

# Análise de assunto: percepção do usuário quanto ao conteúdo de documentos

Eduardo Wense Dias

Professor Titular da Escola de Ciência da Informação da UFMG

Procura-se compreender, numa perspectiva da análise de assunto, a visão dos usuários com relação ao conteúdo de documentos com que têm familiaridade. Para efeito de comparação, utiliza-se um delineamento de natureza longitudinal, com a coleta de dados em dois momentos temporais diferentes. Os resultados mostram alguma convergência mas também discordância dos usuários nos dois momentos referidos, bem como a interpretação dos usuários para eventuais discordâncias. Os resultados sugerem padrões de classificação dos tópicos identificados nos documentos em fundamentais/periféricos e genéricos/específicos. Sugerem também que a natureza do trabalho desenvolvido num determinado momento pelo usuário assim como a evolução terminológica são fatores importantes causadores das divergências na identificação dos tópicos.

**Palavras-chave:** análise de assunto – usuários – identificação de conceitos

Recebido em 12.02.2004

Aceito em 02.09.2004

## Introdução

A análise de assunto é a etapa do tratamento temático da informação em que um determinado documento<sup>1</sup>, a ser incorporado à coleção de um sistema de informação ou de recuperação da informação, é analisado para determinar de qual ou quais assuntos trata e, também, se for o caso, quais desses assuntos deverão ser representados no simulacro (registro catalográfico, metadados etc) que certamente será criado para o documento no sistema. A atividade tem sido considerada crucial naqueles tipos de sistemas, dada a importância da busca por assunto: pesquisas têm mostrado que esse é o tipo de busca mais demandado pelos usuários (TAYLOR, 1995). Entretanto, é também o tipo de busca em que maior é a frustração do usuário (MARTINEZ & PICHININI, 1995). Paradoxalmente, a pesquisa em tratamento temático concentra quase todas as suas forças na etapa seguinte do processo de tratamento temático, o de tradução do resultado da análise de assunto para uma linguagem de indexação. Como pode atestar qualquer pessoa minimamente familiarizada com a literatura de tratamento da informação, a bibliografia que discute o desenvolvimento e o uso de linguagens de indexação é vasta.

Por isso, há um clamor por que mais pesquisas sejam dedicadas à etapa de análise de assunto. Se esta é mal feita, pouco adiantam os cuidados para se desenvolver linguagens de indexação e preparar profissionais habilitados no uso dessas linguagens. É por isso que Langridge (1989) afirmou, apropriadamente, que nenhum sistema de recuperação de informação pode ser melhor que a análise de assunto que nele é realizada. Qualquer conceito ou tópico importante que escape à análise de assunto não poderá, logicamente, se ver traduzido na linguagem utilizada no sistema. Mais e melhores pesquisas sobre o tema certamente vão representar uma importante contribuição no sentido de um melhor conhecimento do processo, providência indispensável a qualquer esforço de aperfeiçoamento do mesmo.

## Revisão de literatura

Os estudos de análise de assunto vêm sendo feitos de três perspectivas principais: a dos profissionais de informação, a dos usuários e a dos autores das obras analisadas. O profissional responsável pela análise de assunto nos sistemas de informação e de recuperação da informação recebe diferentes nomes dependendo do tipo de sistema e do tipo de documento com que lida. Assim, a análise de assunto numa biblioteca (inclusive bibliotecas digitais<sup>2</sup>), e quando o material se trata principalmente de monografias e documentos similares, é feita por dois tipos de profissionais: a) pelo classificador, cuja função é escolher, no sistema de classificação utilizado pela biblioteca, o número de classificação a ser consignado ao documento; e b) pelo catalogador, cuja função é determinar os cabeçalhos de assunto ou descritores a serem utilizados no registro catalográfico que será criado para aquele documento; neste caso, esse catalogador pode ser ainda chamado, mais especificamente, de catalogador de assunto, para diferenciar essa tarefa (catalogação de assunto) da outra tarefa do catalogador, a de identificação dos demais dados bibliográficos (catalogação descritiva.). É claro que, numa biblioteca ou sistema de pequeno porte, essas atividades todas podem ser realizadas por uma única pessoa, que poderá estar desempenhando, ainda, várias outras funções técnicas (seleção,

<sup>1</sup> Desde pelo menos Suzanne Briet que o termo documento é utilizado, na biblioteconomia e ciência da informação, num sentido bastante amplo para designar qualquer tipo de registro material de uma obra de criação intelectual (BRIET 1951). Entretanto, tendo em vista o papel hegemônico desempenhado pelo texto escrito no âmbito dessa grande variedade documental existente, tem-se a tendência de usar o texto como uma espécie de arquétipo dessas formas documentais todas. Por conseguinte, em muitas passagens deste artigo deve-se entender que o termo 'texto' está sendo utilizado como sinônimo de documento.

<sup>2</sup> Em outro artigo (DIAS 2001) faço um paralelo entre a biblioteca tradicional e a biblioteca digital para mostrar que, apesar das grandes diferenças existentes entre a armazenagem de informação no suporte papel e a armazenagem em suportes eletrônicos, todas as funções básicas desempenhadas pelos profissionais de informação na primeira continuam necessárias e importantes na segunda.

referência etc) no sistema e receber uma designação genérica como bibliotecário, documentalista, gestor de informação etc. Em sistemas de grande porte, entretanto, é quase certo que haverá necessidade de se contar com profissionais especializados em cada uma daquelas funções, caso em que poderão receber, então, as designações mais específicas acima identificadas (classificador, catalogador de assunto etc).

Há ainda uma terceira designação para o analista de assunto, a de indexador, quando o trabalho se faz em serviços de indexação e resumo (produtores de bases de dados bibliográficas), sistemas em que o tipo principal de material indexado são os artigos de periódicos científicos.<sup>3</sup>

A atividade do analista de metadados anuncia-se como a função correspondente, no contexto digital, às três funções acima mencionadas (classificador, catalogador de assunto e indexador). Em consonância com análise anterior (DIAS, 2001), a evolução deverá ser no sentido de se replicar, no ambiente digital, essas especializações todas existentes no âmbito dos sistemas tradicionais.

A perspectiva do profissional da informação (analista de assunto) baseia-se no fato de que é ele quem comanda o processo da análise de assunto e, assim, tem um conhecimento e uma vivência que é importante estudar e analisar. Na pesquisa feita da perspectiva do profissional da informação, a abordagem clássica é o estudo desse profissional em sua atividade de análise de assunto, com o objetivo de descrição do processo. Um bom exemplo é o estudo de Chu & O'Brien (1993), que investigou indexadores novatos solicitados a fazerem a análise de assunto de três artigos de revistas populares. Num estudo mais focado, Naves (2001) estudou os analistas de assunto, procurando identificar obstáculos por eles enfrentados.

Um aspecto específico muito estudado nas pesquisas que investigam o profissional é a leitura, etapa inicial da análise de assunto. Um exemplo de pesquisa de compreensão textual é Farrow (1991), que busca conhecer o modelo situacional do indexador, ou seja, compreender como o indexador procede na tarefa de leitura em sua atividade profissional. No Brasil, temos o estudo de Fujita (1999), nessa mesma linha.

A estratégia de investigar os autores dos textos indexados, por sua vez, é fundamentada na lógica de que ninguém melhor que o próprio autor para identificar do que trata o documento que produziu. É claro que alguns vão discordar desse tipo de estratégia, principalmente aqueles que entendem que um texto de certa forma se torna independente do seu autor, depois que se torna público. Ou seja, que os leitores podem dar a um texto significados que o próprio autor não previu. Exemplo de estudo nessa linha é o de Braam e Bruil (1992), que investigaram autores indexados no *Chemical Abstracts online*. O estudo utilizou 270 artigos de periódicos originários de 54 periódicos da área de química. Os autores eram solicitados a opinar sobre os descritores utilizados pelos indexadores.

A perspectiva do usuário, enfim, está sintonizada com a tendência, relativamente recente, de se estudar a informação, sua organização e seus sistemas posicionando o usuário no centro dessas preocupações. É a chamada abordagem centrada no usuário, área em que se destacam autores como Dervin (DERVIN & NILAN, 1986).

A perspectiva do usuário baseia-se no pressuposto de que eles são uma importante fonte para subsidiar ações que visem o aperfeiçoamento do processo de análise de assunto. Afinal, são os destinatários finais dos produtos desenvolvidos nos sistemas de informação e de recuperação de informação

<sup>3</sup> Embora incomum, devido a razões de ordem econômica, o trabalho de indexação e resumos, pode eventualmente ser realizado dentro de uma biblioteca. Nesses casos, observe-se que poderemos ter, então, três instâncias de análise de assunto que, dependendo da dimensão do serviço, poderão ser realizadas por três profissionais diferentes, e cada um com uma designação diferente: classificador, catalogador de assunto e indexador.

com base nessa análise. A idéia subjacente é a de que, sabendo-se como o usuário descreve ou identifica o assunto de uma determinada obra, estaríamos nos aproximando de uma forma muito proveitosa de análise da mesma para fins de tratamento e recuperação da informação. Esta é a perspectiva adotada na pesquisa descrita neste artigo. Com isso, continuamos a adotar estratégia já utilizada em estudo anterior (DIAS, NAVES, MOURA, 2001), onde o foco também era no usuário, no seu comportamento de busca e uso de informação, visto este processo de uma perspectiva de interesse da análise de assunto. Os dois trabalhos adotam formas diferentes de estudar o usuário, mas convergindo para a questão central de interesse, a análise de assunto.

Alguns estudos podem também adotar uma combinação dessas três estratégias básicas acima identificadas (TURNER 1995; SAARTI, 2000). Este último, combina as perspectivas do usuário e do profissional da informação, com o objetivo de verificar como o conteúdo de obras de ficção é sumarizado. Para isso, o autor do estudo pesquisou tanto usuários quanto bibliotecários, de uma biblioteca pública na Finlândia.

O outro aspecto que devemos discutir nesta revisão de literatura diz respeito à substância mesma da pesquisa. Delimitamos o nosso objeto na área da análise de assunto, mas esta área, seja enquanto processo nos sistemas de informação e de recuperação de informação, seja como uma especialidade merecedora de investigação científica no campo da biblioteconomia/ciência da informação, tem uma boa quantidade de aspectos específicos a serem considerados. 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).

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. Todd (1992), citando autores como Cutter, Kaiser, Ranganathan, Coates e Vickery, afirma que, neste campo, há uma considerável confusão terminológica. Cutter define assunto como tema ou tópico, podendo ou não estar no título do documento. Kaiser entende que assuntos são *coisas em geral*, reais ou imaginárias. Ranganathan fala em pensamento contido no documento e para Coates assunto significa uma como que abstração da idéia global personificada no assunto. Vickery, enfim, refere-se ao tema no qual livros, artigos ou parte de artigos são escritos.

Uma evolução rumo ao apuro conceitual nessa matéria ocorre com Fairthorne (1969), que se dedicou a estudar a questão com rigor. É ele quem introduz os conceitos de *atinência extensional* e *atinência intensional*, em que, ao invés do termo *assunto*, usa o termo *atinência* para expressar a idéia do 'de que trata o documento'. A *atinência extensional*, segundo Fairthorne, diz respeito ao assunto inerente ao documento, ao passo que a *atinência intensional* diria respeito à razão ou ao propósito da incorporação do documento ao acervo de uma biblioteca ou sistema de informação, ou de sua procura por um usuário. Essa distinção tem sido reconhecida por outros autores, como Boyce (1982), que utiliza a terminologia *topicalidade* (*atinência extensional*) e *informatividade* (*atinência intensional*), e Beghtol (1986), que usa os termos *atinência* (*atinência extensional*) e *significado* (*atinência intensional*). Em Beghtol, *atinência* significa o assunto intrínseco de um documento, algo que é, até certo ponto,

independente do uso específico que possa ser dado ao mesmo por um determinado leitor. Significado, por outro lado, diz respeito ao assunto ou assuntos que, num determinado documento, têm interesse para um leitor.

Para tornar ainda mais complexa a questão, existem vários outros termos que podem ser utilizados para fazer referência ao conteúdo substantivo de um documento ou informação. Além de assunto, mencionou-se acima termos como atinência e topicalidade, este último remetendo, naturalmente, para o cognato tópico. Assim, quando se fala do tópico de um documento, tópico adquire claramente uma sinonímia com o termo assunto. Finalmente, tem-se termos como tema e conceito. Continua na prática, entretanto, uma tendência a se usar o termo assunto de uma forma bastante genérica, inclusive como sinônimo desses outros termos mencionados. Uma autora importante da área (SVENONIUS, 1995), por exemplo, descreve o objetivo da análise de assunto como sendo *determinar o assunto* (tema, tópico ou atinência) de um documento (SVENONIUS, 1995, p. 3). Faltou somente o termo conceito para que se tivesse, então, toda a gama de termos mencionados representados pelo termo assunto.

## Objeto da pesquisa

Tendo em vista as considerações acima, quanto ao reduzido número de pesquisas na área, muito do que nela se faz em termos de investigação científica há que ter um caráter exploratório. Nesses termos, a pesquisa teve aqueles objetivos clássicos de satisfazer a curiosidade do pesquisador e seu desejo de entender melhor o assunto; testar a viabilidade de realizar um estudo mais cuidadoso; e desenvolver os métodos a serem empregados num tal estudo (BABBIE, 1983, p.74).

Quanto a seu objeto substantivo, pode ser formulado em termos da seguinte questão: se for perguntado a um usuário, em momentos distintos, qual é o assunto de um determinado documento, como se comparam as respostas obtidas? A questão é suscitada por uma desconfiança de que ocorram divergências nesses diferentes momentos. Os sistemas de informação, entretanto, não são propriamente planejados para levar em conta essa dinâmica que o usuário estabelece na relação com um documento. Assim, o objetivo do estudo é comprovar essa hipótese de trabalho e interpretar suas conseqüências, na suposição de que esse conhecimento possa ser de utilidade na tarefa de aperfeiçoamento do processo de análise de assunto.

## Metodologia

Em consonância com a questão substantiva acima enunciada, o projeto foi concebido na forma de um estudo longitudinal, isto é, os mesmos tipos de dados seriam coletados dos mesmos respondentes em momentos distintos. Tendo em vista as limitações existentes para o desenvolvimento desta pesquisa, decidiu-se reduzir esses momentos a duas instâncias de coletas de dados. A diferença temporal entre essas duas coletas não foi fixada *a priori* de forma rígida, mas planejou-se um tempo em torno de doze meses. Na prática, esse tempo terminou por situar-se em cerca de dois anos, na média.

É notório que os problemas de representação lingüística dos conceitos são muito mais sérios nas ciências sociais e humanidades do que nas ciências

exatas e naturais. Assim, decidiu-se que uma amostra de pesquisadores naquelas duas grandes áreas poderia ser muito mais reveladora para os propósitos deste estudo. Por isso, trabalhou-se com oito professores universitários das áreas de ciência da informação, comunicação social, educação, filosofia e história. Trata-se de uma amostra intencional, considerada suficiente, dada a natureza do estudo.

O levantamento de dados foi realizado em duas etapas, conforme explicado acima. Na primeira etapa, os respondentes foram solicitados a preencher um formulário em que se lhes pedia que identificassem um livro e um artigo de periódico com os quais tivessem familiaridade. Os respondentes deveriam, em seguida, identificar cabeçalhos de assunto, palavras-chave, descritores ou terminologia similar que entendessem representar com fidelidade o conteúdo dos itens.

Na segunda etapa, realizada meses depois da primeira etapa, os respondentes foram entrevistados para que:

- ⊙ repetissem o procedimento de identificação do conteúdo dos documentos realizado na primeira etapa, ou seja, identificassem termos para os mesmos documentos analisados naquela etapa;

- ⊙ se pronunciassem sobre os dados coletados nas duas etapas (o entrevistado era lembrado, nesse momento, dos dados coletados na primeira etapa), procurando explicar eventuais convergências e divergências.

Como se pode depreender desses procedimentos, houve o cuidado de que o entrevistado não fosse influenciado pelas respostas dadas por ele quando da primeira coleta.

## Análise dos resultados

A análise dos resultados mostra, de um modo geral, que os respondentes tanto convergiram quanto divergiram no que diz respeito à identificação de assuntos dos documentos nas duas etapas. É possível categorizar esse padrão de convergência/divergência em cinco instâncias principais:

- ⊙ termos idênticos utilizados em ambas as coletas;
- ⊙ nenhum termo da segunda coleta idêntico ao da primeira;
- ⊙ termos utilizados na primeira coleta que foram desprezados na segunda;
- ⊙ termos acrescentados quando da segunda coleta;
- ⊙ alterações no enunciado dos termos, inclusive, por meio da subtração ou acréscimo de subcabeçalhos.

A quantificação da convergência pode ser feita em termos da fidelidade do usuário, ou seja, o número de termos a que ele permanece fiel nos dois momentos. Essa fidelidade está em torno de 2/3 dos termos atribuídos, ou cerca de 66%. Isso significa, tipicamente, um usuário que atribuiu seis termos

de indexação a um documento, na primeira vez, e na segunda substitui dois desses termos por dois outros termos diferentes. Em outras palavras, no conjunto de seis termos ele permaneceu consistente em relação a quatro dos seis termos, ou 2/3 dos termos atribuídos. Um exemplo de resposta que ilustra essa divergência é o que se pode observar na TAB. 1, elaborada com base nos dados obtidos do respondente que escolheu o texto de Line<sup>4</sup>.

TABELA 1 – Assuntos do artigo de LINE segundo o respondente

	<b>Primeira resposta</b>	<b>Segunda resposta</b>
1	Empréstimo interbibliotecário	Empréstimo interbibliotecário
2	Empréstimo entre bibliotecas	Empréstimo entre bibliotecas
3	Cooperação bibliotecária	
4	Acessibilidade ao documento	Acessibilidade ao documento
5	Acesso ao documento	
6	Compartilhamento de acervos	
7		UAP (Universal Availability of Publications)
8	Comutação	

Como se pode observar, o usuário foi fiel aos tópicos *Empréstimo interbibliotecário*, *Empréstimo entre bibliotecas* e *Acessibilidade ao documento*. Por outro lado, desprezou três outros (Cooperação bibliotecária, Acesso ao documento e Compartilhamento de acervos). Finalmente, acrescentou dois outros tópicos: UAP (*Universal Availability of Publications*) e Comutação.

Houve apenas um caso de divergência total entre os termos da primeira e da segunda coleta, o do respondente que trabalhou com o artigo de Buitoni (TAB. 2). O exemplo parece constituir-se num caso radical, pois não houve repetição de nenhum dos termos de uma consulta para a outra. Não é possível, a princípio, identificar sinonímias entre os termos das duas coletas, o que poderia levar o caso a ser classificado numa das outras categorias acima identificadas.

TABELA 2 – Tópicos do artigo de BUITONI segundo um respondente

	<b>Primeira resposta</b>	<b>Segunda resposta</b>
1	Epistemologia do jornalismo	
2	Narrativa jornalística	
3	Ideologia do jornalismo	
4		Métodos jornalísticos
5		Linguagem jornalística
6		Construção da realidade

Poderíamos sugerir ser essa variação precisamente o significado que o texto vai adquirindo para o usuário, de acordo com variados fatores. Os tópicos não variáveis, por outro lado, poderiam ser considerados como constituindo a atinência do documento, seu assunto permanente.

As divergências podem ainda ser analisadas quanto à sua natureza e quanto aos fatores que as determinaram ou podem tê-las determinado. Quanto à sua natureza, pode-se identificar um padrão de especificidade/generalidade. Ou seja, os termos agregados ou são mais específicos do que os anteriores, ou mais genéricos. Por exemplo: o respondente que escolheu o artigo de

<sup>4</sup> As referências completas desses textos podem ser encontradas no APÊNDICE.

Cesarino & Pinto, da primeira vez utilizou os termos Indexação, Análise de assunto e Linguagens de indexação. Da segunda vez, não escolheu Indexação, que é o mais genérico dos três termos. Além disso, atribuiu dois termos mais específicos que todos os outros utilizados anteriormente, Pré-coordenação e Pós-coordenação. Neste caso, por conseguinte, há um claro direcionamento rumo a uma maior especificidade.

TABELA 3 – Assuntos do texto de CESARINO & PINTO segundo um respondente

	Primeira resposta	Segunda resposta
1	Indexação	
2	Análise de assunto	Análise de assunto
3	Linguagens de indexação	Linguagens de indexação
4		Pré-coordenação
5		Pós-coordenação

Um segundo exemplo que mostra o fator especificidade em ação pode ser visto nos dados referentes ao respondente que escolheu o artigo de Buitoni (TAB. 2), já mencionado. Conforme sua interpretação, é possível detectar uma maior especificidade na relação entre os pares *Epistemologia do jornalismo / Métodos jornalísticos* e *Narrativa jornalística / Construção da realidade*.

Uma outra diferenciação que ficou sugerida é entre tópicos fundamentais e tópicos periféricos. A evidência decorre da interpretação da maioria dos respondentes para os dois seguintes tipos de casos:

- ⊙ termos lembrados na primeira consulta não terem sido mencionados na segunda consulta;
- ⊙ termos não lembrados na primeira consulta que o foram na segunda consulta.

Na maioria desses casos, o esquecimento foi atribuído ao fato de o tópico ser considerado um aspecto periférico do documento. Em consequência, devemos interpretar que os demais tópicos – aqueles presentes em ambas as consultas de forma inalterada – referem-se aos aspectos considerados fundamentais no documento. Essa distinção entre assuntos periféricos e fundamentais talvez não se encaixe, de forma perfeitamente ajustada, na teorização em torno dos conceitos de atinência intensional e atinência extensional, conforme discutido na revisão da literatura, mas sem dúvida que guarda com esses conceitos algum tipo de relação bastante consistente. Pode-se mesmo levantar a hipótese de que uma elaboração conceitual mais apurada de atinência intensional deve ser buscada nos assuntos fundamentais de um documento, assim como a de atinência extensional provavelmente estará de alguma forma ancorada nos assuntos periféricos. Em ambos os casos, da ótica do usuário.

Entretanto, parece que em alguns casos, os respondentes tiveram certa dificuldade para distinguir entre o fator generalidade e o aspecto de importância periférica. Assim, para exemplificar, o respondente que escolheu o artigo de Line, tentando explicar porque não utilizou o termo *Cooperação bibliotecária* quando da segunda coleta, diz, inicialmente, que o esquecimento talvez tenha sido devido ao fato de esse tópico ser um aspecto menos importante, ou seja, periférico. Mas diz também que se trata de um termo mais amplo em relação aos demais termos por ele utilizados para descrever o documento.

Outros tipos de divergência seriam causados pela evolução terminológica, conjecturaram os respondentes. Um caso típico é o dos sinônimos. O respondente que trabalhou com o texto de Line considerou que os termos *UAP* e *Comutação*, acrescentados da segunda vez, representam o fator sinonímia ocorrendo de forma distinta. No caso da *UAP*, o termo seria sinônimo de *Empréstimo interbibliotecário*, e depois perdeu esse caráter. Por outro lado, o termo *Comutação* também seria sinônimo de *Empréstimo interbibliotecário*, mas à época de publicação do artigo não seria ainda utilizado. O que não está claro é como essas distinções acabaram por influenciar a atribuição dos termos pelo respondente, já que nada dessa evolução ocorreu entre os dois momentos em que o respondente participou desta pesquisa. Por conseguinte, deve-se sugerir que em futuras pesquisas se trabalhe com conceitos ao invés de simplesmente com termos. Isto é, o respondente deveria não apenas identificar os termos como também os conceitos correspondentes.

Quanto aos fatores que determinam ou podem ter determinado essas divergências, o padrão que se destaca é o da associação com o trabalho que o respondente está realizando naquele determinado momento. Um respondente, por exemplo, divergiu quanto aos termos por ele próprio atribuídos ao texto de Lancaster. Da primeira vez, os termos tinham sido *Indexação* e *Resumos – redação*. Da segunda vez, os termos foram *Indexação* e *resumos*, o que mostra uma coincidência muito grande, mas com uma diferença sutil: o subcabeçalho *redação* utilizado da primeira vez, foi abandonado na segunda etapa. De acordo com o respondente, o que deve ter influenciado sua escolha do subcabeçalho *redação* foi provavelmente o fato de, à época de preenchimento do formulário, estar ele precisamente ocupado em ensinar o tema *Redação de resumos* para seus alunos. Esse fato, fez com que desse ênfase especial ao aspecto de redação de resumos que, no livro, é apenas mais um entre tantos outros aspectos tratados.

## Conclusão

Os resultados desta pesquisa levam a algumas conclusões preliminares quanto à percepção que o grupo particular de usuários estudado tem do conteúdo de um documento. A conclusão mais geral é que essa visão se altera com o tempo, variando de nuances até mudanças bastante significativas. Fatores como a importância dos tópicos, graus de generalidade e especificidade e o tipo de trabalho que o usuário está desenvolvendo num determinado momento parecem guiar a decisão dos usuários quanto ao que constitui o conteúdo de um documento, pensado em termos de seus tópicos mais importantes e expresso na forma de termos ou palavras-chave.

Ressalta nesses resultados uma dinâmica na visão do usuário que contrasta de forma marcante com a forma estática como os documentos são representados nos sistemas de informação e de recuperação da informação. Pode-se dizer que essa dinâmica é algo perfeitamente natural, dada a própria natureza das atividades de uso de informação e geração de conhecimento. Por outro lado, fatores importantes de ordem econômica e financeira impõem muitas limitações à estruturação e funcionamento dos mencionados sistemas de informação. O que se pode conjecturar é que, apesar dessas condições aparentemente difíceis, talvez seja possível pensar-se em estratégias criativas de adaptação dos sistemas de informação à natureza dinâmica do

comportamento de busca de informação dos usuários. É certo que uma tal estratégia terá que contar com avanços na tarefa de análise de assunto para atingir o nível preconizado de adaptabilidade. Esperamos que a pesquisa sobre análise de assunto possa colaborar no sentido de que esses avanços sejam concretizados e, por conseguinte, para a formulação e implementação de uma tal estratégia.

Finalmente, devemos observar que o estudo mostrou ser perfeitamente viável à metodologia experimentada, podendo ser aplicada em amostras mais estruturadas, de forma que se possa investigar os comportamentos das várias categorias de usuários, em seus variados tipos de ambientes sistêmicos e em diferentes áreas do conhecimento. Alguns problemas surgidos durante a condução da pesquisa foram identificados e procedimentos foram sugeridos no sentido de enfrentamento desses problemas em pesquisas futuras.

## Agradecimentos

Agradecimentos são devidos a Demócrito A.N. Manyissa e Andréa Aparecida Ribeiro, alunos do curso de graduação em biblioteconomia da Escola de Ciência da Informação da UFMG que, na condição de bolsistas dos programas de iniciação científica da UFMG apoiados pela FAPEMIG e pelo CNPq, participaram deste projeto em diferentes momentos de sua realização.

## Subject analysis: user's perception of document content

*The study employs a longitudinal design to understand, from the perspective of subject analysis, the users' perception of document content. Data were collected at two different moments. The results show some convergence but also divergence of the users at the two moments, as well as the interpretation of the users for these outcomes. The results suggest standards of classification of the topics identified in the documents in fundamental/peripheral and generic/specific. They also suggest that nature of the work being carried out by the user in a given moment as well as terminology evolution may be important factors for explaining the divergences at the two different moments of data collection.*

**Key-words:** *Subject analysis; Users; Concept identification*

## Referências:

BABBIE, E. *The practice of social research*. 3rd ed. Belmont, CA: Wadsworth, 1983.

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

BOYCE, B. Beyond topicality: a two-stage view of relevance and the retrieval process. *Information Processing and Management*, v. 18, p. 105-109, 1982.

- BRAAM, E.; BRUIJ, J. Quality of indexing information: authors' view on indexing of their articles in Chemical Abstracts online CA-file. *Journal of Information Science*, v. 18, n. 5, p. 399-408, 1992.
- BRIET, Suzanne. *Qu'est-ce que la documentation?* Paris: Éditions Documentaires, Industrielles et Techniques, 1951.
- CHU, C. M.; O'BRIEN, A. Subject analysis: the critical first stage in indexing. *Journal of Information Science*, v. 19, n. 6, p. 439-454, 1993.
- DERVIN, B.; NILAN, M. Information needs and uses. *Annual Review of Information Science and Technology*, v. 21, p. 3-33, 1986.
- DIAS, Eduardo Wense. Contexto digital e tratamento da informação. *DataGramaZero*, Rio de Janeiro, v. 2, n. 5, out. 2001. Publicação eletrônica disponível em [www.dgz.org.br](http://www.dgz.org.br).
- DIAS, E. W., NAVES, M. M. L., MOURA, M. A. O usuário-pesquisador e a análise de assunto. *Perspectivas em Ciência da Informação*. Belo Horizonte, v. 6, n. 2, p.205-221, 2001.
- FAIRTHORNE, R.A. Content analysis, specification, and control. *Annual Review of Information Science and Technology*, v. 4, p. 73-109, 1969.
- FARROW, John F. All in the mind: concept analysis in indexing. *The Indexer*, v. 19, n. 4, p. 243-247, Oct. 1995.
- FUJITA, M. S. A leitura do indexador: estudo de observação. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 4, n. 1, p. 101-116, jan./jun. 1999.
- LANGRIDGE, D. *Subject analysis*. London: Bowker-Saur, 1989.
- MARTÍNEZ, A. M.; PICHININI, M. Indización por materia de monografías en los catálogos en línea de bibliotecas generales. *Investigaciones bibliotecológicas*, México, v. 9, n. 19, p. 3-11, jul./dic. 1995.
- NAVES, M. N. L. 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.
- SAARTI, Jarmo. Taxonomy of novel abstracts based on empirical findings. *Knowledge Organization*, v. 27, n. 4, p. 213-220, 2000.
- SVENONIUS, Elaine. *Subject analysis*. [Belo Horizonte: Escola de Biblioteconomia da UFMG, 1995]. Apostila distribuída durante curso ministrado na Escola.
- TAYLOR, A.G. On the subject of subjects. *The Journal of Academic Librarianship*, v. 21, n. 6, p. 484-491, Nov. 1995.
- TODD, R. T. Academic indexing: what's it all about? *The Indexer*, v. 18, n. 2, p. 101-104, Oct. 1992.
- TURNER, J. Comparing user-assigned terms with indexer-assigned terms for storage and retrieval of moving images: research results. In: AMERICAN SOCIETY FOR INFORMATION SCIENCE ANNUAL MEETING, 58., 1995, Chicago. *Converging technologies; forging new partnerships in information*. Medford, N.J.: Information Today for American Society for Information Science, 1995. v. 32, p. 9-12.

## Apêndice - Documentos escolhidos pelos respondentes

- BARRETO, Marco Heleno. Subjetividade e o novo na arte: reflexões a partir de Adorno. *Kriterion*, Belo Horizonte, n. 85, p. 49-58, jan./jun. 1992.
- BUITONI, Dulcília H. S. Jornalismo: o tecido e o acontecido. *Revista USP*, São Paulo, n. 6, p. 175-182, jun./ago. 1990.
- CAMPELLO, B. S.; CAMPOS, C. M. *Fontes de informação especializada*. 2.ed. Belo Horizonte: Ed. UFMG, 1993.
- CESARINO, M. A. N.; PINTO, M. C. M. F. Análise de assunto. *Revista de Biblioteconomia de Brasília*, v.8, n.1, p.32-43, jan./jun. 1980.
- DALBEN, Ângela I.L.F. Reprovação, avaliação, escola plural. *Educação em Revista*, Belo Horizonte, n. 20/25, p. 132-135, jun. 1997.
- LANCASTER, E.W. *Indexação e resumos*. Brasília: Briquet de Lemos, 1993.

LINE, Maurice. Diretrizes para sistemas nacionais de empréstimo interbibliotecário. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v.10, n.1, p.13-26, mar. 1981.

MARSON, Adalberto. A Locomotiva e a célula: imagens opostas da mesma revolução: Rússia 1918. *Revista Brasileira de História*, São Paulo, v. 10, n. 20, p. 129-146, mar./ago. 1990.