

**PRINCÍPIOS DE ORGANIZAÇÃO E REPRESENTAÇÃO DE  
CONCEITOS EM LINGUAGENS DOCUMENTÁRIAS**  
*PRINCIPLES OF ORGANIZATION AND REPRESENTATION CONCEPTS IN  
DOCUMENTARY LANGUAGES*

Edmeire Cristina Pereira  
[edmeirepereira@pop.com.br](mailto:edmeirepereira@pop.com.br)

Professora Assistente II  
Departamento de Ciência e Gestão da Informação  
Universidade Federal do Paraná (UFPR)

Leilah Santiago Bufrem  
[leilah@ufpr.br](mailto:leilah@ufpr.br)

Professora Titular  
Departamento de Ciência e Gestão da Informação  
Universidade Federal do Paraná (UFPR)

## **RESUMO**

Análise dos princípios e métodos de organização e representação de conceitos para as Linguagens Documentárias (LD), enfocando-se principalmente a Teoria da Classificação Facetada (Ranganathan), a Teoria Geral da Terminologia (Wüester) e a Teoria do Conceito (Dahlberg), considerada sua importância para o avanço dos estudos em Análise Documentária (AD). Propõe uma síntese de princípios/características para a AD. Conclui que a AD comporta em si dois sentidos: o de método interpretativo e o de teoria, cujo conhecimento é produzido a partir do método.

**PALAVRAS-CHAVE:** Análise Documentária (AD). Linguagens Documentárias (LD). Informação Documentária. Terminologia. Lingüística. Semiótica. Classificação Facetada.

## **1 INTRODUÇÃO**

Vive-se uma época marcada por intensas transformações, cuja tônica repousa na informação e nos benefícios dela advindos, especialmente impulsionados pela revolução tecnológica. Nesta chamada “Sociedade da Informação”, “Economia da Informação”, “Sociedade do Conhecimento”, configura-se uma estrutura social associada a um modo de desenvolvimento emergente, organizado em forma de redes. A informação surge como insumo e produto dos mais valorizados principalmente porque, a partir de meados dos anos de 1980, ela começou a ser encarada como “estratégica”, notadamente, pelos países desenvolvidos, embora, a literatura econômica internacional já viesse tratando do assunto desde a década de 1960.

Referindo-se ao cenário mundial, Lastres aponta o alcance das transformações tecnológicas, organizacionais, geopolíticas, informacionais, comerciais e financeiras e suas inter-relações com os âmbitos institucionais, culturais e sociais na atualidade. (1999, p. 72).

Com efeito, todas essas transformações culminaram em um “novo” período na história do capitalismo, denominado por alguns de globalização. É um período de “transição” em todos os sentidos, apesar de que acontecimentos políticos mundiais, como as tentativas drásticas de quebra da hegemonia norte-americana, possam significar para cientistas políticos, economistas e historiadores o marco para o encerramento de um ciclo marcado pelo capitalismo selvagem.

Mas, o que fica evidente nos dias que correm é que “as bases significativas da sociedade, espaço e tempo estão sendo transformadas, organizadas, em torno do espaço de fluxos e do tempo intemporal” (CASTELLS, 1999, p. 504). Nesse cenário, em que a informação e o conhecimento dela resultante são as novas “moedas correntes”, as comunicações instantâneas ocorrem em velocidade sem precedente na história, provocando uma expansão dos seus domínios. Em decorrência, a discussão de questões ligadas a princípios e métodos de organização e representação de conceitos em Linguagens Documentárias (LD) converge para o entendimento de que, mesmo com a velocidade atual das comunicações (impressas ou digitais), não é possível que os profissionais da informação, principalmente os documentalistas que estudam e trabalham com Documentação e Lingüística, fiquem à margem do contexto em que interagem e renovam suas práticas. Mesmo porque a intenção de aperfeiçoamento dos saberes na prática é fundamental para esses profissionais devido à sua intermediação entre pessoas e sistemas de significação e também, pelo fato de que a indexação automática, embora tenha progredido rapidamente, ainda não tenha alcançado os avanços proporcionais à quantidade desmedida de documentos a tratar. Assim, a tentativa de esclarecer os fundamentos teóricos da organização e representação de conceitos para melhor compreensão da percepção dos profissionais sobre suas bases lingüísticas, sociais e culturais poderá favorecer os modos de representação do conhecimento e aperfeiçoar a práxis, pela intermediação entre pessoas(sujeitos)/usuários-clientes e processos.

## **2 LINGÜÍSTICA E DOCUMENTAÇÃO: UMA RELAÇÃO NECESSÁRIA**

A aplicação de teorias contemporâneas de campos científicos limítrofes às práticas em documentação é uma possibilidade que vem despertando interesse dos intelectuais na área. No Brasil, um dos primeiros trabalhos a relacionar Lingüística e Documentação foi o de Wanderley (1973). A partir de 1980, Kobashi, Smit, Lara, Cintra, Tálamo e Cunha são as autoras mais produtivas no campo da representação documentária (DODEBEI, 2002, p. 41).

Com enfoques diversificados, o tema prosperou com Baranow (1983), Medeiros (1986), Miranda (1989), Boulanger (1995), Faulstich (1995 e 1995 a) e Peixoto (1995), que tratam da formação do especialista e cientista da informação, a partir das contribuições da Lingüística e dos Estudos Sociais para a Terminologia e a Socioterminologia.

Mendonça, em estudo bibliométrico sobre a produção de trabalhos que relacionam a Lingüística e as Ciências da Informação, analisa 42 artigos, selecionados em 27 volumes e 59 fascículos da revista *Ciência da Informação* (Brasília), veículo da área que mais publicou artigos sobre o tema, no período de 1972 (lançamento da revista) até o ano de 1998. Após analisar os sete pontos em que ocorre a interseção da lingüística com a Ciência da Informação: o teórico; o quantitativo (pela visão bibliométrica); o temático (pela representação da informação); o aplicativo (pelos métodos diversos); o ensino (pelas relações curriculares); o tecnológico (pela teoria de sistemas) e o normativo (pelas classificações bibliográficas), a autora destaca que “uma das grandes problemáticas reveladas pela pesquisa foi a construção de conceitos e a representação da informação”. A terminologia e a análise documentária são apontadas como as áreas de pico nos estudos mais recentes (2000, p. 50).

Estudos sobre modelagem de domínios de conhecimento têm se destacado na literatura como o de Campos e Campos (2003) que analisam questões teóricas/metodológicas para comparar os mecanismos de abstração presentes em teorias da Ciência da Informação, Ciência da Computação e Teoria da Terminologia. Da Ciência da Informação, analisaram as teorias da Classificação Facetada de Ranganathan e do Conceito de I. Dahlberg, voltadas mais especificamente à representação de sistemas de conceitos e que possibilitam a representação de domínios de conhecimento. Da Ciência da Computação, investigaram os modelos representacionais associados à modelagem de sistemas de bancos de dados, mais especificamente o Modelo Orientado a Objetos (Gray, 1992; Furlan, 1998; Rumbaugh, 1994) e a Ontologia Formal (Guarino, 1998, 1998 a, 1994, 1997; Gruber, 1993; Sowa, 2000), campo que repensa as possibilidades representacionais e de organização de domínios de conhecimento. Da Teoria da Terminologia de E. Wüester (1981) estudaram os princípios estabelecidos para a determinação de conceitos e suas relações.

Ao compararem os mecanismos de abstração presentes nessas teorias e, apoiando-se em Le Moigne, as autoras apontam princípios subjacentes ao processo de modelização, ou seja, “um núcleo comum de conceitos imprescindíveis ao ato de modelar o conhecimento” (2003, p.135). Os “movimentos do ato de modelar” refletem grupos de relações entre conceitos, sendo o primeiro deles a verificação da existência de relações categoriais; o segundo, a verificação dos modos de relação entre objetos de mesma natureza; o terceiro, a

Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005

análise de como “o objeto se constitui”, ou seja, quais são suas partes e elementos; o quarto, em que se verifica como objetos de natureza diferente se relacionam e a forma mais consistente de representar esse relacionamento, a partir da determinação de alguns critérios prescritivos e o quinto verifica um dado tipo de relação que não mais se constitui entre conceitos, mas entre eles e a forma de expressá-los, ou seja, se dá no âmbito da língua, a chamada relação de equivalência.

Pode-se concluir com Campos e Campos (2003, p.143), em prol de nossa hipótese de trabalho, que o diálogo necessário entre a Ciência da Informação, a Ciência da Computação e a Terminologia não prescinde da identificação e compreensão dos fundamentos teóricos dos modos de organizar o conhecimento, do contexto em que se realiza a percepção humana, das bases lingüísticas, sociais e culturais e das estruturas pelas quais se realiza a representação do conhecimento. Assim, ao se propor a “sair do modelo, ultrapassar o método e dominar os meta-conceitos, que devem ser utilizados como mecanismos de abstração” o profissional da informação estará intencionalmente aberto à pesquisa para construir conhecimentos em um campo que integra questões de natureza epistemológica, lógica, cognitiva, terminológica e tecnológica. Esse exercício de produção intelectual ainda não foi superado pela tecnologia, especialmente no que concerne à representação documentária, processo cuja complexidade não pode ser reduzida à intervenção eletrônica.

### **3 REPRESENTAÇÃO DOCUMENTÁRIA**

Resultado de uma prática iniciada com a análise do texto, a representação documentária identifica conteúdos pertinentes e em consonância com os propósitos e características da unidade ou sistema a que serve. A síntese e a representação advindas desse processo balizam-se por critérios como a padronização e a univocidade, qualificadores das principais formas ou produtos que o representam: o resumo e o índice. Considerada um modo de intermediação entre emissor e receptor, a representação é concretizada pelos processos e produtos da condensação de conteúdos informativos e ilustrada pelos componentes documentais que, segundo mapeamento de Dodebei (2002, p. 42-43), traduzem-se em objetos, processos, produtos, instrumentos, usos e campos teórico-metodológicos.

Pode-se entender, sob outro enfoque, a representação como re-apresentação, em que analista, documento e sistema prescindem de mediador, uma vez que partícipes do processo de resignificação. É por ele que a palavra, utilizada livremente, como na linguagem natural, ou controlada por um sistema que a legitima para representar conceitos, converte-se em fim e

Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005

meio para o autor, o leitor e o analista. Seus desdobramentos definem o campo da análise e representação documentais do conteúdo dos textos, seja numa frase, num fragmento de um diálogo, em textos ou contextos específicos.

Mas, mesmo que “não limitado às fronteiras da linguagem verbal, no plano semiótico, de sentido multidimensional, texto ou discurso é sinônimo de processo que engloba as relações sintagmáticas de qualquer sistema de signos. Pode-se então falar de texto ou discurso cinematográfico, teatral, coreográfico, pictórico etc” (GUIMARÃES, 1990, p.14-15). Complementando as duas acepções de texto, quais sejam, a de texto como “enunciado” e a de texto como sinônimo de “discurso”, pode-se denominá-lo, também, seguindo McGarry, um “tecido intelectual de signos”, dotado de uma unidade coerente, enquanto o documento seria o “substrato físico, bem à maneira da computação neural, que considera a mente como o software e o cérebro como o substrato material” (1999, p. 128).

Assim, ao se adotar a noção de documento como “qualquer unidade, impressa ou não”, passível de catalogação ou indexação (NBR-12676/92), se compreende a possibilidade da representação de conteúdos informacionais. Essa, por sua vez, concretiza-se, em plano real, fundamentada nas operações de análise e síntese, que se verificam não apenas em plano lógico, mas também num contexto ontológico. Isso acontece na prática quando são criados códigos de classificação bibliográficos, tesauros, índices de sistemas de informação e resumos (*abstracts*), produtos documentários que podem ser considerados “modelos” de representação de realidades, as mais diversas.

Por definição, temos que os resumos (*abstracts*) têm sua representação por paráfrases, caracterizadas pela contigüidade e semelhança estrutural com o texto original, enquanto a representação dos índices se realiza por tradução: formulação em linguagem que não se confunde com a linguagem do texto; mesmo que os termos apresentem a mesma forma. Podem constituir-se de palavras-chave, cujos elementos são extraídos do título ou do texto; notações classificatórias, ou símbolos de um sistema de classificação de relações hierárquicas ou não-hierárquicas; cabeçalhos de assunto, expressões de listas alfabéticas parcialmente estruturadas e descritores, ou unidades de um tesauro, com relações hierárquicas, não-hierárquicas e de equivalência. Comumente usado para buscas especializadas e não para a indexação, propriamente dita, o tesauro é ferramenta imprescindível em sistemas automatizados de recuperação da informação. O termo *thesaurus*, usado durante muitos séculos para designar léxico ou “tesouro de palavras”, popularizou-se a partir da publicação do *Thesaurus of English Words and Phrases*, de Peter Mark Roget, em Londres, em 1852, cujo subtítulo clarifica o objetivo do instrumento: “palavras classificadas e

Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005

arranjadas para facilitar a expressão das idéias e para ajudar na composição literária”. Ao contrário dos tradicionais dicionários de língua, nos quais se parte de uma palavra para saber os significados que ela pode ter, no Thesaurus de Roget parte-se de um significado, de uma idéia, para se chegar às palavras que melhor a representem. Já, na década de 1970, o Programa UNISIST (DIRETRIZES, 1993, p. 6) define “tesauro” para a área de Ciência da Informação, sob dois aspectos: segundo a estrutura, como “um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento” e segundo a função, como “dispositivo de controle terminológico usado na tradução de linguagem natural dos documentos, dos indexadores ou dos usuários numa linguagem do sistema”. Esta definição vem sendo utilizada na literatura da área até os dias atuais (CAMPOS, 2001, p.90-91).

Cabe destaque, em prol de nossa hipótese de trabalho, a relevância do caráter interpretativo desse dispositivo que sintetiza em cada descritor um leque de significações, pelo que depende da compreensão do contexto, das bases lingüísticas, sociais e culturais e das estruturas pelas quais se realiza a representação do conhecimento. Mas, exatamente devido as suas funções, os tesouros têm sido considerados “formas de controle do discurso”, enquanto os descritores “são sentidos cristalizados, homogeneizados, sedimentados”. A tarefa de indexar implicaria, portanto, em circunscrever os sentidos, prescritos no tesauro – grade interpretativa (LUCAS, 2000, p.71).

Essa grade interpretativa, considerada um sistema de conceitos ou unidades de conhecimento, por sua vez, denotados pelos termos e explicados pelas definições, também se relaciona a conjuntos mais abrangentes de idéias ou conceitos, denominados categorias, cuja estrutura faz parte da Teoria da Classificação e da Teoria Geral de Terminologia. Nessa estrutura destacam-se as relações lógicas (relação genérico-específica, relação analítica e relação de oposição), as relações ontológicas (relação partitiva, relação de sucessão e relação material-produto), as relações de efeito (de causalidade, instrumental e de descendência) e as relações de equivalência. Portanto, a trajetória metodológica da construção dos tesouros requer, por um lado, o reconhecimento dessa estrutura complexa de conceitos e relações e, por outro, a capacidade de controlar ou reordenar significados com base em conhecimentos lógicos e filosóficos, padrões e informações, de modo a permitir ao usuário (indexador ou pesquisador) encontrar os termos mais significativos e relevantes para representar a idéia ou o conceito procurado.

#### **4 HIPÓTESES DE ORGANIZAÇÃO DOS INSTRUMENTOS DE REPRESENTAÇÃO DA INFORMAÇÃO**

As formas textuais, narrativa, descritiva ou dissertativa, são padronizadas em decorrência do propósito e ponto de vista de sua origem. Textos jornalísticos, por exemplo, voltam-se à produção de informações para notícias; dissertações e teses são textos argumentativos, indexados a partir de suas idéias diretrizes ou hipóteses; ensaios, a partir das idéias que defendem e textos descritivos revelam características do objeto que enfocam, sendo o objeto o mais importante para o indexador. Assim, se a estrutura dos textos é de fundamental importância na prática de tratamento, análises/sínteses e representações, ela está intimamente vinculada ao conteúdo e propósito do objeto analisado. A responsabilidade profissional do gestor de documentos ou de informações, em qualquer dos contextos em que atue, será de revelar o conteúdo na estrutura.

Mas, se relativamente à Ciência da Informação os textos são representados em resumos e índices, enseja-se a consideração da análise documentária e do seu produto como resultantes da ideologia do autor/produtor da área. Fundamentado na desmistificação do “bom senso” bibliotecário e seus pré-conceitos relativos à análise documentária, Cunha atribui grande parte dos erros da representação e transferência da informação ao pressuposto de que o texto/discurso/documento é absoluto e passível de uma única leitura e, portanto, de uma única representação (1990, p. 72). Daí a proposta de, aos elementos da análise de texto, incorporarem-se os de análise do discurso (p. 139), com as suas respectivas representações. Reafirma-se, dessa forma, que a análise documentária e o seu produto/informação se encontram irremediavelmente ligados ao conhecimento da área, tanto do produtor como do consumidor. Para a autora, entretanto, a Biblioteconomia e a Documentação passam aleatoriamente pelo texto e pela adequação ao consumidor, este visto de forma assistencialista, pelo que conclui que “controlar o vocabulário de uma língua e de uma cultura é servir o ‘senhor da língua’ e o ‘produtor’ da cultura” (p.145).

Ao questionar a leitura do bibliotecário como mera adequação a métodos de tratamento dos textos, Lucas, por sua vez, refere-se à divisão de trabalho, apontada por Pêcheux (1994), pela qual os leitores de arquivos atuam conforme duas culturas – a literal e a interpretativa (2000, p. 50). Nessa linha de raciocínio, a autora infere que o trabalho do bibliotecário encaminha-se de modo a literalizar sentidos, na crença de que a língua reflete o pensamento com exatidão e de que a sua materialidade tem sido subestimada pelos teóricos da indexação, desde abordagens lingüísticas (Cintra, 2002; Cunha, 1990; Lara, 2002), chegando

às de tratamento informático (Baranow, 1983; Brito, 1992). Acredita a autora que o método de leitura do bibliotecário, “não garante a equivalência de sentido entre o texto-fonte e a sua representação (os termos selecionados como assunto para a teoria da indexação). As operações de análise e síntese para fins de Análise Documentária estão sujeitas a muitos outros fatores para além daqueles que uma metodologia possa disciplinar, administrar, conter” (2000, p. 63). Nesse esforço crítico sobre a teoria e a prática de leitura do bibliotecário, Lucas chega a algumas constatações contundentes: o discurso que descreve a leitura do bibliotecário procura caracteriza-la como científica, logo neutra e apolítica; a leitura do bibliotecário simula o modo de produção industrial, buscando produtividade, rapidez e não dando margem à reflexão e ao acúmulo de conhecimento por parte do bibliotecário. Nesta leitura, que se deseja rigorosa e transparente, o sujeito-leitor não deve interpretar. Ela deve ser literal, apreendendo o conteúdo do texto e produzindo representações ou simulacros do mesmo, dando-lhe unidade por meio de palavras-chave. Deve ser eficaz, administrável, controlada por meio de treinamentos rigorosos, obedecendo sempre a regras de objetividade. Enfim, para a autora, “a leitura do bibliotecário não visa à produção de conhecimentos, à crítica; visa, sim, ao pré-tratamento dos documentos para os leitores das bibliotecas”, da mesma forma que a divisão do trabalho de leitura da Idade Média, entre aqueles indivíduos autorizados a ler, falar e escrever em seus nomes (os cientistas, os intelectuais, os críticos) e os outros (classificadores, indexadores, codificadores) leitores anônimos, cujo trabalho visava sustentar a leitura dos primeiros (2000, p. 85-86).

Ao enfocar o processo de análise documentária, Kobashi (2001) distingue entre o uso de estoques de informação e a capacidade de comunicar, para que se evite a perda de “informação semântica”, devida ao excesso de generalização. Para a autora, o processo é planejado sempre a partir de hipóteses de trabalho, onde se interrelacionam: 1 – esfera da produção (regras de produção documentária); 2 – esfera do consumo (identificação das necessidades dos usuários e as possibilidades de aderência ao sistema documentário); 3 – esfera teórico-pragmática (identificação do objeto da Análise Documentária, suas funções, métodos e metodologias formuladas a partir de parâmetros teóricos e pragmáticos).

Por sua vez, ao argumentar que a informação documentária não existe à priori, Lara a considera como o resultado da segmentação de conteúdos, realizada a partir de hipóteses de organização (2002, p.127), do mesmo modo que as linguagens da biblioteconomia e da documentação partem de hipóteses, pontos de vista que estão na base das formas de organização dos domínios visados. Tais formulações variam histórica e funcionalmente, conforme os pressupostos considerados e os objetivos da classificação: a) os sistemas de Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005

classificação bibliográfica universal, como o *Dewey Decimal Classification* e a Classificação Decimal Universal, têm sua origem na proposta baconiana de divisão do conhecimento, considerada, durante muito tempo, um parâmetro estável para organizar o universo bibliográfico...; b) as listas de Cabeçalhos de Assunto como a *Library of Congress Subject Headings* – LCSH – pautam-se na garantia literária (presença dos termos na literatura) para inventariar sugestões de entradas de assunto, porém, esse procedimento não caracteriza efetivamente uma proposta de segmentação assente num princípio organizador, razão pela qual elas não formam um conjunto uno (2002, p.133). Essa garantia é aleatória e não chega a configurar uma hipótese de organização, do que decorrem dificuldades para o engendramento de mecanismos inferenciais; c) a recuperação de textos feita com base na frequência e na ocorrência de termos também não caracteriza uma segmentação do tipo classificatório, já que o recorte é pulverizado para as palavras retidas no processo de extração (...); d) os tesouros representam uma das formas mais consistentes de apresentar uma proposta de organização de um domínio, já que são formulados segundo princípios lógico-semânticos através dos quais é possível constituir um todo significativo. (...) Todavia, as normas não dizem como encontrar as referências para sedimentar o processo de identificação e de relacionamento entre os termos, razão pela qual há muitos tesouros que o são apenas no nome, pois não formulam claramente os vértices a partir dos quais devem ser organizados os termos; e) os tesouros terminológicos são instrumentos que usam simultaneamente princípios de organização de tesouros e de organização de sistemas de conceitos (2002, p.134). A delimitação de domínios e a operação de seleção dos termos que lhe são próprios fundamentam-se, pois, na terminologia teórica e nas suas aplicações concretas.

## **5 A ARCHÉ DAS LINGUAGENS DOCUMENTÁRIAS**

Assim como os pré-socráticos procuravam o princípio (a *arché*) de todas as coisas, conclui-se que a tentativa de esclarecer os modos de representação do conhecimento passa pela compreensão dos princípios, fundamentos teóricos e elementos constitutivos desse campo específico do saber. Se toda ciência é regida por princípios ou leis, para fundamentar suas práticas, o especialista em Análise Documentária precisa dominar alguns conceitos e conhecer seus princípios básicos.

A explicitação desses conceitos e princípios, portanto, passa pela discussão sobre os modelos de organização do conhecimento que deve acontecer, segundo Dodebei, a partir das interfaces dos conceitos de Conhecimento, Representação, Informação e Memória. O papel

Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005

das Linguagens Documentárias, nesse campo de atuação, corresponde ao de metalinguagens na produção de conhecimentos e as “memórias documentárias”, segundo a autora, são *constructos* do conhecimento gerado pela sociedade (2002, p.19).

Quanto aos princípios em favor dessa construção, para que se tornem amplamente aceitos, devem ser definidos a partir da premissa básica ou princípio geral dos sistemas de recuperação da informação (SRI): possibilitar a seu usuário o acesso à informação/documentos/memória, prevendo-se sua viabilidade e utilidade.

Para que assim seja, a metarrepresentação ou representação documentária é explicitada pelos modelos do Ciclo de Vida da Informação e do Sistema de Recuperação da Informação (SRI) (DODEBEI, 2002, p.37-43).

Mas, se a menção a modelos na área evoca a idéia de paradigmas, tais como os ecológicos, os locacionais (geográficos), os morfológicos, antropológicos/etnológicos, evolucionistas e sistêmicos, entre outros, também pressupõe a existência de uma teoria subjacente que os fundamenta e os tipifica basicamente como físicos/concretos ou mentais/lógicos. Assim, dentre os primeiros, destacam-se os modelos em escala real ou reduzida, dinâmicos ou estáticos e, dentre os segundos, devem-se mencionar os modelos lógico-conceituais, os matemáticos e estatísticos, os gráficos, os sistêmicos, os semânticos e as simulações por computador, para mencionar apenas os mais utilizados no atual contexto (DODEBEI, 2002, p.21). Em Ciência da Informação, segundo a literatura, é o paradigma sistêmico, derivado da “teoria geral dos sistemas” de Von Bertalanffy (1973), o mais empregado e o modelo representativo da comunicação entre dois pólos – o emissor e o receptor - ao longo dos anos, tem sido o de Shannon e Weaver (1949). Em resumo, a Ciência da Informação absorve ou “importa” modelos que, de ângulos distintos, descrevem a sua realidade, razão pela qual tem sido denominada de uma ciência multiparadigmática (SAYÃO, 2001, p. 83).

Para estudar as representações documentárias, o modelo também de caráter sistêmico denominado “ciclo da informação”, ou modelo de transferência da informação, estrutura a representação do conhecimento, em seis etapas: produção, registro, aquisição, organização, disseminação e assimilação. Com algumas modificações à proposta original de Lancaster (1979), Dodebei (2002, p. 25) incorporou o conceito de “memória documentária”.

A estrutura da memória documentária é descrita por Montgomery, citado por Dodebei (1969, p.145-171) em um modelo de concepção sistêmica, denominado “Sistema de Recuperação da Informação” (SRI), no qual os dados de entrada são, ao mesmo tempo, as representações das demandas de informação pelo público e as representações do acervo ou

Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005 30

coleção. O sistema processa essas representações, comparando-as e devolvendo-as sob a forma de novas representações direcionadas a cada demanda. A retroalimentação é garantida pelas novas produções de conhecimento que serão selecionadas para o ingresso no sistema de informação, o que se pode interpretar como um filtro, que prioriza as informações demandadas, nem sempre as mais necessárias.

Campos (2001), ao analisar a Teoria da Classificação Facetada (RANGANATHAN), a Teoria Geral da Terminologia (WÜESTER) e a Teoria do Conceito (DAHLBERG), consideradas as mais representativas da área, aponta seus princípios comuns, cuja síntese concede destaque aos aspectos considerados aqui esclarecedores ao profissional da informação na sua prática em AD.

Em primeiro plano, a indissociabilidade entre conceitos e termos é determinante para a compreensão do processo e do primeiro aspecto comum às três teorias, a forma de abordagem do termo, seja no esquema de classificação, no tesouro ou na terminologia. Os termos são tomados com um significado próprio, dado pelo contexto. A incorporação do referente representado pelo termo é significativa da denominada abordagem onomasiológica (2001, p.117). Como os termos são tomados com um significado próprio, num dado contexto, destaca-se a indissociável relação entre linguagem, pensamento e realidade. Esta evidência se realiza via conceitos, e não entre palavras, dada a relação direta entre o conceito e o termo, isto é, um conceito é representado por um termo (símbolo) e este termo é usado para designar aquele conceito. Wüester incorpora essa idéia ao afirmar que o conceito é uma unidade de pensamento. Do mesmo modo, Ranganathan identifica notação e termo. Porém, nos tesouros terminológicos, com a adoção da Teoria do Conceito, pela qual este é visto como uma tríade referente-características-termo, essa relação ocorre de forma mais consistente (2001, p.118). Outro aspecto, relativo ao ideal inatingível da monossemia absoluta, ou seja, a designação de um conceito único por um termo, apresenta-se como uma exigência da terminologia e do tesouro. Evidentemente, este ideal é concretizável apenas dentro de uma área do conhecimento ou de uma língua especial, pelo que se considera essa monossemia como relativa ou seja, aceita-se a univocidade relativa (2001, p.119). Quanto ao conceito de “termo”, nos tesouros tradicionais, ele é definido como uma palavra ou grupo de palavras que representa o conceito. Da mesma forma, na Teoria Geral da Terminologia, o termo é um representante do conceito. Entretanto, ainda que se procure definir o conceito como uma unidade de pensamento, essa definição não resistiria a uma comprovação científica do tipo positivista, porque o pensamento refere-se aos processos subjetivos diferenciados e individuais. Na Teoria do Conceito, este é considerado uma unidade de conhecimento,

Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005

constituído da tríade referente-características-termo. Se o termo denota o conceito, decorre que sua definição adquire um papel relevante, pois fixa o conteúdo do conceito e o significado do termo. Tesouros e terminologias têm um requisito comum – a precisão do termo - porque um Sistema de Recuperação da Informação também é um Sistema de Comunicação, razão pela qual, a autora considera que até o momento, a Teoria do Conceito oferece o melhor suporte teórico-metodológico para a representação/recuperação da informação (2001, p.120). Possibilita a representação do conhecimento, uma vez que apresenta, como um dos elementos da tríade, as características do conceito. Com isso, se permitem comunicações mais precisas em Ciência e Tecnologia (p.121).

O segundo princípio a se destacar é a importância das características do conceito e seu uso. Se os conceitos são inter-relacionados graças as suas características comuns, essas se destacam para a compreensão do tipo de relações e do seu posicionamento no sistema de conceitos. Ranganathan desenvolveu uma série de cânones para a identificação dessas características comuns, a formação de renques e cadeias, destacando-se a sucessão das características no sistema de conceitos. Já para a Teoria Geral da Terminologia, um de seus princípios é que o termo é visto como um membro de um sistema e não como um objeto isolado. Neste caso, os termos se definem uns em relação aos outros. Os tesouros terminológicos utilizam as características como um elemento essencial para o estabelecimento das relações entre os conceitos e para a formação de definições que são parte integrante deste novo tipo de instrumento (p.122).

O terceiro aspecto a se destacar refere-se às relações apresentadas nas três teorias, não explicitadas ou denominadas da mesma forma. A Teoria da Classificação e a Teoria da Terminologia consideram relações lógicas as relações genérico-específicas que se formam a partir do agrupamento dos conceitos em renques e cadeias. Na Teoria do Conceito, esta relação é denominada de relação paradigmática-material, incluindo a relação parte-todo. A Teoria da Terminologia inclui a relação parte-todo como um tipo de relação ontológica, que se divide em dois tipos: relações de causalidade e relações de contacto. A Teoria da Terminologia e a Teoria do Conceito apresentam, cada uma, uma fundamentação bem estabelecida para a determinação das relações que estão presentes entre os conceitos. Entretanto, apesar da utilização de modos de abordagem diferenciados, destaca-se a ausência de conflitos entre as modalidades de apresentação das relações entre conceitos, especialmente quando se analisam as chamadas relações hierárquicas e as relações lógicas e ontológicas.

As relações hierárquicas referem-se à precedência entre dois conceitos, o maior ou mais abrangente deve ficar acima do menor ou menos abrangente. As relações lógicas e

Enc. BIBLI: R. eletrônica de Bibl. Ci. Inform., Florianópolis, n. 20, 2º semestre de 2005 32

ontológicas, por sua vez, têm de determinar a natureza das relações que ocorrem entre eles (os conceitos) e distinguem-se pelas formas dessas relações.

A apresentação dos sistemas de conceitos parte do princípio de que eles devem estar relacionados de modo a formar um todo coeso e coerente. Uma das grandes contribuições de Ranganathan foi introduzir o pressuposto das Categorias Fundamentais ao processo de classificação. O termo categoria, utilizado para significar caráter, espécie ou natureza, tem origem remota, no termo grego *kategoría* e sua raiz confunde-se com a do verbo grego cujo significado em português é exprimir, enunciar, declarar. As categorias significariam, portanto, as diversas maneiras de expressão e da mesma forma, os modos de se manifestar um ser. O mesmo significado relaciona-se ao termo latino *praedicamenta*, por sua vez ligado ao verbo *praedicare*, para significar apregoar ou enunciar. Assim entendida, a categoria enseja a reunião dos conceitos em classes segundo a sua natureza, isto é, propriedade, entidade, processo etc. A categoria também foi desenvolvida a partir de seu caráter filosófico por Dahlberg em suas *Ontical Structures and Universal Classification* (1978). A Teoria Geral da Terminologia restringe-se simplesmente a uma classe de conceitos quando se refere ao sistema de conceitos. Pode-se dizer que, nesta Teoria, o que se chama de sistema é um sistema parcial, isto é, uma classe (2001, p. 125).

Considera-se a definição como um elemento importante para entender o conceito e posicioná-lo em um Sistema de Conceitos. A Teoria da Classificação não aborda a questão da definição e nos tesouros tradicionais, ela é tomada como um elemento que auxilia o entendimento do termo. É a Teoria da Terminologia, conforme análise de Campos, que propicia à definição o status que ela deve ter no processo de fixação do conceito e de seu posicionamento em um sistema de conceitos. Os tesouros terminológicos já incorporam esta questão, havendo referência à organização de um “novo” e único instrumento de representação e recuperação da informação, o Glossaurus, pelo qual se reuniriam um glossário e um tesouro. (2001, p. 125-126)

Da comparação entre as três teorias, fica evidente para a autora que elas se ocupam de sistemas de conceitos, embora com fins diferentes. A Teoria da Classificação Facetada dá as bases para reunir os conceitos desde cadeias e renques, passando pelas facetas, até as categorias. A Teoria Geral da Terminologia aprimora as questões referentes às relações entre os conceitos, além de dar as bases para um tratamento terminológico, e não lingüístico, ao termo. Inova, ainda, quando considera a língua numa perspectiva sistêmica. A Teoria do Conceito, além de incorporar as bases teóricas anteriores, dá especial destaque à definição.

Esta explicita as características do conceito e é utilizada como um mecanismo para posicionar o conceito nas classes, facetas e até categorias.

Campos e Campos (2003) elencam ainda como princípios o Contexto de Conhecimento, a Natureza dos Conceitos, a Relação entre os Conceitos e o Sistema de Conceitos. Acrescentaríamos, também, aos estudos acima, algumas características da análise documentária, cuja pertinência é percebida em relação ao trabalho intelectual do indexador num sentido mais amplo (QUADRO 1).

QUADRO 1 – SÍNTESE DE PRINCÍPIOS/CARACTERÍSTICAS PARA A ANÁLISE DOCUMENTÁRIA

PRINCÍPIOS	CARACTERÍSTICAS
1. Cultura/Ideologias de pensamentos	Deve-se respeitar a autonomia cultural do “outro” (de outros sistemas).
2. Discursos/Ideologias	Os conceitos/palavras-chave têm uma história e uma ideologia, nos e dos sistemas de informação documentária. Conforme CUNHA (1990, p. 22) “o bibliotecário tem o seu discurso e produz discursos quando analisa, sintetiza, representa e recupera a informação, independentemente dos seus critérios de objetividade e grau de instrumentalização”. Deve-se combater o “mito” da neutralidade do bibliotecário/documentalista e da informação. Ainda, de acordo com CUNHA (1990, p. 71): “na realidade, negar a componente ideologia em análise de texto para fins documentais é não colocar em questão o papel do indivíduo no processo de produção, consumo e tradução do todo conteúdo para o todo/conceito/palavra-chave”.
3. Leitura/Interpretação	As leituras realizadas pelos indexadores não são universais e únicas, portanto, não podem ser neutras.
4. Texto/Documento	Com relação a textos narrativos, descritivos ou dissertativos, de acordo com CUNHA (1990, p. 60), é conseguir que um mesmo texto submetido a analistas diferentes dê lugar a representações semelhantes, condição necessária não só a transmissão da informação como à validade de instrumentos como tesaurus, vocabulários controlados etc.
5. Tradução de Textos	Tradução da síntese dos textos e das perguntas dos usuários. Apreensão do sentido do texto e de sua tradução em linguagens documentárias (de linguagens naturais – LN para linguagens documentárias – LD), de acordo com os seus objetivos específicos de recuperação da informação.
6. Condensação	Faz-se por meio de metodologias. São exemplos: a) resumos (abstracts), normas e metodologias; b) índices (de palavras-chave, de notações classificatórias, de cabeçalhos de assuntos, de descritores).
7. Representação	Análises – sínteses – representações. Por meio de regras e instrumentos, tais como: Códigos, Diretrizes e Normas. Ex.: códigos de classificação; vocabulários controlados (tesaurus, CDD, CDU); linguagens documentárias (tesaurus, terminologias, listas de cabeçalhos); política do sistema.
8. Estrutura	A estrutura básica de uma LD é dada por relações hierárquicas, que podem ser genéricas, específicas ou partitivas. Também, podem ser não-hierárquicas (associativas e de equivalência). Para quem trabalha com indexação de textos, o conhecimento da estrutura textual (dissertativo, descritivo, narrativo) é fundamental.
9. Economia	Por serem sistemas construídos, as LDs são econômicas. Só que, não se trata da aplicação do princípio de economia da LN e sim de uma racionalização de escolhas e procedimentos que permitam uma utilização eficaz do sistema (CINTRA et al., 2002, p. 16).
10. Eficácia	A condição para se obter resultados positivos na busca de informação é que a pergunta e a resposta sejam formuladas no mesmo sistema (CINTRA et al., 2002, p. 39).
11. Sistema Nocional	Parâmetro básico ou “viga-mestra” de sustentação das LDs. A noção ou o conceito, define-se como “unidade de pensamento constituído por propriedades comuns a uma classe de objetos” (ISO 1087) (...) as noções são expressas por termos e símbolos, sendo influenciadas pelo contexto sócio-cultural (CINTRA et al., 2002, p. 49-50).
12. Organização	As LDs viabilizam o compartilhamento de informações produzidas por diferentes instituições.
13. Socialização da Informação	Sem uma linguagem compartilhada, não é possível a comunicação entre serviços de informação e seus usuários.

## 6 CONSIDERAÇÕES FINAIS

O quadro síntese permite encerrar uma discussão complexa sem, no entanto, exaurir-se o intento fosse uma conclusão. Esgotar o entendimento sobre os preceitos da análise documentária e seu funcionamento, seus conceitos e princípios fundamentais seria ambição desmedida.

Ao concordarmos com Campos (2001) sobre os princípios comuns entre a Teoria da Classificação Facetada, a Teoria Geral da Terminologia e a Teoria do Conceito, as mais especificamente representativas dos sistemas de conceitos, percebemo-las relacionadas aos domínios de conhecimento, no âmbito de modelos teóricos de representação.

Assim, o intento de representar o conhecimento exige a compreensão dos princípios, fundamentos teóricos e elementos constitutivos do campo específico do saber. Se toda ciência é regida por princípios ou leis, para fundamentar suas práticas, o analista deve reconhecer esses conceitos e princípios, portanto, refletir sobre os modelos de organização do conhecimento. O papel das linguagens documentárias, nesse campo de atuação, corresponde ao de metalinguagens na produção de conhecimentos e as “memórias documentárias”, de *constructos* do conhecimento gerado pela sociedade (DODEBEI, 2002, p.19).

Ao reconhecermos os níveis descritivos e de análise que introduzem o leitor às fases ou momentos do processo, mais especificamente, o reconhecimento, a redução e a representação, como analisa Moreiro González (2004), reforçamos o papel dos referenciais semânticos para a compreensão do que se entende por texto e por documento, de modo a superar a dicotomia significado/significante por meio da união dos planos sintático, semântico e pragmático no discurso, considerado, uma seqüência de microestruturas. A organização dos textos em macroestruturas que representam seu significado global permite a análise da superestrutura e a identificação das partes do texto, em qualquer de seus esquemas típicos, o da narração e o da investigação.

Logo, acreditamos que o termo análise documentária comporta em si dois sentidos: o de método interpretativo (Análise do Discurso/Análise de Conteúdos) e o de teoria, cujo conhecimento é produzido a partir do método. No entanto, lembrando Paul Otlet ao afirmar que “o pensamento não conhece tudo, a linguagem não expressa tudo, o documento não registra tudo”, cremos, caros leitores, que esse seria o *conspectus*, ou seja, o “pré-texto” para iniciar um texto futuro.

## REFERÊNCIAS

- ABNT. **Métodos para análise de documentos** – determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, ago. 1992. NBR-12676. (Procedimento)
- BARANOW, U.G. Perspectivas na contribuição da lingüística e de áreas afins à ciência da informação. **Ciência da Informação**, Brasília, v.12, n.1, p.23-35, 1983.
- BOULANGER, J-C. Alguns componentes lingüísticos no ensino da terminologia. **Ciência da Informação**, Brasília, v.24, n.3, p.313-318, set./dez. 1995.
- BRITO, Marcílio de. Sistemas de informação em linguagem natural: em busca de uma indexação automática. **Ciência da Informação**. Brasília, v.21, n.3, p.223-232, set./dez., 1992
- CAMPOS, M. L. de A. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói: EDUFF, 2001. 133 p.
- CAMPOS, M.L. de A ; CAMPOS, M.L.M. Princípios para a modelização de domínios de conhecimento: estudo comparativo entre abordagens da ciência da informação, ciência da computação e teoria da terminologia. In: **ISKO Espanha 2003**. Salamanca: Ed.Univ. Salamanca, 2003. p. 135-144.
- CASTELLS, Manuel. **A sociedade em rede**. Rio de Janeiro: Paz e Terra, 1999.
- CINTRA, A. M. M. et al. **Para entender as linguagens documentárias**. 2.ed. São Paulo: Polis, 2002. 92p. (Coleção Palavra-chave, 4).
- CUNHA, I. M. R. F. **Do mito à análise documentária**. São Paulo: Edusp, 1990. 163 p. (Teses, v. 11)
- DAHLBERG, Ingetraut. **Ontical structure and universal classifications**. Bangalore: Sarada Ranganathan Endowment for Library Science, 1978a.
- DIRETRIZES para o estabelecimento e desenvolvimento de tesouros monolíngües. Brasília, DF: IBICT/SENAI, 1993. 86 p. (UNESCO; PGI/81/WS/15).
- DODEBEI, V.L.D. **Tesouro: linguagem de representação da memória documentária**. Niterói: Intertexto; Rio de Janeiro: Interciência, 2002. 119 p.
- FAULSTICH, E. Socioterminologia: mais que um método de pesquisa, uma disciplina. **Ciência da Informação**, Brasília, v.24, n.3, p.281-288, set./dez. 1995.
- FAULSTICH, E. Terminologia: o Projeto Brasilterm e a formação de recursos humanos. **Ciência da Informação**, Brasília, v.24, n.3, p.356-363, set./dez. 1995 a
- GUIMARÃES, E. O texto e suas modalidades. In: \_\_\_\_\_. **A articulação do texto**. São Paulo: Ática, 1990. 87 p.; p. 14-20.
- KOBASHI, N. Y. **Classificação e indexação**. Curitiba: Curso de Mestrado Interinstitucional PUC-Campinas e UFPR, 2001. Apontamentos de aulas.
- KOBASHI, N. Y. . Vocabulário controlado USP: desenvolvimento, implantação e gerenciamento. In: Congresso Internacional de Arquivos, Bibliotecas, Centros de Documentação e Museus, 2002, São Paulo. Integrar: 1. Congresso Internacional de arquivos, bibliotecas, centros de documentação e museus: textos. São Paulo: Imprensa oficial do Estado, 2002. v. 1. p. 225-235.
- LARA, M.L.G. de. O processo de construção da informação documentária e o processo de conhecimento. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.7, n.2, p.127-139, jul./dez. 2002.
- LASTRES, H. M. M. Informação e conhecimento na nova ordem mundial. **Ciência da Informação**, Brasília, v. 28, n. 1, p. 72-78, jan./abr. 1999.
- LUCAS, C. R. **Leitura e interpretação em biblioteconomia**. Campinas, SP: Ed. UNICAMP, 2000. 91 p. (Coleção Pesquisas)
- McGARRY, K. **O conceito dinâmico da informação: uma análise introdutória**. Brasília: Briquet de Lemos, 1999.

- MEDEIROS, M.B.B. Terminologia brasileira em ciência da informação: uma análise. **Ciência da Informação**, Brasília, v.15, n.2, p.135-142, jul./dez.1986.
- MENDONÇA, E. S. A lingüística e a ciência da informação: estudos de uma interseção. **Ciência da Informação**, Brasília, v.29, n.3, p.50-70, set./dez. 2000.
- MIRANDA, L. M. C. de. Terminologia de informática em língua portuguesa: uma análise lingüística e terminológica. **Ciência da Informação**, Brasília, v.18, n.2, p.183-190, jul./dez.1989.
- MOREIRO GONZÁLEZ, José Antonio. **El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural**. Gijón: Ediciones Trea, 2004.
- PEIXOTO, A. S. B.; OLIVEIRA, R. dos A. Terminologia do ensino por computador: abordagem socioterminológica. **Ciência da Informação**, Brasília, v. 24, n.3, p.353-355, set./dez. 1995.
- SAYÃO, L. F. Modelos teóricos em ciência da informação: abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n.1, p.82-91, jan./abr. 2001.

#### **ABSTRACT**

Analysis of the principles and methods of documentary languages organization and representation concepts, considering the importance of Documentary Languages inside Documentary Analysis area. The main theories analyzed were: Colon Classification (Ranganathan); Terminology General Theory (Wüester) and Concept's Theory (Dahlberg). This article presents a synthesis of principles and methods for the Documentary Analysis. It concludes that the Documentary Analysis identifying two senses: interpretative method and theory whose knowledge is produced by the method.

**KEYWORDS:** Documentary Analysis. Documentary Languages. Documentary Information. Terminology. Linguistics. Semiotics. Colon Classification.

*Originais recebidos em 01/06/2005.*