

ANÁLISE DOCUMENTÁRIA: OS LIMITES DA UNIVOCIDADE*

CLAUDIO OMAR IAHNKE NUNES**

RESUMO

O autor discute as implicações e limites teórico-metodológicas do conceito de univocidade, empregado pela Análise Documentária, propondo sua substituição pelo conceito de plurivocidade, que considera mais apropriado para lidar com as múltiplas dimensões do conhecimento.

1 – CONSIDERAÇÕES INICIAIS

Ainda que a expressão *análise de assunto* seja comumente utilizada, nem sempre os estudos que se abrigam sob esta designação efetivamente se ocupam da primeira etapa da indexação. Aliás, é escassa a literatura que trata especificamente da análise conceitual do conteúdo dos documentos, isto é, pouco se publica sobre *como* indexadores e classificadores decidem qual é o assunto de um documento. Uma atenta análise das prescrições constantes das normas existentes, tanto em nível internacional quanto no Brasil, indica seu caráter nitidamente operativo, de regramento do bom senso que as originou. A ISO 5963-1985-E é a norma internacional sobre indexação de assuntos^(1:n.p.), que recomenda um exame cuidadoso das seguintes partes do texto: o título; o resumo, se houver; o sumário; a introdução, as frases e parágrafos de abertura de capítulos e as conclusões; as ilustrações, gráficos, tabelas e respectivas legendas; as palavras ou grupos de palavras que apareçam sublinhados ou impressos em tipos diferentes.

Segundo a NBR 12676, publicada pela ABNT em 1992⁽²⁾, o ato de indexar compreende três momentos ou estágios, ainda que reconheça a ocorrência de uma certa sobreposição entre os mesmos, isto é, de que não se trata de atos estanques:

* Este texto é parte da tese de doutorado defendida pelo autor na ECA/USP em 13.03.2000, intitulada *Consenso e representações sociais sob a perspectiva da Análise Documentária aplicada ao controle de vocabulário na área de Arqueologia Brasileira*.

** Professor do Dep. de Biblioteconomia e História – FURG.

- a) exame do documento e estabelecimento do assunto de seu conteúdo;
- b) identificação dos conceitos presentes no assunto; e
- c) tradução desses conceitos nos termos de uma linguagem de indexação.

As duas primeiras etapas referem-se precisamente ao que se constitui no objeto da Análise Documentária, sendo que a norma recomenda que “o indexador deve adotar uma abordagem sistemática para identificar aqueles conceitos que são os elementos essenciais na descrição do assunto” ^(id.:2). Para tanto, determina que se façam perguntas como as seguintes:

- a) qual o assunto de que trata o documento?
- b) como se define o assunto em termos de teorias, hipótese etc.?
- c) o assunto contém uma ação, uma operação, um processo?
- d) o documento trata do agente dessa ação, operação, processo etc.?
- e) o documento se refere a métodos, técnicas e instrumentos especiais?
- f) esses aspectos foram considerados no contexto de um local ou ambiente especial?
- g) foram identificadas as variáveis dependentes e independentes?
- h) o assunto foi considerado sob um ponto de vista interdisciplinar? (por exemplo, um estudo sociológico da religião).

O restante do conjunto de prescrições constantes da referida norma não foge ao diapasão prescritivo-operativo denunciado pelas transcrições anteriores, o que, aliás, não poderia ser diferente, porquanto o reivindicado corpo teórico-metodológico de qualquer área do conhecimento não se limita ao disposto em normas formais, mas distribui-se pelo conjunto mais amplo dos elementos constitutivos de seu respectivo paradigma. Desse modo, no caso da Análise Documentária, a disponibilidade de procedimentos normalizados é insuficiente para fazer frente aos problemas interpostos pelo desenvolvimento da área.

Devido à sua natureza e complexidade, a análise conceitual sofre a influência direta da pessoa que a executa, a ponto de alguns autores afirmarem que este processo se realiza essencialmente com base no “bom senso” do indexador. Assim, a afirmação de que não há influência do indexador no processo de indexação é discutível, pois se sabe da ocorrência da subjetividade que está presente nesse processo de análise de assunto e que o torna mais complexo, à medida que pode ser usada para justificar possíveis falhas no trabalho do profissional, como inconsistências que possam ser detectadas no sistema, no momento da recuperação de informação.

2 – A CRENÇA NA UNIVOCIDADE

A idéia de que seja possível precisar inequivocamente o sentido de um termo significa aprisionar-lhe o sentido, impedindo que o mesmo continue tendo *vida*, isto é, que venha a incorporar as variações semânticas decorrentes do próprio processo de investigação e produção de novos conhecimentos. Produz-se assim um paradoxo, porquanto, com o objetivo de assegurar a comunicação entre os cientistas, constrói-se uma linguagem pretensamente unívoca, a qual, se atende (ainda que incompletamente) ao desiderato atual por homogeneidade de significado, por outro lado deixa de atender à necessidade futura de heterogeneidade, imprescindível ao progresso da ciência. Como lembra ATLAN, “(...) a ausência de incerteza, de imprevisibilidade, implica também a supressão da novidade e da criatividade”^(3:80), ainda que se reconheça que a fluidez dos conceitos e vocabulários nos diversos domínios científicos, especialmente nas fronteiras de pesquisa, como na Genética, venha tornar mais complicada a recuperação de informações, pois um conceito científico pode ser percebido de diferentes maneiras em tempos também diferentes, como reconhecem CHEN et al.^(4:176).

Se há incerteza no conjunto do conhecimento científico estabelecido e, mais, se a parcela dos resultados de pesquisa comunicados no interior da comunidade científica tipificados como irrelevantes ou incorretos é absurdamente majoritária (estimada por alguns autores em 90% de tudo o que se publica), é muita pretensão exigir-se ou esperar-se que as linguagens documentárias alcancem a univocidade para elas preconizada como condição de sua eficácia e eficiência, enquanto ferramentas intermediárias do processo de produção-consumo de informações científicas, vale dizer, da circulação do conhecimento científico no interior do sistema tecno-científico. Como sugere ATLAN, parece que não é o caso de se

(...) confrontar a experiência subjetiva com a objetividade científica, mas de precisar os limites dos domínios de legitimidade de uma e de outra, tomando uma certa distância cada vez que nos arriscamos a cair na armadilha do dogmatismo, na qual corremos o risco de cair tanto extrapolando ao infinito o subjetivismo quanto à objetividade científica^(3:73).

A representação temática precisa ter em conta os mapas materiais e conceituais de cada uma das áreas do conhecimento a que se aplica. Mais do que isso, esses mapas, por não se constituírem em

entidades estáticas, precisam ser assimilados na sua dinamicidade, porém evitando-se que se percam os *fiões* que fazem sua tessitura orgânica e significativa, enquanto corpus de conhecimentos consensualmente aceitos por uma determinada comunidade científica. A propósito de como atuam as representações significativas, isto é, os mapas conceituais, já em 1939 BLISS (apud BEGHOT) referia-se à maneira como os cientistas percebiam as estruturas de suas respectivas disciplinas:

BLISS (1939) used the term "consensus" for the views that scientists (broadly conceived to include humanists) and educators share about the structures of their disciplines. BLISS maintained that consensus should underlie bibliographic classification systems; but, unlike the concept of literary warrant, consensus has not been widely explored in discussions of subject access. In practice, however, consensus has usually been assumed to be a necessary part of the development and revision of classification systems. For example, the editors of both DDC and UDC regularly consult subject experts for advice on particular classes in the systems. Literary warrant and consensus are not mutually exclusive or opposing concepts, and the relationships between the two have not fully explicated in the literature. To develop a rationale for a quantitative study of the content of the literature of fiction studies, it was helpful to explore some pertinent relationships between them. (...) Literary warrant and consensus can be compared along a number of dimensions. The expression literary warrant in a documentation language would rely primarily on semantic and verbal techniques and concepts. For example, renaming classes in accordance with current usage would be a response to changes in literary warrant. In contrast, consensus may be characterized as governing judgements about how those same topics might be arranged. According to BLISS, bibliographic classification systems "should be organized in consistency with the scientific and educational consensus" (1939, pp. 42-43), original emphasis). Consensus, then, is primarily concerned with how the topics in a particular literary warrant are arranged, and the expression of consensus in a documentation language would rely on syntactic and structural techniques and concepts. (...) Thus, at the most general level, literary warrant emphasizes description of the content of a set of documents and consensus emphasizes description of the structures needed to organize the topics of the same set of documents appropriately^(5:32).

A circulação do conhecimento científico promove a interação de três sujeitos cognoscentes, cada qual possuidor de seus próprios *mapas*, gerando-se o desafio de se estabelecer entre os mesmos *legendas e escalas* comuns, intercambiáveis, sob pena de persistir o

comprometimento da efetividade dos atos comunicativos que tentam realizar. Uma parte desse processo parece atender ao que se designa como *consensibilidade* e a outra parte como *consensualidade*. Para ZIMAN, a consensibilidade refere-se à possibilidade de um mesmo conhecimento científico ser desvelado, obtido, por diferentes pesquisadores, sempre mediante a interveniência dos sentidos, isto é, refere-se à possibilidade de se replicar o procedimento de descoberta, obtendo-se os mesmos resultados. Assim concebida, a consensibilidade constitui-se numa das condições necessárias para a efetividade da comunicação dos conhecimentos científicos, enquanto que a consensualidade é trabalhada por ele com um sentido muito próximo daquele empregado por KUHN^(6:32), ao referir-se aos consensos vigentes na ciência normal. Em estudo recente, ZIMAN discorre sobre a distinção entre o conhecimento científico e os demais saberes acumulados e disponibilizados pela civilização humana, situando-a exatamente no fato

(...) de que seu conteúdo é **consensível**. Com isso quero dizer que cada mensagem não deve ser obscura ou ambígua a ponto de seu receptor ficar incapacitado de dar-lhe um consentimento sincero ou opor-lhe objeções bem fundamentadas. A meta da ciência, além disso, é alcançar o máximo de **consensualidade**. Idealmente, o corpo geral do conhecimento científico deve consistir de fatos e princípios firmemente estabelecidos e aceitos, sem dúvidas sérias, por uma maioria esmagadora de cientistas competentes e bem informados. Tal como veremos, convém fazer uma distinção entre uma mensagem **consensível** com **potencial** para vir a contribuir para um consenso, e uma afirmação **consensual** plenamente testada e objeto de um acordo universal. Pode-se dizer, com efeito, que a consensibilidade é uma condição necessária para qualquer comunicação científica, ao passo que apenas uma pequena proporção de todo o corpo da ciência é inegavelmente consensual em qualquer momento dado^(7:17-48).

Admitindo-se o pressuposto de que os esquemas mentais introjetados na consciência do pesquisador pelo paradigma a que está devotado condicionam fortemente sua *habilidade* para ver, faz sentido a afirmação de REEVES de que "Eu verei quando crer"^(3:134), embora ele o faça tendo em conta o que chama de *condicionamento cultural*, evidentemente um conceito mais abrangente do que *paradigma*. Também Winfried GÖDERT posiciona-se entre aqueles autores que consideram a informação como uma construção cognitiva, sustentando sua posição com os seguintes argumentos:

- There is no information without cognitive structures; knowledge is the knowledge of human beings.
- Information cannot be accumulated outside a cognitive structure.
- Externalization of information into information systems – respectively the reception of information out of information systems by human beings – is not a process of self-evidence, instead it must be explained by an information theoretical model. Externalization and reception of information from any media must be explained by models of cognitive information processing and knowledge acquisition^(8:208).

A linguagem de cada ciência integra o paradigma que a governa, restando algumas dificuldades, como decorrência lógica da discussão anterior, para admitir-se que o mesmo se dê com uma eventual linguagem documentária que tenha sido criada para representar o conteúdo dos conhecimentos gerados por essa mesma ciência. O significado atribuído aos termos de uma dada linguagem documentária contempla apenas o que é consensual entre os integrantes da comunidade científica respectiva, não havendo possibilidade de se falar em consensibilidade nesta situação. Mesmo no que respeita a consenso, restam algumas dificuldades. Na construção da linguagem documentária, o consenso é obtido através do que se convencionou designar como *univocidade*, isto é, cada termo terá na linguagem documentária um único significado, precisamente definido, independente do eventual significado que, enquanto palavra ordinária, possa ter no âmbito da língua natural de que foi extraído.

Até aí não há maior dificuldade, verificando-se uma convergência com a noção de consensualidade, condição indispensável para que se consolide uma dada noção, princípio, lei ou teoria científica. A dificuldade surge no momento em que a univocidade pretendida é estabelecida convencionalmente, enquanto que o consenso é formado socialmente, isto é, ele existe, porém não é o resultado de qualquer procedimento formal, contrariamente ao que ocorre com o processo de formação de uma linguagem documentária. Sendo assim, o consenso não é estático; ao contrário, poderá ser substituído por um novo consenso – e de fato o será, tão logo deixe de ser útil, eficaz para o equacionamento dos problemas propostos pelo que KUHN designa como ciência normal. Quando isto ocorre no âmbito de uma ciência em particular, que eventualmente disponha de uma linguagem documentária própria, ocorrem problemas na circulação das informações, decorrentes do fato de que tal linguagem não dispõe de mecanismos de ajuste, apropriados ao acompanhamento dos desenvolvimentos (supressões, acréscimos etc.) que estão a ocorrer

continuamente no âmbito da disciplina a que pertence. Desde esta perspectiva, pode-se dizer que as linguagens documentárias são instrumentos da ciência normal, e que por essa condição enfrentam dificuldades para lidar com a ciência revolucionária, isto é, com a mudança ou substituição de paradigmas.

Assim como a própria prática científica é autocorretiva, as linguagens documentárias, para bem cumprir seu papel de ferramentas apropriadas ao gerenciamento da circulação da informação científica, também precisam sê-lo. E isto precisa ser feito numa velocidade compatível ao que se verifica no desenvolvimento científico. Não é mais possível trabalhar-se com linguagens que dependem, para sua atualização, de procedimentos burocratizados, administrados à distância. Talvez a velocidade desejável possa ser obtida mediante o trabalho em pequenas unidades, em estreita cooperação com os próprios cientistas, que alimentam e se utilizam da linguagem documentária, estreitamente vinculada aos seus projetos e programas específicos de pesquisa.

A obtenção de consenso, isto é, de linguagens documentárias que cubram campos mais amplos do conhecimento, ou mesmo um campo mais limitado, porém com um alcance espaço-institucional amplo e diversificado, se dará similarmente ao que ocorre com a formação de consensos entre os próprios cientistas: mediante o processo social de intersubjetividade, envolvendo tanto os indexadores quanto os próprios cientistas. O receio de que possa estabelecer-se uma *babel documentária* parece afastado (ou melhor, quem sabe trate-se precisamente de se admitir esta *babel*), se forem considerados os poderosos recursos proporcionados pela informática e telecomunicações, atualmente disponíveis.

Parece-nos que tal proposição, ao contrário de contribuir para *velar* a informação produzida, de fato induzirá a que uma parcela importante de informações que *somem* durante os processos de análise e indexação, devido à inflexibilidade atual das linguagens e procedimentos documentários, possam ser postos à disposição da comunidade científica – ainda que impliquem um processo de recuperação mais interativo e intersubjetivo, contrariando, por um lado, a diretriz algo positivista dos procedimentos tradicionais de armazenagem e recuperação de informações científicas e técnicas, e, por outro, o desiderato de rapidez – o qual, diga-se de passagem, tradicionalmente é aferido apenas no momento da entrada das informações no sistema e muito raramente na ponta da saída.

A este propósito, é útil considerar-se que a entrada de informações no sistema é feita uma única vez, enquanto que sua

recuperação é feita *N* vezes. Assim, é falso considerar-se eficaz um sistema que propicie procedimentos rápidos de análise documentária e indexação, desconhecendo se o mesmo se dá com o processo de busca e recuperação de informações.

De algum modo, precisamos dispor de uma *boa circularidade* no âmbito das linguagens documentárias, de modo a que esse mecanismo possa efetivar a correspondência semântica entre os descritores e os fenômenos ou fatos por eles representados, haja vista a transmutação dos mesmos, como consequência do progresso científico que afeta inapelavelmente a linguagem da ciência, provocando sua variação semântica (criando ou suprimindo termos, alargando ou estreitando seu campo conceitual e modificando sua posição e inter-relacionamentos nos respectivos mapas conceituais). A propósito da dificuldade de os modelos científicos representarem o real, Jean-Pierre DUPUY elaborou a *teoria da complexidade*, que preconiza que o todo resultante da composição das partes retroage sobre elas, redefinindo-as. Explicitamente, este autor diz:

Trata-se de níveis de integração imbricados uns aos outros: os átomos compõem moléculas, que compõem os organismos, os quais compõem as sociedades. (...) Como é que as moléculas se agilizam para produzir macromoléculas biológicas? E como estas se agilizam para produzir o metabolismo celular? Os teóricos da complexidade dos sistemas naturais tiveram de construir, para descrever o que eles compreendiam – ou melhor, não compreendiam –, o conceito de circularidade entre níveis de organização diferentes (...) Todos ouviram falar do “programa genético” – essa hipótese de que o desenvolvimento e a história de um organismo vivo estariam de antemão escritas nos genes. É exatamente o que quer dizer, do ponto de vista etimológico, a palavra “programa”. Mas os biólogos foram obrigados a precisar que esse programa genético é um programa bem curioso na verdade, pois ele “programa-se a si mesmo”! Como assim? Porque os operadores que são as proteínas das células, que se supõe “transcrever” e “traduzir” a mensagem contida nos genes, foram codificados por esses genes. A senhora percebe aqui circularidade, o encavalamento dos níveis no sentido em que falávamos há pouco. Depara-se assim com a auto-referência do programa genético, e vê-se o estreito vínculo que existe entre a noção de complexidade e a de autonomia. Esse paradoxo foi desemaranhado por Henri ATLAN, por meio da noção de “complexidade através do ruído”, que recorre àquilo que não é absolutamente sagrado para os cientistas: o acaso^(3:109-110).

Admitir a idéia de *circularidade* implica aceitar-se uma outra noção igualmente polêmica: a de *ruído*. Ora, o recurso à noção de *ruído* perturba definitivamente os pressupostos de univocidade e precisão

reivindicados anteriormente pelas linguagens documentárias, porquanto estava prescrito como o fator preponderante de sua ineficiência e ineficácia. Recuperá-lo como recurso teórico, atribuindo-lhe foros de fundação para os procedimentos adotados no âmbito das linguagens documentárias, tem o impacto de uma quase ruptura. Por isso, faz-se necessário explorar cuidadosamente sua consistência e implicações metodológicas.

O texto linear foi rompido pelo registro da informação em suporte magnético. Isto ficou materialmente mais evidente com o advento do CD-ROM, com as publicações eletrônicas disponibilizadas na Internet, mas já com as primeiras bases de dados bibliográficos em meio magnético tal fato se prenunciava, de modo a que era possível, com extrema facilidade e sem perder o *fio da meada*, ir de um lado para o outro, juntando os vários pedaços de uma informação, ou relacionando uma informação, localizada em um determinado ponto, com outra, localizada em outro ponto. Isto já podia ser feito, sem dúvida, com os catálogos de fichas, e mesmo com um sistema de remissivas em um índice impresso. Porém, havia uma limitação física que se tornava tão mais acentuada quanto maior fosse a quantidade de informações armazenada, alcançando a inviabilidade operacional a partir de certo ponto, como já percebiam CESARINO & PINTO^(9:284-285), em estudo publicado em 1978.

Se não existe um programa de computador que substitua o cientista em seu que-fazer, é muito possível que resulte incoseqüente ou falho qualquer programa que tente realizar o processo de análise documentária dispensando a intervenção humana. Isto porque a análise documentária é um subprocesso do processo maior de comunicação do conhecimento científico, atuando diretamente sobre o meio pelo qual este se realiza, qual seja, a linguagem científica. Como esta não pode ser reduzida à logicidade subjacente à linguagem da ciência, restam *áreas em branco* das quais um programa, ainda que de inteligência artificial, não tem como dar conta, porquanto os programas de computador se assentam em algoritmos formais, isto é, lógicos em última instância e, por consequência, incapazes de extrapolar criativamente os parâmetros definidos pelo programa.

Ora, se a necessidade da lógica se faz presente apenas na comunicação, isto é, na linguagem através da qual se faz a comunicação, então depreende-se que esta linguagem pode ser reduzida completamente a um algoritmo, a um programa de computador, à inteligência artificial. Parece que aí reside uma dificuldade séria, porquanto também ao nível da linguagem científica registra-se um importante resíduo de indeterminação, de ilogicidade.

Talvez não seja o caso das teorias já consolidadas, devidamente articuladas e assumidas consensualmente pela comunidade – durante a vigência de um determinado paradigma, ressalve-se. Mas, certamente, este não é o caso das teorias, ou hipóteses, que lidam com a fronteira do conhecimento, ou, mesmo, com o mapeamento detalhado de um amplo território que, embora já demarcado, ainda não esteja completamente explorado. A comunicação científica, nessas situações, não é completamente governada pela lógica. Por conseqüência, a linguagem que os cientistas utilizam para comunicar entre si os resultados de suas investigações é variavelmente indeterminada.

É interessante registrar a ocorrência, muito freqüente nas bibliotecas universitárias (em que o acervo costuma ser multidisciplinar), de discordância dos pesquisadores (ou especialistas) com a classificação ou, mais simplesmente, com a colocação (topográfica) de determinadas publicações nas estantes. Uma vez atendido o questionamento de um primeiro usuário, segue-se um novo questionamento, agora de outro usuário, geralmente da mesma área ou de área correlata, o qual, por sua vez, manifesta sua discordância com o arranjo sugerido ou requerido pelo seu colega. É evidente que, por trás desses questionamentos, encontram-se os respectivos mapas conceituais de cada um. Tais mapas, especialmente quando de áreas diferentes (ainda que correlatas), costumam guardar muitos dissensos entre si e, por óbvio, com o próprio mapa conceitual que informa o sistema de classificação adotado pela biblioteca.

3 – DESAFIOS À ANÁLISE DOCUMENTÁRIA

Fazendo eco às palavras de Johanna SMIT^(10:11), pode-se dizer que a Análise Documentária constituiu-se como disciplina a partir de uma utopia universalizante, a qual, vista da perspectiva dos dias de hoje, parece ingênua e dificilmente exeqüível, abrindo passagem para que se imponha uma nova abordagem, agora mais realista e localizada, tanto na perspectiva do tempo quanto do espaço. Reconhece-se a importância crescente da lingüística textual e da lógica, enquanto recursos teórico-metodológicos, assim como os aportes procedimentais da Informática, para a Análise Documentária. As contribuições da Epistemologia e da Psicologia Cognitiva também se agigantam, à medida que assumem relevância as questões da obtenção e comunicação do conhecimento científico.

A propósito, em estudo recente sobre o grau de acordo na nomeação de objetos e conceitos visando a recuperação de informações, COLLANTES afirma que

One important finding is that naming is both user-specific and object – or concept – specific. Furthermore, the user and the object or concept are interrelated. The implication on theory is that naming paradigm should be broadened. This requires the consideration of all aspects of the user, including those that are internal and external. By internal is meant cognitive aspects, for example, knowledge, processing mechanisms, purpose, rules, strategies, etc. By external is meant stimuli, context of a problem, and environment. (...) Do certain people share certain common characteristics that lead them to name things in a similar way? It is possible to group individuals by the naming style they use? This naming style could be a function of an individual's personal characteristic, training and/or experience. It could also relate to a more general cognitive or information style of the particular individual. This may also be a fruitful area for future research. (...) Our knowledge of naming behavior is incomplete at this point. Further studies on naming in other domains may be worthwhile to see if these patterns can be understood, predicted, and sometimes accommodated to improve retrieval of information^(11:131).

A necessidade de se fazer controle de vocabulário – e a contribuição que a Análise Documentária pode fazer neste sentido – vem sendo fortemente questionada, como já se discutiu neste texto, principalmente face à suposição de que alguns aportes da Informática, como as bases de texto integral e o hipertexto, a tornariam supérflua. No nível operativo, das técnicas e procedimentos documentários, sem dúvida, o impacto é devastador. E nos níveis mais complexos, em que se lida com questões como a do significado, da aquisição, compartilhamento e comunicação do conhecimento científico, dá-se o mesmo fenômeno? Parece que é nessa direção que apontam estudos como os de GÖDERT⁽⁸⁾, COLLANTES⁽¹¹⁾ e IIVONEN⁽¹²⁾, entre outros autores que vêm sendo discutidos neste texto.

Contudo, não se deve excluir da discussão os questionamentos associados aos aspectos operativos da Análise Documentária. Por isso, é útil focalizar a questão do custo-benefício, que foi objeto de uma reflexão de Raya FIDEL. Esta autora conclui que

Because the creation and use of controlled vocabularies is very intensive, it has been claimed that thesauri are not cost effective. To date, however, no study has proven this claim. While the study reported here does not prove the opposite (that is, that thesauri are cost effective), it clearly proves the importance of controlled vocabularies to searching^(13:7).

Nessa direção também se deve visualizar o que está sendo feito no sentido de tirar proveito dos aportes procedimentais proporcionados pela Informática. São inúmeros os projetos e sistemas que vêm sendo

desenvolvidos, notadamente dos anos oitenta em diante. Os sistemas de texto integral tendem a se multiplicar, à medida que se aumenta a capacidade de armazenagem e processamento dos suportes, assim como sua portabilidade, como é o caso dos discos óticos. O Projeto ISLA (Information System of Los Angeles), da University of Southern California, é um exemplo desse novo tipo de sistemas, que são concebidos sob a perspectiva do usuário final, como se depreende de suas características:

ISLA involves two components: (1) a library of multi-disciplinary digital materials in various formats about the Los Angeles region, and (2) a specific search and retrieval method involving a space/time/full-text/format index. It differs from other digital library efforts in that:

- All materials are indexed by spatial and temporal, as well as conventional indices;
- Finely defined geographical areas and temporal periods are an option;
- Many of the formats are atypical for spatial indexing (e.g. texts and photographs); and
- A limited number of historically appropriate spatial data layers are available to allow search definition in an historically appropriate setting^(14:441).

É interessante registrar que estes novos sistemas costumam ser construídos e administrados interdisciplinarmente. No caso do Projeto ISLA, seu Comitê Executivo é constituído meio a meio por bibliotecários e usuários, embora se reconheça que a melhor distribuição contaria com um terço de bibliotecários, um terço de usuários e um terço de engenheiros, cientistas da computação e técnicos de outras áreas, que poderiam enriquecer com seu conhecimento o desenvolvimento do projeto.

A idéia de desenvolvimento de procedimentos de ajuda ao usuário final, que também pode ser definida como *amigabilidade* das interfaces dos sistemas com o usuário e que anda muito próximo do que alguns autores chamam de *sistemas inteligentes* ou de *sistemas especialistas*, por óbvio está presente nestes projetos. Parece claro que aqui existe espaço para uma amarração entre os aportes teóricos de disciplinas como a Lingüística, a Psicologia Cognitiva, a Psicologia Social, a Lógica e a Informática, sendo que a abordagem comporta duas vertentes. Uma, mais teórico-metodológica, em que se busca emprestar consistência à Análise Documentária, e outra, mais operativa, em que se busca a criação de ferramentas voltadas ao gerenciamento da circulação do conhecimento nas redes de informação que se expandem com rapidez quase exponencial. Estudando a consistência na seleção

de conceitos e termos de pesquisa, IIVONEN constatou precisamente que

Because searchers seem to select search terms somewhat inconsistently – and they are fairly faithful to their terminological styles – and because the use of many alternative search terms has a positive effect upon search results, we need tools to help searchers produce variety in the selection of search terms. These tools could be an end-user thesaurus and different intermediary systems. The main purpose of these tools is to suggest to a searcher other alternative search terms^(12:188).

Esta sua constatação, que foi argüida a título de justificativa para o desenvolvimento de ferramentas de ajuda, não constitui propriamente uma novidade, porquanto LANCASTER já se referia a isto em 1972, conforme lembrado pela própria autora da citação anterior, no artigo já mencionado^(id., ibid.).

Neste cenário de transformações, dúvidas e questionamentos, há quem se pergunte mesmo sobre o futuro da indexação humana, como o faz Martin L. WARD, em ensaio publicado em 1996, em que conclui que, embora esta tendência não possa ser desconsiderada, ainda não é o caso de se considerar superado o papel da Indexação e, por consequência, da Análise Documentária.

At the moment it does not seem the automatic indexer will replace the human indexer, but when more texts are available in electronic form it may be a useful productivity tool for dealing with large quantities of lowgrade texts (should you want them in your textbase). In his useful paper on indexing models, John Farrow (1991) notes the distinction between “bottom-up”, perceptual operations, and those which are “top-down” and conceptual in nature. Object Analyzer can make a usable abstract by working “bottom-up” within the world-of-the-text. This author wonders, though, whether even in the most bottom-up analysis, a degree of vision and common sense will not produce a better and richer result^(15:224-225).

Invertendo a corrente histórica, talvez seja chegado o momento de se pensar nos pequenos bancos de dados bibliográficos, criados no fundo de um laboratório, junto a uma pequena instituição ou, simplesmente, no disco rígido do microcomputador de um pesquisador solitário. Com a massificação do acesso às redes de computadores, impõe-se um desafio: tornar a Análise Documentária uma ferramenta de processamento da informação nesse primeiro nível de sua produção e consumo. Afinal, qualquer usuário, plugado num provedor, pode dispor de sua *home page* e, eventualmente, intercambiar arquivos ou permitir o

acesso remoto ao seu computador. Tecnicamente, não há mais impedimentos para tanto.

Inspirando-nos no pensamento de autores como Hans-Georg GADAMER^(16:559-589), propomos que a Análise Documentária enfrente este desafio sob uma nova perspectiva, que poderíamos denominar *abordagem fenomenológica da informação científica*, o que possibilitará a superação dos limites do paradigma da univocidade. Somente a ruptura na direção de uma *plurivocidade* permitirá à Análise Documentária abrir-se à compreensão mais completa do fenômeno do conhecimento científico, em suas dimensões epistemológicas, sociológicas e comunicacionais. A manifestação deste rompimento deve incidir tanto sobre seus fundamentos teóricos, quanto metodológicos, alcançando o seu modo de operar sobre o objeto que lhe dá razão de existir: a informação.

REFERÊNCIAS BIBLIOGRÁFICAS

1. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *ISO 5963 : Documentation : Methods for examining documents, determining their subjects, and selecting indexing terms*. Switzerland, 1985. n. p.
2. BRASIL. ABNT. *Métodos para análise de documentos : determinação de seus assuntos e seleção de termos de indexação*. Rio de Janeiro, 1992. 4 p.
3. PESSIS-PASTERNAK, G. *Do caos à inteligência artificial : quando os cientistas se interrogam*. São Paulo: UNESP, 1993. 259 p. (Entrevistas com os cientistas: Ilya Prigogine; Henri Atlan; Jean-Pierre Dupuy; Hubert Reeves, citados ou mencionados neste texto).
4. CHEN, H.; YIM, T.; FYE, D.; SCHATZ, B. Automatic thesaurus generation for an electronic community system. *Journal of the American Society for Information Science*, v. 46, n. 3, p. 175-193, 1995.
5. BEGHTOL, C. Domain analysis, literary warrant, and consensus: the case of fiction studies. *Journal of the American Society for Information Science*, v. 46, n. 1, p. 30-44, 1995.
6. KUHN, T. S. *A estrutura das revoluções científicas*. 2. ed. São Paulo: Perspectiva, 1987. 257 p.
7. ZIMAN, J. *O conhecimento confiável : uma exploração dos fundamentos para a crença na ciência*. Campinas: Papyrus, 1996. 252 p.
8. GÖDERT, W. Information as a cognitive construction: a communication-theoretic model and consequences for information systems. *Knowledge Organization*, v. 23, n. 4, p. 206-212, 1996.
9. CESARINO, M. A. da N.; PINTO, M. C. M. F. Cabeçalho de assunto como linguagem de indexação. *R. Esc. Bibliotecon. UFMG.*, Belo Horizonte, v. 7, n. 2, p. 268-288, set. 1978.
10. SMIT, J. *Análise documentária: breve histórico*. São Paulo: ECA/USP, 1996. 12 p.
11. COLLANTES, L. Y. Degree of agreement in naming objects and concepts for information retrieval. *Journal of the American Society for Information Science*, v. 46, n. 2, p. 116-132, 1996.
12. IIVONEN, M. Consistency in the selection of search concepts and search terms. *Information Processing & Management*, v. 31, n. 2, p. 173-190, 1995.
13. FIDEL, R. Who needs controlled vocabulary? *Special Libraries*, Winter, 1992, p. 1-9.
14. HUNT, L. (Li); WEGNER, L. S. Project ISLA: A space/time/full-text /format search and retrieval system designed by end users. *Journal of Academic Librarianship*, London, v. 22, n. 6, p. 440-449, Nov. 1996.
15. WARD, M. L. The future of the human indexer. *Journal of Librarianship and Information Science*, v. 28, n. 4, p. 217-225, Dec. 1996.
16. GADAMER, H.-G. *Verdade e método: traços fundamentais de uma hermenêutica filosófica*. Tradução de Flávio Paulo Meurer. 2. ed. Petrópolis: Vozes, 1997. p. 559-589.