo por ‘atinência’. Outros traduzem *aboutness* por ‘tematicidade’, ‘temática’, ‘acerca-de’, ‘ser acerca-de’, ‘ser sobre-algo’, etc.

automação de tratamento documental e Inteligência Artificial), além da Teoria Literária, Linguística Textual e Filosofia (na qual é tratado por Intencionalidade), obviamente cada área com suas especificidades. No entanto, foge ao nosso escopo maior compreensão acerca do reconhecimento, significação e utilização do termo em outras áreas de conhecimento.

O aboutness na Ciência da Informação

O termo *aboutness* passou a ser discutido em maior grau na Ciência da Informação a partir da década de 1970 com o trabalho de John Hutchins (GUIMARÃES, MORAES, 2006). Hutchins (1977) definiu “*what documents are about?*” como a pergunta principal para a área.

Mesmo antes de Hutchins, Fairthorne já era creditado pelo termo *aboutness* (JLOUDREY, 2005) em Ciência da Informação, com a intenção de evitar as discussões filosóficas existentes em torno do termo assunto.

De qualquer maneira, para Joudrey (2005), Fairthorne conseguiu apenas criar um sinônimo em Ciência da Informação, para o termo assunto, mudando apenas o nome do conceito, já que logo após, toda a carga de complexas questões em torno do assunto foram transferidas para o termo *aboutness*, e desde então não foi encontrada uma real distinção entre os dois termos.

Dessa forma, vários outros autores também levaram em consideração esse problema, entre os principais, está o trabalho de Hutchins (1977).

John Hutchins, bibliotecário inglês, relata em notas autobiográficas (HUTCHINS, s.d.) que iniciou sua carreira na biblioteconomia após a graduação em 1960 e entre 1963 e em 1964 passou a se interessar pelos trabalhos do linguista norte-americano Noam Chomsky. Ao mesmo tempo, Hutchins passou a trabalhar com Recuperação da Informação (RI) e acreditava que o estudo da Linguística para fins documentais, mais especificamente a Linguística Textual, poderia contribuir com a automação dos sistemas de informação. Hutchins foi o autor que notou a relevância da Linguística Textual para a Ciência da Informação. (BEGHTOL, 1986).

No entanto, até em torno da década de 1990, não havia na área de RI, muito interesse nas questões linguísticas, pois os métodos estatísticos prevaleciam nesse campo. Dessa forma, Hutchins, continuou a trabalhar com sumarização automática, na qual a análise linguística era relevante, entretanto a partir da década de 1980 em diante seu interesse recai primordialmente sobre a tradução automática (*machine translation*) e sua história.

Na concepção de Hutchins (1977, p.17) a primeira pergunta a ser feita é o que se entende por tópico/tema de um documento no contexto de um sistema de informação⁸?

Para o autor, uma resposta superficial é a que o tema é o assunto descrito no índice ou vocabulário relacionado ao documento e admite, no entanto, que raramente há uma verdadeira relação entre a entrada no vocabulário e o que de fato o documento trata, e dá como exemplo os sistemas de indexação pós-coordenados, cuja entrada pode representar somente uma parte do conteúdo do documento, ou seja, a descrição do assunto é apenas uma forma de expressão de parte do tema de um documento.

Uma segunda questão abordada por Hutchins (1977, p.17) é “o que queremos dizer por conteúdo de um documento?”⁹ e afirma que para respondermos a isto devemos estar certos quanto à distinção entre o ‘sentido’ (*sense*) de uma expressão linguística e a ‘referência, relação’ (*reference*) com o objeto, indivíduo que esta expressão representa. De outro modo, podemos dizer que uma expressão só pode fazer sentido e ser encarada como conteúdo relevante dependendo do relacionamento existente entre si e os outros termos que compõem o texto, esse sistema de relacionamento¹⁰ é o que de fato determina o seu sentido e a sua relevância.

Buscando outras abordagens, Beghtol (1986) também abordou o termo *aboutness* na Grã-Bretanha (FUJITA, 2004) e realizou um trabalho no qual

⁸ Trecho original: *The first question to be asked is “what is meant by the topic of a document in the context of an information system?”*

⁹ Trecho original: *“What do we mean by the content of a document?”*

¹⁰ Esta idéia é similar ao conceito de isotopia advindo da Análise do discurso, conforme a obra de Greimas adaptada por Fiorin (1999) trabalhada na análise de textos literários. Entende-se por isotopia, a interpretação ou sentido de um termo somente em acordo com o conjunto em que está inserido, ou seja, uma expressão só será devidamente interpretada e compreendida realizando-se o nivelamento semântico do texto integral do qual faz parte.

apresentou diferentes visões de *aboutness* e relata que por ocasião da realização em 1977 de um colóquio promovido pelo *Aslib Co-ordinate Indexing Group* (CIG) vários pesquisadores puderam expor suas visões sobre o *aboutness* e apesar de utilizarem diferentes termos para expressar a questão, a idéia fundamental de que há um tema intrínseco e outros que podem ser atribuídos está mantida com diferentes nomenclaturas:

Parar Fairthorne, o *aboutness* é **Extensional**, quando inerente ao documento, e **Intensional**, razão ou propósito pelo qual ele é adquirido ou qualquer sentido que lhe é atribuído (*Extensional aboutness* e *Intensional aboutness*).

Boyce denomina como **Topicalidade** (*Topicality*) e **Informativo** (*Informativeness*).

Van Dijk apresentou sua concepção onde para o tema inerente ao texto usou a expressão **Atribuição normal de relevância** (*normal relevance assignment*) e para os sentidos atribuídos **Atribuição diferenciada de relevância** (*differential relevance assignment*).

No entanto, Van Dijk (2002), estava de fato mais interessado em investigar esse problema do ponto de vista da Análise do Discurso e da Psicologia: “Por outras palavras, estamos introduzindo na psicologia cognitiva a conhecida distinção da filosofia entre semântica *intensional* (significado) e semântica *extensional* (referencial)” (Van Dijk, 2002, p. 25).

Dessa forma, para Beghtol (1986) o termo *relevância* utilizado por Van Dijk, deve ser evitado em nosso contexto, pois é um termo já presente na Ciência da Informação e relacionado às discussões em Recuperação da Informação¹¹.

No entanto para Lancaster (2003, p. 14) a *atênência* é um tema bastante próximo ao conceito de relevância “[...] isto é, a relação entre um documento e uma necessidade de informação ou entre um documento e um enunciado de necessidade de informação (uma consulta).”

Lancaster (2003, p.13), buscou não entrar na discussão do termo: “Não pretendo partir para uma discussão filosófica sobre o significado de “trata de” ou

¹¹ O termo relevância, conforme apresenta Foskett (1973) está relacionado ao processo de Recuperação da Informação, e é considerado o grau de correspondência dos resultados apresentados em uma consulta, com as exigências específicas do usuário.

“atênência”. Vários autores já o fizeram. E nem assim conseguiram esclarecer a situação, pelo menos no que tange à atividade de indexação de assuntos.” Mas discute além do *aboutness* em obras textuais, também o *aboutness* de obras de arte, o termo *of-ness* (de-ência). (LAYNES, 2002 citado por LANCASTER, 2003).

Kobashi (1994), também aborda o trabalho de Beghtol (1986) e o considera como um dos trabalhos a trilhar o caminho de sistematização do bom-senso¹² e: “[...] hipotetiza o “*aboutness*” como a proposição de nível mais genérico obtida pela seleção e generalização crescente das proposições de um dado texto.”

Fairthorne (1969) também havia segundo Wanderley (1973, p.187) observado a amplitude do universo relacionado ao *aboutness*:

Problema crucial, mas que transcende o plano estritamente linguístico é, nessa esfera da comunicação, o da “*aboutness*” a que se refere Fairthorne, ou seja, o da “atênência” dos documentos, o da natureza daquilo sobre que versam e a propósito de que os solicitam os usuários. O julgamento do conteúdo tem por substrato um fundo amplo de experiência social e psicológica, mas apóia-se na evidência dos textos quanto às regras de linguagem, conceitos, interpretações, classe e relações.

Além disso, Fairthorne (1969 citado por BEGHTOL, 1986) ainda observou que em geral tem sido maior a preocupação em definir se dois ou mais documentos são referentes a um mesmo assunto, do que determinar o assunto de um documento em si.

Essa prática por sua vez advém da tradição da classificação, pois como lembra Hickey (1976, p.273) o campo da análise de assuntos, acabou dividido em duas direções: catalogação de assuntos e classificação.

Enquanto a catalogação de assunto envolvia a seleção da terminologia para representar o conteúdo do documento singularmente e associada à catalogação descritiva, o processo da classificação é a tentativa de agrupar os materiais sob um mesmo tema, constituindo uma operação distinta.

¹² Para Cunha (1989, p.40 citado por GUIMARÃES, 2003): “[...] os procedimentos envolvidos na localização e identificação dos conceitos no documento, para serem posteriormente representados por meio de linguagens documentárias, eram encarados como operações de bom senso dos bibliotecários, revelando uma vasta diversidade de critérios, de natureza fluida e subjetiva, sem parâmetros que lhe conferissem alguma cientificidade, ou, ainda, a explicitação dos procedimentos ou mecanismos envolvidos.”

Para Beghtol (1986), os sistemas de classificação tradicionais têm sido baseados justamente na hipótese de um *aboutness* relativamente fixo, os esquemas buscam ser receptivos a qualquer *aboutness* que um documento apresente, no entanto, não houve a previsão de que novos significados (*meanings*) pudessem ser adicionados pela necessidade ou uso específico de um leitor.

Beghtol por sua vez, considera segundo Fujita (2004, p.149) que:

[...] o documento tem uma tematicidade que lhe é relativamente permanente, porém um número variado de mensagens ou significados que podem ser medidos, conforme o uso exato do documento para o usuário. Destaca ainda, que o mesmo documento pode ter significados diferentes para o mesmo leitor, em épocas diferentes, mas, como o documento por si mesmo, é imutável, presume-se possuir uma tematicidade que lhe é fundamental.

Dessa maneira Beghtol realiza uma grande contribuição (GUIMARÃES, 2004) ao propor a distinção entre *aboutness* e *meaning*, em lugar dos binômios propostos pelos autores anteriormente citados.

Beghtol pesquisou “[...] a noção de assunto mediante a linguística textual e propôs que a tematicidade em classificação bibliográfica dos documentos fosse determinada pelas cinco macroregras do modelo de leitura de Van Dijk e Kintsch (1983) [...]” (FUJITA, 2004, p.149).

Desta maneira, podemos ver que se não houve entre os autores, de um lado, a concordância quanto à uma definição do *aboutness*, por outro, a idéia de que qualquer documento possui um assunto intrínseco, inalterável e outro, ou outros, que podem ser atribuídos (de acordo com a conveniência do sistema em que esteja inserido ou das necessidades e expectativas de quem o venha consultar), se mostra presente entre a área.

A partir dessas interpretações reafirma Fujita (2003, p. 80) a respeito de Beghtol:

Sob essa influência, Beghtol 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.

Outro trabalho clássico do tema foi o realizado por Maron (1977), e surgiu a partir da constatação de ser esse o ponto principal de um sistema de recuperação de informações, o autor observa: “Desde que o *about* está no coração da indexação, como nós podemos formular qualquer teoria de indexação se nós não podemos explicar o conceito chave de *about*?”¹³ (MARON, 1977, p.38).

Buscando investigar possíveis comportamentos de análise e busca de documentos, após iniciar seu trabalho com a descrição de um sistema de recuperação de informações, o autor simula algumas possíveis situações cotidianas de troca de informação entre indivíduos, buscando exemplificar a dificuldade de exprimir a compreensão sobre um documento. Para Maron (1977), a leitura de uma obra gera uma experiência interior sobre o que ela trata, mas não necessariamente permite que estejamos aptos a relatá-la com precisão, ou instruir alguém a como fazê-lo.

Assim, o esclarecimento de um conceito de *aboutness* é buscado pelo autor, através da proposição de três tipos de *about*: *S-about* (*subjective about*); *O-about* (*objective about*); *R-about* (*retrieval about*).

Esses conceitos de *about* propostos por Maron são definidos pelo autor da seguinte forma:

- *S-about* (*subjective about*): o *about* interior, gerada da experiência psicológica de análise, leitura de uma obra; é a que referimos anteriormente como pouco passível de ser expressa ou descrita pelo indivíduo;
- *O-about* (*objective about*): em contraponto ao anterior, o *O-about* refere-se ao conceito que pode ser observado externamente em uma situação de recuperação de informação, ou seja, refere-se ao atual ou potencial comportamento de busca ou pesquisa de documentos;

¹³ Trecho original: “Since *about* is at the heart of indexing, how are we to formulate any proper theory of indexing if we cannot explicate precisely the key concept of *about*?”

- *R-about (retrieval about)*: este conceito de *about*, também pode ser observado e mensurado, pois refere-se a probabilidade de satisfação dos usuários com um determinado documento a partir de uma determinada busca. Para Joudrey (2005), esta é a tentativa de Maron de operacionalizar o *about* como passível de ser medido quantitativamente.

De qualquer forma, mesmo contribuindo com suas propostas de conceituação, Maron (1977) reconhece a dificuldade dessas definições:

“Primeiro, nós estamos sugerindo que há vários conceitos de *about*. Possivelmente, podemos dizer que pode ser interpretado de vários pontos de vista diferentes e cada novo ponto de vista origina uma interpretação diferente do significado de *about*.” (MARON, 1977, p.40)

Por sua vez, Todd (1992) também buscou analisar esse conceito e apresentou outros autores, dentre os quais Wilson (1985), o qual afirma que a natureza de um assunto é essencialmente indeterminada, e sugere que o processo de análise de assunto é muito mais que um processo subjetivo, a visão do indexador sobre um determinado assunto, é que forma esse processo.

Oliver (citado por TODD, 1992) concorda com a visão de que um mesmo documento pode ter diferentes significados para o indexador ou um leitor, em diferentes ocasiões, de acordo com suas percepções. Todd (1992) aponta então algumas questões para as quais a literatura ainda não indicou algum consenso:

Como os indexadores são capazes de descrever o assunto dos documentos? Que sistemas de orientação internos determinam as suas ações? Qual é o processo cognitivo de identificação do *aboutness* intrínseco dos documentos, que é o processo pelo qual um indexador descreve o assunto intrínseco ou o *aboutness* de um documento? (TODD, 1992, p. 102)¹⁴

¹⁴ Trecho original: How are indexers able to write down the subject documents? What internal guidance systems determine their actions? What is the cognitive process of identifying the intrinsic *aboutness* of documents, that is the cognitive process by which an indexer describes the intrinsic subject or *aboutness* of a document?

Para Todd (1992), há a necessidade de compreendermos esses processos, não somente para aumentar a consistência da indexação entre indexadores, identificando as causas atuais dos baixos índices, mas também para ampliar o nosso conhecimento e compreensão sobre a “operação fundamental” da atividade informacional.

Alguns autores que foram abordados no trabalho de Joudrey (2005), também trazem considerações relevantes quanto às concepções em torno do *aboutness* entre eles Cleveland e Cleveland (2001, citado por JOUDREY, 2005) oferecem a noção de que os textos possuem duas características: “o que as palavras dizem e o que as palavras *significam*”¹⁵.

Bertrand-Gastaldy et al.¹⁶ (citado por JOUDREY, 2005) afirma que toda leitura é subjetiva e orientada a um fim específico, o que também se aplica a leitura no contexto profissional do classificador e indexador que está tentando definir o *aboutness*.

Contrapondo essa visão, outro trabalho retomado por Joudrey (2005) é o de Svenonious (2000), para a autora a visão de que a determinação de assuntos é totalmente subjetiva, é na verdade uma visão deturpada.

Svenonious (citado por JOUDREY, 2005) apresenta ainda um trabalho não-publicado de Fox e Norreault, no qual confirmam a sua idéia, pois consideram que desde que há um nível comum de concordância na utilização da palavra *about*, há então regras que dirigem seu uso. Dessa forma, dizer que a determinação de assuntos é algo totalmente subjetivo, não permitiria a idéia de que erros podem ocorrer nessa análise, mas, no entanto, os erros ocorrem. Por exemplo, seria um grande equívoco dizer que Hamlet é um tratado de termodinâmica.

Para Joudrey (2005), esses pesquisadores assumem dessa forma, uma posição de extremo relativismo, que na sua visão é inaceitável para a organização da informação, pois rejeitam a idéia de um ***aboutness* extensional** (inerente ao

¹⁵ Trecho original: “*what the words say and what they mean.*”

¹⁶ Suzanne Bertrand-Gastaldy et al., “Convergent Theories: Using a Multidisciplinary Approach to Explain Indexing Results,” in *Forging Partnerships in Information, Proceedings of the American Society for Information Science, Vol. 32, 58th Annual Meeting, Chicago, October 9-12, 1995*, ed. by Tom Kinney (Medford, NJ: Information Today, 1995), 56.

documento), e afirmam ser possível que um mesmo leitor encontre diferentes significados em um documento, em diferentes ocasiões.

De fato, essa é a idéia subjacente ao conceito de *aboutness*, e dessa maneira seu raciocínio não nega a existência do *aboutness* inerente, extensional.

Além disso, prossegue Joudrey (2005), os pesquisadores que defendem a idéia de que há um *aboutness* extensional e outros, intensionais, afirmam nitidamente que a atribuição de sentidos (*meanings*) ao texto é além de possível, muito provável.

Joudrey (2005) analisa essas discordâncias, e nos fornece alguma luz quanto à distinção entre *aboutness* e assunto, pois o autor concorda com a visão de que há um *aboutness* extensional e que o intensional, também pode existir:

O pesquisador aceita a distinção entre os dois, e usa os termos *aboutness* e assunto para fazer essa distinção. Assim, acredita que a distinção entre *aboutness* extensional e intensional – ou entre os termos *aboutness* e assunto – é enfim útil, importante, e correto.

Considerando que a identificação do *aboutness*, possui dificuldades intrínsecas, ainda há também a questão relacionada ao gênero documental que está em análise.

Os estudos a seguir, abordam a identificação do *aboutness* em textos literários.

Campbell (2000a citado por Joudrey, 2005) duvida que há um *aboutness* inato esperando a ser descoberto pelos indexadores (contrariando Hutchins (1977) que afirma o inverso).

O trabalho de Campbell (2000a, 2000b), analisa a questão do *aboutness* em textos literários, trabalho no qual aborda o tratamento homossexual na crítica literária. Campbell aponta a subjetividade da determinação do *aboutness* através da discussão da “*queer theory*”¹⁷. Sobre a *aboutness analysis*, Campbell (2000b) observa que:

¹⁷ Demos a tradução para *queer theory* no contexto da pesquisa de Campbell por “teoria homossexual”, considerando as traduções possíveis para *queer*: esquisito, ridículo, adoentado, homossexual (MICHAELIS, 2008). Embora o termo denote certamente uma conotação pejorativa, foi adotado pelo autor, possivelmente com a intenção de denunciar, satirizar ou neutralizar a carga preconceituosa.

[...] o processo pelo qual analistas de assuntos tentam descrever o conteúdo intelectual de um documento para propósitos de recuperação – tem uma longa história na prática e teoria de Ciência da Informação. À primeira vista, sua importância pode parecer confinada à área da teoria da classificação: de qualquer forma, como muitas atividades em ciência da informação, uma atividade análoga e muito importante é desenvolvida fora dessa comunidade profissional.¹⁸

Para Campbell, a teoria literária e a análise de assunto, embora muito diferentes, podem se beneficiar mutuamente.

Esse ponto de vista é compartilhado também por outros autores, em âmbito nacional um dos mais recentes é o trabalho de Moraes e Guimarães (2006), desenvolvidos no âmbito do Grupo de Pesquisa de Análise Documental da UNESP, no qual os autores utilizam recortes teóricos da Linguística Textual e Análise do Discurso para a investigação temática em obras literárias narrativas. O trabalho dos autores, baseia-se no percurso gerativo de sentido, conforme formulado por Fiorin¹⁹.

Aprofundando o problema em relação à análise do conteúdo textual, Hutchins (1977) questiona como seria se estivéssemos também preocupados com as verdades ou inverdades apresentadas pelo autor de uma obra.

Nosso conhecimento dos sentidos das expressões não seria suficiente, um exemplo apresentado por ele, emprestado de Bertrand Russel é o enunciado “O Rei da França é careca”. Sem um contexto referencial não podemos dizer se essa é uma sentença falsa ou verdadeira, no entanto, se é tomado como referência o século XX, a afirmação é falsa, pois não há um rei da França.

Por outro lado, se for aplicado ao século XVIII teremos uma referência a partir da qual poderemos verificar a verdade da afirmação, ou seja, se o rei francês em um dado momento do século XVIII era ou não careca.

Isto é importante para compreendermos que não podemos averiguar a veracidade de um conteúdo por si mesmo, mas apenas testando suas situações e

¹⁸ Trecho original: *Aboutness analysis - the process whereby subject analysts attempt to describe the intellectual content of documents for purposes of retrieval—has a long history in information science theory and practice. At first glance, its importance may seem confined to the area of classification theory: however, like many activities in information science, an analogous and highly important activity is developing outside the information community.*

¹⁹ FIORIN, J. L. **Elementos de análise do discurso**. 7. ed. São Paulo: Contexto, 1999a. 93 p. (Repensando a língua portuguesa).

referências apropriadas. Mas para Hutchins (1977, p. 17):

Indexadores não estão preocupados com o valor-de-verdade dos documentos, nem com as imagens particulares nas mentes dos autores. Eles estão preocupados somente com o sentido dos textos – isto é o que se entende por conteúdo dos documentos.²⁰

A partir disto, Hutchins reafirma a idéia de que o documento possui um sentido, um significado próprio independente do seu autor, e nosso papel enquanto leitores é então, adicionar sentidos.

Dessa forma, o documento atinge uma existência autônoma, parte do que Karl Popper chamou ‘Mundo 3’, o mundo do conhecimento objetivo (POPPER, 1972 citado por HUTCHINS 1977) e:

Como tal, tem propriedades e relacionamentos que seu autor pode não conhecer – e talvez até discordar. O que isto significa em termos práticos é que o indexador deveria não se preocupar com - talvez nunca devesse ter se preocupado com – o que o próprio autor pensa do tópico de seu texto e suas relações com outros textos.²¹

Ainda assim, Hutchins considera que o texto possua um sentido ou significado que possa ser ‘descoberto’ em um estado ‘puro’, ‘abstrato’, pois cada leitor fará a sua própria leitura do documento, e dessa forma, a tarefa do indexador é conseguir uma visão tão abrangente quanto possível do que outras pessoas poderiam encontrar no texto.

Para Ingwersen (1992) a distinção do *aboutness* é ainda mais ampla e dessa forma para o autor o *aboutness* deve ser trabalhado em quatro categorias:

- *aboutness* do autor (*author aboutness*): é o *aboutness* determinado pelo autor através da sua linguagem natural, direta;
- *aboutness* do indexador (*indexer aboutness*): esse implica na transformação da linguagem do autor em uma

²⁰ Trecho original: “Indexers are not concerned with the truth-value of documents, nor with the particular images in the minds of authors (cf. Fairthorne, 1961). They are concerned solely with the sense of texts – this is what is meant by the content of documents.”

²¹ Trecho original: “As such it has properties and relationships that its author may not know of – and perhaps might dispute. What this means in practical terms is that the indexer need not be concerned with – perhaps never should be concerned with – what the author himself think of as the topic of his text and its relationships to other texts.”

linguagem documentária aceita ou criada pelo indexador ou sistema do qual faz parte;

- *aboutness* do usuário (*user aboutness*): é formulado a partir da necessidade do usuário por novos conhecimentos, ou seja, deve criar pontos de contato entre o documento e o que já é conhecido pelo usuário, para a partir daí sugerir quais são as novas informações que podem ser de interesse do usuário;
- *aboutness* da solicitação (*request aboutness*): é o *aboutness* da requisição conforme formulada pelo usuário.

Estas distinções apontadas por Ingwersen implicam que para o autor há na verdade: **a)** um *aboutness* subjacente ao documento independente das intenções do autor, **b)** um *aboutness* ou objetivo do autor, independente da forma pela qual ele exteriorizou em um registro documental, **c)** um *aboutness* do usuário que possui uma necessidade, mas talvez não tenha os subsídios necessários para formulá-la de maneira adequada e **d)** um *aboutness* da requisição, ou seja, da forma pela qual o usuário ‘materializou’ a sua busca.

A partir dessa discussão proposta por Ingwersen verificamos alguns pontos adicionais, que amplificam a discussão ao admitir que estão em jogo outras questões, tais como:

a) Se o *aboutness* do documento é diferente do *aboutness* do autor, considera-se que o autor pode não ter conseguido expor exatamente o que pretendia, ou a sua obra tomou um significado diverso a partir da mudança de um contexto social ou específico de uma área;

b) Se o *aboutness* de um usuário é diferente do *aboutness* de sua formulação, significa que o usuário não pôde expressar de forma conveniente a sua necessidade de informação, talvez justamente por não saber o que está necessariamente procurando, o que traz à tona a idéia a cerca dos ‘estados de conhecimento’ (*states of knowledge*), que como afirma Hutchins (1977), quando alguém consulta um sistema em busca de uma informação particular, pode acabar

não formulando precisamente qual o conteúdo que o documento deve possuir:

Ele pode não especificar qual a 'nova' informação que deveria estar presente em um documento apropriado. Tudo que ele pode fazer é formular suas necessidades em termos do que ele já sabe, seu atual 'estado de conhecimento'. O que ele está procurando, em efeito, é um documento que inicia de um conhecimento básico dentro do qual ele pode fazer pontos de contato, um documento que pressupõe um estado de conhecimento em afinidade com o seu. (HUTCHINS, 1977, p.33).²²

Na área de engenharia de sistemas, o trabalho de Bruza, Song e Wong (1999) busca elaborar um quadro de propriedades fundamentais do *aboutness* que sejam independentes de qualquer modelo de RI, entre essas propriedades os autores propõem: reflexividade (*reflexivity*), transitividade (*transitivity*), assimetria (*asymmetry*), contenção-composição (*containment-composition*), absorção (*absorption*), contenção sem-conflito (*non-conflict containment*) e contenção-exclusão (*containment-preclusion*).

Também é discutido por esses autores o ***non-aboutness***, ou seja, a hipótese de que em uma situação na qual é necessário que haja um filtro de informação que será direcionada aos usuários, seria mais fácil excluir os documentos que não dizem respeito ao tema em questão (o que não exclui o trabalho do profissional de avaliar e determinar a que o documento se refere).

De qualquer maneira, para os autores a importância do estudo do *aboutness* pode trazer “significativas rupturas” na teoria de RI e colaborar com o desenvolvimento de sistemas mais eficientes.

Esse trabalho de Bruza et al. (2000 citado por Lancaster, 2003, p.14), focaliza a questão do *aboutness* de uma perspectiva lógica. Os autores tentam: “formalizar a relevância lógica mediante a formalização de propriedades do senso comum que descrevem a relação de atinência.”

Lancaster (2003) destaca que a proposta dos autores quanto à relação de não-atinência (*non-aboutness*): “[...] constitui realmente uma situação mais simples

²² Trecho original: He cannot specify what 'new' information should be conveyed in an appropriate document. All that he can do is to formulate his needs in terms of what he knows already, his present 'state of knowledge'. What he is looking for, in effect, is a document that starts from a knowledge base with which he can make points of contact, a document which presupposes a state of knowledge with some affinity to his own.

porque a grande maioria dos itens em qualquer base de dados evidentemente não guarda qualquer relação possível com qualquer consulta ou necessidade de informação (isto é, são naturalmente itens ‘não-atinentes’).”

Ainda assim, essas questões não esclarecem como o indexador poderá encontrar o *aboutness* do documento, e justamente em relação a isso a área tem buscado o auxílio de outras disciplinas.

Aboutness e assunto

Conforme afirmamos, a proximidade do *aboutness* com o conceito de assunto (*subject*) ou tema (*subject matter*) apresenta uma profusão que vai além da nomenclatura para a prática e assim, buscamos nesse momento apresentar algumas definições para o termo assunto. A discussão das definições que ora seguem, decorrem do fato de que muitas vezes a área tem tratado esses termos como sinônimos, por isso, consideramos relevante abordar ambos e tentar encontrar traços distintivos – em vez de simplesmente recusar ou acatar essa sinonímia.

Na concepção de Albrechtsen (1993), o termo assunto parece ter ficado “magicamente adormecido” na comunidade científica de Biblioteconomia e Ciência da Informação, mas foi “acordado” para constituir novamente uma área central de estudos. Ainda para a autora, por vezes o conceito “[...] entrou furtivamente no palco para breves performances disfarçado sob o termo ‘*aboutness*’[...]”²³ (1993, p. 219).

O termo assunto é caracterizado de modo amplo pelos lexicógrafos, na visão de Reitz (2004, p.690) o assunto pode ser definido como “qualquer um dos tópicos ou temas de um trabalho, explicitamente declarados no título ou texto ou implícito em sua mensagem”. Para Ferreira (1975, p.69), ele é a “matéria ou objeto do que se trata”, na mesma concepção de Houaiss (2001, p. 41) o define como “aquilo sobre o que se fala” ou “tema que se versa” (OLINTO, 2000, p.85). O assunto pode ser definido ainda como “uma unidade conceitual encontrada em, ou derivada

²³ Trecho original: “[...] the concept sometimes sneaked onto the stage for brief performances disguised under the term ‘*aboutness*’ [...]”

de, manuscrito ou material literário publicado” (HARROD’s..., 1996, p. 619).

Na literatura especializada, no entanto, as questões se ampliam e se faz pertinente o questionamento de Hjørland (2007) quanto ao propósito de se definir conceitualmente o que é o assunto, pois para o pesquisador é interessante partirmos de “[...] como uma compreensão teórica do termo *assunto* pode ser útil na decisão de princípios para a análise de assunto?”²⁴, a partir do que entendemos que o problema principal não está simplesmente na definição do termo, mas que tem sua importância, tendo em vista, como afirma Dias (2004, p. 149) que:

[...] quando alguns autores usam a expressão determinar *de que trata* um documento, estão evitando, propositadamente, o uso da palavra assunto nesse contexto. Isso não é sem razão. Muitos autores consideram o termo assunto inapropriado, para dizer o mínimo.

Para Todd (1992), os principais autores da área não conseguiram esclarecer esse conceito e em lugar disso “[...] apresentam uma considerável confusão terminológica” na qual, para **Cutter** é o *tema* ou *tópico* esteja ou não mencionado no título; para **Kaiser** são as coisas em geral, reais ou imaginárias, e as condições pertencentes a elas, podendo ser chamadas respectivamente de *concretos* e *processos*.

Ranganathan fala sobre o “pensamento contido em um documento”; **Coates** entende como a idéia global incorporada em uma unidade literária; **Vickery**, o *tema* no qual livros e artigos, ou suas partes, são escritos; um complexo agregado de aspectos específicos, uma combinação de termos elementares.

Ainda no levantamento realizado por Todd, **Borko** e **Bernier** definem assunto como o foco principal de um trabalho, o *tema central* para o qual os esforços e a atenção do autor foram direcionados; são aspectos do trabalho que contém idéias, explicações e interpretações.

Segundo Lancaster (2003) apesar da imprecisão das expressões e definições sobre o termo assunto, elas são aceitáveis e compreendidas pela maioria das pessoas. No entanto, essa preocupação em torno da conceituação do termo é algo recente, pois durante décadas os serviços informacionais foram realizados sem

²⁴ Trecho original: “How can a theoretical understanding of the term “subject” be helpful deciding principles of subject analysis?”

que houvesse por parte dos teóricos e profissionais questionamentos sobre esse problema. Para Lancaster (2003, p. 13) essa busca por definições de assunto envolve que:

Nas considerações anteriores não se fez qualquer tentativa para definir a expressão 'de que trata um documento': a expressão 'de que trata' era simplesmente um sinônimo para 'tem por assunto'. Ou seja, usou-se 'de que trata um documento' para designar o mesmo que 'os assuntos de um documento'. Estas expressões talvez não sejam muito precisas e não é fácil definir 'trata de' e 'tem por assunto'. Apesar disso, são expressões que soam aceitáveis para a maioria das pessoas, sendo por elas compreendidas.

Retornando ao trabalho de Hjørland (1992)²⁵: “De um ponto de vista ingênuo o conceito de “assunto” ou “tema”²⁶ não apresenta problema: é bastante óbvio o que é assunto”, mas basta um olhar um pouco mais atento para perceber que nem sempre é fácil definir um assunto, por exemplo, quando o título de um livro não corresponde de forma óbvia ao seu conteúdo, um caso muito comum na literatura romanesca, mas também presente na literatura científica.

Hjørland (1997) realiza um traçado histórico da utilização do conceito partindo da filosofia antiga e medieval onde o significado de *assunto* e *substância* era o mesmo, mas no uso moderno este significado se inverteu e o assunto também pode estar em uma mente individual, ou seja, pode ser interpretado na mente do sujeito, leitor etc. Na linguística o termo denota o portador de uma propriedade e o predicado a propriedade, ou seja, o assunto é aquele sobre o qual algo é dito.

É fato para Hjørland (1997) que o termo, a despeito de sua importância, foi bastante negligenciado, e apresenta algumas exceções como o trabalho de Metcalfe (1973 citado por Hjørland, 1997) que revisou o uso do conceito na Biblioteconomia e Ciência da Informação em quase um século, mas, no entanto “[...] sua conclusão é que o conceito é tão vago que é quase inútil”²⁷.

É também interessante que a concepção de Cutter para assunto é considerada por Hjørland (2007), provavelmente melhor que muitas definições que

²⁵ Trecho original: “From a naïve point of view the concept of ‘subject’ or ‘subject matter’ poses no problem: it is rather obvious what subjects are.” (HJØRLAND, p. 172, 1992).

²⁶ A tradução da expressão *subject matter*, como *tema*, encontra-se no Dicionário Oxford Escolar para estudantes brasileiros de inglês (2002).

²⁷ Trecho original: “[...] his conclusion is that the concept is so vague that it is almost useless.”

dominaram o século XX, incluindo a definição proposta pela norma internacional ISO²⁸.

Para Cutter, a definição de assunto depende de um processo social no qual seu significado é constituído, no entanto Miksa (citado por Hjørland, 2007) demonstra como as limitações da visão de Cutter estão relacionadas com as limitações de suposições filosóficas e apresenta por sua vez, um novo conceito de assunto que chamou de “a idéia moderna de assunto”, que novamente para Hjørland não alcançou uma determinação satisfatória do conceito.

A concepção de Ranganathan (1967 citado por HJØRLAND, 2007) foi até onde se sabe a primeira a oferecer uma definição explícita do conceito de assunto, determinando que o assunto é na verdade a fusão de idéias estruturadas significativas para um determinado campo específico.

Quanto a isso Metcalfe (1973 citado por HJØRLAND, 2007) é bastante crítico ao afirmar que esta definição de Ranganathan é “pseudo-ciência” e a considera influenciada pelo sistema de classificação criado por ele (*Colon Classification*). Hjørland considera essa visão de Metcalfe inaceitável, pois não acredita que Ranganathan tenha direcionado esta definição com intenção de favorecer seu próprio sistema.

Frohmann (1994 citado por HJØRLAND, 1997), ao prosseguir na discussão da concepção de Cutter também apresenta um ponto de interesse, no qual considera que para Cutter há uma interligação entre a mente, a sociedade e os sistemas de organização, estando interligados e se suportando e manifestando a mesma estrutura e construindo conexões discursivas entre mente, educação e sociedade, ou seja, os assuntos são construídos, nomeados e sustentados em conjunto.

Ao contrário, a CDD (Classificação Decimal de Dewey) de Melvil Dewey, enquanto um sistema fechado de classificação corta essas conexões. A partir da discussão de Frohmann, torna-se claro que apesar de entrar tardiamente em debate,

²⁸ "Subject: Anything whatsoever, regardless of whether it exists or has any other specific characteristics, about which anything whatsoever may be asserted by any means whatsoever." ISO 13250-1, retirado de: <http://www1.y12.doe.gov/capabilities/sgml/sc34/document/0446.htm#overview>

as diferentes concepções de assunto, trazem implicações que não são triviais.

Percebemos que, se nesses marcos tão importantes da área há essa divergência fundamental que ainda não foi equacionada, isto revela não apenas uma fragilidade em nossa formação conceitual, mas também a distância que pode haver nos mecanismos de criação e atuação social dos sistemas de informação por nós embasados e desenvolvidos.

Com relação a isso, Dias e Naves (2007, p. 67) afirmam que: “Talvez essa confusão terminológica existente em torno do termo *assunto* possa ser responsável pela complexidade que envolve a representação dos assuntos dos documentos.” Em nossa visão, a complexidade existente na identificação e representação do tema de um documento é que tem causado a multiplicação de nomenclaturas e definições que temos visto na literatura, ou seja, a partir de um processo complexo que é o da compreensão textual e a decisão sobre o conteúdo e como representá-lo, as nomenclaturas refletem na verdade a dificuldade subjacente ao problema e todas as implicações decorrentes de um objeto tão complexo sobre o qual variadas visões e objetivos são lançados.

Como vimos nessa discussão sobre a definição de assunto, não há um consenso a seu respeito, mas existe o reconhecimento histórico da área em uma necessidade e interesse pela sua definição.

A interdisciplinaridade no estudo do *aboutness*

Um método que possa proporcionar a determinação do *aboutness* de um documento depende na verdade de que o analista possua algumas habilidades que deverão atuar em conjunto no momento da análise.

Dessa forma, estamos frente a um problema que possui natureza interdisciplinar e que “[...] vinha sendo estudado há longo tempo por lógicos e linguistas, pois representa um aspecto do significado” (DIAS; NAVES, 2007, p.74).

Lancaster (2003) destaca que na busca em auxiliar a determinação temática no trabalho documental:

[...] Beghtol (1986) e Hutchins (1978) recorrem ambos à linguística do texto ao examinar esta questão; Maron (1977) adota um enfoque probabilístico, e Swift et al.(1978) são cautelosos ao salientar que a atenção na indexação talvez não coincida com a atenção que as pessoas que estão em busca de informações têm em mente. Wilson (1968) chega ao ponto de sugerir que a indexação de assuntos se defronta com problemas 'intratáveis', visto ser tão difícil decidir do que trata um documento.

Assim, vemos que na verdade esta é uma área de estudos, que abrange também outros campos. De qualquer forma, sua abordagem pela Ciência da Informação se torna dependente de sua compreensão em um âmbito maior através da intersecção de outras disciplinas, como relata Fujita (2004), a leitura realizada pelo indexador é na realidade uma leitura documentária que possui finalidade pragmática. Pois:

[...] a leitura do indexador está condicionada a determinados objetivos e ao contexto do sistema de informação. Torna-se, pois, difícil, dissociar esses aspectos da atividade de leitura.

A leitura para fins de indexação difere da leitura tratada de modo geral por possuir finalidades profissionais e pragmáticas. No entanto, os conhecimentos basilares necessários para uma boa compreensão de um texto são comuns a ambas.

Dessa forma, mesmo que seja uma análise realizada por uma leitura com finalidade específica, o analista está dependente de aspectos linguísticos, lógicos e cognitivos (FUJITA, 2004).

Quanto a essa interdisciplinaridade, a questão levantada por Dias e Naves (2007, p.79) é bastante interessante:

Diante das quatro áreas, ciência da informação, linguística, ciência cognitiva e lógica, está claro que a maior interessada na interseção é mesmo a ciência da Informação, disciplina nova, ainda sem um corpo teórico consolidado que possa subsidiar qualquer tentativa de resolver problemas que, sabe-se não são poucos. Haveria interesse, das três outras áreas, em despendar esforços no desenvolvimento dessas pesquisas interdisciplinares? Questiona-se, ainda, qual seria a receptividade de linguistas, psicólogos e filósofos em trabalhar com profissionais da informação?

Independente das possíveis respostas que poderiam ser dadas por esses profissionais, as pesquisas interdisciplinares são realidade na Ciência da

Informação, e já tem gerado propostas de trabalho e mesmo novas subáreas, como a Linguística Documental.

A Linguística Documental

Alguns conceitos da Linguística tem sido abordados a fim de auxiliar na compreensão dos problemas linguísticos presentes no trabalho documental. Conforme apresentado por Mendonça (2000) em estudo sobre a relação entre a Linguística e a Ciência da Informação: “[...] a linguística abre todo um leque de questões no desenvolvimento da linguagem e da tecnologia, que estão cada vez mais presentes no campo documental.”

Fujita (2004, p.151) também afirma essa necessidade e relembra que:

[...] os estudos de interface entre linguística e documentação começaram a surgir em fins dos anos 1960, com a implantação dos computadores e nos trabalhos documentários, indicando que estavam ligados às análises das entidades que se manipulavam, como: as palavras, frases, resumos, descritores.

Para Moreiro Gonzáles (2005) a Ciência da Informação e a Linguística se relacionam tanto pela perspectiva comunicacional em que ambas se enquadram como pela contribuição comum a compreensão dos processos de transferência da informação e a partir da década de 1960 com o surgimento da Linguística Textual que coincidiu com as aplicações linguísticas na Documentação, ocorreu o convencimento de que só havia uma possibilidade de estudo do texto: considerar sua unidade máxima, ou seja, o texto como um todo.

Dessa forma, com o objetivo de automação dos sistemas e de apropriação dos métodos linguísticos, se tornou objeto próprio da Documentação entender os mecanismos lógico-semânticos inerentes ao texto, para identificar inclusive sua carga informativa.

Ainda para Moreiro Gonzales (2005) os documentos enquanto objetos de comunicação apresentam estrutura característica, terminologia própria e construção

particular das frases, e dessa forma resulta impossível seu processamento nem por humanos nem por computadores, excluindo as considerações semânticas.

O trabalho do Grupo TEMMA também tem se mostrado profícuo no sentido de realizar abordagens e experimentações com base nos recortes linguísticos, conforme demonstrado no trabalho de Cunha (1987, p. 50) que ressalta não apenas a importância da Linguística, mas também a liberdade que a Ciência da Informação possui quanto a adotar esses recortes:

[...] é necessário ter em conta que esta utiliza a linguística enquanto universo teórico capaz de subsidiar a sua atividade pragmática. Neste sentido, não lhe interessam as propostas das diferentes correntes linguísticas enquanto escolas de pensamento, com as suas terminologias, modelos e sistemas fechados, mas sim, os elementos e conceitos possíveis de serem utilizados no seu fazer pragmático. Esta questão leva-nos, por um lado, a uma dificuldade de precisão terminológica e conceitual dentro da Análise Documentária quando utiliza o jargão das diferentes correntes linguísticas utilizadas, por outro, confere à Análise Documentária uma mobilidade de procedimentos, permitindo-lhe uma abordagem teórica mais rica e globalizante do objeto em análise.

Kobashi (1994) também reafirma essa questão e aborda o pesquisador francês Gardin, que tem importância fundamental no desenvolvimento da análise documental e na adoção dos procedimentos linguísticos para a solução de problemas ligados à documentação. Gardin cunhou o termo Análise Documentária e a definiu como: “um conjunto de procedimentos efetuados com a finalidade de expressar o conteúdo de documentos científicos, sob formas destinadas a facilitar a recuperação da informação”(1981, p.29).

A Linguística Textual

Retomando a questão da Linguística Textual a partir do trabalho de Hutchins (1977) alguns conceitos são agora apresentados, os quais na visão de Hutchins devem ser compreendidos a fim de resolver os problemas pertinentes ao trabalho de condensação.

Os conceitos apresentados que podem ser de utilidade para a análise de conteúdo, se baseiam principalmente na concepção de tema e rema.

Tema, Rema e Progressão Temática

O dinamismo da comunicação possui mecanismos que alteram a estrutura das sentenças, dessa forma é necessário conhecer seus componentes básicos para compreender como elas funcionam. As idéias de tema e rema são grosso modo: **tema**, sobre o que se está falando em uma dada sentença e **rema** é o que se vai acrescentar de informação a esse elemento, conforme Koch (1997) esses conceitos foram postulados pela Escola Funcionalista de Praga (Danes, Firbas, Sgall), ou seja, “do ponto de vista funcional cada enunciado divide-se em (pelo menos) duas partes, tema e rema”, onde: “[...] tem-se um segmento comunicativamente estático – o tema – oposto a outro segmento comunicativamente dinâmico – o rema, núcleo ou comentário”.

De outro modo, o tema representa os elementos que estão relacionados de alguma forma ao texto precedente ou a características do discurso no qual se ambienta. O rema expressa a informação que é em algum sentido ‘nova’ para o ouvinte ou leitor ou não ‘visível’ a partir do que já foi dito ou escrito e dessa forma, o tema fornece um ponto de partida para o que o emissor da mensagem realmente deseja dizer.

Muitas vezes isso também pode ser realizado por anáforas na sentença através de pronomes e artigos definidos, por exemplo ao citar ‘o homem’ ou ‘ele’ pode estar se referindo a alguém já citado em uma passagem anterior ou no início da mesma. Em um enunciado típico os elementos do tema serão precedidos pelo rema, apresentamos o exemplo de demonstração de Hutchins (1977, p. 19)²⁹:

- (a) O homem de York vendeu um carro para um doutor
- (b) O doutor comprou um carro de um homem de York

²⁹ (a) The man from York sold a car to a doctor
 (b) The doctor was sold a car by a man from York
 (c) The car was sold to a doctor by a man from York

- (c) O carro foi vendido para um doutor por um homem de York

As três sentenças representam na verdade a mesma operação envolvendo os mesmos participantes, no entanto, na primeira somente o 'homem de York' está determinado, na segunda somente o 'doutor' e na terceira somente o 'carro'.

Dessa forma, ao alterarmos a posição dos elementos na frase, alteramos também o seu tema.

Outro exemplo elucidativo é apresentado por Ventura e Lima-Lopes (2002):

É interessante notar que, se mudarmos os elementos que ocupam a posição temática, mudamos também o significado da mensagem: *Junto com o 1o. vídeo, você já recebe grátis a fita Velocidades fantásticas* é diferente de *Você já recebe grátis a fita Velocidades fantásticas junto com o 1o. vídeo*, que é diferente de *A fita Velocidades fantásticas você já recebe grátis junto com o 1o. vídeo*. Na primeira oração, a mensagem é sobre *quando* você recebe a fita; na segunda, sobre *você*; e na terceira, sobre a fita *Velocidades fantásticas*.

Hutchins ainda esclarece que obviamente apenas os conceitos de tema e rema não esclarecem essas questões, existem outras categorias como Agente, Paciente, Destinatário etc., além de outras configurações e regras para esses elementos.

Seguindo a idéia de articulação entre tema-rema, temos também a idéia de 'progressão temática' que seguindo Danes (1974 citado por Hutchins, 1977) pode ser uma 'progressão linear', onde o elemento temático está relacionado a um rema precedente e a 'progressão paralela', no qual o tema permanece o mesmo, exemplificado nas ilustrações abaixo:

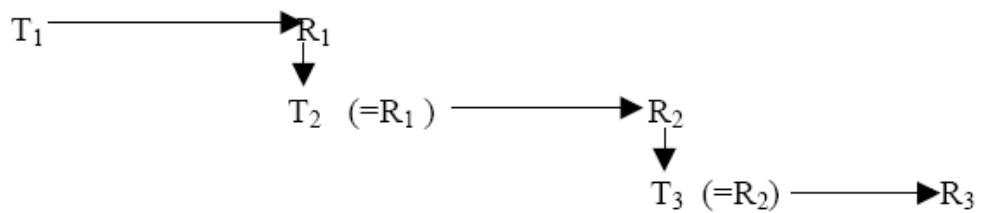


Figura 1: Esquema de progressão temática linear.

Em um exemplo para demonstrar a sequência: “O garoto estava lendo um livro. Era sobre tatus. Eles são encontrados na América do Sul.”³⁰

Cada período possui um tema próprio, mas que está ligado ao rema do período anterior, sem o qual não faria muito sentido.

Na ilustração seguinte, vemos a outra forma de progressão:

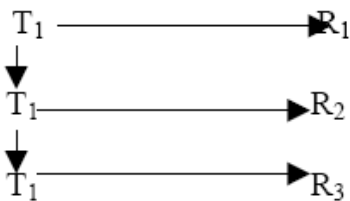


Figura 2: Esquema de progressão temática paralela

Nesse segundo esquema, os diferentes períodos fazem referência a um mesmo tema, e provavelmente utiliza-se de anáforas, como no exemplo seguinte: “O garoto estava lendo um livro. Ele o ganhou em seu aniversário. Ele fez dez anos de idade.”

Essas são sentenças simples, facilmente delimitáveis em relação à sua forma de progressão, no entanto, conforme Hutchins a maioria dos enunciados apresenta uma combinação dos dois tipos, ilustrado na próxima imagem:

³⁰ Trecho original: *The boy was reading a book. It was about armadillos. They are found in South America.*

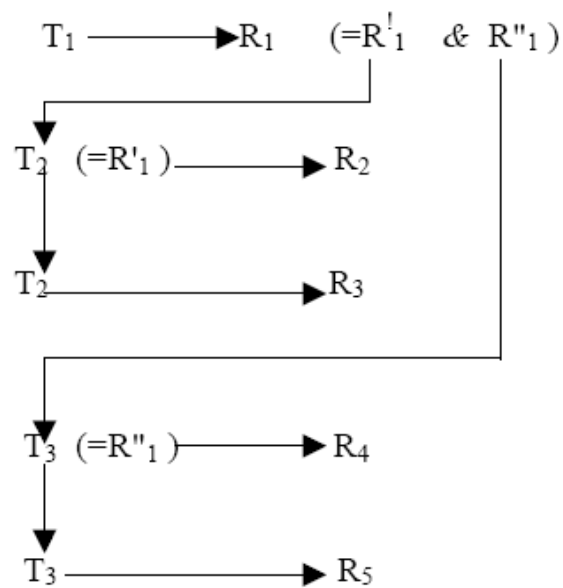


Figura 3: Esquema de progressão temática combinada

Um exemplo para esse esquema: “Todas as substâncias podem ser divididas em duas classes: substâncias elementares e compostas. Substância elementar é uma substância a qual... (etc.). Um composto é uma substância a qual... (etc.)”³¹

Essa abordagem também recebe suas críticas, para Weinberg (citado por JOUDREY, 2005), essa distinção além de ser difícil de ser mantida pelos indexadores é muito pouco útil em um contexto real.

Esses pontos foram até aqui abordados em suas linhas gerais, pois sabe-se que na prática cotidiana essas questões são obviamente mais complexas (HUTCHINS, 1977), principalmente em consideração ao volume de materiais e a complexidade dos conteúdos, e como ressaltado por Koch (1997) a construção dos sentidos depende das escolhas que o leitor realiza, e a articulação tema-rema constitui um dos “leques de escolhas significativas” a serem realizadas pelo leitor.

³¹ Trecho original: All substances can be divided into two classes: elementary substances and compounds. An elementary substance is a substance which... (etc). A compound is a substance which... (etc.).

A abordagem dos aspectos lógicos e cognitivos

Entre os autores elencados ao longo do trabalho, alguns citaram a importância da interlocução com os aportes da lógica e das ciências cognitivas na análise textual.

Dado o caráter interdisciplinar que engloba a questão da análise temática, o viés pesquisado quanto à compreensão cognitiva do processo de leitura e compreensão, é uma lacuna que diz respeito não somente ao processo de compreensão e assimilação da informação e de conteúdos textuais, mas relativo ao próprio entendimento da cognição humana, campo no qual ainda há muitas questões em aberto.

Como destacam Dias e Naves (2007):

“Acredita-se que, junto aos aspectos linguísticos, os aspectos cognitivos e os lógicos formam o tripé de sustentação do processo de análise de assunto. Eles interferem e interagem simultaneamente, não havendo possibilidade de delimitar em que ponto começa ou acaba a influência de cada um dos três.”

Como vimos no decorrer da pesquisa, o processo de análise e determinação do *aboutness* está diretamente ligado às habilidades cognitivas do leitor, onde é a partir de sua interação com o texto, que o *aboutness* passa a ser determinado.

Os autores Cavalcanti³² e Giasson³³ (FUJITA, 1998) entendem a leitura como processo **comunicativo** entre leitor e texto, onde o leitor não se enquadra enquanto um sujeito neutro, mas sim um agente participativo no processo, onde o sentido é “negociado” (FUJITA, 2004) levando em consideração suas experiências anteriores, seus preconceitos e valores e assim: “O sentido não é mais um dado que preexiste ao texto” (LUCAS, p.40, 2000) passa a ser construído. Assim, estes autores enxergam a existência de três fatores no processo da leitura: o **leitor**, o **texto** e o **contexto**.

Dessa forma, compreendemos que a abordagem desse aspecto da

³² CAVALCANTI, M.C. **I-N-T-E-R-A-Ç-Ã-O leitor-texto: aspectos de interação pragmática**. Campinas: UNICAMP, 1989.

³³ GIASSON, J. **A compreensão na leitura**. Lisboa: Asa, 1993.

determinação temática, engloba fatores ainda desconhecidos ou claramente compreendidos. Nisto se enquadram os modelos de leitura de Van Dijk, conforme apresentados por Todd (1992, p. 103), Van Dijk sugere que as ações cognitivas de leitura e compreensão são governadas por macroregras ou macroestratégias ou operações, postuladas em cinco regras da análise de assunto. Apresentamos as cinco regras através da tradução de Fujita (2004, p.150):

- 1) **Regra de deleção fraca:** supressão de informação acidental, isto é, detalhes que não mudam o significado ou influenciam a interpretação das sentenças seguintes.
- 2) **Regra de deleção forte:** supressão de informação importante localmente: a informação que é suprimida especifica associações normais ou esperadas.
- 3) **Regra de deleção zero:** nenhuma redução ocorre. Todas as informações são consideradas relevantes, e são admitidas diretamente na macroestrutura.
- 4) **Regra de Generalização:** referência a diversos objetos ou propriedades da mesma classe superordenada, de forma global, pelo nome da classe superordenada.
- 5) **Regra de construção:** a combinação ou integração de informação que denota propriedades, causas, componentes, consequências, etc. de um fato de nível superior.

Van Dijk (1992) propõe que a identificação da macroestrutura se constitui na explicação teórica do que geralmente chamamos de enredo, trama ou tópico de um texto.

O conceito de Van Dijk (citado por BEGHTOL, 1986) quanto às superestruturas e macroestruturas oferecem uma formulação relativamente específica de como as pessoas, incluindo classificadores, formulam mentalmente o *aboutness* de documentos textuais.

Para Beghtol (1986) os conceitos propostos por Van Dijk podem ser bastante aproveitados na investigação do processo cognitivo de classificação documental. Para a autora, poderiam auxiliar na descrição e explicação dos 1) relacionamentos entre os documentos classificados, 2) relacionamentos entre sistemas de classificação e documentos classificados e 3) o processo cognitivo de classificação pode se apoiar nos conceitos de processamento *top-down* e *bottom-up*, apresentados na teoria de identificação do *aboutness* proposta por Van Dijk.

Para Dias e Naves (2007, p.89): “Além dos aspectos cognitivos, o uso da

linguagem cria vários problemas do ponto de vista lógico, como, por exemplo, a homonímia (várias coisas com o mesmo nome), a sinonímia (vários nomes para coisa), e recorre-se à lógica para resolver esses tipos de obstáculos.”

Para Fujita (2004, p. 152): “No momento da leitura do documento para indexação, o indexador realiza esforços mentais classificados dentro de uma das divisões da “Ciência Normativa Lógica”, que por sua vez pertence a uma ciência maior, a Filosofia.”

Encerrando essas considerações quanto aos aportes interdisciplinares: “A opinião de Cunha (1987) é de que a lógica, enquanto ciência, encontra-se bem distante das preocupações dos bibliotecários/analistas da documentação, mesmo quando se utiliza da automação.” (DIAS, NAVES, 2007).

Na seção seguinte, realizamos a análise da discussão nas revistas brasileiras selecionadas.

III PROCEDIMENTOS METODOLÓGICOS

Nesta seção descrevemos os procedimentos metodológicos utilizados para o desenvolvimento desse estudo, que teve como objetivo geral analisar na literatura da área a discussão e conceituação relativa ao termo *aboutness* e sua inserção na literatura brasileira, de modo a contribuir para a sedimentação teórica da área.

Dessa maneira, a metodologia aplicada foi de natureza bibliográfica exploratória, possuindo caráter investigativo, no qual teve por base o levantamento bibliográfico da literatura especializada, buscando estudos que abordassem o conceito “*aboutness*” e a definição temática de documentos.

No segundo momento da pesquisa realizamos a coleta de dados a partir dos periódicos nacionais com o objetivo de verificar a inserção do conceito *aboutness* em âmbito nacional, para tanto se levou em consideração nessa etapa as variações de tradução do termo para português.

Corpus da pesquisa

Delimitamos como *corpus* de pesquisa quatro dos principais periódicos brasileiros especializados em Ciência da Informação, a partir do seu destaque e reconhecimento no cenário brasileiro, além das suas respectivas colocações conforme o conceito Qualis, atribuído pela CAPES³⁴. Buscando, deste modo, analisar e compreender a inserção do tema “*aboutness*” na discussão na comunidade brasileira de Ciência da Informação.

Para tanto, definimos os seguintes periódicos nacionais³⁵ para a

³⁴ O conceito Qualis sofreu alterações de critérios e níveis de conceituação em abril de 2008, período no qual essa pesquisa já estava em andamento. Dessa forma, mantivemos a seleção inicial de periódicos, na qual, todos os títulos selecionados, recebiam a atribuição de conceito A. Mais informações sobre o Qualis e alterações estão disponíveis em < <http://www.capes.gov.br> >.

³⁵ Todos estes periódicos estão disponíveis para acesso público nos respectivos endereços

constituição do *corpus* da pesquisa:

- Revista Ciência da Informação;
- Perspectivas em Ciência da Informação;
- Revista Transinformação;
- DataGramaZero.

Revista Ciência da Informação (RCI)

A Revista Ciência da Informação (ISSN 1518-8353) foi criada em 1972, e possui periodicidade é quadrimestral. Publicada pelo IBICT (Instituto Brasileiro de Informação em Ciência e Tecnologia) tem por escopo a publicação de artigos inéditos sobre a Ciência da Informação a partir dos fenômenos ligados à produção, organização, difusão e utilização de informações em todos os campos do saber, ciência e tecnologia. Em 2008 recebeu Qualis B1 na área de Ciências Sociais Aplicadas I.

A RCI utiliza o Sistema Eletrônico de Editoração de Revistas (SEER³⁶) e fornece acesso gratuito pelo endereço eletrônico: <http://www.ibict.br/cienciadainformacao> e é indexada por bases de dados nacionais e internacionais. Sua publicação está licenciada sob *Creative Commons 3.0*.

Revista Perspectivas em Ciência da Informação (PCI)

A revista Perspectivas em Ciência da Informação (ISSN 1413-9936) é publicada pela Escola de Ciência da Informação da Universidade Federal de Minas Gerais desde 1996 e surgiu em substituição a extinta Revista da Escola de

eletrônicos.

³⁶ Mais informações sobre o sistema estão disponíveis na página eletrônica do IBICT (Instituto Brasileiro de Informação, Ciência e Tecnologia) < <http://seer.ibict.br/> >.

Biblioteconomia da UFMG. Sua periodicidade é quadrimestral, e utiliza o Sistema Eletrônico de Editoração de Revistas com acesso gratuito pelo endereço eletrônico: <http://www.eci.ufmg.br/pcionline/>. Possui Qualis B1 na área de Ciências Sociais Aplicadas I.

Revista Transinformação (TR)

A revista Transinformação (ISSN 0103-3786) é publicada pela Pontifícia Universidade Católica de Campinas (SP) desde 1989, sua periodicidade é quadrimestral. Tem por escopo publicar trabalhos inéditos que contribuam para o estudo e o desenvolvimento científico nas áreas da Ciência da Informação. Em 2008 recebeu Qualis B2 na área de Ciências Sociais Aplicadas I.

Desde 2002 a TR tem acesso gratuito a partir do Sistema Eletrônico de Editoração de Revistas, no endereço eletrônico: <http://revistas.puc-campinas.edu.br/transinfo/>. As publicações anteriores a essa data podem ser acessadas através da biblioteca digital CVA-RICESU no endereço <http://biblioteca.ricesu.com.br/>, no qual foram digitalizadas.

Revista Datagramazero (DGZ)

A revista Datagramazero (ISSN 1517-3801) é publicada pelo IASI³⁷ (Instituto de Adaptação e Inserção na Sociedade da Informação), desde 1999, sua periodicidade é bimestral. Em 2008 recebeu Qualis B3 na área de Ciências Sociais Aplicadas I.

Tem por escopo divulgar e promover estudos em áreas interdisciplinares da Ciência da Informação, tais como Informação e Sociedade, Informação e Políticas

³⁷ O IASI é uma organização não-governamental, sem qualquer vinculação político-partidária ou religiosa, estabelecida em novembro de 1998 e dedicada a estudos e pesquisas sobre Sociedade da Informação para promover maior inclusão digital < <http://www.e-iasi.org/> >.

Públicas, Informação e Filosofia ou Informação e Comunicação. Seu acesso é gratuito e disponibilizado pelo endereço eletrônico: <<http://www.datagramazero.org.br>>.

Período cronológico abordado

Para seleção dos artigos a serem analisados, delimitamos o período integral (análise da coleção completa) de publicação dessas revistas, com o propósito de abranger o período mais amplo possível, verificando também a possibilidade de delimitação cronológica da discussão.

Por isso, foram analisados 37 anos de publicação da Revista Ciência da Informação, referente aos anos de 1972 a 2009; 13 anos da revista Perspectivas em Ciência da Informação referente aos anos de 1996 a 2009; 20 anos da Revista Transinformação relativo aos anos de 1989 a 2009; e 10 anos da revista DataGramaZero referente aos anos de 1999 a 2009.

Termos definidos para busca

Para coleta dos artigos nos periódicos selecionados delimitamos um domínio conceitual composto pelos seguintes termos:

- *Aboutness*
- Atinência
- Sobrecidade
- Tematicidade

Esse domínio conceitual foi selecionado a partir do seu uso corrente na literatura nacional e por serem as variações de tradução identificadas do termo *aboutness*.

Assim, esses termos foram utilizados como objetos de busca nos artigos. Destaca-se, que durante esse momento de busca foram excluídas da análise as seções: introdução, instruções para autores, editoriais, anúncios, resenhas, resumos de teses e dissertações e resenhas, por não possuírem a estrutura formal de artigos científicos.

Busca dos termos e procedimentos de seleção dos artigos

Os procedimentos de busca e seleção dos artigos se sucederam da seguinte maneira a partir de dois momentos: a) busca pelo SEER e b) busca manual.

Nesse primeiro momento a busca por termos ocorreu nos respectivos periódicos a partir dos seus próprios mecanismos de busca interna, no qual utilizamos os mecanismos de busca disponibilizados por cada periódico, sendo que dos quatro periódicos selecionados, três utilizam o SEER disponibilizado pelo IBICT.

O sistema SEER possui ferramenta de busca interna, permitindo a busca nos campos por Autor, Resumo, Título, Termos indexados e Texto completo. Ressalta-se que esse sistema têm se tornado o padrão para publicação de periódicos eletrônicos no Brasil. Todavia, percebemos que ao buscar o artigo por texto completo o sistema apresentava falhas³⁸ na recuperação, por isso optou-se em fazer a transferência de todos os arquivos e realizar a busca manual.

Nesse segundo momento para essa busca manual utilizamos o sistema de busca oferecido pelo *software Adobe Reader* (versão 9.0) que permite a pesquisa por expressões exatas ou por termos no texto completo. Esse sistema possibilitou maior precisão e exatidão nas buscas dos artigos.

Destaca-se que a estratégia adotada de busca por termos pelo sistema SEER ocorreu somente nas revistas RCI, PCI e TR. Quanto à revista Datagramazero, que não é publicada por esse sistema, utilizamos o sistema de

³⁸ Quanto às falhas mencionadas, ocorreu que artigos já conhecidos de levantamentos anteriores não foram apresentados nos resultados das buscas, através do sistema.

busca do próprio navegador (*Mozilla Firefox* versão 3.5), que também atendeu adequadamente às necessidades de busca.

Todavia essa forma de busca também analisa as referências do documento, dessa forma, realizou-se também a análise integral do documento para verificar se o termo buscado estava presente somente nas referências e quando isso ocorreu o artigo foi desconsiderado (quando o restante do seu conteúdo não apresentava menção aos termos desejados).

Análise dos resultados

A partir das estratégias de buscas adotadas foram analisados um total de 83 volumes, 229 números e 1708 artigos, como descritos na *Tab. 1* apresentada a seguir:

PERIÓDICOS NACIONAIS ANALISADOS				
Periódicos	Volumes	Números	Artigos	Artigos selecionados
<i>Revista Ciência da informação</i>	38	90	719	06
<i>Perspectiva em ciência da informação</i>	14	32	274	03
<i>Revista Transinformação</i>	21	48	361	01
<i>Datagramazero</i>	10	59	354	03
TOTAL	83	229	1708	13

Tabela 1: Periódicos nacionais analisados.

Desse total de 1708 artigos analisados identificamos apenas 13 artigos que apresentam em seu contexto menção aos termos da busca, isso corresponde

aproximadamente a 0,175% o que demonstra a carência de literatura nacional a respeito dessa temática relevante para área de representação da informação. Podemos inferir também que o termo assunto, embora com imprecisões quanto a sua definição, seja ainda o mais usual nos estudos de análise temática.

Desses 13 artigos identificados 46% foram publicados pela Revista Ciência da Informação, como mostra o percentual no Gráfico 1 a seguir:

Artigos selecionados



Gráfico 1: Total de artigos selecionados por percentagem de revistas.

Ao observarmos o percentual relativo à publicação a partir de cada periódico também podemos verificar que essa abordagem temática é bastante inexpressiva no escopo da coleção. Embora sejam números aproximados apenas 0,834% refere-se ao total de artigos publicados, número bem insignificante, como mostra o Gráfico 2 a seguir:

Total de artigos analisados e selecionados

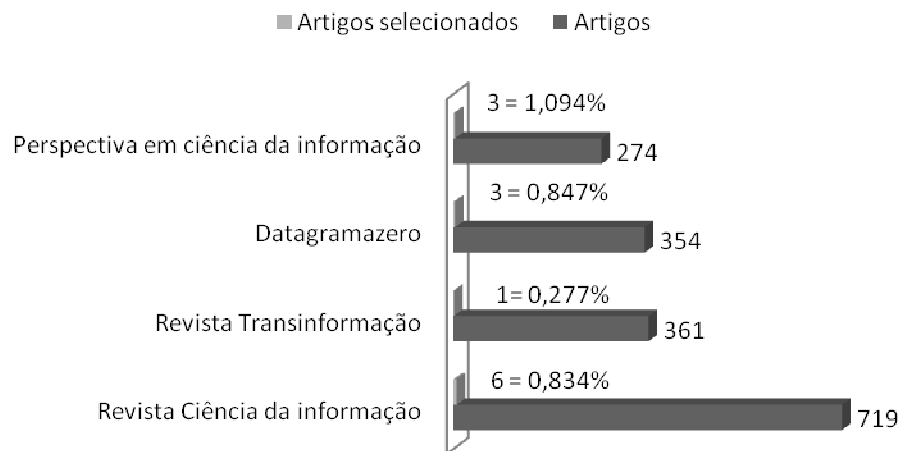


Gráfico 2: Percentual de artigos identificados a partir de cada periódico.

Como já mencionado esses 13 artigos selecionados foram identificados a partir dos termos *aboutness*, atinência, tematicidade e sobrecidade.

Nos 06 artigos recuperados na Revista Ciência da Informação foi possível identificar todos os termos de buscas, todavia observa-se a predominância dos termo *aboutness*, já que 5 artigos foram recuperados a partir desse termo. No total destes 6 artigos eles foram recuperados a partir dos temos:

- 1) 1 apenas pelo termo *aboutness*;
- 2) 1 por *aboutness* e sobrecidade;
- 3) 1 por *aboutness* e atinência;
- 4) 1 por *aboutness* e tematicidade;
- 5) 1por *aboutness*, tematicidade e sobrecidade;
- 6) 1 apenas pelo termo atinência.

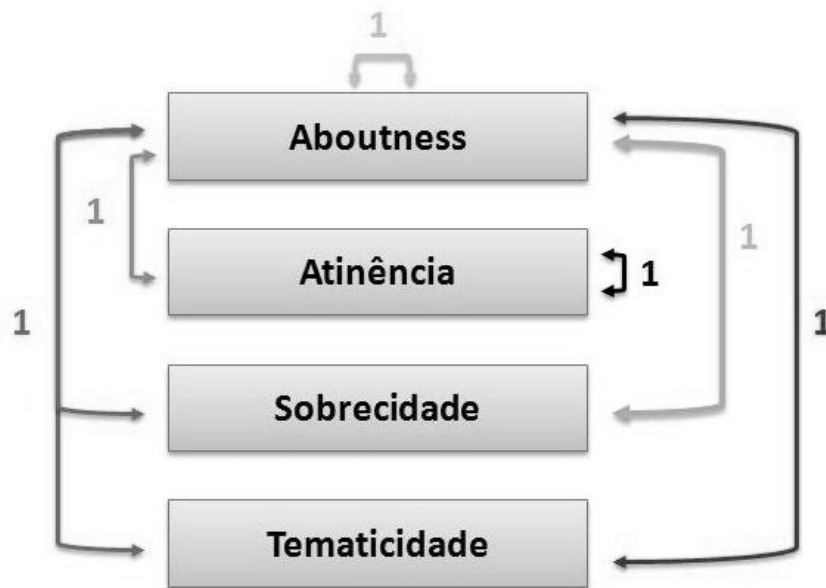


Figura 4: Relação dos termos de pesquisas identificados nos artigos na RCI

A Figura 4 ilustra bem a relação entre os termos. Essas seis referências dos artigos localizados são apresentadas a seguir pela ordem de data de publicação:

- WANDERLEY, Manoel Adolpho. Linguagem Documentária: Acesso à Informação Aspectos do problema. *Ciência da Informação*, Rio de Janeiro, 2(2): 175-217, 1973
- BARANOW, Ulf Gregor. Perspectivas na contribuição da linguística e de áreas afins à ciência da informação. *Ciência da Informação*, Brasília, 12 (1): 23-35, 1983.
- MEDEIROS, Marisa Bräscher Basílio. Terminologia brasileira em ciência da informação: uma análise. *Ciência da Informação*, Brasília, 15(2): 135-42, jul./dez. 1986
- MARCONDES, Carlos Henrique. Representação e economia da informação. *Ciência da Informação*, Brasília, v. 30, n. 1, p. 61-70, jan./abr. 2001
- SOUZA, Renato Rocha; ALVARENGA, Lídia. A Web Semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 132-141, jan./abril 2004.
- CORDEIRO, Rosa Inês de Novais; AMÂNCIO, Tunico. Análise e representação de filmes em unidades de informação. *Ciência da Informação*, Brasília, v. 34, n. 1, p.89-94, jan./abr. 2005

Na revista *Transinformação* o único artigo encontrado foi recuperado pelos termos *aboutness*, *atinência* e *tematicidade*. Sendo a referência:

- FUJITA, Mariângela Spotti Lopes. A prática da indexação: análise da evolução de tendência teóricas e metodológicas. *Transinformação*, Campinas, v.16, n. 2, p.133-161, maio/ago.2004.

Na Datagramazero 2 artigos foram recuperados pelos termos *aboutness* e 1 artigo pelo termo *atinência*.



Figura 5: Relação dos termos de pesquisas identificados nos artigos na DGZ

As referências desses três artigos localizados são:

- ALVARENGA, Lígia. A Teoria do Conceito Revisitada em Conexão com Ontologias e metadados no Contexto das Bibliotecas Tradicionais e Digitais. *DataGramaZero*, Rio de Janeiro, v.2, n.6, dez. 2001.
- LARA, Marilda Lopes Ginez de. Novas relações entre Terminologia e Ciência da Informação na perspectiva de um conceito contemporâneo da informação. *DataGramaZero*, Rio de Janeiro, v.7, n.4, ago. 2006
- MARCONDES, Carlos H.; MENDONÇA, Marília A. R.; MALHEIROS, Luciana R. Uma proposta de modelo de representação do conhecimento contido no texto de artigos científicos publicados na web em formato legível por programas. *DataGramaZero*, Rio de Janeiro, v.7, n.5, out. 2006

Na revista *Perspectivas em Ciência da Informação* dos 3 artigos identificados, 2 artigos foram recuperados pelos termos *aboutness* e *atinência* e 1 artigo pelo termo *atinência*.



Figura 6: Relação dos termos de pesquisas identificados nos artigos na PCI.

As referências desses três artigos localizados são apresentadas em ordem de publicação:

- DIAS, Eduardo Wense; NAVES, Madalena Martins Lopes; MOURA, Maria Aparecida. O usuário-pesquisador e a análise de assunto. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 6, n. 2, p. 205 - 221, jul./dez. 2001
- NAVES, Madalena Martins Lopes. Estudo de fatores interferentes no processo de análise de assunto. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 6, n. 2, p. 189 - 203, jul./dez. 2001.
- DIAS, Eduardo Wense. Análise de assunto: percepção do usuário quanto ao conteúdo de documentos. *Perspectivas em Ciência da Informação*, Belo Horizonte, v.9 n.2, p. 146-157, jul./dez. 2004

A partir do levantamento desses periódicos foi possível verificar que a discussão em torno “*do que trata um documento*” em sintonia com os estudos internacionais já era abordada na década de 1970 com o estudo de Manoel Adolpho Wanderley, que também considerava aportes teóricos da Linguística e da Semântica para resolver problemas das linguagens documentais. Todavia, apesar da tentativa de discussão a partir desse autor, só após uma década que foram publicados novos artigos, dando maior destaque a problemática a partir da década de 2000.

Foi possível verificar também a partir do estudo que o termo mais utilizado pela literatura nacional é o *aboutness*, pois ele foi recuperado em dez dos treze artigos, seguido do termo *atênência*, que foi recuperado sete vezes, como mostra a *Figura 7*.

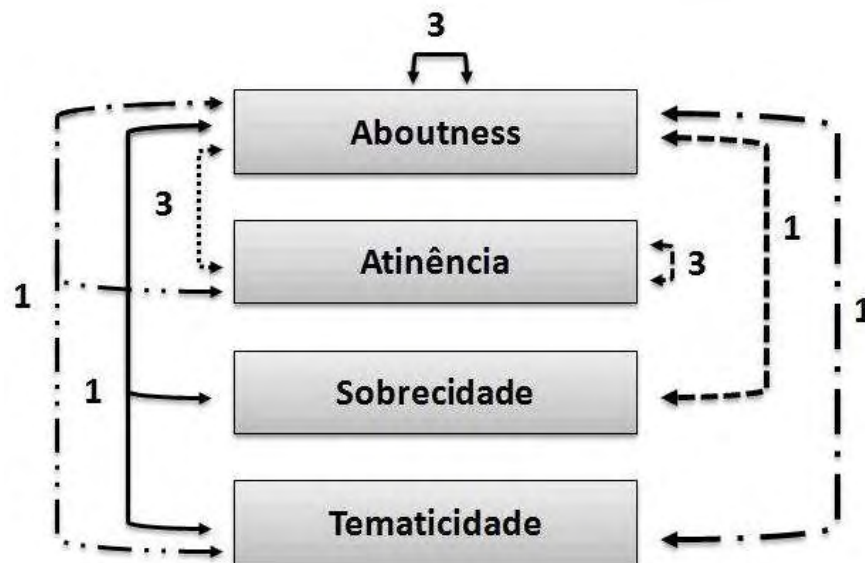


Figura 07: Relação dos termos de pesquisas identificados nos artigos.

Isso reflete o que já foi mencionado no referencial teórico devido à complexidade de se conceituar o termo assunto, muitos pesquisadores optaram em utilizar o termo *aboutness*. No Brasil isso também caracteriza a falta de consenso na tradução do termo para o português. Como é possível verificar no *Quadro 01* o termo *aboutness* é utilizado para especificar “o que se trata” ou “assunto” do documento.

BARANOW (1983)	<i>[...] Existe uma problemática análoga, muito familiar para o indexador: o primeiro passo na indexação é determinar do que trata um determinado documento. É o que poderíamos chamar de 'concernência' (do verbo concernir, adj. concernente), termo mais adequado ao vernáculo do que o esdrúxulo anglicismo 'sobrecidade' (aboutness) [...]</i>
MARCONDES (2001)	<i>[...] Todos estes componentes são interpretados pelo usuário e, assim, constituem um dispositivo de inferência ou uma tecnologia cognitiva conforme estabelecido por Pierre Lévy, que permite inferir o aboutness e a relevância do documento [...]</i>
(MARCONDES (2006)	<i>[...] o processo de análise/representação temática empreendida pela Ciência da Informação visa somente à sua recuperação em sistemas automatizados e à identificação superficial de seu conteúdo, mais especificamente sobre o quê é o texto, seu "aboutness"[3]; este "sobre o quê" é expresso nas linguagens de representação temática por descritores [...]</i>

Quadro 01: Comparativo de citações sobre a utilização do termo *aboutness*

Embora não haja consenso na literatura sobre sua definição e tradução, sendo o *aboutness* traduzido como; atinência (WANDERLEY, 1973; NAVES, 2001; ALVARENGA, 2001; DIAS 2004), concernência (BARANOW, 1983), sobrecidade ou tematicidade (MEDEIROS, 1986; FUJITA, 2004), o que se observa é a tendência a partir desses periódicos em utilizar o termo atinência para designar o *aboutness* do documento, como mostra as citações no Quadro 2:

WANDERLEY (1973)	<i>[...] Problema crucial, mas que transcende o plano estritamente linguístico é, nessa esfera da comunicação, o da "aboutness" a que se refere Fairthorne (34), ou seja, o da "atinência" dos documentos, o da natureza daquilo sobre que versam e a propósito de que os solicitam os usuários. [...]</i>
NAVES (2001)	<i>[...] O processo de análise de assunto, do ponto de vista do indexador, é iniciado com a fase de leitura do texto. Para isso, é necessário que se conheçam tipos e estruturas de textos para iniciar-se a sua leitura com fins específicos. Após essa leitura, passa-se à fase da extração de conceitos que possam representar o conteúdo temático do texto, para se chegar ao momento da fase de representação da atinência (aboutness), em que são definidos os termos em linguagem natural [...]</i>
DIAS (2004)	<i>[...] Devemos identificar o nosso interesse na vertente do que veio a ser chamado a partir de um determinado momento de atinência (aboutness, em inglês), ou seja, aquilo de que trata o conteúdo substantivo de uma obra (não sendo importante, a princípio, aspectos como a forma ou o suporte em que essa obra está registrada).[...]</i>

Quadro 02: Comparativo de citações sobre a utilização dos termos *aboutness* e atinência.

Como visto os dois termos são entendidos e tratados com o mesmo significado. Todavia em alguns estudos o termo atinência já não faz menção ao termo *aboutness* em inglês e ganha autonomia de significância na literatura da área, em relação ao “o que se representa” em um documento a partir do processo de análise temática, como destaca as citações no Quadro 3:

ALVARENGA (2001)	<i>[...] A parte substancial dos documentos que se refere a seu conteúdo, à sua atinência, ao seu significado, os enunciados que compõem os conceitos neles contidos, tudo isso continua invariável; [...]</i>
DIAS; NAVES; MOURA	<i>[...] FAIRTHORNE (1969) dedicou-se a estudar o assunto com rigor conceitual, chegando a uma distinção entre atinência extensional e atinência intensional. A primeira é o assunto inerente ao documento, ao passo que a segunda seria a razão ou o propósito da incorporação do</i>

(2001)	<i>documento ao acervo de um sistema de informação, ou de sua procura por um usuário. [...]</i>
SOUZA; ALVARENGA (2004)	<i>[...] A dificuldade de determinar os contextos informacionais tem como consequência a impossibilidade de se identificar de forma precisa a atinência dos documentos. [...]</i>

Quadro 3: Comparativo de citações da utilização do termo atinência.

Percebe-se que a utilização do termo sobrecidade não teve adesão na literatura, fato justificado pela inadequação em língua portuguesa (MEDEIROS, 1986), o que tem impulsionado o uso do termo tematicidade como sinônimo de *aboutness* (FUJITA, 2004; CORDEIRO; AMANCIO, 2005) pela sua adequação gramatical, garantindo seu respaldo no uso científico. Como mostra o Quadro 04 com as citações:

MEDEIROS (1986)	<i>[...] Assim, o termo sobrecidade constitui-se num neologismo por tradução de morfemas erroneamente combinados em vernáculo, uma vez que, de acordo com as regras gramaticais de derivação, as preposições em português (ex. sobre) não são passíveis de derivação lexical (ingl. aboutness). Já o neologismo tematicidade é mais apropriado para representar o conceito expresso pelo termo inglês aboutness, pois foi criado observando-se os padrões gramaticais da língua portuguesa (derivação do adjetivo temático com o sufixo — (d)ade). [...]</i>
CORDEIRO; AMANCIO (2005)	<i>[...] O objeto principal de estudo era a análise do conteúdo dos filmes. Para tal, realizaram-se pesquisas na literatura referentes aos procedimentos usados em unidades de informação, para ampliar a indexação e a recuperação da tematicidade (aboutness) dos filmes.[...]</i>
FUJITA (2004)	<i>[...] Como já verificado, o objetivo da indexação é o de representar o conteúdo informacional do documento, tendo em vista sua recuperação, para tanto, realiza-se um exame do documento a fim de identificar conceitos pelos quais a tematicidade de um documento estará representada.[...] acreditamos que a tematicidade sempre será o conteúdo relevante do documento [...]. No idioma português, há divergências entre os pesquisadores para se referir a aboutness; para alguns, o referido termo, pode ser tematicidade, por se considerar como um substantivo ligado ao termo temático, enquanto outros adotam "atinência".[...]</i>

Quadro 04: Comparativo de citações da utilização dos termos *aboutness* e tematicidade.

A “impropriedade do conceito *aboutness*” (LARA, 2006) suscita uma série de questões relativas à representação temática da informação. Como já havíamos observado, para Naves (1996) essa complexidade inerente ao processo de análise

em representar o documento a partir do seu assunto, pode ter relação direta com a “imprecisão terminológica” quanto a esse processo. O que gera diferentes termos, tais como assunto, tema, conteúdo, rótulos, conceitos etc., e acaba sendo muito vago, pois torna-se difícil a delimitação quanto ao uso de cada um dos termos.

Por isso, esses distintos termos nos revelam a dificuldade apresentada no processo de categorização de uma “ação” ou “objeto” em virtude de defini-lo de um modo “conceitual”, que seja preciso e inequívoco. Isso demonstra que a área de análise temática necessita ainda de maior investigação quanto aos seus conceitos e definições, minimizando possíveis ambiguidades e imprecisões conceituais nocivas à organização da informação. Dessa maneira, entendemos a importância conceitual do termo *aboutness*, em virtude da necessidade de entender e identificar corretamente o que extrair do documento para representá-lo, pois é mediante esse processo de extração e representação da informação, que se tornará possível a busca e recuperação pelos usuários.

IV CONSIDERAÇÕES

Realizamos nesse trabalho a abordagem de um tema bastante complexo, e que é um dos mais inconstantes quanto à sua eleição como objeto de pesquisa em Ciência da Informação. O conceito *aboutness*, atinência ou tematicidade, como pudemos perceber, tem um papel de importância inequívoco entre os estudos documentais.

Assim, ao invés de caminharmos para frente na abordagem dos produtos documentais finais, ou seja, as ferramentas através das quais os usuários chegarão aos documentos, nossa pesquisa baseou-se em uma espécie de “retrocesso” necessário, que visou à discussão do problema considerado como o núcleo da área.

Vimos nesse percurso que a discussão a respeito do *aboutness* não se resume a indagações de ordem prática sobre como os profissionais realizam a leitura de um documento e assim chegam à sua representação. Essa questão na verdade, encontra-se na base de todo o processo de tratamento documental.

No entanto, porque tão poucos estudos (e como pudemos corroborar, ainda menos em âmbito nacional) são dedicados à parte mais importante do processo do tratamento temático?

As correntes de tratamento temático (Análise documental, Indexação e Catalogação de assuntos), possuem suas especificidades de acordo com o contexto em que foram originadas, e assim obtiveram uma evolução mais pragmática ou mais voltada aos questionamentos teóricos e conceituais.

Mas de qualquer forma, possuem a mesma função: prover a representação do conteúdo informativo de um documento de forma a proporcionar o acesso a esse conteúdo por quem se faça necessário.

Sob esse ponto de vista compreendemos que independente das relações epistemológicas convergentes ou discrepantes, os problemas fundamentais para todas se assemelham, e o seu núcleo principal lhes torna ainda razoavelmente tímidas. Pois ainda não puderam explicitar como deve ser realizada a análise de um conteúdo documental com o intuito de sintetizá-lo e representá-lo sob alguns termos

ou em alguns casos sob apenas um único termo.

Essa questão demonstra que se por um lado as ferramentas de representação e acesso já estão consideravelmente avançadas, por outro suas questões conceituais mais basilares permanecem um campo profícuo, mas no qual as explorações já realizadas não puderam dar as respostas necessárias.

No entanto, cabe ressaltar que em relação à questão específica por nós levantada, não é um desagradável privilégio da Ciência da Informação que esse problema seja ainda um verdadeiro “buraco negro” no nosso conhecimento.

O problema do *aboutness* pode ser encontrado em diversas áreas, mesmo sob nomes e prismas diferentes, ele se impõe a partir do momento que entra em cena a habilidade dos indivíduos de analisarem e compreenderem o conteúdo de um documento e assim, fazerem julgamentos e escolhas quanto às maneiras de representá-lo.

Em uma analogia, quando um tradutor precisa decidir entre uma expressão ou outra para a correta tradução de uma passagem literária, ele certamente irá se perguntar: o que o autor quis dizer com determinada expressão?

Se o tradutor entender que o autor quis ser meramente irônico escolherá uma palavra correspondente no idioma para o qual está traduzindo, provavelmente uma palavra de linguagem coloquial, talvez uma gíria, mas caso julgue e decida que pelo contrário, o autor quis ser o mais claro ou formal possível, decidirá por outra palavra, provavelmente obedecendo à forma culta.

Ou seja, para um mesmo original, o tradutor terá que julgar e realizar uma escolha, e na maior parte das vezes tendo que arcar com as consequências dessa decisão, seja ela correta ou equivocada.

Não julgamos necessário citar o paralelo existente entre o tradutor acima descrito e o bibliotecário, mas consideramos interessante manter a idéia de avaliação>julgamento>consequências.

Em um sistema de informação, avaliações e julgamentos mal realizados, levarão à consequências imprevisíveis, como o desperdício de recursos e em um último estágio, a impossibilidade de o documento ser acessado, o que obviamente compromete a própria razão de ser do sistema.

Apesar da formação e da atuação específicas, nesse momento da avaliação do conteúdo, o bibliotecário, o classificador, o indexador não são tão diferentes de qualquer outro indivíduo.

O que pode torná-lo diferente é a sua preparação prévia, relacionada a uma formação acadêmica e profissional que leve em consideração esses aspectos, além do conhecimento do sistema em que atua e com grande relevância o conhecimento da sua comunidade usuária, quando relembramos Cutter (1876), que o júri supremo da classificação de uma obra deve ser o público usuário.

Obviamente, nos é vedado censurar a falta de avanços práticos nesse sentido, pois se têm claramente a dificuldade subjacente ao problema. Como vimos no decorrer do trabalho, são conjugados nessa questão aspectos linguísticos, lógicos e cognitivos, de alta complexidade, e muitos dos quais ainda pouco pesquisados.

Essas dificuldades ficaram bastante evidentes no decorrer do trabalho, verificou-se que as discordâncias quanto à definição consensual do *aboutness* ou mesmo do significado de assunto, permanece sem uma perspectiva de esclarecimento em curto prazo. Arriscamos a dizer, mesmo em longo prazo.

O mesmo vale para os processos mentais de análise textual, compreensão e síntese do conteúdo.

Ao fim desse trabalho, pudemos concluir que se por um lado pudemos contribuir com a bibliografia do tema em língua portuguesa. Por outro, pouco pudemos participar na elucidação das divergências conceituais em relação aos conceitos explorados. No entanto, cabe ressaltar que essa não foi a intenção da pesquisa, porquanto nosso intento se limitou a buscar uma compreensão da questão pelo diálogo entre os autores e novas referências na discussão abordada.

Pudemos evidenciar a necessidade da discussão e investigação do *aboutness*, tendo em conta, que mesmo que possa ser para muitos uma espécie de problema virtual, considerando a sua intrínseca subjetividade, seus desafios e consequências são reais. E aprofundam-se com a própria transformação das necessidades e ferramentas de informação.

Apesar das propostas como de Maron, de uma possibilidade de mensuração do *aboutness*, ou dos trabalhos de Bruza et al. para sistemas de

informação baseados no *aboutness* e no *non-aboutness*, essas propostas ainda não foram levadas a cabo. E assim, continuaremos no campo da discussão teórica, obviamente de extrema importância no subsídio de avanços pragmáticos.

Em âmbito nacional, verificamos através da análise dos periódicos abordados, que a discussão não teve maiores incursões em nossa literatura, e restringiu-se a pouquíssimos artigos que mencionaram ou abordaram o problema.

Além disso, verificou-se que a adoção do termo em parte desses trabalhos ocorreu de forma imprópria, além de não apresentarem considerações quanto à sua definição, o que demonstra a dificuldade em lidar com o conceito, além da própria fragilidade terminológica presente na área, conforme também destacado por Tálamo (2009, informação verbal)³⁹.

Uma possibilidade de continuação desse trabalho é o levantamento dos demais periódicos que não foram analisados nesse momento, além da busca por outros termos. Nesse trabalho delimitamos os termos: atinência, sobrecidade e tematicidade, pela hipótese inicial de que eram os termos que mais se relacionavam ao conceito de *aboutness*, devido a sua tradução direta.

De qualquer forma, em nossa opinião, concordamos que a distinção entre *aboutness* (tematicidade) e assunto, pode ser de grande utilidade para uma formalização da análise temática, porquanto se considera o *aboutness*, como identificado no documento a partir de uma análise metódica e passível das propostas de análise sugeridas pelos aportes interdisciplinares. Além de permitir a distinção entre o tema intrínseco da obra e os sentidos atribuídos, através das diferentes nomenclaturas propostas.

Mesmo assim, não se escapa com essa posição de se definir quais seriam as nomenclaturas definidas para cada caso (intrínseco ou extrínseco). Dessa maneira, a distinção de Beghtol (1986) (*aboutness e meanings*), demonstra ser bastante relevante, e provavelmente suficiente para a nomenclatura em Ciência da Informação.

Essa distinção também evita as possíveis confusões causadas pelos “outros tipos de *aboutness*”, tais quais propostos por Maron (1977), Bruza et al.

³⁹ Realizada durante a arguição desse trabalho na defesa pública em 23 de setembro de 2009.

(1999) e Ingwersen (1992). Fujita (2009, informação verbal)⁴⁰, destaca que Ingwersen ao propor o *aboutness* do autor, acaba ainda por desconsiderar a linha de Análise do discurso, que propõe a abordagem do contexto do autor, a partir do qual poderá colaborar com a compreensão da obra. Tálamo (2009, informação verbal)⁴¹ também aponta que o *aboutness* do indexador é por sua vez, a linguagem documentária do sistema. Fujita e Tálamo concordam que há dessa forma uma descaracterização do conceito original de *aboutness* e concordamos com essa constatação.

Lancaster, por sua vez, apesar de seu destaque e pesquisa na área demonstra não possuir grandes preocupações com a questão, apresentando um direcionamento pragmático que questiona (ou desconsidera) a necessidade de definição do *aboutness* para a melhoria dos trabalhos de indexação. No entanto, tais definições irão colaborar com a sedimentação da Ciência da Informação enquanto campo científico, e dessa forma, subsidiar seu próprio avanço, conforme já observado por Lucas (1997)⁴².

Consideramos de destaque também a idéia compartilhada pela maioria dos autores abordados, de que para se chegar a entender como o conteúdo de um documento será compreendido e representado a grande questão está de fato no entendimento do processo cognitivo humano. Ou seja, como um indivíduo processa uma informação, a compreende e em seguida a sintetiza e representa? Essa questão na verdade está implícita em nosso dia a dia e essa atividade é realizada de forma tão natural que poucas vezes nos preocupamos a esse respeito.

Entender como funciona nossa cognição e como apreendemos algo, ou são feitas as associações durante a leitura de um documento, certamente seria o fundamental para se atingir uma proposta mais consistente que pudesse ser mais amplamente abordada na área.

O processo de compreensão de um texto é basicamente subjetivo (mas como vimos, não entregue ao acaso) e isso ocorre, tendo em vista que na maior parte dos casos, mesmo tendo por auxílio uma linguagem documentária específica, a escolha quanto à representação é de fato realizada por um indivíduo, e esse

⁴⁰ idem.

⁴¹ idem.

⁴² Citação presente na pg.30.

poucas vezes, possui as condições ideais para um trabalho de representação de forma a atender a um público específico satisfatoriamente.

Pois pode possuir um público altamente especializado (bibliotecas corporativas, centros de informação especializados etc.) ou de outro lado, possuir públicos e objetivos tão heterogêneos que fica dessa forma impossibilitado de entender as necessidades específicas desse público, considerando não apenas a sua heterogeneidade, mas a quantidade e complexidade do material a ser analisado.

Ressaltamos ainda que o direcionamento teórico dessa pesquisa a conduziu por um viés que buscou muito mais realizar a revisão de literatura e reflexões sobre o tema sem o compromisso de apontar respostas. De qualquer modo, em nossa área essa questão possui um fim prático e objetivo: ou seja, fornecer à sociedade instrumentos de representação e acesso à informação, que sejam de qualidade e possibilitem a satisfação de suas necessidades de acesso e geração de conhecimento; evidenciando dessa forma a importância e relevância da Ciência da Informação, e suas temáticas implícitas, para a sociedade.

REFERÊNCIAS

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

ALVARENGA, Lígia. A Teoria do Conceito Revisitada em Conexão com Ontologias e metadados no Contexto das Bibliotecas Tradicionais e Digitais. **DataGramaZero** - Revista de Ciência da Informação, Brasília, v.2, n.6, dez. 2001.

AULETE, Caldas. Dicionário contemporâneo da Língua Portuguesa. 2009. Disponível em: < <http://www.auletedigital.com.br> >

BARANOW, Ulf Gregor. Perspectivas na contribuição da Lingüística e de áreas afins à Ciência da Informação. **Ciência da Informação**, Rio de Janeiro, v. 12, n. 1, p. 23-35, 1983.

BEGHTOL, Claire. 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, June. 1986.

BRUZA, Peter. SONG, Dawei. WONG, Kam-Fai. Fundamental properties of aboutness. **Research and Development in Information Retrieval**, p. 277–278, 1999.

CAMPBELL, Grant. (2000a). Queer Theory and the Creation of Contextual Subject Access Tools for Gays and Lesbian Communities, **Knowledge Organization**. V. 27, n.3, p.123. 2000.

CAMPBELL, Grant. (2000b). Aboutness and Meaning: How a Paradigm of Subject Analysis Can Illuminate Queer Theory in Literary Studies. IN: **CAIS 2000**. Canadian Association for Information Science: Proceedings of the 28th Annual Conference. <http://www.slis.ualberta.ca/cais2000/campbell.htm>

CASTAÑON MORENO, B. Análisis temático documental. **Investigación bibliotecológica**, Mexico, D.F., v.6, n.12, p:4-10, Ene./Jun. 1992.

CHRIST, G. Elgie (Ed.). **The Nuttall dictionary of english synonyms and antonyms**. Rio de Janeiro: Artenova, 1975. 305 p.

CIÊNCIA DA INFORMAÇÃO. Brasília: IBICT. Disponível em: < <http://www.ibict.br/cienciadainformacao/> >. Acesso em 20 jul. 2007.

CINTRA, Ana Maria Marques. Estratégias de leitura em documentação. In: Smit, Johanna W. (Coord.). **Análise documentária: a análise da síntese**. 2.ed. Brasília : IBICT, 1987.

CLEVELAND D. B.; CLEVELAND, A. D. **Introduction to Indexing and Abstracting**, 3 ed. Englewood, Colo.: Libraries Unlimited, 2001.

CORDEIRO, Rosa Inês de Novais; AMÂNCIO, Tunico. Análise e representação de filmes em unidades de informação. **Revista de Ciência da Informação**, Brasília, v. 34, n. 1, p.89-94, jan./abr. 2005

CUNHA, Isabel Maria R. Ferin. **Análise documentária**: considerações teóricas e experimentações. São Paulo: FEBAB, 1989.

CUTTER, C. A. **Rules for a Dictionary Catalog**. W. P. Cutter, ed. 4th ed. Washington: Government Printing Office, 1904. (Repr. London: The Library Association, 1962.)

DATAGRAMAZERO. Rio de Janeiro: IASI. Disponível em: < <http://www.dgz.org.br/> >. Acesso em: 20 jul. 2007.

DIAS, Eduardo Wense. Análise de assunto: percepção do usuário quanto ao conteúdo de documentos. **Perspectiva em Ciência da Informação**, Belo Horizonte, v.9 n.2, p. 146-157, jul./dez. 2004

DIAS, Eduardo Wense; NAVES, Madalena Martins Lopes. **Análise de assunto**: teoria e prática. São Paulo: Thesaurus, 2007.

DIAS, Eduardo Wense; NAVES, Madalena Martins Lopes; MOURA, Maria Aparecida. O usuário-pesquisador e a análise de assunto. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 6, n. 2, p. 205 - 221, jul./dez. 2001

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

FERREIRA, Aurélio Buarque de Holanda. **Novo Dicionário Aurélio da Língua Portuguesa**. Rio de Janeiro: Nova Fronteira, 1975.

FOSKETT, A. C. **A abordagem temática da informação**. Tradução de Antonio Agenor Briquet de Lemos. São Paulo: Polígono, 1973.

FROHMANN, Bernd. The social construction of knowledge organization: the case of Melvil Dewey. **Advances in Knowledge Organization**, v. 4, 109-117, 1994.

FUJITA, Mariângela Spotti Lopes. A leitura em análise documentária. **Transinformação**, v. 10, n. 3, p. 13-31, set./dez. 1998.

FUJITA, Mariângela Spotti Lopes. **A análise documentária no tratamento da informação**: as operações e os aspectos conceituais interdisciplinares. Marília: Departamento de Ciência da Informação, FFC/UNESP, 2003a. 15 f.

FUJITA, Mariangela Spotti Lopes. A prática da indexação: análise da evolução de tendência teóricas e metodológicas. **Transinformação**, Campinas, v.16, n. 2, p.133-161, maio/ago., 2004.

FUJITA, Mariângela Spotti Lopes. A identificação de conceitos no processo de análise de assunto para indexação. **Revista Digital de Biblioteconomia e Ciência**

da Informação, Campinas, v. 1, n. 1, p.60-90, dez. 2003. Disponível em:<
http://dici.ibict.br/archive/00000239/>. Acesso em: 05 ago. 2005.

FUJITA, Mariângela Spotti Lopes. A Leitura Documentária na perspectiva de suas variáveis: leitor-texto-contexto, . **DataGramaZero - Revista de Ciência da Informação**, v.5, n.4, ago 2004. Disponível em:
<http://www.dgzero.org/ago04/Art_01.htm> . Acesso em: 20 jan. 2007.

GARDIN, Jean-Claude. et al. La logique du plausible: essais d'epistemologie pratique. Paris: Maison de Sciences de L'Homme, 1981. Citado por 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.

GUIMARÃES, José Augusto Chaves. A Análise Documentária no âmbito do tratamento da informação: elementos históricos e conceituais. In: RODRIGUES, G. M. ; LOPES, I. L. (org.). **Organização e Representação do Conhecimento na Perspectiva da Ciência da Informação**. Brasília: Thesaurus, 2003.

HARROD'S Librarian's Glossary of terms used in librarianship, Documentation and the Book Crafts and Reference Book. 6th edition, Aldershot: Gower, 1989, p. 281.

HICKEY, D. H. Subject Analysis: an interpretative survey. **Library Trends**, July 1976. 273-291.

HJØRLAND, Birger. **Information Seeking and Subject Representation: An Activity-theoretical approach to Information Science**. Westport & London: Greenwood Press, 1997.

HJØRLAND, Birger. **Subject**. 2007. Disponível em:
<http://www.db.dk/bh/core%20concepts%20in%20lis/articles%20a-z/subject.htm>.
Acesso em: 18 set. 2008.

HJØRLAND, Birger. The concept of 'subject' in Information Science. **Journal of Documentation**. v.48. no. 2 June 1992.p p. 172-20

HOUAISS, Antônio. **Minidicionário Houaiss da Língua Portuguesa**. Rio de Janeiro: Objetiva, 2001.

HUTCHINS, John William. On the problem of 'aboutness' in document analysis. **Journal Of Informatics**, East Anglia, v. 1, n. 1, p.17-35, 1977. Disponível em:
<http://www.scils.rutgers.edu/~muresan/551_IR/Resources/Docs/jinfHutchins1977.pdf>. Acesso em: 20 jan. 2007.

HUTCHINS, John William. **History of my publications in machine translation**. Disponível em: < <http://www.hutchinsweb.me.uk/publications-history.pdf>> Acesso em: 18 set. 2008. s.d.

INGWERSEN, Peter. **Information Retrieval Interaction: definitions, references & subject index**. 1992. Disponível em: <www.db.dk/pi/iri>. Acesso em: 20 set. 2008.

JOUDREY, Daniel N. **Building puzzles and growing pearls: a qualitative exploration of determining aboutness**. 2005. 475p. Phd Thesis – School of Information Sciences, University of Pittsburgh.

HOUAISS, Antonio; VILLAR, Mauro de Salles; FRANCO, Francisco Manoel de Mello. **Dicionário Houaiss da língua portuguesa**. Rio de Janeiro: Objetiva, 2001.

KOBASHI, Nair Yumiko. **A elaboração de informações documentárias: em busca de uma metodologia**. 1994. 163 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, 1994.

KOCH, Ingedore Grunfeld Villaca. **O texto e a construção dos sentidos**. São Paulo: Contexto, 1997.

LANCASTER, Frederick Wilfrid. **Indexação e resumos: teoria e prática**. Tradução de Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos/Livros, 2003.

LANGRIDGE, Derek. **Classificação: abordagem para estudantes de biblioteconomia**. Rio de Janeiro: Interciência, 1977.

LARA, Marilda Lopes Ginez de. Novas relações entre Terminologia e Ciência da Informação na perspectiva de um conceito contemporâneo da informação. **DataGramaZero**, Rio de Janeiro, v.7, n.4, ago. 2006

LUCAS, Clarinda Rodrigues. Biblioteconomia: produção e administração da interpretação. **Ciência da Informação**. Inf., Brasília, v. 26, n. 1, 1997. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19651997000100007&lng=en&nrm=iso>. Acesso em: 20 jan. 2007.

LUCAS, Clarinda Rodrigues. **Leitura e interpretação em biblioteconomia**. Campinas: UNICAMP, 2000. 91 p. (Pesquisas).

MARCONDES, Carlos Henrique; MENDONÇA, Marília A. R.; MALHEIROS, Luciana R. Uma proposta de modelo de representação do conhecimento contido no texto de artigos científicos publicados na web em formato legível por programas. **DataGramaZero** - Revista de Ciência da Informação - v.7 n.5 out/06

MARCONDES, Carlos Henrique. Representação e economia da informação. **Revista de Ciência da Informação**, Brasília, v. 30, n. 1, p. 61-70, jan./abr. 2001

MARON, M. E. On Indexing, retrieval and the meaning of about. **Journal of the American Society for Information Science**, n. 28, n. 1, p. 38-43, 1977.

MEDEIROS, Marisa Bräscher Basílio. Terminologia brasileira em ciência da informação: uma análise. **Revista Ciência da Informação**, Brasília, v.15, n.2, p.135-142, jul./dez.1986

MENDONÇA, Ercília Severina. A lingüística e a ciência da informação: estudos de uma interseção. 2000. **Revista Ciência da Informação**, v. 29, n.3, p.50-70.

METCALFE, Jennifer. When a subject is not a subject? In C. H. Rawski (Ed.), **Towards a theory of librarianship**. New York: Scarecrow Press, 1973.

MICHAELIS. Universo On Line Uol. **Dicionário Michaelis**. São Paulo, 2008. CD-ROM.

MORAES, João Batista Ernesto de; GUIMARÃES, José Augusto Chaves. Análise documental de conteúdo de textos literários narrativos: em busca do diálogo entre as concepções de aboutness/meaning e de percurso temático/percurso figurativo. In: ENCUEENTROS INTERNACIONALES SOBRE SISTEMAS DE INFORMACIÓN Y DOCUMENTACIÓN, 11., 2006, Zaragoza. **Anales...** Zaragoza: Facultad de Filosofía y Letras da Universidad de Zaragoza, 2006. 13f.

MOREIRO GONZÁLEZ, José Antonio. **Conceptos introductorios al estudio de La información documental**. Salvador: EDUFBA/Fondo Editorial de la Pontificia Universidad Católica del Perú, 2005.

MOREIRO GONZALES, José Antonio. Criterios e indicadores para evaluar la calidad del análisis documental de contenido. Revista **Ciência da Informação**, Brasília, v. 31, n. 1, p. 53-60, jan./abr. 2002.

NAVES, Madalena Martins Lopes. Estudo de fatores interferentes no processo de análise de assunto. **Perspectivas em ciência da informação**, Belo Horizonte, v. 6, n. 2, p. 189 - 203, jul./dez. 2001. Disponível em: <<http://www.eci.ufmg.br/pcionline/viewarticle.php?id=361>>. Acesso em: 20 jan. 2007.

NOVELLINO, Maria Salet Ferreira. Instrumentos e metodologias de representação da informação. **Informação & Informação**, Londrina, v. 1, n. 2, p. 37-45, jul./dez. 1996.

OLINTO, Antonio. **Minidicionário Antonio Olinto da Língua Portuguesa**. São Paulo; Moderna, 2000.

OLIVER. C. One eyed king: automatic indexing. **Canadian Library Journal**, v.46, n.5, p.312-316, 1989.

PERSPECTIVAS EM CIENCIA DA INFORMACÃO. Belo Horizonte: UFMG. Disponível em: < <http://www.eci.ufmg.br/pcionline/> >. Acesso em: 20 jul. 2007.

PINTO MOLINA, M. **Análisis documental**: fundamentos y procedimientos. 2. ed. rev. ampl. Madrid: Eudema Universidad; Manueles, 1993.

RANGANATHAN, S. R. **Prolegomana to library classification**. Bombain, Asia Publishing House, 1967.

REITZ, Joan M. **Dictionary for library and Information Science**. Lodon: Libraries, 2004.

SMIT, Johanna Wilhelmina. (Coord.). **Análise documentária**: a análise da síntese.

Brasília: IBICT, 1987.

SOUZA, Renato Rocha; ALVARENGA, Lídia. A Web Semântica e suas contribuições para a ciência da informação. **Revista de Ciência da Informação**, Brasília, v. 33, n. 1, p. 132-141, jan./abril 200

SVENONIUS, E. **The intellectual foundation of information organization**. Cambridge, MA: MIT Press, 2000.

TODD, R. T. Academic indexing: what's it all about? **The Indexer**, London, v. 18, n. 2, p. 101-104, abr. 1992.

TOSCANA, David. **O último leitor**. Tradução Ana Lúcia Pelegrino e Magali Pedro. Rio de Janeiro: Casa da palavra, 2004.

TRANSINFORMAÇÃO. Campinas: Puccamp. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/>> Acesso em: 15 abr. 2008.

VAN DER LAAN, Regina Helena. Análise de assunto e lingüística textual: uma tentativa de aproximação. *In*: Congresso Brasileiro de Biblioteconomia e Documentação, 19., 2000, Porto Alegre. **Anais...** Porto Alegre: PUC-RS, 2000. Disponível em: <<http://dici.ibict.br/archive/00000800/>>. Acesso em: 20 jan. 2007.

VENTURA, Carolina Siqueira Muniz; LIMA-LOPES, Rodrigo Esteves. **O Tema**: caracterização e realização em português. São Paulo/Liverpool 2003 (artigo publicado - DIRECT Papers 47).

WANDERLEY, Manoel Adolpho. Linguagem Documentária: Acesso à Informação Aspectos do problema. **Revista de Ciência da Informação**, Brasília, Rio de Janeiro, 2(2): 175-217, 1973