

Estratégia de Busca

Search strategy

JEANNETTE M. KREMER *

Descrição das etapas do processo de elaboração de estratégias de busca em qualquer sistema de recuperação da informação. Apreciação crítica dos principais problemas identificados em buscas de informação, e discussão sobre avaliação de resultados de buscas.

1. SISTEMAS DE RECUPERAÇÃO DA INFORMAÇÃO

Segundo Lancaster, «a principal função de qualquer serviço de informação é servir de interface entre uma determinada população de usuários e o universo de recursos bibliográficos ou não-bibliográficos da informação. A comunidade de usuários é geralmente definida por área geográfica, filiação institucional, área de interesse ou uma combinação desses fatores (...) Um serviço de informação eficiente e moderno deveria ser capaz de garantir que qualquer documento do universo da literatura disponível, ou quaisquer dados contidos em documentos, fossem acessíveis aos membros da comunidade de usuários a que serve». (1)

* Professora da Escola de Biblioteconomia da UFMG.

Os sistemas de recuperação da informação exercem várias atividades, que podem ser classificadas conforme suas funções de **input** (entrada) e **output** (saída) da informação (2). Na sua função de **input** estão compreendidas atividades que se sucedem na seguinte ordem:

a) Os documentos são inicialmente adquiridos, o que implica a existência de critérios e políticas de seleção baseados no conhecimento das necessidades dos usuários servidos pelo sistema;

b) Os documentos são organizados e controlados de forma que possam ser identificados e localizados em resposta às demandas dos usuários. As atividades de organização e controle dos documentos incluem sua catalogação, classificação, indexação e resumo. O processo de indexação envolve dois passos distintos: a análise do conteúdo temático dos documentos e a tradução do resultado dessa análise para o vocabulário do sistema de recuperação da informação.

Na maioria dos sistemas de recuperação da informação é utilizado um vocabulário controlado para a indexação dos documentos. Esse pode ser uma lista de cabeçalhos de assuntos, um tesauro, um sistema de classificação ou simplesmente uma lista de palavras-chaves ou frases, que devem obrigatoriamente ser usados para representar os assuntos contidos nos documentos indexados. Quando é utilizado um vocabulário não controlado, ou linguagem natural, são usadas palavras ou frases na forma como aparecem nos documentos, sem nenhuma restrição aos termos que o indexador possa adotar;

c) Completado o processo de indexação, os documentos são armazenados de acordo com algum tipo de organização adotado pelo sistema. Esses documentos (e os assuntos que contêm) poderão ser recuperados através

das suas representações (referências bibliográficas, resumos, representações de assuntos) reunidas em bases de dados bibliográficos. Essas bases podem ser catálogos de bibliotecas em fichas ou automatizados, índices impressos ou índices legíveis (recuperáveis) através de computador;

d) As bases de dados são organizadas de forma que as informações sobre os documentos ali incluídas possam ser recuperadas em atendimento às demandas dos usuários. Essas podem ser expressadas de várias maneiras, sendo as mais comuns por assuntos, autores e títulos dos documentos. No caso de bases de dados automatizadas, qualquer dado incluído pode-se transformar num ponto de acesso para a recuperação da informação.

Na função de **output**, o serviço de informação atende às demandas dos seus usuários e para essa finalidade devem ser elaboradas **estratégias de busca de informação** que, obrigatoriamente, devem ser coerentes com o tipo de indexação adotado no processo de entrada dos documentos no sistema.

2. INFORMAÇÃO E INFORMAÇÃO BIBLIOGRÁFICA

«O que é informação? É aquilo que reduz a incerteza. É aquilo que assiste na tomada da decisão. Pode existir como dados em livros, computadores, pessoas, arquivos e milhares de outras fontes. Essas fontes têm de ser consideradas simplesmente como dados crus até que sejam usadas para resolver incertezas. O que muitas vezes chamamos informação é geralmente uma coleção aleatória de dados que não se transforma em informação até que seja utilizada por alguém para alcançar um fim específico. No

sentido mais amplo, qualquer estímulo oferece o potencial de prover informação, mas uma melhor forma de visualizar a informação é considerá-la como representações simbólicas da realidade (palavras — faladas e escritas; gráficos; figuras; números e combinações de todos esses). Um critério básico, entretanto, é sua capacidade de reduzir incertezas». (3)

Entre as inúmeras definições existentes para a palavra «informação», esta de Faibisoff e Ely é uma das mais pertinentes. Deve-se, no entanto, distinguir entre uma «informação» ou um «dado» no sentido descrito acima, e a «informação bibliográfica» (descrição/representação de documentos) que apenas indica **onde** poderá ser encontrada a informação propriamente dita. Conforme essa distinção, há também dois tipos de bases de dados, de acordo com sua capacidade de conter as informações propriamente ditas ou, como geralmente acontece, de permitir apenas a recuperação de informações bibliográficas. Nesse último caso, o usuário ainda terá de localizar os documentos que contêm as informações procuradas, o que pode ser problemático, pois nem sempre isso será imediatamente possível. Nenhuma biblioteca do mundo, por maior e melhor que seja, consegue hoje armazenar todos os materiais bibliográficos ou não-bibliográficos existentes.

Como solução para o problema de acesso aos materiais que contêm as informações, alguns teóricos da Ciência da Informação (4) prevêm que num futuro próximo esse acesso a todos os dados existentes em qualquer área do conhecimento será possível através de redes de computadores. A tecnologia para tanto já existe hoje; o que falta é conseguir multiplicar milhares de vezes a capacidade de memória dos computadores.

Entretanto, esse não é o pior problema, pois as grandes barreiras para a comunicação técnico-científica eram, são e continuarão sendo as diferenças políticas e a competição comercial. Apesar de tudo isso, o que hoje é ficção científica, amanhã será realidade. Prova disso é o surgimento ininterrupto de novas bases de dados que contêm as informações propriamente ditas graças, principalmente, à difusão dos microcomputadores. Deve-se ressaltar, entretanto, que a maioria dessas bases de dados é de propriedade particular (pessoas, empresas) ou de organizações governamentais.

Mesmo sendo a informação bibliográfica apenas uma etapa na busca da informação propriamente dita, ela se torna a cada dia mais imprescindível para os usuários, principalmente em áreas especializadas do conhecimento. Da explosão bibliográfica dos últimos anos (que também pode ser chamada de «poluição bibliográfica») emergiu um novo problema: se no passado havia falta de informação, hoje estamos nos sufocando e nos perdendo entre enormes quantidades de trabalhos publicados. Esse problema tornou-se tão grave a ponto de cientistas, às vezes, alegarem que levam menos tempo fazendo um experimento, do que tentando descobrir se ele já foi realizado. Naisbitt mostra alguns dados que retratam bem o fenômeno atual da explosão bibliográfica:

- «— São escritos a cada dia entre seis e sete mil artigos científicos.
- A informação científica e técnica aumenta agora 13% ao ano, o que significa que duplica a cada cinco anos e meio.
- A taxa de aumento de informações deverá brevemente pular para 40% ao ano por causa de novos e mais poderosos sistemas de informação, e por causa da população crescente de cientistas. Isso

significa que os dados deverão duplicar a cada vinte meses.

- Em 1985 o volume de informações deverá estar entre quatro e sete vezes o que era há apenas poucos anos passados». (5)

Diante desse quadro, é fácil concluir que o grande problema atual reside na criação e na utilização de mecanismos eficientes de seleção de documentos. Os atuais recursos tecnológicos, que incluem os computadores e os sistemas de comunicação, tornam essa tarefa possível. As bases de dados bibliográficos são elementos imprescindíveis para o controle bibliográfico e para a recuperação eficiente da informação. Quando automatizadas, oferecem a possibilidade de se poder localizar e selecionar eficientemente os documentos que melhor atendam às necessidades dos usuários. Essa eficiência dependerá, obviamente, da qualidade e dos recursos oferecidos pelas bases de dados, mas também dependerá da **qualidade das estratégias de busca** utilizadas.

É importante, entretanto, ressaltar que os «sistemas de recuperação da informação» (usando-se esse termo já bastante difundido) não recuperam, na verdade, «informações». Na realidade, costumam apenas recuperar documentos e suas representações, na forma de referências bibliográficas e resumos. Há ainda sistemas que recuperam apenas as representações dos documentos.

3. NECESSIDADES, DEMANDAS E BUSCAS DE INFORMAÇÃO

Buscas de informação bem sucedidas dependem, em primeiro lugar, de uma eficiente determinação das necessidades dos usuários. Nem sempre as demandas (isso é, aquilo que os usuários pedem) refletem as suas

reais necessidades, muitas vezes porque eles mesmos não são capazes de identificá-las com precisão. O problema da diferença entre demanda e necessidade também costuma ocorrer por causa de ruídos na comunicação entre os usuários e os serviços de informação, causados por deficiências ou incorreções no uso da terminologia técnico-científica, ou porque as pessoas tendem a pedir somente aquilo que pensam poder obter, sem conhecer realmente todos os recursos informacionais que se encontram à sua disposição. Outro problema é que as demandas podem expressar apenas uma parcela do que é necessário, ou então acontece justamente o contrário: os usuários não especificam qual é a sua real necessidade, fazendo o pedido de uma forma ampla e geral, quando apenas uma parte dessa demanda atenderia à necessidade. Conclui-se, portanto, que, muitas vezes, será necessário que o serviço de informação possa negociar com o usuário os termos da sua demanda de informação. Esse é o primeiro passo para a elaboração de uma estratégia de busca de informação.

Segundo Lancaster (6), as principais necessidades e demandas de usuários de serviços de informação se classificam em dois tipos:

- «1. A necessidade de se localizar e obter a cópia de um determinado documento do qual se conhece o autor ou o título.
2. A necessidade de se localizarem documentos tratando de assuntos determinados ou capazes de responder a uma questão específica».

O primeiro tipo implica uma busca de um item conhecido, e essa demanda poderá ser atendida através do uso de um catálogo de biblioteca, catálogo coletivo, ou catálogo de obras de um editor ou livraria. A eficiência

no atendimento a esse tipo de pedido dependerá da qualidade da coleção de obras disponível, e da capacidade de recuperação de documentos do serviço de informação (que depende, por sua vez, da eficácia do seu catálogo, da forma de armazenamento da sua coleção, ou da sua capacidade em obter noutra instituição as obras demandadas que não constam do seu acervo).

O segundo tipo trata de uma busca de assunto, que geralmente é mais complicada de que uma busca de um item conhecido, por exigir que os documentos apropriados ao atendimento da demanda sejam selecionados em meio à coleção de uma biblioteca, ou mesmo em meio a toda a literatura disponível num determinado campo de especialização. É nesse caso que todos os problemas da explosão bibliográfica se fazem sentir, como também pode ocorrer o contrário, pois algumas áreas de especialização ainda apresentam uma literatura escassa e difícil de se localizar.

As necessidades de obtenção de informação sobre determinados assuntos podem ser de dois tipos:

- «1. A necessidade de se obter informação para auxiliar na solução de um determinado problema, ou facilitar a tomada de uma determinada decisão.
2. Informação sobre novos desenvolvimentos em determinado campo de especialização». (7)

No primeiro tipo, o usuário deverá ser capaz de expressar bem sua necessidade, para que seja possível encontrar a informação adequada para resolver seu problema. Essa demanda poderá ser satisfeita quase sempre através de uma busca retrospectiva na literatura.

Seu atendimento pode ser possível: a) através de um único dado ou fato (possivelmente encontrado numa obra de referência ou manual especializado); b) através de um ou mais documentos que tratam do assunto; c) através de uma busca exaustiva na literatura, incluindo o máximo possível de obras publicadas sobre o assunto num determinado período de tempo.

No caso de busca de informação sobre desenvolvimento corrente de uma certa área do conhecimento, o usuário pode solicitá-la, periodicamente, ou então a iniciativa poderá ser do próprio serviço de informação. Um exemplo disso é a Disseminação Seletiva da Informação, que exige do serviço de informação um perfeito conhecimento (nem sempre possível) das necessidades dos seus usuários. Através desse tipo de serviço, os usuários são informados constantemente sobre assuntos de seu interesse, sem terem de apresentar repetidas solicitações desse tipo. A eficiência desse serviço depende principalmente da exatidão dos perfis dos usuários, que devem refletir seus interesses correntes.

Podem-se realizar, portanto, buscas por itens conhecidos e por assuntos. As buscas por assunto podem ser retrospectivas ou buscas por desenvolvimentos atuais e correntes. Elas podem resultar na prestação da informação propriamente dita ou no fornecimento de um ou mais documentos que contêm a informação procurada. Ou então, elas podem resultar numa lista de referências bibliográficas, que permitirá ao usuário selecionar as obras de que realmente necessita. A busca pode ser efetuada pelo próprio usuário, ou através de um intermediário (bibliotecário ou outro funcionário de um serviço de informação). Em qualquer um desses casos, uma busca de informação dependerá de uma **estratégia de busca adequada e eficiente** para ser bem sucedida.

4. ESTRATÉGIA DE BUSCA

Estratégia de busca «é o meio pelo qual o pesquisador se comunica com o sistema, e é muitas vezes a chave para uma busca bem sucedida». (8)

«Técnicas de busca são métodos, heurísticas, táticas, estratégias ou planos que podem ser usados nas buscas em sistemas de informação manuais ou automatizados». (9)

Portanto, estratégia de busca é o procedimento através do qual se procuram documentos determinados ou dados sobre um assunto. Pressupõe sempre a existência de um sistema de recuperação da informação, constituindo um elemento essencial da função de **output** (saída) da informação nesse sistema. Para que se possa recuperar uma obra ou uma informação anteriormente armazenada, é preciso que haja uma perfeita identidade entre a linguagem adotada pelo sistema na entrada dos dados e a linguagem utilizada durante a busca (saída dos dados). Segundo Rogers (10), «O processo de busca provê uma «ponte bibliográfica» entre a demanda e o sistema», e inclui os seguintes passos:

- «1. Esclarecer a questão (entrevista).
2. Estabelecer os parâmetros da busca baseados na entrevista.
3. Identificar sistema(s) onde deverá ser feita a busca.
4. Traduzir (indexar) a questão para a linguagem do sistema.
5. Realizar a busca.
6. Fornecer a informação». (11)

Em outras palavras, isso significa que é preciso, em primeiro lugar, estabelecer claramente qual é a necessi-

dade de informação do usuário que deverá ser atendida, procurando-se, portanto, aproximar ao máximo possível a demanda que está sendo expressada da necessidade que deverá ser identificada. No segundo passo, o pedido do usuário é analisado, procurando-se reconhecer as várias facetas dos conceitos envolvidos. Como terceiro passo, é preciso identificar qual ou quais sistemas de recuperação da informação é mais indicado para se efetuar a busca da informação solicitada. No caso de busca de um item conhecido, por exemplo, a melhor fonte poderá ser o catálogo de uma determinada biblioteca, cuja coleção possivelmente inclua a obra desejada, ou então um catálogo coletivo. No caso de busca por assunto, esse pode também ser procurado em catálogos de bibliotecas ou então a busca pode ser bem mais abrangente, incluindo até mesmo toda a literatura produzida em determinado campo de especialização. Esse último tipo de busca é realizado em índices ou **abstracts** da literatura, que podem ser impressos ou legíveis (recuperáveis) por computador, nas modalidades **on-line** ou **off-line**.

Depois da análise conceitual da demanda de informação e da escolha do sistema de recuperação da informação mais adequado à busca pretendida, é preciso traduzir os termos resultantes dessa análise conceitual para a linguagem adotada pelo sistema escolhido. Obviamente, é preciso antes conhecer essa linguagem. Portanto, os dois passos distintos na elaboração de uma estratégia de busca são:

- «1. A análise conceitual daquilo que é realmente desejado, e
2. A tradução desta análise para o conjunto de termos usados para representar os conceitos procurados em determinada base de dados a ser utilizada». (12)

Esses passos são semelhantes ao processo de indexação dos documentos que entram numa base de dados. É óbvio que esses dados só poderão ser mais tarde recuperados se os termos incluídos nas estratégias de busca corresponderem aos termos da indexação.

As estratégias de busca podem ser bastante simples ou muito complexas, variando conforme o tipo de demanda, os recursos de busca oferecidos e o grau de sofisticação da indexação e armazenamento dos dados em determinado sistema de recuperação da informação. Geralmente, os catálogos de fichas das bibliotecas são os que oferecem menos recursos de busca, e os recursos mais sofisticados são encontrados em bases de dados automatizadas, principalmente as que são **on-line**. Entretanto, os passos a serem seguidos em buscas de informação em qualquer tipo de sistema são basicamente os mesmos.

5. BUSCAS EM BASES DE DADOS ON-LINE

Os principais índices e **abstracts** relativos à literatura especializada são apresentados em duas versões, que são a tradicional impressa e publicada periodicamente e a versão com os mesmos dados, mas legíveis por computador, cuja atualização pode ser bastante freqüente (até mesmo diária).

No caso de índices e **abstracts** impressos, o processo de busca fica estritamente limitado aos pontos de acesso pré-determinados na indexação, que classificam os documentos sob um certo número de termos retirados de um vocabulário controlado, além de entradas sob nomes de autores. É então imprescindível o uso de listas de cabeçalhos de assuntos ou tesouros (ou outro instrumento de controle de vocabulário) na elaboração de uma estratégia de busca. No caso de buscas de assunto retrospectivas, será preciso manusear um certo número de

volumes desses índices, conforme o número de meses ou anos incluídos no período de interesse. Essas buscas, portanto, podem tornar-se muito trabalhosas e cansativas, e muitas informações relevantes poderão deixar de ser recuperadas.

Deve-se mencionar ainda que existem também índices impressos, porém gerados por computador, que utilizam uma indexação automática e bem simples, baseada em palavras dos títulos dos documentos (índices KWIC — Key-Word in Context, ou KWOC — Key-Word out of Context, além de outros). Esses índices são ainda mais limitados em recursos de busca e são importantes apenas para o fornecimento de informação rápida e corrente sobre a literatura de determinado campo.

Bases de dados bibliográficos **on-line** são as únicas que permitem total versatilidade de recursos de busca, número ilimitado de pontos de acesso aos dados e buscas simultaneamente em todos os anos desde a implantação do sistema. Permitem ainda o uso de vocabulário controlado e/ou linguagem livre. Enquanto nos índices impressos a indexação dos assuntos dos documentos é necessariamente pré-coordenada (isto é, as relações entre os assuntos e a ordem de citação dos termos que representam os assuntos são estabelecidos pelo indexador e só assim os dados podem ser recuperados), nos sistemas automatizados a indexação pode ser totalmente pós-coordenada (isto é, as relações entre os termos são estabelecidas livremente durante a elaboração da estratégia de busca).

É interessante enfatizar que, quanto maior for a sofisticação de recursos existentes num sistema de recuperação da informação, maior será a importância das estratégias de busca utilizadas. As buscas de informação bibliográfica mais precisas, ou mais completas, são as

realizadas em bases de dados **on-line**. Entretanto, só é possível realizar eficientemente essa tarefa através do conhecimento e perfeito domínio de todos os recursos de busca disponíveis. Para isso é preciso familiarizar-se bem com o manual de uso de cada base de dados e consultar sempre a lista de cabeçalhos de assunto ou tesouro (ou outro instrumento adotado para indexação dos dados) ao elaborar uma estratégia de busca.

São muitos os recursos que podem ser incluídos em estratégias de busca de assuntos nessas bases de dados **on-line**. Esses recursos diferem conforme o tipo de indexação adotado, podendo constituir-se de: a) apenas vocabulário controlado; b) vocabulário controlado mais linguagem livre; c) apenas linguagem livre. É interessante notar que o reconhecimento da importância e maior difusão da linguagem livre tem aumentado muito nos últimos anos. Pode ser que isso indique uma tendência para o futuro desaparecimento dos vocabulários controlados, que complicam e atrasam o processo da indexação. Essa pode ser automática através da adoção de apenas linguagem livre. Entretanto, o uso da linguagem livre pode ser mais complicado, pois exige muitas vezes a identificação e inclusão de sinônimos dos termos utilizados.

As informações bibliográficas, que representam os documentos nas bases de dados, são armazenadas em arquivos na memória de um computador. Essas informações são introduzidas obedecendo a formatos padronizados em campos específicos, constituindo parágrafos. Os formatos, os campos, os tipos de informações incluídas, suas etiquetas (**labels**) e seus recursos de busca podem variar em cada base de dados e em cada sistema de recuperação da informação que inclui essas bases de dados. As informações mais comuns, encontradas nas principais bases de dados atualmente disponíveis internacionalmente, são: número de acesso ao documento,

título, autor(es), fonte bibliográfica, ano de publicação, língua, notas, tipo de publicação, descritores, resumo.

O seguinte exemplo mostra a recuperação de uma informação bibliográfica sobre um determinado artigo de periódico indexado pelo ERIC — Educational Resources Information Center, em busca realizada no ERIC/ECE Computer Search Service:

EJ 120612

TITLE: CHILDREN'S SELF-CONCEPT AND ATTITUDE TOWARD SCHOOL IN OPEN AND TRADITIONAL CLASSROOMS

PERSONAL AUTHOR: KLAFF, FRANCES R.; DOCHERTY, EDWARD M.

PUBLICATION DATE: SUM 75

JOURNAL CITATION: JOURNAL OF SCHOOL PSYCHOLOGY; 13; 2; 97-103

DESCRIPTORS: * OPEN EDUCATION; * SELF CONCEPT; * STUDENT ATTITUDES; * CLASSROOM ENVIRONMENT; * SEX DIFFERENCES; ELEMENTARY EDUCATION; SELF DIRECTED CLASSROOMS; RESEARCH PROJECTS; EDUCATIONAL ENVIRONMENT.

ABSTRACT: CHILDREN FROM A TRADITIONAL SCHOOL AND TWO OPEN SCHOOLS WERE COMPARED, USING THE PICTORIAL SELF-CONCEPT SCALE AND THE ATTITUDES TOWARD SCHOOL SCALE. ATTITUDE TOWARD SCHOOL WAS HIGHER FOR FEMALES AND DECREASED FROM KINDERGARTEN TO GRADE 2. SELF CONCEPT INCREASED FROM KINDERGARTEN TO GRADE 2, WITH NO SIGNIFICANT SEX DIFFERENCES.
(AUTHOR)

Um segundo exemplo mostra como foi recuperada a informação bibliográfica sobre uma monografia indexada também no ERIC, mas em busca realizada em outro sistema, o BRS — Bibliographic Retrieval Services, Inc.:

AN ED 129229.
IN SYRACUSE UNIV. N. Y. SCHOOL OF
INFORMATION STUDIES (BBB 13596).
TI SUBJECT ACCES PROJECT; QUARTERLY
PROGRESS REPORT.
SO SEP 76.
IS RIEFEB 77.
CH IR 004005.
PR EDRS PRICE MF — \$ 0.83 HC — \$ 1.67
PLUS POSTAGE.
PT K.
NT 16P. INCLUDES OCCASIONAL NEWSLETTER
NO. 1. AUGUST 1976.
YR 77
MJ CATALOGING EXPERIMENTAL-PROGRAMS
INDEXING SUBJECT-INDEX-TERMS.
MN HUMANITIES INFORMATION-RETRIEVAL
LIBRARY-RESEARCH SOCIAL-SCIENCES
UNIVERSITY-LIBRARIES.
ID MACHINE READABLE CATALOGING; MARC;
SUBJECT ACCESS.
AB AN ATTEMPT IS BEING MADE TO IMPROVE
SUBJECT ACCESS TO MONOGRAPHS BY
AUGMENTING MARC (MACHINE READABLE
CATALOGING) RECORDS. WORKING WITH A
SAMPLE OF BOOKS DRAWN FROM THE
COLLECTIONS OF THE UNIVERSITY OF
TORONTO AND COMPRISING A NUMBER OF
SUBJECT AREAS IN THE HUMANITIES AND

SOCIAL SCIENCES, THE PROJECT PLANS TO ENLARGE THE MARC DESCRIPTION BY USING A SET OF DECISION RULES FOR SELECTING WORDS AND PHRASES FOUND IN THE INDEX AND/OR THE TABLE OF CONTENTS. THIS FIRST QUARTERLY REPORT FOR THE PERIOD JUNE TO AUGUST 1976 DESCRIBES THE SAMPLE OF SELECTED MONOGRAPHS, SUMMARIZES THE PROJECT BUDGET TO DATE, AND PROJECTS ACTIVITIES FOR THE FUTURE.

Nota-se, nesses dois exemplos, que há variação nas etiquetas que identificam os parágrafos. No segundo exemplo, os parágrafos indicam o número de acesso ao documento (AN), nome da instituição onde o documento foi produzido (IN), título (TI), fonte (SO), número do índice **RIE** onde documento foi indexado (IS), código da Clearinghouse (CH), preço (PR), código do tipo de documento (PT), nota descritiva (NT), data da indexação (YR), descritores gerais (MJ), descritores específicos (MN), identificadores (ID), resumo (AB). As mesmas referências que aparecem nas versões legíveis por computador podem ser recuperadas nos dois índices impressos do ERIC: **Current Index to Journals in Education — CIJE** (que indexa artigos de periódicos — Ver primeiro exemplo), e **Resources in Education — RIE** (que indexa monografias, relatórios, trabalhos apresentados em congressos, etc. — ver segundo exemplo).

A organização dos dados em parágrafos garante a flexibilidade na busca, pois cada um deles pode exercer uma função específica. A busca pode ser por autor e/ou instituição, por título, ou por assunto. No caso de buscas por assunto, nos exemplos apresentados, elas podem ser por vocabulário controlado ou por linguagem livre. No

caso de buscas por vocabulário controlado, são chamados os parágrafos dos descritores e identificadores e será preciso consultar primeiramente o **Thesaurus of ERIC Descriptors** para se obter uma perfeita identificação entre a linguagem do sistema e a linguagem da busca (senão a recuperação da informação será impossível). Muito mais flexível, apesar de obedecer a algumas regras, é a busca que utiliza a linguagem natural, não atrelada ao **Thesaurus**, na qual qualquer termo utilizado pode resultar na recuperação da informação. Os parágrafos que permitem esse último tipo de busca são as áreas do título e do resumo.

A busca que utiliza descritores é bem mais precisa do que a outra por linguagem natural, pois pode-se ter certeza da existência ou não dos assuntos procurados. Ela pode utilizar um único descritor, que pode resultar num número exagerado e indesejável de recuperações, ou então utilizam-se combinações de descritores, segundo a Lógica Booleana, que permite delimitar e tornar mais precisa uma busca por assunto. Os operadores lógicos que permitem combinações de dois termos de busca são: OR (um ou ambos os termos devem aparecer no documento), AND (ambos os termos devem estar na mesma descrição de documento), NOT (um termo não deve constar de uma descrição de documento), XOR (um termo ou o outro deve aparecer na descrição do documento, mas não ambos). Não há, geralmente, restrições para o número de termos e operadores lógicos incluídos numa estratégia de busca, mas há limites para o número de caracteres e espaços utilizados (nos exemplos apresentados o limite era de 234 caracteres).

Muitos outros recursos existem, sendo que alguns podem ser utilizados tanto nas buscas por vocabulário controlado como naquelas por linguagem natural. Alguns dos mais importantes são: truncar palavras (que permite

recuperar todos os termos com a mesma raiz comum), alocar pesos aos termos conforme sua incidência nos parágrafos ou usar a proximidade dos termos num mesmo parágrafo. Nesse último caso, os operadores podem ser: SAME (os termos devem estar no mesmo parágrafo na descrição de um documento), WITH (os termos devem ocorrer na mesma sentença na descrição de um documento), e ADJ (os termos devem estar imediatamente adjacentes um ao outro, nessa ordem). Há portanto muitas formas de realizar buscas usando vários termos compostos, por exemplo:

- | | |
|-----------------------------|---|
| library-science | como um termo de vocabulário controlado |
| library and science | permitindo que dois termos ocorram em qualquer parte da descrição do documento |
| same | |
| library with science
adj | colocando limitações na proximidade em que os termos ocorrem na descrição de um documento |

Também é possível limitar as buscas de acordo com alguns critérios que, como nos outros tipos de recursos, variam para cada base de dados ou sistema de recuperação da informação que fornece esse tipo de serviço. Nos exemplos dados, os parágrafos que servem para limitar as buscas são: o número de acesso ERIC (pode-se escolher qual dos dois índices deve ser pesquisado, **CIJE**, ou **RIE**), data, código da Clearinghouse e tipo de publicação. São os seguintes os operadores utilizados nessas buscas no sistema BRS:

Operador	Símbolo	Significado
eq	=	igual a
ne		diferente de
lt	<	menor que
gt	>	maior que
wl		dentro dos limites de
ol		fora dos limites de

Por exemplo, assim será possível recuperar descrições de documentos publicados antes de certa data (< 1945) ou depois de determinado período de tempo (> 1980). Pode-se limitar ainda a buscas de documentos escritos em determinada(s) língua(s), formas de publicação, ou qualquer outro critério incluído no sistema. O segundo exemplo de descrição de documentos apresentado anteriormente foi um dos 112 documentos recuperados no ERIC, usando-se a seguinte estratégia:

```
information-retrieval or information adj retrieval
Result 2215
limit / yr = 77
Result 112
```

Portanto, havia 112 documentos indexados em 1977 sobre recuperação da informação, entre os 2.215 sobre o assunto sem restrição de data.

Esses são os principais recursos de busca em bases de dados **on-line**, mas muitos outros existem ou poderão ser ainda desenvolvidos no futuro. É interessante o fato de o pesquisador poder «conversar» com o sistema, e ir desenvolvendo e refinando sua estratégia de busca enquanto ela vai sendo realizada. A maioria dos sistemas **on-line** elabora, separadamente para cada base de dados, um arquivo dicionário apresentando em ordem alfabética todos os termos que podem ser pesquisados, incluindo estatísticas a respeito do número de documentos onde

esses termos ocorrem. Quando a busca é feita através de combinações de termos, o computador verifica em quantos documentos eles coincidem. Na busca indicada anteriormente, o assunto «recuperação da informação» estava representado em 2.215 documentos, o que foi considerado um número exagerado pelo usuário. Então, decidiu-se que havia necessidade apenas de recuperar os trabalhos publicados em 1977, e com essa limitação o sistema encontrou 112 documentos, número considerado razoável pelo usuário. É claro que esse número poderia ter sido baixado mais ainda, através de recursos da Lógica Booleana ou outras limitações possíveis. Quando o usuário concorda com o número de documentos indicados, pode pedir que sejam impressas as informações bibliográficas. Mesmo nesse caso, ainda há delimitações possíveis, pois pode pedir para imprimir todos ou um número determinado (por exemplo, as primeiras vinte referências), ou então qualquer uma (por exemplo, pode pedir aleatoriamente o documento número 10 para amostragem).

Os recursos de busca em bases de dados bibliográficos estão se tornando a cada dia mais sofisticados e, ao mesmo tempo, demasiadamente complicados para serem usados na sua plenitude pela maioria dos seus potenciais usuários. Espera-se que os novos desenvolvimentos tecnológicos na construção de computadores e os estudos sobre inteligência artificial resultem, num futuro próximo, em sistemas de recuperação da informação mais transparentes e mais simples na sua utilização. Também seria bastante proveitoso se os fornecedores de serviços nessa área pudessem chegar a um consenso a respeito de um formato único para a representação dos dados bibliográficos, permitindo assim a busca em várias bases de dados, consecutivamente, através de uma única estratégia.

6. TREINAMENTO EM ESTRATÉGIAS DE BUSCA

Tanto os bibliotecários quanto os usuários devem ser treinados em estratégias de busca, para que os recursos oferecidos pelos sistemas de recuperação da informação sejam plenamente utilizados.

O ideal (nem sempre alcançado) seria que os usuários fossem capazes de realizar suas próprias buscas de informação, pois eles, melhor do que ninguém, podem conhecer suas verdadeiras necessidades. Outra vantagem é que eles conhecem (ou deveriam conhecer) a terminologia em suas áreas de especialização melhor do que os não-especialistas. É, portanto, importante que sejam treinados, procurando-se transmitir-lhes o conhecimento das fontes de informação disponíveis e dos recursos de busca existentes nos sistemas de recuperação da informação. Mesmo que eles próprios não realizem as buscas, esse treinamento continua importante, pois os usuários precisam ser conscientizados a respeito da necessidade de explicarem detalhadamente o assunto de suas buscas de informação. Se falhar a comunicação entre os usuários e o serviço de informação, será impossível alcançar a eficiência no seu atendimento.

Os bibliotecários (ou outros intermediários utilizados entre os usuários e os recursos informacionais existentes) devem conhecer perfeitamente todas as fontes de informação disponíveis, e saber utilizá-las eficientemente. Mas, isso apenas não basta. Devem também conhecer as características dos usuários, suas principais demandas e necessidades. E, principalmente, serem capazes de discutir com os usuários a respeito das estratégias de busca mais adequadas a cada tipo de demanda.

No caso de uso de bases de dados **on-line** é imprescindível conhecer bem o manual de uso e o vocabulário utilizados em cada uma delas. Deve-se também

saber construir uma estratégia de busca adequada para cada demanda/necessidade de informação. Além disso, é importante saber realizar a busca da forma mais econômica possível, para se obter o mais alto custo/benefício, sem desperdício de tempo de conexão com o computador.

Segundo Lancaster, os fatores que influem no sucesso da interação usuário-sistema na etapa da solicitação da informação são: «a) a habilidade do usuário em definir sua necessidade para si mesmo; b) a habilidade do usuário em se expressar; c) as expectativas do usuário a respeito das capacidades do sistema — há uma forte tendência para o usuário pedir, não o que realmente quer, mas aquilo que ele pensa que o sistema poderá dar-lhe — e, d) a quantidade e tipo de assistência que lhe é prestada pelo sistema». (13) O mesmo autor, na sua avaliação do sistema MEDLARS, verificou que grande parte dos fracassos nas buscas era devida a uma interação usuário-sistema inadequada. Outra conclusão interessante desse estudo foi que houve maior número de sucessos em buscas para as quais as solicitações tinham sido enviadas por escrito pelos usuários. Isso provavelmente ocorreu porque escrever uma solicitação exige maior disciplina do que expressá-la oralmente, e também porque esses usuários se preocuparam mais com suas necessidades de informação e menos com conjecturas sobre se o sistema poderia fornecer-lhes o que queriam. (14)

Tenopir (15) discute os principais erros cometidos por pesquisadores principiantes em bases de dados **on-line**. Sua lista inclui erros no uso da Lógica Booleana, principalmente com os operadores OR e AND, que podem ser confundidos. Erros de digitação podem ser desastrosos, como deixar de incluir os espaços exigidos, ou colocar as letras «L» minúscula ou «O» maiúscula em lugar dos números 1 (um) e 0 (zero). Um problema sério é ignorar as diferenças existentes entre as diversas bases de dados,

confundindo seus recursos e terminologia. Alguns pesquisadores têm ainda uma confiança exagerada no computador, e não se preocupam suficientemente em construir estratégias de busca eficientes, pois pensam que o computador resolverá todos os seus problemas. A solução para evitar esses erros é um treinamento adequado para os bibliotecários. Entretanto, também seria necessário que os serviços que oferecem bases de dados procurassem diminuir as diferenças existentes entre elas, facilitando assim seu uso e as buscas sucessivas em várias bases de dados, utilizando-se uma única estratégia (em vez de várias, como costuma acontecer). Outro grande problema é que diferentes sistemas usam linguagens e recursos de busca diferentes, confundindo seus usuários. Por exemplo, um pesquisador com experiência no uso do sistema BRS terá dificuldades iniciais se mudar para DIALOG ou ORBIT, precisando aprender a lidar com as diferenças que encontrará para usá-los.

Ladendorf (16) aborda os problemas que ocorrem quando os bibliotecários depositam uma fé exagerada nos recursos de busca de informação fornecidos pelos computadores. Aponta quatro falsos juízos dos bibliotecários a respeito da interação usuário-computador, todos derivados do fato de que os computadores são ótimos para lidar com referências bibliográficas, mas não são bem equipados para lidar com pessoas. O primeiro engano é acreditar que os usuários precisam dos computadores. Isso realmente ocorre nas buscas de assuntos complexos ou buscas retrospectivas e exaustivas, mas esses são casos que não acontecem com freqüência na rotina de uma biblioteca. O segundo falso juízo é pensar que os usuários entendem os computadores. Na verdade, os usuários não têm, geralmente, muita vontade de aprender as regras dos sistemas de recuperação da informação e um sistema automatizado pode representar mais uma barreira e

frustração entre eles e a informação desejada. O terceiro engano é acreditar que os usuários são compreendidos. O problema, nesse caso, é que foram desenvolvidos sistemas de recuperação da informação adaptados aos computadores, e não às pessoas, sem a devida compreensão das suas diferenças em comunicação. Geralmente os usuários conseguem explicitar suas necessidades apenas a outras pessoas, sendo mal interpretados pelas máquinas, que não admitem falta de precisão de linguagem. O último engano é pensar que todos os usuários são iguais. Um computador responde da mesma forma a uma mesma pergunta, não importando quem solicitou. Por exemplo, uma criança, um adulto não-especialista, e um especialista em determinado assunto, podem fazer um mesmo pedido, mas precisam de respostas diferentes. Nesse caso, apenas a sensibilidade humana poderá distinguir a necessidade em fornecer dados diferentes a cada usuário. Finalizando, Ladendorf mostra o papel do bibliotecário nesse contexto:

«O equipamento de processamento de dados eletrônico tem ajudado a colocar a explosão da informação sob controle bibliográfico; mas, fazendo isso, acrescentou mais um nível de complexidade ao problema do usuário de biblioteca. Isso significa que os bibliotecários, que podem ajudar os usuários a tirar a máxima vantagem dos recursos disponíveis, são mais necessários do que jamais foram. Nessa tarefa, um ser humano pode realizar um trabalho que máquina nenhuma poderá igualar. A programação mais elaborada e o equipamento mais caro não são substitutos para anos de vida numa sociedade humana». (17)

7. AVALIAÇÃO DE RESULTADOS DE BUSCAS DE INFORMAÇÃO

O grau de sucesso nas buscas de informação depende igualmente do processo de indexação dos documentos, da linguagem de indexação, da interface usuário-sistema, e das estratégias de busca empregadas. Conseqüentemente, caso sejam detectadas falhas na recuperação da informação, é importante avaliar qual desses fatores deve ser responsabilizado e sanado. Muitas vezes a eficiência do sistema de recuperação da informação pode ser imediatamente aumentada através de uma melhor comunicação com os usuários, e aperfeiçoamento nas estratégias de busca utilizadas.

Com o objetivo de avaliar o processo de busca, Lancaster propôs algumas questões que devem ser consideradas:

- «1. Quais são as exigências dos usuários a respeito de revocação e precisão?
2. As estratégias de busca podem ser elaboradas para atender pedidos de alta revocação ou alta precisão?
3. Com que eficiência os pesquisadores podem selecionar resultados de buscas?
Qual é o efeito dessa seleção no índice de revocação e precisão?
4. Quais são as modalidades mais promissoras de interação usuário-sistema?
 - a) Maior contato na etapa da solicitação de informação.
 - b) Maior contato na etapa da elaboração da estratégia de busca.

c) Um processo de busca iterativo que apresenta ao usuário uma amostra de referências bibliográficas recuperadas numa busca de «primeira aproximação» com o sistema, permitindo-lhe reformular seu pedido à luz desses resultados.

5. Qual é o efeito desses vários tipos de interação sobre a rapidez de atendimento?» (18)

Os usuários, obviamente, «esperam que o sistema seja capaz de recuperar um ou mais documentos que contribuam para satisfazer alguma necessidade de informação (**documentos relevantes**). Presume-se que todos os usuários apresentem alguma necessidade de informação — do contrário não se teriam aproximado do sistema». (19) Para se analisar quantitativamente o grau de sucesso do sistema em recuperar documentos relevantes, Lancaster (20) propôs uma estimativa de índice de revocação (**recall ratio**):

$$\frac{\text{Nº de documentos relevantes recuperados pelo sistema}}{\text{Nº total de documentos relevantes contidos no sistema}} \times 100$$

Nesse tipo de avaliação, solicita-se ao usuário que aponte, entre todos os documentos ou representações de documentos recuperados durante uma busca, aqueles que são relevantes. A dificuldade, muitas vezes, é descobrir o número total de documentos relevantes contidos no sistema, e isso nem sempre é possível. O problema é que será preciso utilizar um segundo estratagema, diferente do processo de busca utilizado, para detectar esses documentos. Por exemplo, no caso da busca ter sido realizada numa base de dados bibliográficos **on-line**, poder-se-ia comparar os resultados obtidos com os resultados de uma busca no índice correspondente na sua versão impressa.

Bem mais simples é realizar uma estimativa do índice de precisão (**precision ratio**) de resultados de buscas de informação, definida por Lancaster (21) como:

$$\frac{\text{Nº de documentos relevantes recuperados pelo sistema}}{\text{Nº total de documentos recuperados pelo sistema}} \times 100$$

É interessante mencionar que se pode verificar uma relação inversa entre os índices de revocação e precisão. Quanto mais alta for a recuperação, menor será a precisão e vice-versa. Por causa disso, é necessário conhecer a preferência dos usuários.

Outra questão refere-se à novidade das informações recuperadas. Um documento, ou sua representação, pode ser considerado relevante por um usuário, mas não lhe interessar por já conhecê-lo. A capacidade do sistema em fornecer documentos relevantes, e ao mesmo tempo novos, pode ser avaliada quantitativamente através da estimativa de índice de novidade (**novelty ratio**) proposta também por Lancaster (22), podendo-se escolher entre duas fórmulas:

$$\frac{\text{Nº de novos documentos relevantes recuperados}}{\text{Nº de documentos relevantes recuperados}} \times 100$$

ou

$$\frac{\text{Nº de novos documentos relevantes recuperados}}{\text{Nº de documentos recuperados}} \times 100$$

É também importante analisar o fator tempo utilizado nas buscas, que pode envolver os seguintes aspectos:

- a) tempo gasto em buscas em relação ao número de documentos relevantes recuperados;

- b) tempo gasto em buscas em bases de dados **on-line**, com o objetivo de verificar o custo médio dessas buscas (tempo de conexão com o computador, incluindo gastos com telecomunicação e **royalties** pelo uso das bases de dados). Nesses casos, estratégias bem elaboradas tornam as buscas mais econômicas;
- c) capacidade do sistema em fornecer os documentos propriamente ditos (correspondentes às representações dos documentos recuperados nas buscas) num período de tempo considerado adequado pelos usuários. Esse tempo é mais subjetivo do que físico, pois enquanto um usuário pode achar intolerável esperar uma hora pela informação ou documento solicitado, outro pode ficar satisfeito se puder ser atendido dentro de uma semana, ou mesmo dentro de um mês. O tempo nesses casos dependerá da capacidade do sistema em possuir o documento em coleção própria, ou da capacidade em conseguir localizá-lo e obtê-lo (empréstimo ou cópia) em outra instituição.

Concluindo, é importante enfatizar que avaliações periódicas dos resultados obtidos em buscas de informação são um instrumento poderoso para detectar falhas no atendimento aos usuários. Conhecendo-se os problemas existentes, será possível procurar caminhos para promover um contínuo aperfeiçoamento de um sistema de recuperação da informação.

Description of the steps in the development process of search strategies within any kind of information retrieval system. Critical appraisal of the main problems occurring in information searches, and discussion on the evaluation of information search results.

REFERÊNCIAS BIBLIOGRÁFICAS

1. LANCASTER, F. Wilfrid. **Information retrieval systems: characteristics, testing and evaluation.** 2. ed. New York, Wiley — Interscience, c1979. p. 4-5.
2. Ibid., p. 7-10.
3. FAIBISOFF, Sylvia G. & ELY, Donald P. Information and information needs. In: KING, Donald W., ed. **Key papers in the design and evaluation of information systems.** White Plains, Knowledge Industry Publications, c1978. p. 270.
4. LANCASTER, F. Wilfrid. **Toward paperless Information systems.** New York, Academic Press, c1978. 179 p.
5. NAISBITT, John. **Megatrends; ten new directions transforming our lives.** New York, Warner Books, c1984. p. 16-17.
6. LANCASTER. **Information retrieval systems.** Op. cit., p. 6-7.
7. Ibid.
8. HAWKINS, Donald T. Online information retrieval systems. In: WILLIAMS, Martha E., ed. **Annual Review of Information Science and Technology.** Washington, American Society for Information Science, c1981. v. 16, p. 182-83.
9. BATES, Marcia J. Search techniques. In: WILLIAMS, Martha E., ed. **Annual Review of Information Science and Technology.** Washington, American Society for Information Science, c1981. v. 16, p. 139.
10. ROGERS, Sharon. Research strategies: bibliographic instruction for undergraduates. **Library Trends**, Urbana, **29** (1): 72, Summer 1980.
11. BENSON, James & MALONEY, Ruth K. Principles of searching. **RQ**, Chicago, **14**: 316, Summer 1975. APUD ROGERS, op. cit., p. 72.
12. LANCASTER, F. Wilfrid. **Indexing and searching of data bases.** Urbana, Illinois, Graduate School of Library and

information Science, University of Illinois [1977]. 25 p.
apostila.

13. LANCASTER. **Information retrieval systems**. Op. cit., p. 147.
14. Ibid., p. 147-152.
15. TENOPIR, Carol. To err is human: seven common searching mistakes. **Library Journal**, New York, 109 (16): 635-6, Apr. 1984.
16. LADENDORF, Janice M. Four misconceptions of machine fanatics in judging direct human-computer interation. **Wilson Library Bulletin**, New York, 48 (7): 561-64, March 1974.
17. Ibid., p. 564.
18. LANCASTER. **Information retrieval systems**. Op. cit., p. 124.
19. LANCASTER, F. Wilfrid. Evaluation and testing of information retrieval systems. In: KENT, Allen & LANCOUR, Harold, ed. **Encyclopedia of Library and Information Science**. New York, Marcel Dekker, c1972. v. 8, p. 236.
20. Ibid.
21. Ibid.
22. LANCASTER. **Information retrieval systems**. Op. Cit., p. 133.

BIBLIOGRAFIA RECOMENDADA

- ADAMS, Arthur L. Planning search strategies for maximum retrieval from bibliographic databases. **Online Review**, 3 (4): 373-79, Dec. 1979.
- BATES, Marcia J. Search techniques. In: WILLIAMS, Martha E., ed. **Annual Review of Information Science and Technology**. Washington, American Society for Information Science, c1981. v. 16, p. 139-69.
- BONTA, Bruce D. Online searching in the reference room. **Library Trends**, Urbana, 31 (3): 495-510, Winter 1983.
- R. Esc. Bibliotecon. UFMG, Belo Horizonte, 14(2):187-220, set. 1985

- BRANDHORST, Wesley, T. & ECKERT, PHILIP F. Document retrieval and dissemination systems. In: CUADRA, Carlos A., ed. **Annual Review of Information Science and Technology**. Washington, American Society for Information Science, c1972. v. 7, p. 379-437.
- CLINIC ON LIBRARY APPLICATIONS OF DATA PROCESSING, Urbana, 1975. **The use of computers in literature searching and related reference activities in libraries**. Urbana — Champaign, University of Illinois, c1976. 159 p.
- CONNOR, Judith Holt. Selective dissemination of information: a review of the literature and the issue. **The Library Quarterly**, Chicago, 37 (4): 373-91, Oct. 1967.
- EAST, Harry. The development of SDI services. **Aslib Proceedings**, London, 20 (11):483-91, Nov. 1968.
- FOSKETT, A. C. **The subject approach to information**. 3. ed. London, Clive Bingley, 1977. 476 p.
- GECHMAN, Marvin C. Machine-readable bibliographic data bases. In: CUADRA, Carlos A., ed. **Annual Review of Information Science and Technology**. Washington, American Society for Information Science, c1972. v. 7, p. 323-78.
- HARTER, Stephen P. & FENICHEL, Carol H. Online searching in library education. **Journal of Education for Librarianship**, State College, 23 (1): 3-22, Summer 1982.
- HAWKINS, Donald T. Online information retrieval systems. In: WILLIAMS, Martha E., ed **Annual Review of Information Science and Technology**. Washington, American Society for Information Science, c1981. v. 16, p. 171-208.
- INFORMATION retrieval research. London, Butterworths, c1981. 389 p.
- KING, Donald W. **Key papers in the design and evaluation of information systems**. White Plains, N. Y., Knowledge Industry, c1978. 405 p.
- KNOX, Douglas R. & HLAVA, Marjorie M. K. Effective search strategies. **Online Review**, 3 (2): 148-52, June 1979.
- LADENDORF, J. M. Four misconceptions of machine fanatics in judging direct human-computer interaction. **Wilson Library Bulletin**, New York, 48 (7): 561-4, March 1974.

- LANCASTER, F. Wilfrid. Evaluation and testing of information retrieval systems. In: KENT, Allen & LANCOUR, Harold, ed. **Encyclopedia of Library and Information Science**. New York, Marcel Dekker, c1972. v. 8, p. 234-59.
- LANCASTER, F. Wilfrid. **Information retrieval systems: characteristics, testing and evaluation**. 2. ed. New York, John Wiley, c1979. 381 p.
- LANCASTER, F. Wilfrid. **The measurement and evaluation of library services**. Washington, Information Resources Press, c1977. 395 p.
- LANCASTER, F. Wilfrid. **Toward paperless information systems**. New York, Academic Press, c1978. 179 p.
- LANCASTER, F. Wilfrid. **Vocabuiary control for information retrieval**. Washington, Information Resources Press, c1972. 233 p.
- LANCASTER, F. Wilfrid & FAYEN, E. G. **Information retrieval on-line**. Los Angeles, Melville Publishing, c1973. 597 p.
- LANCASTER, F. Wilfrid & OWEN, Jeanne M. Information retrieval by computer. In: HAMMER, Donald P., ed. **The information age: its development, its impact**. Metuchen, N. J., Scarecrow Press, c1976. p. 1-33.
- OLDROYD, B. K. Study of strategies used in online searching 5: differences between the experienced and the inexperienced searcher. **Online Review**, 8 (3): 233-44, June 1984.
- PITERNICK, Anne B. Searching vocabularies: a developing category of online search tools. **Online Review**, 8 (5): 441-49, Oct. 1984.
- ROGERS, Sharon J. Research strategies: bibliographic instruction for undergraduates. **Library Trends**, Urbana, 29 (1): 69-81, Summer 1980.
- TEDD, Lucy A. Education, training and marketing for online retrieval systems. **Online Review**, 3 (2): 205-212, June 1979.
- TENOPIR, Carol. To err is human: seven common searching mistakes. **Library Journal**, New York, 109 (6): 635-6, Apr. 1984.
- R. Esc. Bibliotecon. UFMG, Belo Horizonte, 14(2):187-220, set. 1985

VICKERY, B. C. Thesaurus — a new word in documentation. **Journal of Documentation**, London, **16** (4): 181-89, Dec. 1960.

WANGER, Judith. Education and training for online systems. In: WILLIAMS, Martha E., ed. **Annual Review of Information Science and Technology**. Washington, American Society for Information Science, c1979. v. 14, p. 219-45.

WILDE, Daniel U. Generation and use of machine-readable data bases. In: WILLIAMS, Martha E., ed. **Annual Review of Information Science and Technology**. Washington, American Society for Information Science, c1976. v. 11, p. 267-98.