

ANÁLISE DA REVISTA CIÊNCIA DA INFORMAÇÃO DISPONIBILIZADA NA SciELO A PARTIR DO SEU VOCABULÁRIO CONTROLADO¹

ANALYSIS OF REVISTA CIÊNCIA DA INFORMAÇÃO AVAILABLE AT SciELO SINCE ITS CONTROLLED VOCABULARY

Fábio Mascarenhas e SILVA²

RESUMO

Pesquisa realizada no vocabulário controlado da Revista Ciência da Informação disponível na SciELO. Procurou-se inferir resultados a partir da busca feita no sistema de recuperação da informação. Diante das frequências que apresentaram maiores relevâncias quantitativas, foram expostas opiniões no intuito de contextualizar os documentos produzidos pela Revista em formato digital, com o período em que foram publicados.

Palavras-Chave: revista ciência da informação, análise de conteúdo, análise do discurso, SciELO, vocabulário controlado.

ABSTRACT

Research made in the controlled vocabulary of the journal Ciência da Informação available in the SciELO. The aim was to infer results from the search carried out through the information retrieval system. Before the frequencies that had presented greater quantitative relevances, opinions in the intention of contextualizing documents produced for the journal in digital format with the period they had been published.

Key words: journal information science, analysis of content, analysis of speech, SciELO, burst vocabulary.

INTRODUÇÃO

Um dos projetos de pesquisa que contribuiu significativamente para o progresso da ciência brasileira nos últimos anos foi a *Scientific*

Electronic Library Online (SciELO – <http://www.scielo.br>), desenvolvido pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) em parceria com o Centro Latino-Americano e do Caribe de Informação em Ciências da

⁽¹⁾ Trabalho apresentado à disciplina Fundamentação Lingüística no Tratamento da Informação em junho de 2001, ministrada pela Profa. Dra. Else B.M. Válio do Programa de Pós-graduação em Biblioteconomia e Ciência da Informação da PUC-Campinas.

⁽²⁾ Mestrando do Programa de Pós-graduação em Biblioteconomia e Ciência da Informação, PUC-Campinas.

Saúde (BIREME). Trata-se de uma biblioteca virtual que abrange uma coleção selecionada de periódicos científicos brasileiros, que além de propiciar o acesso à parte da produção científica nacional, desenvolve metodologia referente a disponibilização destes documentos em meio eletrônico na *Internet*.

Até o mês de maio de 2001 estavam disponíveis no acervo da SciELO 57 títulos de periódicos, entre os quais a Ciência da Informação (CI) online publicada pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), que foi para esta pesquisa o objeto de análise.

Os conceitos referentes aos sistemas que gerenciam informações em ambientes digitais (comumente chamados de bibliotecas eletrônicas, virtuais ou digitais) ainda não podem ser vistos como definitivos, pois as opiniões quanto às taxonomias do referido tema apresentam-se ainda em desenvolvimento. Assim, visando apenas expor uma noção elementar do assunto, foram inseridas algumas considerações básicas sobre estas.

Para Marchiori (1997), na biblioteca eletrônica os processos básicos são de natureza eletrônica, implicando a utilização de computadores na construção de índices online e a busca nos documentos e na recuperação e armazenagem de registros. Já a biblioteca virtual, baseia-se em tecnologias de realidade virtual, dotados de ambientes artificiais gerados a partir de um computador. Indiferentes a tais definições, a nova modalidade de biblioteca resulta da evolução das formas de se oferecer melhores acessos aos estoques, minimizando-se as barreiras impostas pelo tempo e espaço.

Exposto isto, salienta-se que no acervo eletrônico da SciELO o usuário tem a sua disposição ferramentas que facilitam a recuperação da informação, sendo uma delas o vocabulário controlado. De acordo com Lancaster (1993) um vocabulário controlado é basicamente uma lista de termos autorizados semanticamente estruturados, em três tipos: esquemas de classificação bibliográfica, lista de cabeçalhos de assuntos e tesouros. Todos são compostos de termos (tanto alfabeticamente como sistematicamente), mas, a lista de cabeçalhos (interesse deste trabalho) incorpora uma estrutura hierárquica imperfeita e

não distingue claramente as relações hierárquicas das associativas.

Os vocabulários na SciELO estão disponíveis de duas formas: um conjunto dos termos referentes a apenas um periódico ou de toda a coleção. Nesta pesquisa foi feita uma análise do conteúdo da CI *online*, enfocando as relações existentes entre os termos do seu índice de assunto e respectivos artigos disponíveis, contextualizando-os com os períodos em que foram publicados.

O VOCABULÁRIO CONTROLADO DA CI ONLINE

Neste ponto não se pretende apresentar revisões conceituais que proporcionem embasamentos teóricos para discussões sobre os vocabulários controlados, entretanto traçou-se considerações encontradas na literatura especializada que abordam os princípios básicos sobre o tema.

Segundo Robredo (1994, p. 202), o índice é “[...] num sentido mais amplo, um conjunto ordenado de códigos, representativos de assuntos, tópicos ou conceitos [...] os quais podem servir como critério de busca relacionados com alguma chave de acesso que permita localizar os documentos - ou suas partes ou representações - relativos a cada assunto”.

Na análise dos assuntos mais representativos da CI online utilizou-se a busca automatizada disponível no site. Talvez este recurso não proporcionasse uma eficácia na recuperação da informação se não houvesse uma política para o gerenciamento do estoque deste sistema. Entre algumas ações previstas desta política, há no conjunto de procedimentos do tratamento da informação a análise documentária que visa a criação de uma interface entre a linguagem natural utilizada pelo autor e usuários, e a linguagem artificial usada pelo sistema de recuperação da informação (van DERLAAN, 2000).

Segundo Cunha (1987, p.39), uma análise manual ou automática “tem procurado desenvolver léxicos de base ou léxicos intermediários destinados a preencher o papel de sistemas de conversão entre conceitos apresentados de forma independente nas diversas linguagens e conceitos de leitura “universal” definidos pela própria análise documentária”.

A busca pela representação do conteúdo de um documento em limitadas palavras-chave requer do bibliotecário uma observação neutra e objetiva, e o uso de estratégias de leitura pré-determinadas, pois é necessário colocar-se como um leitor que “*sabe ler sem interpretar, a sua subjetividade, se existe, prende-se a letra, e deve ser conduzida através de procedimentos técnicos que assegurem a atribuição de assuntos que representam o texto fielmente [...]*” (LUCAS, 1998, p. 36).

No caso da SciELO, seu compromisso era disponibilizar os documentos enviados pelos respectivos periódicos, portanto, no caso da CI *online*, a indexação dos documentos era previamente feita pelos profissionais designados pela revista. De acordo com Packer (1998), a SciELO adotou uma metodologia composta por um conjunto de princípios que fundamentaram todo o projeto. Alguns destes princípios foram: preservação da identidade dos periódicos, obediência a padrões, e uso de tecnologias compatíveis às realidades da América Latina e Caribe. Compreende-se que uma das características latentes na estrutura da SciELO é a flexibilidade em se ajustar aos acervos (considerou-se o plural em razão de cada revista possuir suas particularidades).

O tratamento dos artigos científicos na SciELO era composto de cinco módulos, mas para esta pesquisa será apresentado apenas a **Definição de Tipos de Documentos** (DTD). As DTD's baseiam-se nas normas ISO 8879/86 e ISO 12083/94, bem como em formatos usados pela *Elsevier Science* e *European Group on SGML*. Descrevem os documentos, definindo suas estruturas, elementos bibliográficos e contexto em que aparecem.

As DTD's SciELO descrevem em conjunto os elementos-chave dos documentos dos periódicos, classificados como *Serial*, *Article* e *Text*, e representam respectivamente: o fascículo de periódico como um todo; os elementos bibliográficos de um artigo científico; outros tipos de textos como editoriais, cartas ao leitor. Na DTD *Article* há o bloco *Front* que contém além do título e autor, informações bibliográficas complementares (resumos, palavras-chave, histórico) (PACKER, 1998).

Nesta última DTD são evidenciadas as palavras-chave que irão representar o documento durante o processo de busca, podendo ainda haver a inclusão de um ou mais termos no vocabulário controlado, caso este ainda não tenha sido incorporado através de uma inclusão anterior. Por se tratar de um processo automático, o sistema se encarrega de evitar repetições de palavras no índice, muito embora algumas representações compostas por mais de uma palavra estejam apresentadas por inteira e também por partes tidas como todos isolados.

MATERIALE MÉTODOS

Esta pesquisa se desenvolveu em maio de 2001, até então, estavam disponibilizados 12 números da CI *online* na SciELO, correspondentes aos fascículos do volume 26 (número 1 de 1997) até o volume 29 (número 3 de 2000). Os procedimentos para a coleta destes documentos foram:

1) Na primeira etapa acessou-se o site da SciELO visando encontrar o *link* para a revista da CI *online* (havia opções por ordem alfabética, por assunto, ou pesquisa de títulos). No site da Revista, escolheu-se a opção de pesquisa por assunto, no qual foi possível encontrar a alternativa para visualizar o índice do periódico.

2) A segunda etapa consistiu em selecionar todos os termos individualmente (só foram consideradas as palavras de outro idioma que não tivessem similares na língua portuguesa) para que fosse requisitada a busca automática ao sistema. Assim, os resultados obtidos para cada palavra foram registrados separadamente. Salienta-se que não foram utilizados operadores booleanos nem estratégias de busca com mais de um descritor. No entanto, houve uma particularidade que acarretou resultados repetidos na busca de alguns termos. Observou-se, por exemplo, que ao se utilizar o termo INFORMACAO foram recuperados 85 registros, no entanto, neste montante estavam inclusos artigos que não foram indexados pelo descritor INFORMACAO e sim por termos compostos como CIENCIA DA INFORMACAO ou INFORMACAO TECNOLOGICA³.

⁽³⁾ A grafia aqui utilizada corresponde a mesma disponível no índice da revista.

Tabela 1. Termos que apresentaram as maiores incidências.

A-B		C-G		H-L	
Automação	5	Ciência	16	Indicadores	6
Bibliometria	5	Ciência da Informação	8	Indicadores Científicos	6
Biblioteca	20	Conhecimento	16	Informação	85
Biblioteca Digital	7	Desenvolvimento	5	Informação Tecnológica	6
Biblioteca Universitária	6	Digital	9	Informações	6
Biblioteca Virtual	8	Eletrônica	11	Inteligência	9
Brasil	5	Fontes	5	Inteligência Competitiva	7
		Gestão	9	Internet	13
		Globalização	6	Literatura	5
M-Q		R-T		U-Z	
Negócios	5	Recuperação	6	Tecnologia	9
Periódicos	6	Recuperação da Informação	6	Tecnologias	10
Política	5	Redes	6	Tecnológica	9
Produção	6	Serviço	5	Universitária	6
Qualidade	5	Serviços	5	Usuário	7
		Sistema	9	Virtual	12

3) Depois de coletadas todas as incidências, foram selecionadas aquelas que apresentaram os maiores resultados. Devido a limitação do tempo desta pesquisa considerou-se excluídos os descritores que obtiveram menos de 5 registros como resposta (Tabela 1).

4) A quarta etapa envolveu o processo de contextualização com o período em que foram publicados os números da Revista. Entretanto, as considerações não se limitaram apenas aos anos de publicação dos artigos (1997 a 2000), pois a análise dos anos precedentes foi fundamental para se compreender melhor os interesses e convergências das pesquisas publicadas na Revista. Porém, como este trabalho não tinha a pretensão de aprofundar nas discussões a respeito de fatos isolados destas épocas, seus limites foram fixados em acontecimentos mais generalizados, pois os eventos específicos exigiriam uma análise mais detalhada, extrapolando assim as propostas deste estudo.

Identificar estas relações implicou em obter mais do que dados quantitativos, pois o enfoque voltou-se às influências às quais os autores estavam

suscetíveis, o que poderia refletir diretamente nas delimitações temáticas abordadas nos artigos.

A priori, a expectativa indicava algumas hipóteses quanto ao conteúdo da Revista. Contudo, elas se baseavam meramente em uma avaliação empírica, destituída de quaisquer métodos científicos. A seguir estão três delas:

a) A frequência dos termos representaria a linha de publicação adotada pelo corpo editorial da revista no período estudado;

b) Os termos que apresentariam as maiores incidências estariam diretamente relacionados ao processo do desenvolvimento da eletrônica e das telecomunicações;

c) Ênfase na informação como objeto de estudo e discussão.

RESULTADOS E DISCUSSÃO

Finalizada a terceira fase da pesquisa, identificou-se as incidências que serviram como indicadores sobre a linha editorial da *CI online*. Salienta-se que as opiniões aqui apresentadas

basearam-se em resultados quantitativos obtidos na busca efetuada no vocabulário controlado da SciELO, o que não garante certamente a opinião do corpo editorial da Revista.

Assim, dentre algumas observações, estão listados abaixo àquelas que mais chamaram a atenção:

- O termo “INFORMAÇÃO”, mesmo que apresentado em diferentes contextos e utilizados para fins distintos, obteve a maior representação entre todos. Isto evidencia que o foco principal do periódico é a informação, porém não há como assegurar que o referido objeto seja compreendido igualmente por todos os autores;

- A ocorrência do termo “BIBLIOTECA” foi consideravelmente superior a outros que também poderiam expressar os estoques de conhecimento, ou outros sistemas de informação pertinentes aos interesses das Ciências da Informação. O descritor “ARQUIVO” (mais voltado a área de Arquivologia) por exemplo, obteve uma incidência bastante pequena nos resultados. Neste caso a justificativa pode estar no pequeno número de cursos (Graduação e Pós-Graduação) oferecidos no Brasil, bem como a pequena quantidade de profissionais que atuam neste ramo.

- Os números da CI *online* disponíveis na SciELO compreendem os anos de 1997 a 2000. Que mudanças ocorreram neste período que afetaram os meios de registro e disseminação do conhecimento? A década de 90 foi marcada por transformações influenciadas pelo avanço das tecnologias da informação, promovendo mudanças que possivelmente atingiram grande parte dos profissionais da informação. O fim da reserva de mercado da informática no Brasil, a popularização da *World Wide Web* (WWW) e o desenvolvimento das telecomunicações propiciaram uma rápida assimilação de novos conceitos e posturas, e exauriram ícones das antigas rotinas de trabalho (máquinas de escrever, mimeógrafos, Telex, Fax).

Este processo desencadeou o interesse de vários segmentos da sociedade para um possível entendimento e compreensão da dimensão do que acontecia e o que poderia vir a ser em anos vindouros. Conseqüentemente, pesquisadores e cientistas passaram a dividir as suas atenções aos

ambientes digitais e não mais apenas aos analógicos, resultando em uma produção intelectual bastante significativa sobre o assunto. E os números parecem contribuir com esta afirmativa, pois os termos de 10 maiores incidências estão claramente relacionados ao contexto exposto sobre a década de 1990.

Como o objetivo desta pesquisa limitava-se a análise da CI *online* disponível apenas na SciELO, não foram consideradas as publicações anteriores a estas datas. Portanto, não foi possível estabelecer parâmetros de comparação entre épocas diferentes.

CONCLUSÃO

Os periódicos científicos, assim como quaisquer outros produtos editoriais, acompanham as mudanças que vão se desenvolvendo no ambiente social. Esta é uma das condições básicas para assegurarem suas existências e continuarem a despertar o interesse de seus leitores. O conteúdo da CI *online* disponível na SciELO, ao menos em um enfoque quantitativo, apresenta-se alinhado ao contexto histórico compreendido entre os anos de 1997 a 2000.

Um dos focos principais da Revista, a informação, está bem representado em setores bem distintos em que este se apresenta como elemento primordial. Para ser mais preciso, pode-se afirmar que, dos termos que obtiveram as maiores incidências, poucos não estão intrinsecamente relacionados com a informação. A maior parte abrangeu recursos tecnológicos, estratégicos, políticos, econômicos, humanos, entre outros, voltados à informação.

Contudo, há uma proporção explicitamente representativa da influência da Biblioteconomia na Revista. Por um lado isto pode ser interpretado como um fato positivo, já que historicamente há uma estreita relação entre a informação e o universo bibliotecário. Mas por outro, esta tendência pode ser vista como a parcialidade dos profissionais atuantes no Corpo Editorial da Revista.

É provável que os próximos números da Revista acompanharão a evolução das reflexões e práticas da área da Ciência da informação. As

mídias digitais tendem a expandir-se nos campos do registro e disseminação do conhecimento, inflando novos pressupostos para que profissionais e pesquisadores produzam novas propostas, idéias e discussões, que alimentem ainda mais o ensejo de se continuar este dinâmico processo da construção do conhecimento.

REFERÊNCIAS

- CUNHA, I.M.R.F. Análise documentária. *In*: SMIT, J.W. (Coord.). **Análise documentária**. Brasília: IBICT, 1987. p.39.
- LANCASTER, F.W. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 1993.
- LE COADIC, Y – F. **A ciência da informação**. Brasília: Briquet de Lemos, 1996.
- LUCAS, C.R. Relação do sujeito com a linguagem: a teoria e a prática da indexação. **Transinformação**, Campinas, v.10, n.3, p.33-44, 1998.
- MARCHIORI, P.Z. “Ciberteca” ou biblioteca virtual: uma perspectiva de gerenciamento de recursos de informação. **Ciência da Informação**, Brasília, v.26, n.2, p.159-164, 1997.
- PACKER, A.L. SCIELO: uma metodologia para publicação eletrônica. **Ciência da Informação**, Brasília, v.27, n.2, p.109-121, 1998.
- ROBREDO, J.; CUNHA, M.B. **Documentação de hoje e de amanhã**: uma abordagem informatizada da biblioteconomia e dos sistemas de informação. São Paulo: Global, 1994. p.202.
- Van DER LAAN, R.H. Análise de assunto e lingüística textual: uma tentativa de aproximação. *In*: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 19., 2000, Porto Alegre. **Anais...** Porto Alegre: PUC-RS, 2000. CD-ROM.