

05/04/2001 amara

já tem na base no Spt. 0276902

no adm. 277028



## Estudo comparativo de algumas linguagens de indexação : eficácia e tempo de pesquisa

Maria Antonietta Requião Piedade

Departamento de Biblioteconomia e Ciência  
da Informação, Centro de Documentação e  
Letras, Universidade Santa Úrsula.

Assembléia Legislativa do Rio de Janeiro.

Dissertação apresentada ao Instituto  
Brasileiro de Bibliografia e Documentação/  
Universidade Federal do Rio de Janeiro  
para obtenção de grau de mestre em  
Biblioteconomia e Documentação.

*Orientador:* Professor F. Wilfrid Lancaster  
University of Illinois, Urbana, Ill., E.U.

A Lidia Sambaqui

Nossa gratidão ao Professor Wilfrid Lancaster pela valiosa orientação, sem a qual dificilmente este trabalho seria realizado.

Nossos agradecimentos à Professora Gilda Braga, por ter aceito iniciar a orientação desta dissertação, interrompida devido a sua viagem para os Estados Unidos.

Agradecemos ao IBBD, especialmente a Ida Maria Cardoso Lima, pela preparação do índice KWIC.

Desejamos assinalar a boa vontade e a eficiência das bibliotecárias da Biblioteca do IBBD.

Não podemos deixar de lembrar, ainda, as alunas da Universidade Santa Ursula, que realizaram as pesquisas necessárias à dissertação, e as amigas bibliotecárias, que formularam as perguntas utilizadas nos testes.

## SUMÁRIO

0 - INTRODUÇÃO	6
1 - OBJETIVO	22
2 - MATERIAL	23
3 - MÉTODO	24
3.1 - Armazenagem	24
3.1.1 - KWIC	24
3.1.2 - Uniterm	25
3.1.3 - Cabeçalhos de Assunto	26
3.1.4 - Classificação Decimal de Dewey	27
3.1.5 - Classificação Decimal Universal	28
3.1.6 - Classification of Library and Information Science	30
3.2 - Perguntas	31
3.3 - Recuperação	31
3.4 - Avaliação	32
4 - RESULTADOS	34
5 - INTERPRETAÇÕES E CONCLUSÕES	65
6 - CITAÇÕES BIBLIOGRÁFICAS	71

## ANEXOS

## SINOPSE

Relata um teste comparativo de recuperação da informação através das linguagens KWIC, Uniterm, Cabeçalhos de Assunto, Classificação Decimal de Dewey, Classificação Decimal Universal e Classification of Librarianship and Information Sciences (uma classificação facetada de autoria do Classification Research Group).

Para o teste foram selecionados 100 documentos brasileiros sobre biblioteconomia e ciência da informação e preparados 6 índices, segundo as linguagens indicadas no parágrafo anterior.

Oito bibliotecários prepararam 10 perguntas respondíveis por um ou mais documentos da coleção-teste.

Sessenta alunas da Faculdade de Biblioteconomia e Documentação da Universidade Santa Úrsula, divididas em 6 grupos de 10, pesquisaram nos 6 índices, para encontrar documentos que respondessem às 10 perguntas formuladas.

Cada pesquisador anotou: tempo de pesquisa, entradas procuradas, documentos selecionados para exame e documentos que de fato respondem ao perguntado. A esses dados foi acrescentado o número de documentos registrados sob cada uma das entradas procuradas.

Computaram-se os dados coletados e obtiveram-se os números relativos a cada sistema.

Pelo exame dos dados observou-se que os sistemas alfabéticos (KWIC, Uniterm e Cabeçalhos de Assunto) tiveram melhor desempenho do que os sistemas classificados e que entre estes a melhor atuação coube à classificação facetada.

Quanto ao Sucesso na Pesquisa, os sistemas se escalonaram: Cabeçalhos de Assunto (96%), Uniterm (94%), KWIC (85%), Classification of Librarianship and Information Sciences (84%), Classificação Decimal Universal (75%) e Classificação Decimal de Dewey (73%).

O exame das falhas de recuperação mostrou que foram devidas à falhas humanas, na pesquisa ou na indexação, e não originárias de defeitos das próprias linguagens.

A preferência dos pesquisadores favoreceu Cabeçalhos de Assunto.

Os resultados finais revelaram Cabeçalhos de Assunto como o melhor desempenho em 5 das 6 variáveis estudadas e como o segundo na sexta.

## 0 - INTRODUÇÃO

A Ciência da Informação é a disciplina que investiga as propriedades, o comportamento, as forças que regem a comunicação do conhecimento, bem como os objetivos, as propriedades e o desempenho dos métodos de controlar a informação.

A avaliação do desempenho das várias linguagens de indexação é um dos temas que vem preocupando os cientistas da informação e nas últimas duas décadas iniciaram-se as pesquisas sobre os métodos de controle da informação.

Vários dos mais destacados profissionais da informação no estrangeiro vêm se dedicando ao assunto, como Cleverdon, Mills e Keen, levando a efeito testes experimentais sobre linguagens de indexação. Lancaster, preocupado em testar sistemas de informação em funcionamento, Salton investigando a indexação mecanizada; Carlisle e Fetter estudando o comportamento do usuário diante dos sistemas mecanizados.

As pesquisas neste campo podem ser divididas em dois grandes grupos:

- 1) pesquisas sobre o desempenho de sistemas de recuperação da informação em funcionamento;
- 2) pesquisas experimentais sobre o desempenho de linguagens de indexação.

Entre os estudos do primeiro grupo, destaca-se a avaliação do sistema MEDLARS, sob a responsabilidade de Wilfrid Lancaster, e, entre os testes experimentais, merecem especial menção os testes da ASTIA, os projetos Cranfield I e II, complementados pelos testes da Western Reserve University e da English Electric Co.; o teste de Aberystwyth e as experiências feitas por James Carlisle e Robert Fetter, sobre a interação homem-computador.

### 0.1 - Teste da Astia (Armed Services Technical Information Agency)

O primeiro teste experimental foi o ASTIA-Uniterm, levado a efeito em 1953, comparando o desempenho do sistema Uniterm de indexação, de Mortimer Taube, e o catálogo alfabético de assuntos.

Na primeira etapa do trabalho, o pessoal da Documentation Incorporation, firma fundada por Taube, indexou 15 000 documentos pelo sistema Uniterm, enquanto o pessoal da ASTIA (Armed Services Technical Information Agency, hoje Defense Documentation Center) fazia a mesma tarefa utilizando cabeçalhos de assunto, segundo uma lista compilada na própria organização.

Na segunda etapa, cada um dos grupos pesquisou no seu índice 93 perguntas, que haviam sido feitas ao serviço de documentação da ASTIA, e analisou os resultados, pretendendo determinar a relevância dos documentos recuperados.

A etapa final seria a comparação entre os resultados obtidos pelos dois sistemas de indexação, porém os dois grupos de indexadores não chegaram a acordo sobre as respostas corretas às perguntas, divergindo quanto ao significado de perguntas ou a relevância de certos documentos recuperados. Nunca foi publicado o resultado do teste, a não ser um breve informe apresentado por D.C.Gull<sup>1</sup>.

Uns consideraram o sistema Uniterm como aquele que maior sucesso obteve no experimento, tendo o grupo da ASTIA, a mesma impressão quanto ao desempenho dos cabeçalhos de assunto.

Desta experiência ficou a lição da necessidade de prever um método de atribuir a relevância dos documentos em relação às perguntas, que não deixe dúvidas na hora da avaliação dos resultados.

## 0.2. - Projeto Cranfield I

Graças a uma subvenção da National Science Foundation, a ASLIB programou, no início de 1957, um teste sobre a eficiência de sistemas de indexação, conhecido como Projeto Cranfield I ou Projeto ASLIB-Cranfield, executado sob a direção de Cyril W. Cleverdon, bibliotecário do Royal College of Aeronautics, em Cranfield, Inglaterra.

Esta experiência investigou o desempenho de quatro linguagens de indexação e para isso compilou: (1) um catálogo sistemático, tendo por base a Classificação Decimal Universal, com um índice alfabético comum; (2) um catálogo sistemático, segundo u

ma classificação facetada de aeronáutica, da autoria de B.C. Vickery e J. Farradane, acompanhado de um índice em cadeia; (3) um catálogo alfabético de assuntos, segundo uma lista de cabeçalhos de assuntos de aeronáutica; e (4) um índice coordenado Uniterm, seguindo uma lista de termos autorizados compilada durante a indexação.

Para o teste foram indexadas três coleções de 6 000 documentos sobre aeronáutica, incluindo relatórios e artigos de periódicos.

O trabalho de indexação, que durou dois anos, foi feito por três indexadores sendo que um deles possuía experiência em indexação e conhecimento de aeronáutica; outro tinha experiência em indexação, mas não possuía especial conhecimento de aeronáutica; e o terceiro, vindo de uma biblioteca pública, não tinha experiência em indexação ou conhecimento de aeronáutica.

Cada um deles indexou, em épocas diferentes, cada uma das coleções de 6 000 documentos, para ser verificado se a experiência teria influência na qualidade da indexação.

Os resultados obtidos na recuperação em cada uma das coleções mostraram uma superioridade de 4% na última coleção indexada, parecendo indicar uma melhor indexação.

Não houve diferenças marcantes entre os trabalhos dos três indexadores, mas as falhas humanas, na indexação e na pesquisa, foram o principal fator dos fracassos nas pesquisas.

Para cada uma das linguagens, a indexação foi feita 5 vezes por cada indexador, controlando-se o tempo dispendido, sendo este limitado a 2, 4, 8, 12 e 16 minutos por item, variando o número de entradas por item com o tempo reservado à tarefa.

Os resultados da recuperação não justificaram o dispêndio de mais de 4 minutos, pois o tempo adicional não melhorou proporcionalmente o desempenho.

Para os testes de recuperação reuniram-se 1 200 perguntas, enviadas por pessoas ligadas a diferentes organizações, e cerca de 400 pesquisas foram feitas em cada um dos índices, pelos diferentes pesquisadores. Participaram dos testes de recupere



ração 4 bibliotecários

e 5 alunos graduados do College of Aeronautics.

Não houve diferenças significativas entre os resultados obtidos pelos vários pesquisadores.

Os resultados finais do teste<sup>2</sup> indicaram um desempenho levemente superior para o sistema Uniterm (82%), seguido pelos cabeçalhos de assunto (81,5%), pela Classificação Decimal Universal (75,6%) e pela classificação facetada (73,8%).

Um subexperimento revelou que o principal problema da classificação em facetas foi a ordem de citação fixa, associada a um índice em cadeia, pois a utilização de entradas múltiplas, acompanhadas por um índice alfabético simplificado, melhorou o desempenho dessa linguagem, elevando-o para 83%, passando a ser levemente superior ao do índice Uniterm.

A análise das causas das falhas na recuperação mostrou que 17% foram devidas a defeitos das perguntas, 60% a falhas de indexação, 17% a erros na pesquisa e só 6% a deficiências do sistema de indexação.

A CDU demonstrou ser o sistema que levou a maior número de erros na indexação. Os cabeçalhos de assunto e a classificação facetada foram responsáveis por maior número de erros na pesquisa.

As perguntas foram feitas tendo por base determinado documento, assegurando assim o julgamento da relevância ou precisão do documento recuperado, pois só foram consideradas como bem sucedidas as pesquisas que revelaram o documento-fonte.

O Projeto Cranfield I incluiu também vários programas suplementares, utilizando pequenas coleções de documentos, mais dois dos seus experimentos merecem especial menção: os testes da English Electric Co. e o teste da W.R.U. (Western Reserve University).

A biblioteca da English Electric Co. em Whetstone, onde foi realizado este novo teste, utiliza uma classificação facetada e possuía, na época, 1961, 36 000 documentos catalogados e um catálogo com 76 000 fichas, sendo 48 000 na parte classificada e 28 000 no índice em cadeia.

O pessoal da biblioteca selecionou 186 perguntas que, pesquisadas no catálogo pelo pessoal de Cranfield, atingiram um sucesso de 77,4%, numa média de 5 2/3 minutos por pesquisa, percentagem esta bem semelhante à obtida no cômputo geral do projeto Cranfield I, que foi de 73,8% para a classificação facetada.

A percentagem de falhas relativas à indexação foi de 43%, aquelas devidas à pesquisa de 32% e as atribuídas ao sistema de 13%, causadas estas pelo índice em cadeia.

As percentagens acima citadas são bem semelhantes às encontradas no teste Cranfield I próprio, incluindo os erros ocasionados pelos defeitos das perguntas, no teste de Whetstone, que foram os mais elevados, atingindo 12%, enquanto haviam sido só 4% no cômputo geral do Cranfield I.

Ainda em 1961 o pessoal do Projeto Cranfield executou novo teste, em colaboração com a Western Reserve University (W.R.U.), tendo por base uma coleção de 1 300 documentos, sobre metalurgia.

Esta Universidade remeteu a Cranfield uma coleção de 1 300 documentos e uma fita contendo a indexação da coleção, feita segundo uma técnica especial, que consiste em decompor cada conceito num conjunto de conceitos fundamentais, chamados fatores semânticos, codificados e combinados de modo a representar detalhadamente os assuntos, para processamento por computador.

Em Cranfield, os mesmos documentos foram indexados de modo bastante completo por J. Aitchison, utilizando a English Electric Facet Classification.

Doze engenheiros metalúrgicos, de diferentes entidades, prepararam 137 perguntas, derivadas de documentos-fonte, para serem pesquisadas nos dois índices, ficando o sucesso da recuperação na dependência do aparecimento do documento-fonte.

Neste teste a classificação facetada atingiu 90% de eficiência, contra 79% da indexação feita na W.R.U.

Algumas das conclusões a que chegou Cleverdon<sup>3</sup>, são:

- 1) os fatores mais importantes a avaliar, num sistema de recuperação da informação, são revocação e precisão;
- 2) a forma física do catálogo não influi na eficiência

do sistema, quanto à revocação e à precisão;

3) a linguagem de indexação pouco influi no desempenho de um sistema de recuperação, cabendo a principal participação ao trabalho intelectual de reconhecimento dos conceitos;

4) partindo da mesma identificação de conceitos, do mesmo perfil do documento, dois ou mais tipos de linguagens se são potencialmente capazes de desempenho semelhante, quanto a revocação e a precisão;

5) a revocação depende da exaustividade e a precisão da especificidade;

6) há uma relação inversa entre revocação e precisão;

7) há um nível ótimo para a exaustividade, acima do qual passa a prejudicar a precisão;

8) mais de 4 minutos para a indexação não melhora o seu desempenho;

9) indexadores treinados são capazes de boa indexação, mesmo sem conhecer o assunto;

10) há indicação de que os sistemas de indexação operam normalmente entre 70 e 90% de revocação e 8 a 20% de precisão;

11) as 4 linguagens de indexação estudadas operam mais ou menos no mesmo nível de recuperação.

Os resultados dos testes abalaram muitas das convicções correntes na época, mostrando que:

1) não é fato que a indexação pós-coordenada seja muito superior a pré-coordenada;

2) o índice em cadeia não funciona tão bem como o esperado;

3) não é necessário grande número de entradas para garantir a recuperação.

Como resultado do Projeto Cranfield I, Cleverdon considerou que nenhum dos sistemas investigados demonstrou decisiva superioridade, que justifique o seu emprego em todas as condições, considerando que o tamanho da coleção, o número de usuários, o assunto, são elementos que devem influir na escolha de determinado sistema<sup>2</sup>.

O teste da W.R.U. mostrou, segundo Cleverdon, Mills e Keen<sup>4</sup>, que é possível utilizar coleções de documentos relativa -

mente pequenas para experiências.

0.3 - Projeto Cranfield II

Ainda sob o patrocínio da National Science Foundation, a Aslib projetou um novo teste a ser executado em Cranfield, para estudar linguagens de indexação, no qual participaram Cyril Cleverdon<sup>4</sup>, Jack Mills, Michael Keen e Wilfrid Lancaster.

Este novo experimento destinou-se a examinar o papel dos vários artifícios das linguagens de indexação na recuperação

Partiram da premissa de que todas as linguagens de indexação são um amálgama de artifícios destinados a promover a revocação, isto é, tendentes a elevar a percentagem de documentos a recuperar, alargando o âmbito da classe, e artifícios destinados a melhorar a precisão, isto é, tendentes a assegurar a exatidão dos documentos recuperados em resposta à determinada necessidade bibliográfica, restringindo à extensão da classe.

São considerados artifícios destinados a aumentar a precisão, coordenação, ponderação (weighting), elos (links) e papéis (roles).

Destinam-se a aumentar a recuperação: (1) a reunião dos vários sinônimos e quase-sinônimos; (2) a reunião das várias palavras com um mesmo radical; (3) os relacionamentos hierárquicos do tipo gênero/espécie; (4) os relacionamentos derivados de análises em facetas; (5) o acoplamento bibliográfico; (6) o vocabulário reduzido, tal como "L'Unité" de Te Nuyl.

Para este teste foram selecionados 271 documentos sobre aerodinâmica e estruturas de aeronaves, em língua inglesa, publicados entre 1962 e 1963, que trouxessem pelo menos duas citações bibliográficas, em inglês, posteriores a 1954.

Enviaram questionários a cada um dos autores dos 271 documentos solicitando que:

- 1) indicassem, em forma de pergunta, qual o problema básico, que dera origem à pesquisa relatada no documento e mais três problemas, surgidos durante aquela pesquisa, que ocasionaram ou poderiam ter ocasionado consultas a serviço de documentação;

2) escalonassem de 1 a 5 a relevância dos documentos citados no seu trabalho, em relação ao problema envolvido.

Rasponderam aos questionários 182 autores, apresentando 641 questões, entre as quais foram escolhidas 361 para os estudos.

Foi reunida uma coleção de 1 400 documentos, formada por 173 dos documentos originais, 1 018 documentos provenientes do exame das citações e mais alguns documentos.

Um grupo de estudantes de pós-graduação do College of Aeronautics examinou cada documento da coleção, comparando-o com uma das perguntas, para verificar se, além dos documentos indicados pelos autores dos trabalhos originais, existiriam outros escritos relevantes em relação ao assunto das perguntas.

Procedeu-se também a um estudo das citações bibliográficas, pela técnica de acoplamento, e foram selecionados aqueles trabalhos citados 7 ou mais vezes. Nesta pesquisa foram localizados 213 documentos, dos quais só 15 já constavam da coleção.

Os documentos selecionados pelos estudantes e obtidos pelo acoplamento bibliográfico foram enviados aos autores das perguntas, para que os escalonassem, quanto à relevância, conforme a escala anteriormente utilizada.

Na mesma oportunidade lhes foi solicitado que reexaminassem as perguntas que haviam proposto, indicando a importância relativa de cada termo ou conceito empregado, e que informassem sobre possíveis sinônimos. Foi-lhes facultado reformular a pergunta, se assim o desejassem.

Responderam a esta segunda consulta 144 autores e obtiveram, assim, 279 perguntas, para as quais os documentos da coleção teste estavam devidamente escalonados, quanto a relevância, mas só 221 dessas perguntas foram utilizadas nas várias fases do projeto.

O documento-fonte de cada pergunta foi eliminado, posteriormente, dos resultados da recuperação, a fim de atender às críticas relativas a utilização de documentos-fonte, devidas especialmente a Sawson<sup>7,8</sup>, mas os responsáveis pelo teste Cranfield II afirmaram a sua convicção de que podem ser utilizados satisf

toriamente perguntas derivadas de documentos-fonte em testes<sup>4</sup>.

Considerando que os artifícios destinados à recuperação só podem ser medidos se a indexação tiver sido exaustiva, procuraram indexar da mesma forma a coleção de documentos, mas, para ser possível avaliar o efeito da exaustividade nos diversos artifícios, foi estabelecida uma escala de três valores para cada descritor designado para cada documento.

Foram estudados três tipos de linguagens documentárias:

1) termos únicos, em linguagem natural, retirados dos próprios documentos, observados nas 5 formas abaixo:

- a) termos na linguagem natural;
- b) termos na linguagem natural mais os sinônimos;
- c) termos na linguagem natural mais as palavras do mesmo radical;
- d) termos na linguagem natural mais sinônimos e palavra do mesmo radical;
- e) termos na linguagem natural mais sinônimos, palavras do mesmo radical e quase-sinônimos.

2) linguagens a base de conceitos únicos, estudadas em 15 agrupamentos diferentes;

3) linguagens à base de termos controlados (6 linguagens)

Além disso foram testadas 4 linguagens em que a indexação incluía todas as palavras chaves dos títulos e dos resumos.

Este teste da Cranfield mostrou que o emprego de palavras isoladas, provenientes da linguagem natural, porém com controle de sinônimos e reunião das palavras com o mesmo radical, é perfeitamente aceitável.

Uma coleção de 200 documentos e 42 perguntas da coleção Cranfield II vem sendo utilizada em outros testes, como no SMART e na Cambridge Language Research Unit.

#### 0.4 - Teste de Aberystwyth

Entre os testes sobre linguagens de indexação, merece atenção especial o teste levado a efeito no College of Librarianship Wales, em Aberystwyth, Gales, citado às vezes como ISILT, siglas de "Information Science Index Language Test", sob

a responsabilidade de Michael Keen e J.A. Digger<sup>9,10</sup>.

As linguagens testadas foram as seguintes:

1) "Compressed term index language", isto é, um vocabulário limitado a 300 termos, preparado inicialmente por Alan Gilchrist e Kathleen Gaster, para a Aslib, acrescido de termos correlatos, utilizado pós-coordenadamente;

2) "Uncontrolled index language", isto é, linguagem natural, utilizando 1 200 palavras encontradas nos próprios docu - mentos, sem qualquer controle, mesmo de sinônimos ou palavras com o mesmo radical, utilizada pós-coordenadamente;

3) "Hierarchically structured index language, post-coordinate", isto é, a Classification of Library Science, da autoria do Classification Research Group, edição de 1965, utiliza-da pós-coordenadamente, sem emprego de síntese, sem ordem de citação fixa, mas conservando a notação;

4) "Hierarchically structured index language, pre-coordinate", isto é, a Classification of Library Science, empre-gada pré-coordenadamente;

5) "Relational indexing index language", isto é, a indexação relacional de J.Farradane, utilizada pré-coordenadamente.

A coleção-teste foi constituída de 800 documentos sobre Biblioteconomia e Ciência da Informação, sendo 408 documen - tos da coleção do projeto SMART, consistindo de artigos de periódicos e relatórios publicados entre 1961 e 1963, e 392 documentos indicados pelos autores das perguntas, publicados em 1968.

Os documentos selecionados foram escalonados em rela - vantes, não-relevantes e parcialmente relevantes por 20 membros do pessoal do College of Librarianship, examinando toda a coleção em relação a cada uma das perguntas; geralmente utilizando os títulos e os resumos, mas eventualmente lendo o próprio texto.

A indexação foi feita por Keen e Digger e a indexação relacional por S.Datta, da Aslib, pessoas que não participaram do estabelecimento da relevância ou da pesquisa.

A fim de evitar variações nas interpretações, o traba - lho de indexação pelas várias linguagens foi feito tendo por base um perfil do documento, indicando os tópicos tratados, em lin

guagem natural, a ser traduzida para as diferentes linguagens de indexação sob teste, empregando todas as suas potencialidades.

Os indexadores foram instruídos para:

- 1) escolher os tópicos tratados substancialmente no documento e ignorar aqueles só ligeiramente mencionados;
- 2) não levar muito tempo na indexação de cada documento;
- 3) não escolher tantos tópicos para indexar, de modo que o documento recebesse número excessivo de entradas.

Procurou-se fazer a indexação num nível de exaustividade de igual ou semelhante aquele normalmente observado nos serviços de documentação.

A coleção de perguntas constituiu-se de 63 questões, obtidas por questionários enviados com a colaboração da Aslib e da Library Association a 76 bibliotecários e cientistas da informação, mas que representassem reais necessidades de informação.

As pesquisas foram executadas por estudantes dos últimos níveis do College of Librarianship Wales.

As pesquisas foram feitas duas vezes, uma livre e a outra controlada. Esse controle destinou-se a evitar a consulta mais exaustiva ou mais específica em um índice do que em outro, sendo o controle feito pelo estabelecimento de uma linguagem intermediária de pesquisa, traduzida depois para as várias linguagens de indexação sob teste.

Para o teste livre, os pesquisadores foram informados dos números dos documentos relevantes e instruídos a dar como terminada a pesquisa ao serem os mesmos localizados. Foi estabelecido o limite máximo de tempo de 40 minutos para cada pesquisa.

Exigiu-se o registro de todas as etapas da pesquisa e, a fim de controlar esse registro, cada pesquisa foi repetida por funcionários burocráticos, seguindo os passos descritos pelos pesquisadores originais.

Alguns dos resultados foram:

- 1) não houve diferenças significativas entre o vocabulário não-controlado e a linguagem hierárquica específica;



2) as linguagens bem específicas tiveram tão bom desempenho quanto as pouco específicas;

3) a linguagem sem controle não teve pior desempenho do que a pior linguagem controlada, nem tão bom quanto o da melhor, mas as diferenças não foram significativas;

4) a utilização de vocabulário limitado não deu bom resultado na recuperação;

5) a exaustividade não influenciou na precisão;

6) a redundância na indexação não melhorou a revocação;

7) a indexação exaustiva resultou em melhor revocação;

8) o emprego de artifícios de precisão, além da coordenação, não influíram na recuperação;

9) as linguagens mais específicas não revocaram significativamente menos do que as menos específicas;

10) foram virtualmente iguais as falhas de recuperação nas linguagens pré e pós-coordenadas.

Os resultados dos testes sugeriram que "a receita para a melhor linguagem de indexação é uma linguagem tão específica quanto possível, sem emprego de artifícios de precisão mais sofisticados do que a coordenação, e com pouca ou nenhuma coordenação. Parece que as palavras isoladas da linguagem natural inglesa está próxima de atingir o nível ótimo de especificidade".

Os relatores do teste estão convictos de que os resultados dessa experiência aplicam-se ao assunto testado, bem como a outros campos do conhecimento, em que a terminologia não é muito firme, as ditas "soft sciences", e a coleções de vários milhares de itens.

0.5 - Carlisle e Fetter

Uma experiência diferente foi tentada em 1970-71 por J.H. Carlisle<sup>11,12</sup> e R.B. Fetter, destinada a verificar o comportamento do usuário na obtenção de informações computadorizadas "online", utilizando o vídeo e o teletipo.

Foram utilizados um Teletype Model 33 e um vídeo modelo CC-30, a Computer Communications Inc., sendo a transmissão pelo vídeo de 30 caracteres por segundo e a pelo teletipo de 10 caracteres por segundo.

A experiência foi feita no Mead Data Central, consistindo os dados armazenados nos textos completos das súmulas da Ohio Supreme Court, desde 1853.

A pesquisa esteve a cargo de 12 alunos do Yale Law School, trabalhando 6 com o vídeo e 6 com o teletipo.

Cada estudante deveria processar a mesma pergunta e as instruções foram localizar o maior número de processos pertinentes ao tema da causa em estudo.

A relevância dos documentos recuperados em relação às questões propostas foi estabelecida por um advogado, que examinou todos os itens e os graduou segundo a seguinte escala:

10 pontos	Exatamente o mesmo tipo de demanda
4 pontos	Referente ao assunto
1 ponto	Na área certa, mas não relevante
0 ponto	Sem valor. Falsa recuperação

Comparou-se o comportamento dos dois grupos de pesquisadores verificando-se:

- a) o número de documentos recuperados e examinados;
- b) a relevância dos documentos selecionados;
- c) o tempo gasto na pesquisa;
- d) os erros cometidos;
- e) as estratégias individuais de pesquisa;
- f) a satisfação pessoal dos pesquisadores com o terminal utilizado.

Constatou-se que a pesquisa feita pelo vídeo levou 60% mais tempo, mostrando que o tempo de transmissão dos caracteres não influenciou no tempo da pesquisa e demonstrando maior tolerância do usuário, apresentou erros, em termos percentuais, na ordem de 300%, porém atingiu 40% mais relevância e recuperou 42% mais processos, verificando-se que os pesquisadores que utilizaram o vídeo mostraram-se mais dispostos a fazer pesquisas em profundidade.

Atribuiu-se o maior volume de erros nos utilizadores do vídeo a menor familiaridade com o teclado deste equipamento.

A análise das pesquisas revelou:

- 1) uma relação entre o número de erros e o tempo gasto

na pesquisa: quanto maior o tempo dispendido, maior o número de erros;

2) uma relação entre o grau de relevância e o número de erros: quanto maior o número de erros, maior também o número de documentos relevantes recuperados;

3) uma relação entre o grau de relevância e o número de processos recuperados: quanto maior o número de documentos recuperados maior o número de documentos relevantes localizados.

A opinião pessoal dos pesquisadores favoreceu antes o vídeo do que o teletipo - 281 pontos para o vídeo contra 263 para o teletipo.

### 0.6 - Medidas de Avaliação

Os principais critérios para julgamento de um sistema de recuperação da informação foram apresentados por Cleverdon, citado por Lancaster<sup>13</sup>, como sendo os seguintes:

1. Cobertura do assunto
2. Revocação
3. Precisão
4. Tempo de resposta
5. Esforço do usuário
6. Forma da resposta (output)

A maioria das avaliações de desempenho de sistemas segue a orientação de Cleverdon e procura determinar os coeficientes de revocação (recall ratio) e de precisão (precision ratio ou relevance ratio). O primeiro aferindo a capacidade do sistema em recuperar todas as referências relevantes e o segundo medindo a habilidade do sistema em só oferecer itens relevantes.

O coeficiente de revocação de um sistema é obtido pela equação:

$$\frac{\text{Nº de documentos relevantes recuperados}}{\text{Nº total de documentos relevantes existentes na coleção}} \times 100$$

O coeficiente de precisão é obtido pela equação:

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

Vários autores, além de Cleverdon, Keen e Mills, preo-

cuparam-se em encontrar meios de medir a eficiência e a eficácia das linguagens de indexação e dos sistemas de recuperação, como Goffman e Newill<sup>14</sup>, Swets<sup>15,16</sup>, Fels<sup>17</sup>, Miller<sup>18</sup>, Robertson<sup>19,20</sup>, Lancaster<sup>13</sup> e Swanson<sup>7,8</sup>.

Goffman e Newill consideram que a avaliação de um sistema deve ser medida em função de duas variáveis: eficácia (effectiveness) e eficiência (efficiency). A primeira representando a habilidade em desempenhar a tarefa e a segunda o custo desse desempenho.

Esses autores consideram que a eficácia é demonstrada pela sensibilidade (sensitivity) e pela especificidade (specificity). Esta medindo a capacidade do sistema de não recuperar documentos que não sejam relevantes e aquela a sua habilidade em revelar os documentos relevantes.

A sensibilidade (Se) e a especificidade (Es) podem ser obtidas pelas equações abaixo:

$$Se = \frac{\text{Nº de documentos relevantes recuperados}}{\text{Nº de documentos relevantes existentes na coleção}}$$

$$Es = \frac{\text{Nº de documentos não relevantes não recuperados}}{\text{Nº de documentos não relevantes existentes na coleção}}$$

Pelas equações acima conclui-se que a sensibilidade de Goffman e Newill é a mesma revocação de Cleverdon, mas a medida de especificidade é uma nova medida proposta.

A eficácia (Ef) será a soma da sensibilidade e da eficiência menos um.

$$Ef = Se + Es - 1$$

Essas medidas foram utilizadas por Goffman e Newill no Comparative Systems Laboratory, do Center for Documentation and Communication Research, School of Library Science, Case Western Reserve University.

Lancaster sugeriu o "novelty ratio" obtido pela relação entre os documentos relevantes e novos para o usuário revelados pelo sistema e a totalidade dos documentos recuperados, ou seja

$$\frac{\text{Nº de documentos novos, relevantes e recuperados}}{\text{Nº de documentos relevantes recuperados}}$$

Miller apresenta uma medida de avaliação semelhante ao "novelty ratio" de Lancaster, denominando-a "extension ratio" , isto é, a relação entre o número de documentos relevantes que o usuário fica conhecendo após a pesquisa em relação ao número de documentos que conhecia antes.

Já Robertson acha que a medida da eficácia seja, tanto quanto possível, independente da generalidade (generality), descrita em termos de revocação e "fallout" e capaz de medir a curva de desempenho, isto é, seja definível por uma série de pontos, bem como por um único ponto.

Por "fallout" Robertson entende a proporção de documentos irrelevantes recuperados e por "generality" a proporção de documentos existentes na coleção que são relevantes.

Swanson propôs uma medida de avaliação representada pela equação:

$$M = R - pI$$

onde R é a soma dos pesos de relevância dos documentos recuperados dividida pela soma dos pesos de relevância de todos os documentos existentes na coleção; I é o número de documentos irrelevantes e p um valor arbitrário.

O valor de I é obtido subtraindo-se do total de documentos recuperados (N) o total de documentos relevantes da coleção (L), multiplicado por R, isto é:

$$I = N - LR$$

Borko modificou a medida proposta por Swanson, passando I a ser o número de documentos não-relevantes recuperados dividido pelo número de itens recuperados, mas não sendo empregado o valor arbitrário p.

Carlisle e Fetter<sup>11</sup> e Pollock<sup>21</sup> procuram encontrar um meio de comparar o desempenho de diversos sistemas de indexação, considerando as diferentes variáveis estudadas.

Carlisle e Fetter analisaram seis variáveis - tempo , erros, relevância, diálogo homem-computador, recuperação e satisfação do usuário - utilizando a análise discriminatória múltipla.

Fairthorne<sup>22</sup> salienta que, antes de pretender testar ou

avaliar um sistema de recuperação da informação é necessário de terminar os seus objetivos.

A razão de ser, o objetivo primeiro de um sistema de armazenagem e recuperação da informação, é fornecer ao usuário documentos que respondem as suas necessidades de informação, no entanto, é necessário que esta tarefa seja executada eficiente - mente.

Lancaster<sup>23</sup> encontra cinco etapas no desenvolvimento de um programa de avaliação de sistemas de recuperação da informação:

- 1) decidir o que avaliar;
- 2) desenhar um teste que possibilite responder as questões propostas;
- 3) levar a efeito o teste programado;
- 4) analisar os resultados do teste;
- 5) interpretar os resultados.

## 1 - OBJETIVO

O presente trabalho procura estudar, em condições de laboratório, o desempenho na solução de pesquisas, das seguintes linguagens de indexação:

- 1) o sistema KWIC, representando a linguagem natural;
- 2) o sistema Uniterm, apresentando uma linguagem à base de conceitos simples, com controle de sinônimos e reunião das palavras com o mesmo radical;
- 3) quatro linguagens estruturadas, à base de conceitos compostos:
  - a) cabeçalhos de assunto, representando a indexação alfabética;
  - b) Classificação Decimal de Dewey (CDD), um sistema de classificação geral, considerado pouco detalhado;
  - c) Classificação Decimal Universal (CDU), um sistema de classificação geral, detalhado, permitindo a combinação de símbolos para obtenção de especificidade;
  - d) Classification of Library and Information Sciences (1971), da autoria do Classification Research Group, utilizada

nos Library & Information Science Abstracts (LISA), citada neste trabalho como CRG, representando a classificação especializada e facetada.

Justifica a escolha destas linguagens o fato de serem o KWIC, o Uniterm, os Cabeçalhos de Assunto, a CDD e a CDU largamente empregados no Brasil e a classificação facetada ser objeto de especial atenção do nosso meio biblioteconômico, já havendo mesmo um sistema de classificação deste tipo em uso, a classificação de Odontologia, de autoria de Malvina Rosa.

Pretende-se examinar o grau de sucesso na recuperação, a precisão dos índices, a precisão dos pesquisadores, quanto à seleção dos documentos e aos acertos nas respostas e o tempo dispendido nas pesquisas.

Deseja-se, também, observar a aceitação pelos pesquisadores, representando os usuários, dos vários métodos de armazenagem e recuperação da informação.

## 2 - MATERIAL

O material básico constituiu-se de 100 documentos sobre Biblioteconomia e Ciência da Informação, incluindo 18 teses de mestrado, apresentadas ao IBBB, e 82 artigos publicados nas revistas Ciência da Informação, Revista de Biblioteconomia de Brasília e Revista da Escola de Biblioteconomia da Universidade Federal de Minas Gerais, nos anos de 1972-1974 (Anexo 1).

Foram excluídos os artigos em linguas estrangeiras, a queles considerados como pertencentes a domínios conexos e os de caráter puramente biográfico.

O relatório Cleverdon-Mills-Keen, sobre o Projeto Granfield II, justifica o assunto escolhido lembrando que "num teste experimental, a prática comum, natural, é escolher trabalhar com uma coleção de documentos com a qual o grupo tenha alguma familiaridade"<sup>4</sup>.

A maioria dos testes relatados, diz Miller<sup>18</sup>, envolvem índices de 200 a 1000 referências. Este teste, uma experiência piloto, limitou-se a 100 documentos.

### 3 - MÉTODO

#### 3.1 - Armazenagem

Para possibilitar a experiência foram compilados:

- 1) um índice KWIC e respectiva lista de citações bibliográficas;
- 2) um índice Uniterm, acompanhado de uma lista dos documentos indexados;
- 3) um catálogo alfabético de assuntos;
- 4) um catálogo sistemático, pela Classificação Decimal de Dewey (CDD), acompanhado de um índice em cadeia;
- 5) um catálogo sistemático, pela Classificação Decimal Universal (CDU), acompanhado por um índice em cadeia;
- 6) um catálogo sistemático, pela Classification of Library and Information Sciences (CRG), acompanhado por um índice em cadeia.

A indexação para os catálogos 3-6 foi feita à base do mesmo perfil do documento, isto é, a mesma análise dos conceitos presentes no documento.

Procurou-se representar o mais completamente possível o perfil de cada documento, utilizando-se todas as possibilidades oferecidas pela linguagem empregada, porém os indicativos de forma de apresentação só foram incluídos quando pertencente à representação do assunto.

As entradas destinadas aos vários índices e catálogos foram feitas em forma de referência bibliográfica, reproduzidas em xerox e montadas em fichas.

##### 3.1.1 - Indexação KWIC

Preparou-se um índice KWIC utilizando o programa do IBBD, com 120 caracteres por linha.

Verificou-se que, devido ao limite de caracteres por linha, os títulos de 9 documentos (Doc. n. 4, 11, 29, 44, 49, 57, 75, 79 e 91) não puderam ser integralmente incluídos, mas foram feitas supressões e abreviaturas de partes dos títulos, a fim de tornar possível entradas por todas as palavras significativas.



Constatou-se que foram feitas 466 entradas neste índice, assim distribuídas:

2 documentos exigiram	1 entrada
17 documentos exigiram	2 entradas
18 documentos exigiram	3 entradas
15 documentos exigiram	4 entradas
18 documentos exigiram	5 entradas
11 documentos exigiram	6 entradas
6 documentos exigiram	7 entradas
7 documentos exigiram	8 entradas
1 documento exigiu	9 entradas
2 documentos exigiram	10 entradas
3 documentos exigiram	11 entradas

Além das entradas acima, o índice KWIC exige uma relação das referências bibliográficas: 100 neste caso. Pode-se, portanto, dizer que foram feitas 100 entradas completas e 466 entradas de índice, no total de 566.

O índice KWIC representa, neste trabalho, a indexação em linguagem natural.

### 3.1.2 - Indexação Coordenada Uniterm

A fim de organizar o índice Uniterm foi necessário compilar um vocabulário. Seguiu-se nesta tarefa a experiência de John A. Sanford e Frederick R. Theriault<sup>24</sup>, decidiu-se "deixar os próprios documentos gerarem seus unitermos".

Procurou-se seguir os preceitos fundamentais da indexação Uniterm, expressos pela seguinte regra: "Num índice coordenado Uniterm, faça de cada palavra uma palavra de ordem, numa ficha Uniterm independente. Sempre que num determinado índice Uniterm, uma palavra é empregada numa única locução descritiva, faça dela uma palavra de ordem, numa ficha independente, seguida pela palavra ou palavras restantes da frase. A palavra ou as palavras que seguem a palavra de ordem em qualquer ficha serão, e las próprias, palavras de ordem em outras fichas"<sup>25</sup>.

"As palavras que funcionam como única palavra descritora ou como elemento de várias locuções são chamadas livres e as

que não funcionam sós ou não aparecem senão em uma única locução são ditas não livres. Todo termo livre deve ser um unitermo"<sup>25</sup>.

Assim tornaram-se unitermos independentes BIBLIOTECONOMIA, DOCUMENTAÇÃO, bem como CIÊNCIA, INFORMAÇÃO, SERVIÇO e REFERÊNCIA, encontrados em CIÊNCIA DA INFORMAÇÃO, SERVIÇO DE INFORMAÇÃO e SERVIÇO DE REFERÊNCIA.

Reuniram-se numa única ficha as palavras com o mesmo radical, como ÍNDICES/INDEXAÇÃO, SISTEMAS/SISTEMÁTICOS/SISTÊMICOS.

Controlaram-se os sinônimos, mas não foram feitas fichas de referência ligando os assuntos gerais aos específicos ou relacionando assuntos correlatos.

Os 100 documentos indexados geraram 262 unitermos e 54 remissivas de sinônimos.

Foram feitos 316 assentamentos de números de documentos nas fichas Uniterm, numa média de 3,16 entradas por documento.

Essa indexação exigiu, também, uma relação numérica de todos os documentos indexados - 100 entradas completas.

Somando-se as entradas completas, as fichas uniterm e as remissivas temos 416 entradas para o índice Uniterm.

A indexação Uniterm representa, neste trabalho, uma linguagem pós-coordenada, à base de conceitos simples, com pequeno controle de vocabulário, porém permitindo entradas específicas.

### 3.1.3 - Cabeçalhos de Assunto

Na falta de uma lista de cabeçalhos de assunto em português, foi escolhida a Lista de Encabezamientos de Materia para Bibliotecas (incluindo os suplementos 1 e 2), da autoria de Carmen Rovira e Jorge Aguayo, publicada pela União Pan Americana.

Justifica a escolha o fato de a língua espanhola ser mais semelhante à portuguesa do que a inglesa e o fato de a lista espanhola ser quase uma tradução dos Subject Headings Used in the Dictionary Catalogs da Library of Congress.

Partindo da premissa de que essa lista é o maior repo-

sitório geral de cabeçalhos de assunto com que contam os bibliotecários brasileiros, do mesmo modo como não foram criadas subdivisões aos sistemas de classificação, procurou-se seguir a lista escolhida, traduzindo-a, mas evitando ampliá-la e modificá-la.

Mesmo assim, considerou-se indispensável incluir 8 cabeçalhos novos como remissivas, para melhorar a recuperação, tendo em vista o fato de a lista não estar atualizada com a nova terminologia do assunto.

Acreditou-se perfeitamente válido este procedimento, porque os índices compilados para as classificações também incluíram a nova terminologia, que não constava das tabelas.

Procurou-se obter cabeçalhos específicos e assim foram frequentemente utilizadas as listas de subdivisões comuns.

Nesta parte do trabalho foram feitas 166 entradas completas, 186 remissivas "ver" e "ver também", num total de 352 entradas para 100 documentos, numa média de 3,5 por documento.

Foi feita uma média de 1,08 entradas completas por título, assim distribuídas:

- 3 documentos exigiram 4 entradas
- 15 documentos exigiram 3 entradas
- 27 documentos exigiram 2 entradas
- 55 documentos exigiram 1 entrada

Os Cabeçalhos de Assunto representam, neste trabalho, uma linguagem controlada, pouco específica, utilizando conceitos compostos, expressos em palavras, empregada pré-coordenadamente.

3.1.4 - Classificação Decimal de Dewey (CDD)

Foi utilizada a 18ª edição de 1971.

Procurou-se obter símbolos de classificação tão específicos quanto possível, aplicando-se as tabelas de subdivisões comuns, especialmente a tabela de subdivisões standard e a tabela de áreas.

A CDD não é suficientemente enumerativa na classe 020 e reflete a insegurança da terminologia moderna deste campo do conhecimento, como, por exemplo, reservando 029.9 para centros de documentação especializados e deixando a biblioteca especiali

zada em 026.

Muitos temas atuais ainda não encontraram guarida na CDD - como ocorre com Bibliometria, Disseminação Seletiva da Informação e Classificação Facetada - obrigando a composição de símbolos com o auxílio de subdivisões standard ou a classificação em símbolos mais gerais.

Assim surgiram os símbolos seguintes:

025.5            para Disseminação Seletiva da Informação  
 010.182        para Bibliometria  
 010.723        para levantamentos bibliométricos  
 016.550723    para levantamentos bibliométricos no campo da literatura de geologia

Nem sempre foi possível obter símbolos suficientemente específicos, recorrendo-se então a duas ou mais entradas. Ex.:

025.60285      Empréstimo mecanizado  
 027.78150285 Mecanização do empréstimo em bibliotecas universitárias de Minas Gerais.

Acompanhou o catálogo pela CDD um índice em cadeia, tão específico quanto possível, recorrendo-se à extensão verbal, como faz a British National Bibliography, porém sem acréscimos aos símbolos de classificação.

Foram feitas 219 entradas completas, numa média de 2,19 entradas por documento, assim distribuídas:

1 documento exigiu	5 entradas
15 documentos exigiram	4 entradas
16 documentos exigiram	3 entradas
38 documentos exigiram	2 entradas
30 documentos exigiram	1 entrada

Foram feitas fichas para o índice alfabético, totalizando 516 fichas para a indexação pela CDD.

Considerou-se que a indexação pela CDD representa, neste trabalho, uma indexação pré-coordenada pouco específica.

### 3.1.5 - Classificação Decimal Universal (CDU)

Foi utilizada a edição desenvolvida em língua portuguesa

sa das Classes 0 e 8 e, para as demais classes, a edição abreviada em língua portuguesa.

Tendo em vista que a Classe 0 em português foi publicada em 1961, comparou-se as suas subdivisões com as apresentadas na edição média, divulgada pela IBBD/CDU em 1966, verificando-se não haver modificações.

Intencionalmente não se recorreu a edições em outras línguas ou às Extensions and Corrections to UDC, porque são publicações inacessíveis à maioria dos bibliotecários brasileiros.

A CDU não se mostrou suficientemente enumerativa, quanto a conceitos simples. Não há símbolos próprios para muitos dos temas atuais, tais como Automação em Bibliotecas, Bibliometria, Citações Bibliográficas, Classificação Facetada, Índices de Citações, Índice em Cadeia, Linguagem Documentária.

Recorreu-se quase sempre ao relacionamento ou a síntese para obter classificações específicas, assim foram formados:

- 31:311.17 para Bibliometria
- 01.001.5 para Pesquisa Bibliográfica

Foi feita uma entrada por cada um dos símbolos relacionados para apresentação dos assuntos dos documentos, mas procurou-se manter juntos àqueles relacionamentos que expressavam um conceito, não sendo feita uma pura rotação, muito menos uma permutação. Ex.:

- 016:55(81):311.17            025.45CDU:025.347:65.011.5
- 55(81):016:311.17            025.347:025.45CDU:65.011.5
- 311.17:016:55(81)            65.011.5:025.347:025.45CDU

Procurou-se compilar um índice em cadeia bem completo, recorrendo à extensão verbal, mas sem acréscimos aos símbolos de classificação.

A indexação pela CDU, considerada como exemplo de entradas múltiplas, exigiu 289 entradas completas, numa média de 2,89 fichas completas por título, assim distribuídas:

1 documento	exigiu	12 entradas
1 documento	exigiu	8 entradas
1 documento	exigiu	7 entradas
10 documentos	exigiram	5 entradas
18 documentos	exigiram	4 entradas
16 documentos	exigiram	3 entradas
39 documentos	exigiram	2 entradas
14 documentos	exigiram	1 entrada

Foram feitas 666 fichas para o índice alfabético, totalizando 955 fichas para a indexação pela CDU, numa média de 9,55 fichas por documento.

Considerou-se que a indexação pela CDU representa, neste trabalho, uma indexação pré-coordenada específica, por entradas múltiplas.

### 3.1.6 - Classification of Library and Information Sciences (CRG)

Classificar pela Classification of Library and Information Sciences, da autoria do Classification Research Group, ora citada como CRG, é um trabalho que satisfaz um bibliotecário. O sistema está atualizado com os novos temas de Biblioteconomia e Ciência da Informação, procurando resolver certas inseguranças da terminologia, reunindo assuntos similares com terminologias diferentes, como é o caso das listas de cabeçalhos de assunto e dos thesaurus, e permitindo a especificidade necessária, pela combinação dos símbolos das várias facetas.

As ciências afins, no entanto, parecem exigir melhor tratamento, pois o recurso à CDU, permitido pelo sistema, nem sempre satisfaz. Neste caso está o Livro, como veículo de comunicação.

Foi utilizada a ordem de citação inversa àquela apresentada nas tabelas da classificação.

Foram feitas entradas únicas e só 7 documentos receberam duas entradas, totalizando 107 entradas completas.

Completou-se a indexação com um índice em cadeia, para o qual foram feitas 304 entradas, totalizando, portanto, 411 entradas para o catálogo sistemático pela CRG.

A CRC representou, neste trabalho, a classificação especializada, em facetas, uma indexação pré-coordenada e específica, com entradas únicas.

### 3.2 - Perguntas

Solicitou-se a 8 bibliotecários que fizessem perguntas que pudessem ser respondidas satisfatoriamente por algum dos artigos indexados (Anexo 2).

A escolha das perguntas foi feita pela leitura do texto dos documentos e os documentos-fonte constituíram o alvo da recuperação, se bem que tenham sido consideradas com certas respostas que indicaram outro documento.

Entre as perguntas apresentadas, selecionaram-se 10 para serem utilizadas no teste (Anexo 3). Procurou-se escolher aquelas que estivessem claramente expostas, não deixando dúvida quanto à resposta desejada.

Acrescentou-se uma consulta para a qual não havia resposta nos documentos da coleção-teste, a Pergunta nº 7, procurando-se criar uma situação semelhante àquela que enfrentam os usuários, quando se dirigem a um serviço de documentação. Esta pergunta foi considerada como não existente na avaliação.

### 3.3 - Recuperação

O trabalho de pesquisa ficou a cargo de 60 alunas da Faculdade de Biblioteconomia e Documentação da Universidade Santa Úrsula, assim distribuídas: 54 cursando a disciplina Catalogação III (4º período letivo), 3 cursando a disciplina Catalogação I (2º período letivo) e 3 alunas cursando a disciplina Teoria da Classificação (6º período letivo).

As pesquisadoras foram divididas em 6 grupos de 10, seguindo-se a ordem alfabética dos prenomes para a escolha dos 5 primeiros grupos e formando-se o último com 4 alunas de Catalogação III e seis voluntárias, das turmas de Catalogação I e Teoria da Classificação, encarregado este grupo da pesquisa no índice KWIC, o último a ficar pronto.

Conseguiu-se, assim, grupos diferentes para indexação (feita pela autora da dissertação), a formulação das perguntas e

a pesquisa, evitando-se incorrer em um dos defeitos apontados por Swanson - o papel da memória no auxílio à pesquisa.

As pesquisadoras foram informadas de que nem todas as perguntas poderiam ser respondidas pelos documentos indexados e salientou-se que eram os índices que estavam sendo testados e não as pesquisadoras, procurando levá-las a agir normalmente, como o usuário comum.

Cada pesquisadora recebeu uma folha de "Instruções" sobre o preenchimento dos formulários (Anexo 4), uma explicação sobre o modo de consultar o índice (Anexos 5 A-D), onze "Folhas de Pesquisa" (Anexo 6) e uma "Folha de Avaliação do Sistema" (Anexo 7).

Pediu-se que indicassem o tempo de duração de cada pesquisa, as entradas procuradas, os documentos selecionados para exame, o documento que de fato responde à consulta e a opinião pessoal sobre o método de indexação utilizado.

Considerou-se como certa a resposta que indicou o documento-fonte, mas também assim foram consideradas aquelas que indicaram outros documentos que de fato respondem ao perguntado.

A fim de testar a clareza dos formulários, foi feito um pré-teste, com 6 alunas da mesma Faculdade, que não participaram do teste definitivo.

À luz dos resultados do pré-teste, os formulários foram ligeiramente modificados e foi trocada uma das perguntas, por ter sido verificado que não estava precisa e poderia levar a diferentes interpretações quanto à resposta, podendo ser respondida por vários documentos da coleção.

### 3.4 - Avaliação

O desempenho de cada uma das linguagens testadas foi examinado, neste trabalho, em função do usuário, verificando o seu sucesso na pesquisa, as respostas certas, a precisão na seleção dos documentos, a precisão do índice, segundo o número de referências bibliográficas examinadas para chegar à resposta, e o tempo dispendido na pesquisa.

Computaram-se os dados referentes a cada pesquisador,



somaram-se os resultados encontrados em cada linguagem de indexação e essas somas serviram de base aos cálculos abaixo.

O Sucesso na Pesquisa, a percentagem de respostas certas, deu origem ao coeficiente de Sucesso na Pesquisa (CSP), obtido pela equação:

$$CSP = \frac{N^{\circ} \text{ de sucessos na pesquisa}}{N^{\circ} \text{ de pesquisas propostas}}$$

A precisão na solução dos problemas apresentados foi examinada nos três diferentes níveis, obtendo-se os dados abaixo:

1) Coeficiente de Precisão do Índice (CPI), resultado do montante de referências bibliográficas, que o pesquisador teve que examinar para selecionar os documentos que lhe pareciam responder a pergunta, segundo a equação

$$CPI = \frac{N^{\circ} \text{ de respostas certas}}{N^{\circ} \text{ de referências bibliográficas examinadas}}$$

2) Coeficiente de Precisão do Pesquisador (CPP), mostrando a relação existente entre os documentos selecionados para exame e aqueles que de fato respondem ao perguntado, obtido pela equação

$$CPP = \frac{N^{\circ} \text{ de respostas certas}}{N^{\circ} \text{ de documentos selecionados para exame}}$$

3) Coeficiente de Precisão das Respostas (CPR), itens recuperados, confrontando o número de respostas certas com o número de documentos indicados como respostas, conforme a equação:

$$CPR = \frac{N^{\circ} \text{ de respostas certas}}{N^{\circ} \text{ de documentos indicados como resposta}}$$

Computou-se o tempo dispendido na solução de cada uma das perguntas, somaram-se os resultados e dividiu-se pelo número de perguntas pesquisadas (100), obtendo-se o tempo médio de pesquisa em cada índice, donde

$$TM = \frac{N^{\circ} \text{ de minutos dispendidos nas pesquisas}}{N^{\circ} \text{ de pesquisas feitas}}$$

A fim de estudar a aceitação pelo usuário (pesquisadores) dos métodos de indexação utilizados foram examinadas as respostas aos itens 1, 10 e 11 da "Folha-de-Avaliação do Sistema" e computados os resultados do item 10, com as correções devidas às

respostas ao item 11.

#### 4 - RESULTADOS

Apresentamos a seguir os Quadros 1-4 mostrando os dados obtidos pelos 60 pesquisadores, na utilização dos índices compilados pelas 6 diferentes linguagens de indexação.

## RESPOSTAS

PESQUISADORES	Nº DAS PERGUNTAS - Nº DOS DOCUMENTOS RESPOSTA											TOTAL DE DOCUMENTOS RECUPERADOS	TOTAL DE RESPOSTAS CERTAS	TOTAL DE SUCESSOS NA PESQUISA	
	Nº DOS DOCUMENTOS RECUPERADOS														
	1	2	3	4	5	6	8	9	10	11	11				
	92	20 74	100	5 78	44	10 61	56	21	9	10	11	45			
1	92	74	100	5	44	10	56	21 64	35	45	11	9	9		
2	92	20	100	5	44	10 61	56	21	17	45	11	10	9		
3	92	20	100	5	44	61	56	21	17	45	10	9	9		
4	92	20	100	5	44	61	56	21	17	45	10	9	9		
5	84	-	93	5	44	61	56	21	-	45	8	6	6		
6	92	20	-	5	44	61	56	21	-	45	8	8	8		
7	92	20	-	5	44	61	56	21	-	45	8	8	8		
8	92	20	-	5	44	10	56	21	-	45	8	8	8		
9	92	20	100	5 78	44	10 61	56	21	9	45	12	12	10		
10	92	20	100	5	44	61	56	21	-	45	9	9	9		
TOTAL DE RESPOSTAS CERTAS	9	9	6	11	10	12	10	10	1	10	95	88	85		



RESPOSTAS

ESQUISADORES	Nº DAS PERGUNTAS - Nº DOS DOCUMENTOS RESPOSTA										TOTAL DE DOCUMENTOS RECUPERADOS	TOTAL DE RESPOSTAS CERTAS	TOTAL DE SUCESSOS NA PESQUISA
	Nº DOS DOCUMENTOS RECUPERADOS												
	1	2	3	4	5	6	8	9	10	11			
	92	20	100	78	44	10	56	21	9	45			
		74			61			11					
1	92	20	100	5	44	10	56	21	9	45	12	12	10
				78		61							
2	92	20	100	5	44	10	56	21	9	45	12	12	10
				78		61							
3	92	20	100	5	44	61	56	21	75	45	10	9	9
4	92	20	100	5	44	10	56	21	9	45	12	12	10
				78					11				
5	92	-	100	78	44	61	56	21	17	45	10	8	8
										7			
6	92	20	100	5	44	61	56	21	75	45	10	9	9
7	92	20	100	5	44	10	56	21	9	45	12	12	10
				78		61							
8	92	20	100	5	44	10	56	21	9	45	13	12	10
				78		61			35				
9	92	20	100	5	44	10	56	21	9	45	10	10	10
10	92	20	100	5	44	61	56	21	9	45	12	11	10
				78					35				
TOTAL DE RESPOSTAS CERTAS	10	9	10	16	10	14	10	10	8	10	113	107	96

RESPOSTAS

PESQUISADORES	Nº DAS PERGUNTAS - Nº DOS DOCUMENTOS RESPOSTA											TOTAL DE DOCUMENTOS RECUPERADOS	TOTAL DE RESPOSTAS CERTAS	TOTAL DE SUCESSOS NA PESQUISA
	Nº DOS DOCUMENTOS RECUPERADOS													
	1	2	3	4	5	6	8	9	10	11	11			
92	20	100	5	44	10	56	21	9	10	11	45			
	74		78		61									
1	92	-	5	44	-	56	21	11	11	45		7	7	7
2	92	20	-	26-56 44-89	10 61	-	21	-	-	22		10	6	5
3	92	-	-	44	10	56	21	-	-	45		6	6	6
4	92	20	-	44	10	56	21	11	11	45		9	9	9
5	92	20	-	44	-	56	21	11	11	45		8	8	8
6	92	-	-	44	-	56	21	11	11	45		7	7	7
7	92	20	-	5 78	61	56	21	75	75	45		10	9	8
8	92	20	100	5	10	56	21	-	-	45		8	8	8
9	92	-	-	44	-	56	21	11	11	45		7	7	7
10	92	20	-	44	-	56	21	11	11	45		8	8	8
TOTAL DE RESPOSTAS CERTAS	10	6	1	9	9	6	10	6	6	9		80	75	73

RESPOSTAS

	Nº DAS PERGUNTAS - Nº DOS DOCUMENTOS RESPOSTA										TOTAL DE DOCUMENTOS RECUPERADOS	TOTAL DE RESPOSTAS CERTAS	TOTAL DE SUCESSOS NA PESQUISA	
	1	2	3	4	5	6	8	9	10	11				
ESQUISADORES	92	20	100	5	44	10	56	21	9	9	45			
	74	78			61			11						
1	92	20	100	5	44	10	56	21	11		45	10	10	10
2	92	74	100	5	44	10	56	21	9		45	13	10	10
				78	61				11					
3	92	8	8	5	9	10	64	21	9		45	10	6	6
4	92	74	100	5	44	23	56	21	29		29	10	7	7
5	92	8	8	5	44	61	64	21	11-73 29-32			12	6	6
6	92	20	100	78	44	61	56	21	9 35		45	11	10	10
7	92	1	1	8	36	10	56	21	25		45	10	5	5
8	92	20	100	5	44	61	56	21	9		45	10	10	10
9	92	1	1	8	36	10	56	21	25		45	10	5	5
10	92	8	-	5	9	10	64	21	9		45	9	6	6
TOTAL DE RESPOSTAS CERTAS	10	5	5	9	6	10	7	10	8	8	78	105	78	75

LINGUAGEM DE INDEXAÇÃO: CRG

RESPOSTAS

	Nº DAS PERGUNTAS - Nº DOS DOCUMENTOS RESPOSTA											TOTAL DE DOCUMENTOS RECUPERADOS	TOTAL DE RESPOSTAS CERTAS	TOTAL DE SUCESSOS NA PESQUISA	
	1	2	3	4	5	6	8	9	10	11	11				
ESQUISADORES	92	20	100	5	44	10	56	21	9	45					
		74		78		61			11						
1	92	20	100	5	44	61	81	21	88	45		10	8		8
2	92	96	100	4	44	10	56	21	9	45		13	9		8
		99		77		61									
3	92	20	100	5	44	62	56	21	1	45		10	8		8
4	92	20	100	5	44	61	81	21	88	45		10	8		8
5	92	20	-	5	44	10	56	21	9	45		10	10		9
						61									
6	92	20	100	78	44	10	56	21	-	29		10	9		8
						61									
7	92	20	100	5	44	10	81	21	88	45		10	8		8
8	92	20	100	5	44	61	56	21	88	45		10	9		9
9	92	20	100	5	44	10	56	21	9	45		10	10		10
10	92	20	100	5	44	61	81	21	88	45		10	8		8
TOTAL DE RESPOSTAS CERTAS	10	9	9	9	10	12	6	10	3	9		103	87		84



LINGUAGEM DE INDEXAÇÃO : KWIC  
PRECISÃO DO ÍNDICE

PESQUISADORES	REFERÊNCIAS BIBLIOGRÁFICAS EXAMINADAS											TOTAL DE REFERÊNCIAS EXAMINADAS	TOTAL DE RESPPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	3	3	19	3	5	5	38	1	6	7		88	9
2	2	2	14	3	7	7	1	2	9	7		56	10
3	7	7	9	3	9	7	1	1	9	4		51	9
4	7	7	9	3	10	7	1	1	9	4		55	9
5	27	27	4	3	6	26	9	1	10	39		153	6
6	3	3	1	3	1	2	1	1	9	3		26	8
7	14	14	14	3	15	12	9	22	18	44		153	8
8	2	2	1	3	5	5	1	9	2	9		38	8
9	30	30	5	14	2	7	10	1	12	3		85	12
10	3	3	24	3	2	5	9	1	16	3		67	9
TOTAIS	98	98	100	41	62	83	80	40	100	123		772	88

LINGUAGEM DE INDEXAÇÃO : UNITERM  
 PRECISÃO DO ÍNDICE

PESQUISADORES	REFERÊNCIAS BIBLIOGRÁFICAS EXAMINADAS											TOTAL DE REFERÊNCIAS EXAMINADAS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	1	1	4	4	4	3	3	1	7	1		29	12
2	1	1	4	4	3	2	3	1	7	1		27	12
3	3	1	4	4	4	3	3	1	7	1		31	11
4	1	2	4	4	3	3	2	1	10	3		33	12
5	1	2	4	4	4	2	3	1	11	1		33	10
6	1	1	4	4	3	3	3	1	7	1		28	12
7	1	2	4	4	4	3	3	1	8	15		45	13
8	1	2	3	4	3	2	1	1	1	1		19	10
9	1	1	4	4	3	3	-	1	6	1		24	8
10	1	1	4	4	3	2	3	1	7	1		27	12
TOTAIS	12	14	39	40	34	26	24	10	71	26		296	112

## LINGUAGEM DE INDEXAÇÃO : CAB. ASS.

## PRECISÃO DO ÍNDICE

ESQUISADORES	REFERÊNCIAS BIBLIOGRÁFICAS EXAMINADAS											TOTAL DE REFERÊNCIAS EXAMINADAS	TOTAL DE RESPOSTAS CERTAS	
	1	2	3	4	5	6	8	9	10	11	11			
1	1	1	2	3	2	5	1	1	2	1			19	12
2	1	1	2	3	2	5	1	1	2	1			19	12
3	2	8	2	4	4	4	1	1	1	4			31	9
4	2	1	2	3	2	5	2	1	2	1			21	12
5	1	5	2	3	5	5	1	1	2	4			29	8
6	1	6	2	3	4	8	2	1	3	2			32	9
7	1	1	2	3	2	5	2	1	2	1			20	12
8	1	1	2	3	2	5	5	3	2	1			25	12
9	1	1	2	3	2	5	1	2	2	1			20	10
10	1	1	2	3	2	5	1	1	2	1			19	11
TOTALS	12	26	20	31	27	52	17	13	20	17			235	107

LINGUAGEM DE INDEXAÇÃO : CDD  
 PRECISÃO DO ÍNDICE

ESQUISADORES	REFERÊNCIAS BIBLIOGRÁFICAS EXAMINADAS											TOTAL DE REFERÊNCIAS EXAMINADAS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	1	1	-	3	4	-	3	1	8	1		22	7
2	1	1	1	-	4	4	-	1	-	2		14	6
3	1	1	-	-	3	5	1	1	-	1		13	6
4	3	1	3	3	4	6	1	1	8	1		31	9
5	1	2	-	3	4	-	3	1	8	1		23	8
6	1	1	-	3	4	-	3	1	8	1		22	7
7	5	2	3	3	9	7	3	1	10	1		44	9
8	6	1	8	3	4	7	1	1	4	1		36	8
9	1	1	-	3	4	-	3	1	8	1		22	7
10	1	2	-	3	4	-	3	1	8	1		23	8
TOTAIS	21	13	15	24	44	29	21	10	62	11		250	75

LINGUAGEM DE INDEXAÇÃO : CDU  
PRECISÃO DO ÍNDICE

INVESTIGADORES	REFERÊNCIAS BIBLIOGRÁFICAS EXAMINADAS											TOTAL DE REFERÊNCIAS EXAMINADAS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11	12		
1	1	1	11	1	18	13	2	1	10	12		70	10
2	3	20	11	1	24	8	4	9	30	3		113	13
3	12	11	11	1	10	2	1	1	10	1		60	6
4	6	1	2	1	11	14	6	1	27	10		79	7
5	12	1	11	1	2	2	1	1	10	2		43	6
6	1	1	2	1	2	2	1	1	20	1		32	10
7	12	11	11	13	2	2	1	1	10	1		64	5
8	1	2	14	1	2	4	1	1	10	1		37	10
9	13	11	11	3	2	2	1	1	10	1		55	5
10	12	2	2	1	10	2	1	1	10	1		42	6
TOTAIS	73	61	86	24	83	51	19	18	147	33		595	78

LINGUAGEM DE INDEXAÇÃO : CRG  
 PRECISÃO DO ÍNDICE

ESQUISADORES	REFERÊNCIAS BIBLIOGRÁFICAS EXAMINADAS											TOTAL DE REFERÊNCIAS EXAMINADAS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	4	1	4	2	2	2	2	7	3	1		27	4
2	4	1	2	2	2	1	1	1	2	1		18	9
3	1	2	4	2	2	2	1	1	3	1		20	8
4	1	1	4	2	2	2	1	1	2	1		17	8
5	1	1	2	2	2	2	1	1	2	1		14	10
6	1	2	2	6	5	2	2	4	1	1		26	9
7	1	1	1	4	2	2	2	1	1	1		16	8
8	1	1	1	4	4	5	1	1	2	1		21	9
9	1	1	1	2	2	2	2	1	4	1		17	10
10	1	4	1	4	2	2	2	1	2	1		20	8
TOTAIS	12	20	12	36	25	23	17	19	22	10		196	87

LINGUAGEM DE INDEXAÇÃO: KWIC  
 PRECISÃO DOS PESQUISADORES

PESQUISADORES	Nº DE DOCUMENTOS SELECIONADOS PARA EXAME										TOTAL DE DOCUMENTOS SELECIONADOS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11		
1	1	2	1	1	2	1	5	1	3	2	19	9
2	2	2	2	1	1	2	1	1	1	1	14	10
3	1	2	2	1	3	2	1	1	1	1	15	9
4	2	2	2	1	3	2	1	1	1	1	16	9
5	2	3	2	2	3	1	3	1	2	2	21	6
6	1	1	1	1	1	1	1	1	2	1	11	8
7	1	1	1	2	2	2	1	1	2	1	14	8
8	1	1	1	1	1	2	2	1	-	1	11	8
9	1	1	1	3	1	2	1	1	1	1	13	12
10	1	1	1	2	1	1	1	1	-	1	10	9
TOTAIS	13	16	14	15	18	16	17	10	13	12	144	88

LINGUAGEM DE INDEXAÇÃO: UNITERM  
 PRECISÃO DOS PESQUISADORES

PESQUISADORES	Nº DE DOCUMENTOS SELECIONADOS PARA EXAME											TOTAL DE DOCUMENTOS SELECIONADOS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	1	1	4	4	2	3	3	1	7	1		27	12
2	1	1	4	4	2	2	2	1	7	1		25	12
3	1	1	4	4	2	3	3	1	7	1		27	11
4	1	2	3	3	2	2	2	1	5	3		24	12
5	1	1	4	4	3	2	2	1	11	1		30	10
6	1	1	4	4	2	2	2	1	7	1		25	12
7	1	1	4	4	5	3	3	1	6	15		43	13
8	1	2	3	4	3	2	1	1	1	1		19	10
9	9	1	4	4	3	2	-	1	6	1		31	8
10	1	1	4	4	2	2	2	1	7	1		25	12
TOTAIS	18	12	38	39	26	23	20	10	64	26		276	112



LINGUAGEM DE INDEXAÇÃO: CAB. ASS.  
 PRECISÃO DOS PESQUISADORES

PESQUISADORES	Nº DE DOCUMENTOS SELECIONADOS PARA EXAME										TOTAL DE DOCUMENTOS SELECIONADOS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11		
1	1	1	1	2	1	2	1	1	1	1	12	12
2	1	1	1	2	1	2	1	1	1	1	12	12
3	1	1	2	2	1	3	1	1	2	1	15	9
4	1	1	1	2	1	1	1	1	2	1	12	12
5	1	3	2	1	1	1	1	1	1	2	14	8
6	1	2	2	1	1	3	1	1	1	1	14	9
7	1	1	1	2	1	2	1	1	1	1	12	12
8	1	1	1	2	1	2	1	1	2	1	13	12
9	1	1	1	1	1	2	1	2	2	1	13	10
10	1	1	1	2	1	1	1	1	2	1	12	11
TOTAIS	10	13	13	17	10	19	10	11	15	11	129	107

LINGUAGEM DE INDEXAÇÃO: CDD  
 PRECISÃO DOS PESQUISADORES.

PESQUISADORES	Nº DE DOCUMENTOS SELECIONADOS PARA EXAME											TOTAL DE DOCUMENTOS SELECIONADOS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	1	-	-	1	1	-	1	1	1	1	1	7	4
2	1	1	1	-	4	2	-	1	-	-	2	12	6
3	1	-	-	-	1	1	1	1	-	-	1	6	6
4	1	1	1	1	1	1	1	1	1	1	1	10	9
5	1	1	-	1	1	1	1	1	1	1	1	9	8
6	1	-	-	1	1	-	1	1	1	1	1	7	7
7	3	2	3	3	3	2	1	1	9	9	3	30	9
8	3	1	3	1	1	3	1	1	2	2	1	17	8
9	1	-	-	1	1	-	1	1	1	1	1	7	7
10	1	1	-	1	1	-	1	1	1	1	1	8	8
TOTAIS	14	7	8	10	15	10	9	10	17	13	113	75	

LINGUAGEM DE INDEXAÇÃO: CDU  
 PRECISÃO DOS PESQUISADORES

PESQUISADORES	Nº DE DOCUMENTOS SELECIONADOS PARA EXAME											TOTAL DE DOCUMENTOS SELECIONADOS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	1	1	1	1	1	1	1	1	1	1	1	10	10
2	3	5	2	4	4	2	3	1	3	1	1	28	13
3	1	1	1	1	1	1	1	1	1	1	1	10	6
4	1	1	2	1	2	1	2	1	3	1	1	15	7
5	1	1	1	1	2	2	1	1	5	-	-	15	6
6	1	1	1	1	1	1	1	1	2	1	1	11	10
7	1	2	1	2	2	2	1	1	3	1	1	16	5
8	1	1	2	1	1	1	1	1	1	1	1	11	10
9	1	1	1	1	1	1	1	1	1	1	1	10	5
10	1	1	1	1	1	1	1	1	1	1	1	10	6
TOTAIS	12	15	13	14	16	13	13	10	21	9	9	136	78

LINGUAGEM DE INDEXAÇÃO: CRG  
 PRECISÃO DOS PESQUISADORES

PESQUISADORES	Nº DE DOCUMENTOS SELECIONADOS PARA EXAME											TOTAL DE DOCUMENTOS SELECIONADOS	TOTAL DE RESPOSTAS CERTAS
	1	2	3	4	5	6	8	9	10	11			
1	3	1	2	1	1	2	1	1	1	1		14	8
2	2	1	2	1	2	1	1	1	1	1		13	9
3	1	1	1	2	1	2	1	1	2	1		13	8
4	1	1	2	1	1	2	1	1	1	1		12	8
5	1	1	-	1	1	2	1	1	1	1		10	10
6	1	1	1	2	3	2	1	1	-	1		13	9
7	1	1	1	2	1	1	2	1	1	1		12	8
8	1	1	2	2	2	3	1	1	1	1		15	9
9	1	1	1	1	1	1	1	1	5	1		14	10
10	1	3	1	2	1	1	2	1	1	1		14	8
TOTAIS	10	15	10	18	13	15	15	10	14	10		130	87

LINGUAGEM DE INDEXAÇÃO: KWIC

TEMPO DE PESQUISA

PESQUISADORES	MINUTOS DESPENDIDOS NA PESQUISA											TEMPO TOTAL EM MINUTOS	TEMPO MÉDIO EM MINUTOS
	1	2	3	4	5	6	8	9	10	11			
1	15	18	3	2	6	4	18	4	15	12		97	9,7
2	5	8	10	2	4	4	3	3	2	1		42	4,2
3	1	10	6	3	5	3	3	2	3	3		39	3,9
4	5	7	10	2	5	3	3	4	5	2		46	4,6
5	30	20	10	15	20	10	20	10	15	15		165	16,5
6	10	3	4	8	5	4	5	4	13	5		61	6,1
7	5	12	20	5	15	10	5	10	12	8		102	10,2
8	5	5	10	5	10	10	10	3	12	4		74	7,4
9	10	10	10	10	5	5	10	5	15	5		85	8,5
10	10	10	15	8	7	5	7	8	10	5		85	8,5
TOTALS	96	103	98	60	82	58	84	53	102	60		796	7,96

LINGUAGEM DE INDEXAÇÃO: UNITERM

TEMPO DE PESQUISA

PESQUISADORES	MINUTOS DESPENDIDOS NA PESQUISA											TEMPO TOTAL EM MINUTOS	TEMPO MÉDIO EM MINUTOS
	1	2	3	4	5	6	8	9	10	11			
1	15	5	10	10	5	10	5	10	5	5	5	80	8
2	11	5	5	9	11	12	7	8	15	10	9,3	93	9,3
3	5	9	6	7	10	5	7	4	9	8	7	70	7
4	15	5	10	10	10	10	5	5	5	5	8	80	8
5	5	6	10	5	5	4	7	3	16	3	6,4	64	6,4
6	10	10	10	10	5	5	5	10	10	10	8,5	85	8,5
7	5	10	6	7	10	9	8	3	6	10	7,4	74	7,4
8	5	8	5	7	10	10	5	5	5	5	6,5	65	6,5
9	10	10	5	5	3	4	5	5	10	5	6,2	62	6,2
10	10	10	5	8	10	10	8	5	15	5	8,6	86	8,6
TOTAIS	91	78	72	78	79	79	62	58	96	66	759	7,59	

TEMPO DE PESQUISA

PESQUISADORES	MINUTOS DESPENDIDOS NA PESQUISA											TEMPO TOTAL EM MINUTOS	TEMPO MÉDIO EM MINUTOS	
	1	2	3	4	5	6	8	9	10	11				
1	5	5	5	10	5	15	5	10	5	5			70	7
2	5	5	4	4	5	4	4	5	5	4			45	4,5
3	5	20	10	15	5	5	4	5	15	5			89	8,9
4	5	10	5	15	5	5	10	10	10	5			80	8
5	9	35	10	15	10	4	9	7	10	10			119	11,9
6	10	30	15	28	1	5	3	5	24	10			131	13,1
7	5	5	10	10	5	15	5	10	10	5			80	8
8	5	5	10	10	5	10	20	15	20	5			105	10,5
9	5	7	1	4	7	1	15	15	10	5			70	7
10	10	5	12	5	7	5	4	7	10	5			70	7
TOTALS	64	127	82	116	55	69	79	89	119	59			859	8,59

LINGUAGEM DE INDEXAÇÃO - CDD :

TEMPO DE PESQUISA

PESQUISADORES	MINUTOS DESPENDIDOS NA PESQUISA											TEMPO TOTAL EM MINUTOS	TEMPO MÉDIO EM MINUTOS
	1	2	3	4	5	6	8	9	10	11			
1	10	30	5	5	10	15	10	5	5	10	10	105	10,5
2	5	3	6	3	5	3	4	2	7	6	4,4	4,4	4,4
3	20	15	5	5	5	10	10	10	5	5	90	9	9
4	12	10	15	5	3	3	5	3	7	4	67	6,7	6,7
5	15	20	5	5	10	10	10	5	5	5	90	9	9
6	10	25	5	5	10	15	10	5	5	10	100	10	10
7	20	10	24	7	5	6	2	1	25	5	105	10,5	10,5
8	15	35	30	12	15	18	17	10	15	15	182	18,2	18,2
9	18	22	5	5	10	10	10	5	5	10	100	10	10
10	20	20	5	5	10	10	10	5	5	5	95	9,5	9,5
TOTAIS	145	190	105	57	83	100	88	51	84	75	978	9,78	9,78



## LINGUAGEM DE INDEXAÇÃO: CDU.

## TEMPO DE PESQUISA

PESQUISADORES	MINUTOS DESPENDIDOS NA PESQUISA											TEMPO TOTAL EM MINUTOS	TEMPO MÉDIO EM MINUTOS
	1	2	3	4	5	6	8	9	10	11	11		
1	5	9	5	3	13	6	11	2	12	5		71	7,1
2	22	17	9	5	24	9	12	5	8	15		126	12,6
3	5	10	5	5	10	5	5	5	5	5		60	6
4	10	5	10	5	10	15	10	5	10	10		90	9
5	15	15	5	10	5	5	5	2	5	5		72	7,2
6	3	10	10	15	10	5	5	10	15	5		88	8,8
7	5	5	5	5	7	5	4	4	10	10		60	6
8	10	5	10	5	10	10	5	5	5	9		74	7,4
9	15	15	10	13	17	7	8	22	3	9		119	11,9
10	20	10	20	10	15	5	5	5	5	10		105	10,5
TOTAIS	110	101	89	76	121	72	70	65	78	83		865	8,65

TEMPO DE PESQUISA

PESQUISADORES	MINUTOS DESPENDIDOS NA PESQUISA											TEMPO TOTAL EM MINUTOS	TEMPO MÉDIO EM MINUTOS
	1	2	3	4	5	6	8	9	10	11			
1	10	10	15	15	5	5	15	10	15	10		110	11
2	5	10	5	10	10	5	5	10	5	5		70	7
3	5	8	10	10	5	5	8	3	15	3		72	7,2
4	10	5	7	15	5	5	15	8	5	10		85	8,5
5	3	10	13	8	5	18	2	1	10	3		73	7,3
6	4	19	11	26	24	10	5	8	10	5		122	12,2
7	10	5	15	30	7	10	5	10	5	10		107	10,7
8	10	5	20	40	5	15	5	20	5	20		145	14,5
9	2	3	5	3	3	4	3	5	25	3		56	5,6
10	10	10	5	12	5	5	15	10	5	5		82	8,2
TOTAIS	69	85	106	169	74	82	78	85	100	74		922	9,22

Os totais obtidos nas 6 linguagens de indexação apresentam-se como se segue:

RESULTADOS TOTAIS DAS PESQUISAS

Linguagens de Indexação	Referências bibliográficas examinadas	Documentos selecionados para exame	Documentos indicados no resumo - posta	Respostas certas	Sucessos na Pesquisa	Tempo despendido na pesquisa (em minutos)
KWIC	772 <sup>+</sup>	144	95	88	85	796
UNITERM	296 <sup>++</sup>	276 <sup>++++</sup>	125	112	94	759
CAB. ASS.	235	129	113	107	96	859
CDD	250	113	80	75	73	978
CDU	595 <sup>+++</sup>	136	105	78	75	865
CRG	196	130	103	87	84	922

+ Foram contadas todas as entradas existentes no índice sob a palavra localizada pelos pesquisadores.

++ Contaram-se as referências bibliográficas anotadas nas fichas Uniterm localizadas pelos pesquisadores, quando ocorreu coordenação entre elas.

+++ Computaram-se todas as referências bibliográficas com o mesmo símbolo de classificação que a entrada localizada pelo pesquisador no índice, mas contaram-se também as referências seguintes ao símbolo, quando verificou-se, pelas respostas, que o pesquisador as consultou também.

++++ Verificou-se que quase todos os pesquisadores indicaram como documentos selecionados para exame todas as referências bibliográficas resultantes da coordenação, em lugar dos itens escolhidos na relação de citações bibliográficas, para pesquisa no texto. O fato parece não afetar o resultado, pois muitas entidades que utilizam o sistema Uniterm não mantêm relações de citações e o usuário é remetido do índice direto para a coleção de documentos.

Utilizando as medidas propostas no item Avaliação e a

plicando-as aos dados expostos obtiveram-se os coeficientes de Precisão do Índice, Precisão do Pesquisador e Precisão das Respostas, apresentados a seguir, juntamente com os coeficientes de Sucesso na Pesquisa e as medidas de Tempo de Pesquisa, apresentados abaixo:

COEFICIENTES DE PRECISÃO, SUCESSO NA PESQUISA E MÉDIA DE TEMPO

Linguagens de Indexação	Coefficiente de Precisão do Índice (CPI)	Coefficiente de Precisão do Pesquisador (CPP)	Coefficiente de Precisão das Respostas (CPR)	Coefficiente de Sucesso nas Pesquisas (CSP)	Tempo Médio de Pesquisa (em minutos) (TM)
KWIC	0,113	0,611	0,926	0,850	7,96
UNITERM	0,378	0,405	0,896	0,940	7,59
CAB. ASS	0,455	0,829	0,946	0,960	8,59
CDD	0,300	0,663	0,937	0,730	9,78
CDU	0,131	0,573	0,742	0,750	8,65
CRG	0,443	0,669	0,844	0,840	9,22

A finalidade de uma pesquisa bibliográfica é encontrar documentos que de fato respondam a uma indagação, esperando-se, portanto, que a linguagem de indexação permita sucesso nessa busca.

Segundo o Sucesso na Pesquisa, as linguagens de indexação se escalonaram do seguinte modo:

1º Cab. Ass.	coeficiente 0,960
2º Uniterm	" 0,940
3º KWIC	" 0,850
4º CRG	" 0,840
5º CDU	" 0,750
6º CDD	" 0,730

Considerando que quanto menos referência bibliográfica é necessário consultar para encontrar o documento desejado (Precisão do Índice) melhor é para o usuário, verificando-se que os coeficientes mais altos são aqueles que representam menor número de pesquisas. As linguagens documentárias se escalonaram co

mo se segue:

1º Cab. Ass.	coeficiente	0,455
2º CRG	"	0,443
3º Uniterm	"	0,378
4º CDD	"	0,300
5º CDU	"	0,131
6º KWIC	"	0,113

Admitindo-se que quanto menor o número de documentos o pesquisador escolheu para examinar (Precisão do Pesquisador) , mais elevada foi a sua capacidade de perceber os documentos realmente relevantes para sua consulta, verificou-se que os coeficientes mais altos são aqueles que representam as melhores seleções. Escalonaram-se as linguagens de indexação como se segue:

1º Cab. Ass.	coeficiente	0,829
2º CRG	"	0,669
3º CDD	"	0,663
4º KWIC	"	0,611
5º CDU	"	0,573
6º Uniterm	"	0,405

Reconhecendo-se que tanto melhor é uma pesquisa quanto menor for o número de documentos irrelevantes recuperados, obteve-se, quanto à Precisão das Respostas, o escalonamento seguinte:

1º Cab. Ass.	coeficiente	0,946
2º CDD	"	0,937
3º KWIC	"	0,926
4º Uniterm	"	0,896
5º CRG	"	0,844
6º CDU	"	0,742

Entendendo-se que quanto menor for o tempo dispendido na pesquisa, melhor porta-se o sistema em relação ao usuário. As linguagens de indexação escalonaram-se, quanto ao Tempo Médio de Pesquisa, como se segue:

1º Uniterm	tempo médio	7,59
2º KWIC	" "	7,96
3º Cab. Ass.	" "	8,59
4º CDU	" "	8,65
5º CRG	" "	9,22
6º CDD	" "	9,78

Quanto ao trabalho exigido para a armazenagem da informação nos vários índices, apresentamos a seguir os dados obtidos:

Linguagens de Indexação	Entradas completas	Entradas de índice	Assentamentos no índice Uniterm	Remissivas	Total
KWIC	100	466.			566
Uniterm	100	262	316	54	416
Cab. Ass.	166			186	352
CDD	219	297			516
CDU	289	666			955
CRG	107	304			411

A comparação entre os trabalhos de indexação pelos seis sistemas estudados é impossível, devido à grande diferença existente entre as tarefas de indexar pelo KWIC, Uniterm e catálogos alfabéticos e sistemáticos, a não ser medindo o tempo dispendido, o que não foi feito.

Admitindo-se que quanto maior é o número de fichas a fazer, maior é o trabalho de indexação. Podemos escalonar os sistemas tradicionais - Cabeçalho de Assunto, CDD, CDU e CRG - como se segue:

Cabeçalhos de assuntos	352	entradas
CRG	411	"
CDD	516	"
CDU	955	"

Considerando-se o concenso geral, diríamos que, quanto ao trabalho de indexação, os sistemas poderão ser escalonados:

- 1º KWIC
- 2º Uniterm
- 3º Cab. Assunto
- 4º CRG
- 5º CDD
- 6º CDU

Computaram-se os dados relativos aos itens 1-3, 5-6, 8-11 da "Folha-de-Avaliação do Sistema", apresentados no Quadro 5,

notando-se que, ao responderem os itens 10 e 11, as pesquisadoras expressaram as seguintes opiniões:

1) respondendo ao item 10 disseram sim 2 pesquisadoras do KWIC e 7 do Uniterm, mas aquelas ligadas ao primeiro não indicaram o método que prefeririam e as que utilizaram o Uniterm optaram por Cabeçalhos de Assunto ou Catálogo Dicionário;

2) entre as pesquisadoras do Cabeçalho de Assunto, quatro prefeririam o Catálogo Dicionário, o que significa a mesma coisa, uma optou por CDD ou CDU e uma gostaria de um índice com resumos dos documentos, sem indicar o tipo;

3) seis pesquisadoras destinadas ao índice CDD informaram preferir Cabeçalhos de Assunto;

4) dentre as pesquisadoras que usaram a CDU duas prefeririam a CDD, uma os Cabeçalhos de Assunto, uma a CDD ou os Cabeçalhos de Assunto, sendo que duas se queixaram da extensão dos símbolos da CDU;

5) das pesquisadoras que utilizaram o CRG, uma informou preferir a CDD e uma considerou não estar em condições de julgar, as demais (8) aceitaram bem o sistema.

Os dados coletados nos dão a seguinte percentagem de aceitação para os 6 sistemas:

Cab. Assunto	80%
CRG	80%
KWIC	80%
CDU	50%
CDD	40%
Uniterm	30%

Nota-se que entre as pesquisadoras que utilizaram outros sistemas, 19 são favoráveis à indexação por Cabeçalhos de Assunto, 5 pela CDD e 1 pela CDU.

Considerando-se as preferências indicadas por pesquisadoras de outros sistemas e os índices de aceitação de cada sistema utilizado, podemos afirmar que Cabeçalhos de Assunto mereceram a predileção dos usuários.

DADOS DA FOLHA DE AVALIAÇÃO DO SISTEMA

ITENS	KWIC		UNITERM		CAB. ASS.		CDD		CDU		CRG	
	SIM	NÃO	SIM	NÃO	SIM	NÃO	SIM	NÃO	SIM	NÃO	SIM	NÃO
1	1	9	-	10	-	10	2	8	2	8	-	10
2	10	-	8	2	8	2	4	6	9	1	9	1
3	9	1	7	3	8	2	10	-	6	4	9	1
5	1	9	3	6*	2	8	-	10	2	8	-	9*
6	2	8	2	8	3	7	1	9	2*	7	1	9
8	7	3	5	5	3	7	7	3	7	3	10	-
9	7	3	6	4	9	1	5	5	8	2	8	2
10	2	8	7	3	6	4**	6	4	5	5	2	8
11	Não indicou preferência 2		Cat. Dic. 5 Cab. Ass. 2		Cat. Dic. 4 CDD ou CDU 1 Com resumos 1		Cat. Dic. 6		Cat. Dic. 1 CDD 2 Cab. Ass. ou CDD 1 Não indicou preferência 1		CDD 1 Outro 1	
PREFERÊNCIA	80%		30%		80%		40%		50%		80%	

\* Uma pesquisadora não respondeu.

\*\* Note-se que 4 indicaram Catálogo Dicionário como o sistema preferido.



## 5 - INTERPRETAÇÕES E CONCLUSÕES

Os resultados encontrados favorecem os sistemas alfabéticos - KWIC, Uniterm e Cabeçalhos de Assunto - no que se refere ao Sucesso na Pesquisa e ao Tempo de Pesquisa, medidas relacionadas diretamente com a linguagem de indexação.

Estes resultados lembram John Metcalf, o grande apolo-gista dos Cabeçalhos de Assunto, que tão veementemente defende a indexação alfabética ao afirmar: "Toda indexação depende direta ou indiretamente da ordem alfabética, porque, qualquer que seja o número, notação, símbolo ou código a que leve, é preciso começar por nomes, escritos em letras do alfabeto ... A ordem alfabética é o mais importante instrumento de precisão do mundo ..." <sup>27</sup>.

É fato que os Cabeçalhos de Assunto, que surgiram em primeiro lugar em Sucesso na Pesquisa, atingindo 96%, são, pela sua rede de remissivas, uma verdadeira classificação disfarçada.]

Excluindo-se a indexação baseada na simples rotação de títulos, todos os outros sistemas, baseados em conceitos, exigem um processo de classificação. É o que afirma B.C. Vickery <sup>28</sup> quando diz: "quase sempre, na armazenagem é recuperação da informação, na fase de indexação ou na fase de pesquisa, ou em ambas, a classificação está implícita".

Verificou-se que o KWIC situou-se entre os 3 primeiros sistemas, possibilitando 85% de recuperação, certamente devido ao fato de 8 das questões propostas conterem palavras do título dos documentos, pois o sistema falhou exatamente nas questões 3 e 10, aquelas em que isto não ocorreu.

O desempenho do sistema Uniterm, recuperando 94% dos documentos-fonte e ocupando o 1º lugar em tempo de pesquisa, foi muito bom, fato que já havia sido constatado no Projeto Cranfield I, quando atingiu o melhor desempenho (82%).

Este sistema, permitindo grande especificidade, possibilitando a indexação exaustiva no grau desejado e, especialmente, graças à facilidade de indexação que oferece, merece ser largamente utilizado, quando as condições locais permitirem, quando não houver necessidade de grande volume de pesquisas simultâneas.

Entre os sistemas de classificação, o do Classification Research Group teve o melhor desempenho em Sucesso na Pesquisa, atingindo 84%. O seu desempenho neste teste foi superior àquele atingido no teste Cranfield I, porém deve-se notar que foi outro o sistema utilizado.

A CDU situou-se sempre no segundo grupo e em Sucesso na Pesquisa (75%) atingiu, provavelmente por coincidência, praticamente o mesmo nível alcançado no teste Cranfield I (75,6%).

A CDD, com 73%, ocupou o último lugar em Sucesso na Pesquisa.

Resta saber se o sucesso do KWIC foi devido ao fato de as perguntas serem originárias de documentos-fonte ou se o mesmo ocorreria normalmente, uma vez que o título de um trabalho é, geralmente, um resumo do seu conteúdo.

A eficiência na recuperação do sistema KWIC depende dos próprios títulos dos documentos. Vários estudos já foram feitos sobre a qualidade informativa dos títulos de documentos e a relação entre títulos e descritores.

Swanson<sup>7</sup> estudando 100 questões utilizadas no teste Cranfield só encontrou 15% que não continham palavras usadas nos títulos dos documentos-fonte.

Kraft<sup>29</sup>, examinando títulos de documentos no campo do Direito, verificou que só 10,5% não continham pelo menos um dos termos usados como descritores e que 64,5% continham todos os descritores.

Maizell<sup>30</sup>, comparando 25 títulos de documentos sobre Física com os descritores utilizados no Physics Abstracts e no Chemical Abstracts, encontrou 63% contendo todos os descritores.

Ruhl<sup>31</sup>, estudando 84 títulos relacionados no Chemical Abstracts, achou 57% contendo todos os descritores usados no índice da publicação.

Montgomery e Swanson<sup>32</sup>, comparando duas amostras de 4 770 e 451 títulos com as entradas do Index Medicus, encontraram 86% das palavras dos títulos como descritores.

Pode-se, assim, inferir que o sistema KWIC, mesmo não sendo usado com perguntas derivadas de documentos-fonte, será capaz de recuperar cerca de 85% dos documentos sobre o assunto.

A necessidade de consultar primeiro o índice alfabético para descobrir o símbolo de classificação, é, certamente, a razão pela qual os sistemas classificados exigiram mais tempo à pesquisa, se bem que a diferença entre Cabeçalhos de Assunto (8,59 minutos) e CDU (8,65 minutos), neste teste, não tenha sido significativa.

Quanto à Precisão do Índice, duas das linguagens alfabéticas - Cabeçalhos de Assunto e Uniterm - aparecem entre os três primeiros sistemas, acompanhadas pela Classification of Library and Information Sciences (CRG).

Era de esperar que um sistema de classificação especializado, possibilitando uma indexação detalhada e precisa, atingisse melhor desempenho na recuperação do que sistemas de classificação gerais. Esta superioridade evidenciou-se na etapa de indexação, permitindo à CRG classificações adequadas, enquanto a CDD e CDU constituíram experiência frustrante para um bibliotecário.

A Precisão do Pesquisador, liderada pelos Cabeçalhos de Assunto, não demonstrou diferença entre os sistemas alfabéticos e classificados, mas o sistema Uniterm exigiu maior esforço do usuário.

O grande sucesso do teste coube aos Cabeçalhos de Assunto, que lideraram 5 das 6 avaliações e mesmo naquela em que não ocuparam o primeiro lugar, Tempo de Pesquisa, situaram-se no 3º, com pouco mais de meio minuto de diferença para o primeiro colocado.

Feitas estas constatações, passamos a examinar as causas dos fracassos na recuperação, estudando aqueles que ocorreram 3 ou mais vezes na mesma questão.

KWIC As questões nº 3 e 10, com 4 e 9 falhas respectivamente, foram aquelas que constituíram problema no KWIC, pois os demais erros foram casos isolados, atribuíveis aos pesquisadores.

Acreditamos que estas falhas deveram-se ao sistema de indexação, baseado na representação das palavras do título dos do

cumentos e não ao reconhecimento de conceitos, pois os documentos que respondiam às perguntas nº 3 e 10 não continham nos seus títulos palavras dos enunciados das perguntas.

**Uniterm** As falhas ocorridas no índice Uniterm não podem ser atribuídas ao sistema.

A questão nº 8 apresentou 3 falhas, todas devidas aos pesquisadores. O pesquisador nº 5 localizou o documento, mas não o indicou como resposta; o pesquisador nº 9 parece não ter localizado a ficha MEDICINA/MÉDICOS; e o pesquisador nº 4 não procurou os termos adequados para a coordenação.

A pergunta nº 10 levou a 3 falhas, duas devidas ao indexador, que não anotou o número dos documentos-fonte nas fichas para ESTATÍSTICA e PRODUÇÃO, e uma ao pesquisador, que localizou a resposta, mas não a incluiu no item 5a.

**Cab.Ass.** A única pergunta, que apresentou dificuldade de solução na indexação por Cabeçalhos de Assunto, foi a nº 10, com 3 falhas.

Todos que a responderam o fizeram pelo cabeçalho BIBLIOGRAFIA - ESTATÍSTICA, mas o sistema possibilitava a existência de LIVROS - PRODUÇÃO ou LITERATURA - PRODUÇÃO e os pesquisadores que falharam procuraram em LIVROS e em LITERATURA.

Diríamos que as falhas deveram-se ao indexador, que não previu a possibilidade de entradas em LITERATURA - PRODUÇÃO, LIVROS - PRODUÇÃO ou LIVROS - ESTATÍSTICA, LITERATURA - ESTATÍSTICA.

Notou-se, aliás, uma freqüente exigência dos pesquisadores em localizar entradas em LITERATURA ou LITERATURA CIENTÍFICA para estudos da produção literária, tais como crescimento e dis-tribuição.

**CDD** A pesquisa pela CDD apresentou 4 fracassos nas questões nº 2 e 10, 5 na questão nº 6 e 9 na questão nº 3.

A resposta à questão nº 2 encontrava-se no texto de um documento sobre a Biblioteca do Senado Federal e o artigo não foi indexado em profundidade suficiente para representar PRODASEN ou LEGISLAÇÃO. Os pesquisadores que não a responderam limitaram-se a procurar por PROCESSAMENTO DE DADOS, não examinando outros ângu

los da pergunta.

Quanto à pergunta nº 3, a falha foi do indexador, que não incluiu no índice INFORMÁTICA ou CIÊNCIA DA INFORMAÇÃO, entra das procuradas pelos pesquisadores, remetendo para 020, uma das classificações atribuídas ao documento-fonte.

O único pesquisador que localizou esta resposta o fez procurando no índice pela entrada BIBLIOGRAFIA.

Os erros ocorridos na pergunta nº 6 só podem ser atribuídos aos pesquisadores, porque localizaram BIBLIOTECAS NACIONAIS em 027.5 e não encontraram os documentos nºs 10 e 61, relacionados logo a seguir, em 027.581.

Estudando as falhas à pergunta nº 10 nota-se que os pesquisadores esperavam entradas no índice em LITERATURA CIENTÍFICA, PRODUÇÃO ou ANÁLISE QUANTITATIVA, enquanto aqueles que conseguiram localizar a resposta o fizeram pela entrada ESTATÍSTICA. Pesquisadores ou indexador podem ser responsabilizados pelas falhas nesta questão.

**CDU** Falharam na CDU as questões nºs 2 e 3 (5 falhas), nº 5 (4 falhas) e nº 8 (3 falhas).

A resposta à pergunta nº 2 deveria ser localizada num artigo sobre a Biblioteca do Senado, mas os pesquisadores que não atingiram o alvo concentraram-se nas entradas INFORMÁTICA, PROCESSAMENTO DE DADOS e PRODASEN, não generalizaram a pesquisa, para atingir BIBLIOTECAS PARLAMENTARES.

Os erros nas respostas à pergunta nº 3 cabem aos pesquisadores, pois procuraram em INFORMÁTICA, a mesma entrada que levou outros pesquisadores a localizarem a resposta, mas não selecionaram o documento certo.

Quatro pesquisadores não responderam corretamente a questão nº 5, dois deles chegaram à classificação do documento fonte, mas não o reconheceram, e os outros procuraram numa entrada improvável.

Quanto à pergunta nº 8, devem-se as falhas ao fato de o documento-fonte haver sido classificado em 027.6:362.11 (Biblio -

tecas destinadas a hospitais) e não haver sido feita entrada em 026:61 ou 61:026 (Bibliotecas especializadas em medicina).

Evidencia-se falha do pesquisador, que não tentou BIBLIOTECAS HOSPITALARES, e falha do indexador, que não incluiu entrada em BIBLIOTECAS ESPECIALIZADAS EM MEDICINA ou uma ficha de referência.

CRG As falhas na CRG ocorreram na questão nº 8 (4 falhas) e na questão nº 10 (7 falhas).

Todos os pesquisadores que erraram a questão nº 8 localizaram o documento-fonte nº 56 e o documento nº 81, mas indicaram o último, em lugar do primeiro, como resposta, quando este não satisfaz o perguntado.

Quando à pergunta nº 10, os pesquisadores que não a responderam satisfatoriamente pretenderam chegar à solução pelo cabeçalho LIVROS: COMERCIALIZAÇÃO, PRODUÇÃO, LITERATURA CIENTÍFICA, ANÁLISE DE CITAÇÕES, CITAÇÕES ou REVISÕES DA LITERATURA, enquanto aqueles que acertaram tentaram BIBLIOGRAFIA - ESTATÍSTICA ou ESTATÍSTICA.

O documento-fonte havia sido indexado em Kb (Bibliometria) e uma melhoria do índice, remetendo de LITERATURA - ESTATÍSTICA para BIBLIOMETRIA teria evitado as falhas de recuperação.

Inferre-se do exposto que, com exceção das falhas no sistema KWIC, pesquisadores ou indexador podem ser responsabilizados pelas deficiências de recuperação ocorridas nas várias linguagens de indexação.

Também este foi um fato constatado nos testes Cranfield I e W.R.U., quando as falhas humanas na pesquisa e na indexação foram da ordem de 77 e 84%.

Evidencia-se o grande acerto de Cleverdon<sup>2</sup>, quando afirma que, partindo do mesmo reconhecimento de conceitos, todas as linguagens são potencialmente capazes de produzir a mesma recuperação.

Mas é ainda Cleverdon<sup>2</sup> quem diz: "A única razão válida para preferir determinada ordem para um índice é a conveniência dos usuários".

Neste teste os usuários preferiram, sem deixar dúvida, um índice por Cabeçalhos de Assunto.

Pode-se argumentar que estivessem mais familiarizados com este tipo de indexação, mas é um fato comum à grande maioria dos usuários, habituados que estão ao uso de dicionários, enciclopédias, índices, listas, etc. em ordem alfabética.

Os resultados deste teste mostraram que a indexação alfabética é capaz de excelente desempenho na recuperação e que recebe a preferência dos usuários.

## 6 - CITAÇÕES BIBLIOGRÁFICAS

1. GULL, D.C. Seven years of work on the organisation of materials in the special library. American Documentation, Washington, 7(4):320-329, Oct. 1956.
2. CLEVERDON, C.W. Aslib Cranfield research project; a report on the testing and analysis of an investigation into the comparative efficiency of indexing systems. Cranfield, 1962. 305, /6/ p.
3. CLEVERDON, C.W. The Cranfield hypotheses. Library Quarterly, Chicago, 35(1):121-124, Jan. 1965.
4. CLEVERDON, C.W.; MILLS, J.; KEEN, M. Aslib Cranfield research project; factors determining the performance of indexing systems. Cranfield, 1966. 2 v.
5. CLEVERDON, C.W. & MILLS, J. The testing of index language devices. Aslib Proceedings, London, 15(4):106-130, Mar. 1963.
6. CLEVERDON, C.W. The Cranfield tests on index language devices. Aslib Proceedings, London, 19(6):173-194, June, 1967.
7. SWANSON, D.R. The evidence underlying the Cranfield results. Library Quarterly, Chicago, 35(1):1-20, Jan. 1965.
8. SWANSON, D.R. Some unexplained aspects of the Cranfield test of indexing performance factors. Library Quarterly, 41(3):223-228, July, 1971.
9. KEEN, E.M. & DIGGER, J.A. Report of an information science index languages test. Aberystwyth, College of Librarianship Wales, Dept. of Information Retrieval Studies, 1972. 2 v.

10. KEEN, E.M. The Aberystwyth index languages test. Journal of Documentation, London, 29(1):1-35, Mar. 1973.
11. CARLISLE, J.H. Comparing behavior at various computer display consoles in time-shared legal information. Santa Monica, Rand Corp., 1970. 43 p.
12. FETTER, R.B. & CARLISLE, J.H. Man-computer interaction in a decision-making environment. New Haven, Conn., Yale University, Dept. of Administrative Sciences, 1971. 56 p.
13. LANCASTER, F.W. & FAYEN, E.G. Information retrieval on-line. Los Angeles, Melville /1973/ 597 p.
14. GOFFMAN, W. & NEWILL, V.A. A methodology for test and evaluation of information retrieval systems. Information Storage and Retrieval, Oxford, 3(1):19-25, Aug. 1966.
15. SWETS, J.A. Information retrieval systems. Science, Washington, 141(3577):245-250, July 1963.
16. SWETS, J.A. Effectiveness of information retrieval methods. American Documentation, Washington, 20(1):72-89, Jan. 1969.
17. FELS, E.M. Evaluation of the performance of an information retrieval system by modified Mooers plan. American Documentation, Washington, 14(1):28-34, Jan. 1963.
18. MILLER, W.L. The extension of users' literature awareness as a measure of retrieval performance and its application to MEDLARS. Journal of Documentation, London, 27(2):125-135, June 1971.
19. ROBERTSON, S.E. The parametric description of retrieval tests. Part I: The basic parameters. Journal of Documentation, London, 25(1):1-27, Mar. 1969.
20. ROBERTSON, S.E. The parametric description of retrieval tests. Part II: Overall measures. Journal of Documentation, London, 25(2):93-107, June 1969.
21. POLLOCK, S.M. Measures for the comparison of information retrieval systems. American Documentation, Washington, 19(4):387-397, Oct. 1968.
22. FAIRTHORNE, R.A. Some basic comments on retrieval testing. Journal of Documentation, London, 21(4):267-270, Dec. 1965.
23. LANCASTER, F.W. Information retrieval systems; characteristics, testing and evaluation. New York, J. Wiley, 1968. 222 p.



24. SANFORD, J.A. & THERIAULT, F.R. Problems in the application of uniterm coordinate indexing. In: TAUBE, M. et alii. Studies in coordinate indexing. /Washington/ Documentation Inc., 1956. v. 3, p. 101-113.
25. TAUBE, M. et alii. Uniterms in coordinate indexing. In: ——— Studies in coordinate indexing. /Washington/ Documentation Inc., 1953. v.2, p. 37-46.
26. DOCUMENTATION INC. O sistema uniterm de indexação. Trad. de Washington José de Almeida Moura. /Rio de Janeiro, 195-/ 41 f. Datilografado.
27. METCALFE, J. Information indexing and subject cataloguing. New York, Scarecrow, 1957. 338 p.
28. VICKERY, B.C. Classificatory principles in natural language indexing systems. In: MALTBY, A. Classification in the 1970's. London, C. Bingley /1972/. p. 169-191.
29. KRAFT, D.H. A comparison of keyword-in-context indexing of titles with a subject heading classification system. American Documentation, Washington, 15(1):48-52, Jan.1964.
30. MAIZELL, R.E. Value of titles for index purposes. Revue de la Documentation, La Haye, 27(3):126-127, 1960.
31. RUHL, M.J. Chemical documents and their titles: human concept indexing vs. KWIC machine indexing. American Documentation, Washington, 15(2):136-141, Apr. 1964.
32. MONTGOMERY, C. & SWANSON, D.R. Machine-like indexing by people. American Documentation, Washington, 13(4):359-366, Oct. 1962.

## ANEXOS

- 1 - Relação dos documentos da coleção-teste
- 2 - Relação das perguntas apresentadas
- 3 - Relação das perguntas utilizadas no teste
- 4 - Instruções
- 5A- Como consultar o índice KWIC
- 5B- Como consultar um índice Unitermo
- 5C- Como consultar o catálogo alfabético de assuntos
- 5D- Como consultar o catálogo sistemático
- 6 - Folha de Pesquisa
- 7 - Folha de Avaliação do sistema

## ANEXO 1

## RELAÇÃO DOS DOCUMENTOS DA COLEÇÃO - TESTE

1. ABREU, D. Livros para os países em desenvolvimento. Ciência da Informação, Rio de Janeiro, 1(1):37-39, 1972.
2. ANDRADE, A.M.C. de. A mecanização do serviço de empréstimo da Biblioteca do Instituto de Ciências Exatas da U.F.M.G. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(2): 238-250, set. 1973.
3. ASSUNÇÃO, J.B. de. Organização de um serviço de informação em bioquímica. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2): 132-152, set. 1972.
4. ASSUNÇÃO, J.B. de. Projeto de um sistema de classificação bibliográfica analítico-sintético (ou facetado) para a indexação e recuperação de informações em biologia. Belo Horizonte, Instituto de Ciências Biológicas da UFMG, 1972. xxxi, 110 f. Dissertação de Mestrado.
5. BARBOSA, A.P. Classificação facetada. Ciência da Informação, Rio de Janeiro, 1(2):73-81, 1972.
6. BÔA MORTE, L. da. Utilização do índice rotativo na informação legislativa. R. Bibliotecon. Brasília, 2(1): 85-89, jan./jun. 1974.
7. BORDA, J.C. da S. Disseminação seletiva de informações: revisão bibliográfica e projeto para a Companhia Vale do Rio Doce. R. Bibliotecon. Brasília, 1(2): 181-191, jul./dez. 1973.
8. BOTELHO, T.M.G. A documentação como sistema. R. Bibliotecon. Brasília, 2(1): 57-70, jan./jun. 1974.

9. BRAGA, G.M. Informação, ciência, política científica: o pensamento de Derek de Solla Price. Ciência da Informação, Rio de Janeiro, 3(2): 155-177, 1974.
10. BRAGA, G.M. Projeto de microfilmagem de jornais na Biblioteca Nacional: aspecto bibliométrico. Ciência da Informação, Rio de Janeiro, 2(2): 233-237, 1973.
11. BRAGA, G.M. Relações bibliométricas entre a frente de pesquisa (research front) e revisões da literatura: estudo aplicado à ciência da informação. Rio de Janeiro, 1972. 39 p. Dissertação de Mestrado.
12. BURNS, R.W., jr. Biblioteca e enfoque sistêmico. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2): 164-183, set. 1972.
13. CALDEIRA, P. da T. Crescimento da literatura brasileira de doenças de Chagas; análise bibliométrica. Rio de Janeiro, 1974. 61 f. Dissertação de Mestrado.
14. CAMPOS, A. O nascer de uma utopia: ainda e sempre o problema da classificação bibliográfica. R. Bibliotecon. Brasília, 1(1): 15-19, jan./jun. 1973.
15. CARVALHO, A. de O. Técnicas de ensino: três experiências. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(1): 32-42, mar./set. 1972.
16. CARVALHO, C.P. de. A biblioteca e os estudantes. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2): 196-211, set. 1972.
17. CARVALHO, M. de L.B. de. Índice de citações: uma revisão da literatura. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(2): 207-217, set. 1973.
18. CESARINO, M.A. da N. O ensino de biblioteconomia: um currículo a ser mudado. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(1): 43-59, mar. 1973.

19. CHASTINET, Y.S. Metodologia para implantação de um banco de dados em ciência e tecnologia; projeto piloto aplicado à química. Rio de Janeiro, 1973. 81 f. Dissertação de Mestrado.
20. COELHO, A.L. Histórico da Biblioteca do Senado Federal. R. Bibliotecon. Brasília, 1(2):233-240, jul./dez. 1973.
21. CORUJEIRA, L.A. Métodos de prevenção e eliminação de fungos em materiais bibliográficos. R. Bibliotecon. Brasília, 1(1): 59-65, jan./jun. 1973.
22. COSTA, J.F. da. O Sistema Nacional de Informação Científica e Tecnológica (SNICT). R. Bibliotecon. Brasília, 1(2): 95-107, jul./dez. 1973.
23. CUNHA, M.B. da. Necessidades atuais de bibliotecários no Brasil. R. Bibliotecon. Brasília, 2(1):15-24, jan./jun 1974.
24. DIAS, A.C. Futurologia - doença infantil da Biblioteconomia. Ciência da Informação, Rio de Janeiro, 2(1):55-58, 1973.
25. DI GIORGI, M.L.A. Análises da comunicação entre autores no campo da literatura brasileira de tecnologia de alimentos. Rio de Janeiro, 1974. 133 f. Dissertação de Mestrado.
26. ESPÍRITO SANTO, A. do. Programa de relações públicas em bibliotecas universitárias. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(1):73-85, mar. 1973.
27. EYRE, J.J. Características de um serviço de informação para a indústria. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(2):176-195, set. 1973.
28. FARINAS, V.H.P. Sobre biblioteconomia. R. Bibliotecon. Brasília, 1(2):141-144, jul./dez. 1973.

29. FERNANDEZ, R.P. Análises bibliométricas da produção científica dos grupos de pesquisa sobre física do estado sólido na América Latina. Rio de Janeiro, 1973. 114, ix, |23| p: Dissertação de Mestrado.
30. FERREIRA, M.L.A. de G. Seminário sobre "A formação do bibliotecário face às exigências profissionais da atualidade." R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(2):251-263, set. 1973.
31. FIGUEIREDO, A. de. Uma introdução à biblioteconomia comparada: sumário de pontos importantes. R. Bibliotecon. Brasília, 1(2) 133-140, jul./dez. 1973.
32. FIGUEIREDO, L.M. de. Distribuição da literatura geológica brasileira: estudo bibliométrico. Rio de Janeiro, 1972. 32 p. Dissertação de Mestrado.
33. FIGUEIREDO, N. Evolução e avaliação do serviço de referência R. Bibliotecon. Brasília, 2(2): 175-198, jul./dez. 1974.
34. FONSECA, E.N. da. Bibliografia brasileira corrente: evolução e estado atual do problema. Ciência da Informação, Rio de Janeiro, 1(1): 9-14, 1972.
35. FONSECA, E.N. da. Bibliografia estatística e Bibliometria: uma reivindicação de prioridades. Ciência da Informação, Rio de Janeiro, 2(1): 5-7, 1973.
36. FONSECA, E.N. da. Biblioteca Central da Universidade de Brasília: história com um pouco de doutrina e outro tanto de memórias. R. Bibliotecon. Brasília, 1(1):35-42, jan/jun. 1973.
37. FOSKETT, D.J. Alguns aspectos sociológicos dos sistemas formais de comunicação do conhecimento. R. Bibliotecon. Brasília, 1(1):3-14, jan./jun. 1973

38. FOSKETT, D.J. Teoria dos sistemas gerais e a organização de bibliotecas. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(1):9-22, mar. 1973.
39. FREITAS, A.D. de. Processamento de informações de registros médicos-hospitalares. Rio de Janeiro, 1972. 54 f. Dissertação de Mestrado.
40. GALVÃO, C.M. Documentação na região Amazônica. R. Bibliotecon Brasília, 1(2):203-206, jul./dez. 1973.
41. GARCIA, M.L.A. A pesquisa em biblioteconomia. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(1):7-11, mar./set. 1972.
42. GARCIA, M.L.A. Uso da biblioteca entre professores do Instituto de Ciências Exatas da UFMG. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2):113-124, set. 1972.
43. GARCIA, M.L.A. Uso da informação bibliográfica entre professores do Instituto de Ciências Exatas da UFMG. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(2):196-206, set. 1973.
44. GELFAND, M.A. Política de aquisição em bibliotecas universitárias: planos e programas, individuais e cooperativos, de formação de acervos. R. Bibliotecon. Brasília, 2(2):155-165, jul./dez. 1974.
45. GOMES, T.F. & MARQUES, A. Seleção de periódicos científicos para a área de Física. Ciência da Informação, Rio de Janeiro, 3(2):105-146, 1974.
46. GONÇALVES, J.B. CICS, Centro de Informações em Ciências Sociais, FGV. Rio de Janeiro, 1972. 32 | 6 | p. Dissertação de Mestrado.

- 47. GONÇALVES, N.A.R. A ISBD e sua utilização nas bibliotecas. R. Bibliotecon. Brasília, 1(2):159-168, jul./dez. 1973.
- 48. GORE, D. Contra os dogmáticos: uma opinião cética sobre as bibliotecas. R. Bibliotecon. Brasília, 2(1):43-55, jan/jun. 1974.
- 49. IPPOLITO, C. Análise comparativa da aquisição e circulação de periódicos em bibliotecas da Universidade de São Paulo, na área médica e afim: uma metodologia bibliométrica. São Paulo, 1973. 163 f. Dissertação de Mestrado.
- 50. JACKSON, W.V. Um plano nacional para desenvolvimento de bibliotecas e centros de documentação. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(1):23-42, mar. 1973.
- 51. LEMOS, A.A.B. de. Estado atual do ensino de biblioteconomia no Brasil e a questão da ciência da informação. R. Bibliotecon. Brasília, 1(1):51-58, jan./jun. 1973.
- 52. LEMOS, A.A.B. de & MACEDO, V.A.A. A posição da biblioteca na organização operacional da universidade. R. Bibliotecon. Brasília, 2(2):167-174, jul./dez. 1974,
- 53. LENTINO, N. Considerações gerais sobre a 18ª edição da Classificação Decimal de Dewey. R. Bibliotecon. Brasília, 1(2):211-219, jul./dez. 1973.
- 54. LIMA, E. et alii. Biblioteca Central da Universidade Federal de Minas Gerais. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2):125-131, set. 1972.
- 55. LIMA, E. O bibliotecário brasileiro na década dos 70. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2):212-218, set. 1972.
- 56. LIMA, E. Bibliotecas de hospitais. R. Esc. Bibliotecon, UFMG, Belo Horizonte, 2(2):141-159, set. 1973.



57. LIMA, M.L. de A. Usuários de uma biblioteca universitária ; estudo realizado no Instituto de Filosofia e Ciências Humanas da Universidade Federal de Pernambuco. Rio de Janeiro, 1974. 70 f. Dissertação de Mestrado.
58. LION, M.L. da C. Assessoramento técnico na Câmara dos Deputados: depoimento de uma bibliotecária como assessora. R. Bibliotecon. Brasília, 2(1):71-84, jan./jun. 1974.
59. LOPEZ ROBLERO, E.L. Estudo do desenvolvimento da literatura sobre física do estado sólido no México, no período 1959-1971. Rio de Janeiro, 1974. 188 p. Dissertação de Mestrado.
60. MAIA, A.B. Centro de análise da informação: requisitos mínimos para seu funcionamento. Rio de Janeiro, 1972. 31 p. Dissertação de Mestrado.
61. MONTE-MÓR, J. Microfilmagem de jornais da Biblioteca Nacional. R. Bibliotecon. Brasília, 2(2):143-153, jul./dez. 1974.
62. MONTE-MÓR, J. Reforma da Biblioteca Nacional. Ciência da Informação, Rio de Janeiro, 1(1):15-23, 1972.
63. MORAES, R.B. de. Relatório do Diretor da Divisão de Preparação da Biblioteca Nacional ao Ministro da Educação e Saúde (março de 1945). R. Bibliotecon. Brasília, 2(1) : 91-106, jan./jun. 1974.
64. MOURA, W. Programa educacional da Biblioteca Regional de Medicina da Organização Pan Americana da Saúde. Ciência da Informação, Rio de Janeiro, 1(1):25-35, 1972.
65. NEVES, A.M.C.Q. et alii. Sistema de informações para o Ministério do Interior: Projeto SIPLAN. R. Bibliotecon. Brasília, 1(2):193-201, jul./dez. 1973.

66. OHMAN, E. Índices Alfabéticos da Classificação Decimal Universal. Ciência da Informação, Rio de Janeiro, 2(2) : 169-174, 1973.
67. OLIVEIRA, A.L. de. Escola e biblioteca. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2):184-195, set. 1972.
68. OLIVEIRA, E. de A. Automação dos índices das tabelas da Classificação Decimal Universal. Rio de Janeiro, 1973. 1 v. e 3 pastas. Dissertação de Mestrado.
69. OLIVEIRA, E.A. UNIDEK; o sistema e seu preparo para a automação de bibliografias. Ciência da Informação, Rio de Janeiro, 1(2):97-111, 1972.
70. PERES, O.C. & FULGÊNCIO, C.M. de O. Pesquisa sobre os usuários da Biblioteca Pública de Minas Gerais "Prof. Luis de Bessa". R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2):101-112, set. 1972.
71. POLÍTICA do Instituto Nacional do Livro para o setor bibliotecas (MEC-Plano setorial de educação e cultura, 1972 - 1974. Projeto 33: Programa Nacional do Livro). R. Bibliotecon. Brasília, 1(1):91-94, jan./jun. 1973.
72. POLKE, A.M.A. A biblioteca escolar e o seu papel na formação de hábitos de leitura. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(1):60-72, mar. 1973.
73. POLKE, A.M.A. Pesquisa bibliográfica. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(1):43-54, mar./set. 1972.
74. RAULINO, P.C. Um sistema de disseminação seletiva de informação para o Congresso Nacional. R. Bibliotecon. Brasília, 1(2):169-179, jul./dez. 1973.
75. ROBREDO, J. et alii. Metodologia para a elaboração da lista básica dos periódicos nacionais em ciências agrícolas e estudo da dispersão da literatura agrícola brasileira. R. Bibliotecon. Brasília, 2(2):119-142, jul./dez. 1974.

76. ROCHA, J.F. Publicações oficiais brasileiras: um tema a ser estudado. R. Bibliotecon. Brasília, 2(1):1-14, jan./jun 1974.
77. ROSA, M.V. Classificação facetada em odontologia. Porto Alegre, 1972. 114, 29 f. Dissertação de Mestrado.
78. ROSA, M.V. O papel da classificação na recuperação da informação: classificação facetada. R. Bibliotecon. Brasília, 1(2):207-210, jul./dez. 1973.
79. ROSSETE, L.M.T. Investigação sobre parâmetros de projeto de sistemas de recuperação da informação; abordagem teórica e aplicação prática em um ambiente universitário. /Rio de Janeiro, 1973/107 p. Dissertação de Mestrado.
80. SALGADO, C. Biblioteca: definição legal. R. Esc. Bibliotecon, UFMG, Belo Horizonte, 1(1):59-61, mar./set. 1972.
81. SAMBAQUY, L. de Q. A biblioteca do futuro. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(1):62-68, mar./set. 1972.
82. SANTOS, I.R. dos. A biblioteca escolar e a atual pedagogia brasileira. R. Bibliotecon. Brasília, 1(2):145-149, 1973.
83. SARACEVIC, T. Tecnologia da informação, sistemas de informação e informação como utilidade pública. Ciência da Informação, Rio de Janeiro, 3(1):57-67, 1974.
84. SHEPARD, M.D. A biblioteca moderna na era da tecnologia. R. Bibliotecon. Brasília, 1(2):121-132, jul./dez. 1973.
85. SOUSA, G.H.B.P. de. Classificação automática: revisão da literatura. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(2): 218-237, set. 1973.

86. SOUZA, S. de. Discografia da literatura brasileira. R. Bibliotecon. Brasília, 1(1):67-75, jan./jun. 1973
87. SPERANDIO, L. Histórico da Biblioteca Pública do Paraná. R. Bibliotecon. Brasília, 2(2):199-202, jul./dez. 1974.
88. STEPANENKO, A. Recursos humanos na produção e comercialização do livro no Brasil: problemas e possibilidades de solução. Ciência da Informação, Rio de Janeiro, 3(1):79-86, 1974.
89. STEVENS, N.D. O eterno conflito: biblioteca versus escola de biblioteconomia. R. Bibliotecon. Brasília, 1(2):151-158, jul./dez. 1973.
90. TEIXEIRA, I.L.R. Uma linguagem de busca para sistemas de recuperação de informação. Ciência da Informação, Rio de Janeiro, 3(1):21-50, 1974.
91. TSUPAL, R. Modelo para inclusão da ciência da informação nos currículos das escolas e cursos de graduação de biblioteconomia e documentação no Brasil. Rio de Janeiro, 1973. 51 | 3 | f. Dissertação de Mestrado.
92. VEIGA, E.A. & JAKOBSON, S.R. Índice em cadeia. Ciência da Informação, Rio de Janeiro, 3(1):69-78, 1974.
93. VICENTINI, A.L.C. Informática agrícola. Ciência da Informação, Rio de Janeiro, 1(2):83-90, 1972.
94. VICENTINI, A.L.C. et alii. Mecanização da Classificação Decimal Universal: o projeto LEMME. R. Bibliotecon. Brasília, 1(1): 21-33, jan./jun. 1973.
95. VIEIRA, A. da S. A automação no currículo de biblioteconomia. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(1):12-31, mar./set. 1972.

96. VIEIRA, A. da S. A biblioteca automatizada na Inglaterra. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 1(2):153-163, set. 1972.
97. VOLPINI, E.E. A Biblioteca Central da Universidade de Brasília e o planejamento de seu novo edifício. R. Bibliotecon. Brasília, 1(1):43-50, jan./jun. 1973.
98. WANDERLEY, M.A. Linguagem documentária; acesso à informação aspectos do problema. Ciência da Informação, Rio de Janeiro, 2(2):175-217, 1973. Dissertação de Mestrado.
99. WANDERLEY, M.A. Utilização de processos de automação na Biblioteca Nacional. Ciência da Informação, Rio de Janeiro, 2(1):41-54, 1973.
100. ZAHER, C.R. & GOMES, H.E. Da bibliografia à ciência da informação: um histórico e uma posição. Ciência da Informação, Rio de Janeiro, 1(1):5-7, 1972.

## ANEXO 2

## RELAÇÃO DAS PERGUNTAS APRESENTADAS

INQUIRIDOR: Thereza de Magalhães Requião

1. Qual é a equivalência livraria/habitante na cidade do Rio de Janeiro ?

Resposta: STEPANENKO, A. Recursos humanos na produção e comercialização do livro no Brasil. Ciên. Inform., Rio de Janeiro, 3(1):79-86, 1974.

2. Quantos títulos de periódicos possuía a Biblioteca do Instituto de Filosofia e Ciências Humanas da Universidade Federal de Pernambuco em 1974 ?

Resposta: LIMA, M.L.A. Usuários de uma biblioteca universitária, estudo realizado no Instituto de Filosofia e Ciências Humanas da UFP. Ciên. Inform., Rio de Janeiro, 3(1):51-66, 1974.

3. Para que tipo de leitor o índice em cadeia é indicado ?

Resposta: VEIGA, E.A. & JACKOBSON, S.R. Índice em cadeia. Ciên. Inform., Rio de Janeiro, 3(1):69-78, 1974.

INQUIRIDOR: Celeste Magalhães

1. Preciso de uma informação sobre a implantação do Centro de Processamento de Dados da Biblioteca do Senado Federal (PRODA-SEN)

Resposta: COELHO, A.L. Histórico da Biblioteca do Senado Federal. R.Bibliotecon. Brasília, 1(2):233-240, 1973.

2. Qual é o acervo atual da Biblioteca do Senado Federal ?

Resposta: COELHO, A.L. Histórico da Biblioteca do Senado Federal. R.Bibliotecon. Brasília, 1(2):233-240, 1973.

3. Qual a situação atual da biblioteca e do bibliotecário no nosso meio escolar ?

Resposta: SANTOS, I.R. dos. A biblioteca escolar e atual pedagógica brasileira. R.Bibliotecon. Brasília, 1(2):145-149, 1973.

4. Qual a situação real e verdadeira das instituições onde funcionam juntas a Escola de Biblioteconomia e a Biblioteca ?

Resposta: STEVENS, N.D. O eterno conflito: biblioteca versus escola de biblioteconomia. R. Bibliotecon. Brasília, 1(2):151-158, 1973.

INQUIRIDOR: Cecilia Dornelles

1. Segundo Mikailov qual é o objetivo da informática ?

Resposta: ZAHER, C. & GOMES, H.E. Da bibliografia à ciência da informação: um histórico e uma posição. Ciên. Inform., Rio de Janeiro, 1(1):5-7, 1972.

2. O que são facetas e subfacetas ?

Resposta: BARBOSA, A.P. Classificação facetada. Ciên. Inform., Rio de Janeiro, 1(2):73-81, 1972.

INQUIRIDOR: Rosy Bleggi Peixoto

1. Bibliografia sobre política de aquisição em bibliotecas universitárias e o que se está tentando fazer nos Estados Unidos a respeito.

Resposta: GELFAND, M.A. Política de aquisição em bibliotecas universitárias: planos e programas, individuais e cooperativos, de formação de acervos. R. Bibliotecon. Brasília, 2(2):155-165, jul./dez. 1974.

2. O que a Biblioteca Nacional do Rio de Janeiro faz para preservar sua coleção de jornais ?

Resposta: MONTE-MÓR, J. Microfilmagem de jornais da Biblioteca Nacional. R. Bibliotecon. Brasília, 2(2):143-153, jul./dez. 1974.

3. Como o leitor pode ter acesso a uma coleção completa de jornais, já que foram microfilmados também os números dos jornais que a Biblioteca Nacional não possui ?

Resposta: MONTE-MÓR, J. Microfilmagem de jornais da Biblioteca Nacional. R. Bibliotecon. Brasília, 2(2):143-153, jul./dez. 1974.

INQUIRIDOR: Maria Thereza Parente Napoleão

1. As funções educativas de um bibliotecário escolar não entram em choque com os professores ?

Resposta: POLKE, A.M.A. A biblioteca escolar e o seu papel na formação de hábitos de leitura. R. Esc. Bibliotecon, UFMG, Belo Horizonte, 2(1):60-72, mar. 1973.

2. O bibliotecário escolar deve exercer outras funções além das administrativas e das técnicas ?

Resposta: POLKE, A.M.A. A biblioteca escolar e o seu papel na formação de hábitos de leitura. R. Esc. Bibliotecon, UFMG, Belo Horizonte, 2(1):60-72, mar. 1973.

INQUIRIDOR: Walkyria Augusta Rocha

1. Como é vista a integração de bibliotecas para médicos e bibliotecas para enfermeiros ?

Resposta: LIMA, E. Bibliotecas de hospitais. R. Esc. Bibliotecon, UFMG, Belo Horizonte, 2(2):141-159, set. 1973.

2. Qual a posição ideal da biblioteca na estrutura do hospital ?

Resposta: LIMA, E. Bibliotecas de hospitais. R. Esc. Bibliotecon, UFMG, Belo Horizonte, 2(2):141-159, set. 1973.

INQUIRIDOR: Nizeth Cohen

1. Desejo informações sobre métodos de conservação de documentos, principalmente prevenção e eliminação de fungos.

Resposta: CORUJEIRA, L.A. Métodos de prevenção e eliminação de fungos em materiais bibliográficos. R. Bibliotecon, Brasília, 1(1):59-65, jan./jun. 1973.

2. Uma bibliografia sobre programas de relações públicas em bibliotecas universitárias.

Resposta: ESPÍRITO SANTO, A. do. Programa de relações públicas em bibliotecas universitárias. R. Esc. Bibliotecon, UFMG, Belo Horizonte, 2(1):73-85, mar. 1973.

INQUIRIDOR: Ada Maria Coaracy

1. Onde encontrar uma análise quantitativa de produção da literatura



tura científica ?

Resposta: BRAGA, G.M. Informação, ciência, política científica: o pensamento de Derek de Solla Price. Ciên. Inform., Rio de Janeiro, 3(2):155-177, 1974.

2. Onde encontrar uma lista selecionada de periódicos dedicados ao campo da física, que sirva de orientação na escolha dos periódicos para uma biblioteca de ciência e tecnologia ?

Resposta: GOMES, T.F. & MARQUES, A. Seleção de periódicos científicos para a área de Física. Ciên. Inform., Rio de Janeiro, 3(2):105-146, 1974.

RELAÇÃO DAS PERGUNTAS UTILIZADAS NO TESTE

1. Para que tipo de leitor o índice em cadeia é indicado ?  
Resposta: Doc. 92 - VEIGA, E.A. & JACKOBSON, S.R. Índice em cadeia. Ciên. Inform., Rio de Janeiro, 3(1):69-78, 1974.
2. Preciso de uma informação sobre a implantação do Centro de Processamento de Dados da Biblioteca do Senado Federal (PRODASEN)  
Resposta: Doc. 20 - COELHO, A.L. Histórico da Biblioteca do Senado Federal. R.Bibliotecon. Brasília, 1(2):233-240, 1973.
3. Segundo Mikailov qual é o objetivo da informática ?  
Resposta: Doc. 100 - ZAHER, C. & GOMES, H.E. Da bibliografia à ciência da informação: um histórico e uma posição. Ciên. Inform., Rio de Janeiro, 1(1):5-7, 1972.
4. O que são facetas e subfacetadas ?  
Resposta: Doc. 5 - BARBOSA, A.P. Classificação facetada. Ciên. Inform., Rio de Janeiro, 1(2):73-81, 1972.
5. Bibliografia sobre política de aquisição em bibliotecas universitárias e o que se está tentando fazer nos Estados Unidos a respeito.  
Resposta: Doc. 44 - GELFAND, M.A. Política de aquisição em bibliotecas universitárias: planos e programas, individuais e cooperativos, de formação de acervos. R.Bibliotecon. Brasília, 2(2):155-165, jul./dez.1974.
6. O que a Biblioteca Nacional do Rio de Janeiro faz para preservar sua coleção de jornais ?  
Resposta: Doc. 61 - MONTE-MÓR, J. Microfilmagem de jornais da Biblioteca Nacional. R.Bibliotecon. Brasília, 2(2):143-153, jul./dez. 1974.
7. Preciso de uma explicação sobre o sistema de indexação coordenada Unitermo.  
Resposta: Não há resposta nos documentos indexados.

8. Como é vista a integração de bibliotecas para médicos e bibliotecas para enfermeiros ?

Resposta: Doc. 56 - LIMA, E. Bibliotecas de hospitais. R. Esc. Bibliotecon. UFMG, Belo Horizonte, 2(2):141-159, set. 1973.

9. Desejo informações sobre métodos de conservação de documentos, principalmente prevenção e eliminação de fungos.

Resposta: Doc. 59 - CORUJEIRA, L. A. Métodos de prevenção e eliminação de fungos em materiais bibliográficos. R. Bibliotecon. Brasília, 1(1):59-65, jan./jun. 1973.

10. Onde encontrar uma análise quantitativa da produção da literatura científica ?

Resposta: Doc. 9 - BRAGA, G.M. Informação, ciência, política científica: o pensamento de Derek de Solla Price. Ciên. Inform., Rio de Janeiro, 3(2):155-177, 1974.

11. Onde encontrar uma lista selecionada de periódicos dedicados ao campo da física, que sirva de orientação na escolha dos periódicos para uma biblioteca de ciência e tecnologia ?

Resposta: Doc. 45 - GOMES, T.F. & MARQUES, A. Seleção de periódicos científicos para a área de Física. Ciên. Inform., Rio de Janeiro, 3(2):105-146, 1974.

ANEXO 4

INSTRUÇÕES

1. Este é um teste da eficiência de cada modo de indexar por as sunto. Não é o pesquisador que está sendo testado. Deste, es pera-se, somente, que preencha com exatidão a FOLHA DE PESQUI SA, descrevendo passo a passo cada pesquisa, e que responda sinceramente a FOLHA DE AVALIAÇÃO DO SISTEMA.
2. Leia a explicação sobre o modo de consultar este índice antes de iniciar a pesquisa.
3. Para preencher a FOLHA DE PESQUISA:
  - a) marque a hora de início de cada pesquisa no item 1;
  - b) leia a pergunta apresentada no item 2;
  - c) analise a pergunta e verifique quais os conceitos (assun - tos) que apresenta;
  - d) procure no índice alfabético pelo conceito mais importante. Caso não o encontre, tente sinônimos ou procure pelos demais conceitos. Lembre-se de que uma obra sobre um assunto mais geral pode conter o assunto que procura;
  - e) relacione no item 3, na ordem em que fizer a consulta, to das as entradas procuradas no índice alfabético, mesmo aque las que não resultaram em resposta;
  - f) examine os documentos relacionados sob o cabeçalho adequado ao assunto da pergunta e selecione aqueles possíveis de res pondê-la;
  - g) anote no item 4 o número de ordem dos documentos localiza - dos que parecem responder à pergunta;
  - h) procure nas revistas os documentos localizados, leia-os e veja qual deles responde à pergunta;
  - i) anote no item 5 o número de ordem do documento que responde à pergunta ou indique o fato de não ter encontrado a respos ta;
  - j) marque no item 6 a hora em que terminou cada uma das pesqui sas;
  - k) assine.

4. Se estiver consultando um catálogo classificado que utilize símbolos decimais, não aceite documentos cujos símbolos de classificação não tenham pelo menos os três primeiros algarismos iguais ao símbolo localizado no índice alfabético.  
Se estiver consultando o catálogo organizado pelo sistema de classificação do Classification Research Group (CRG), só aceite documentos cujos símbolos de classificação comecem pela mesma letra daquele localizado no índice alfabético.
5. Entre as questões apresentadas podem existir algumas para as quais não haja resposta.
6. Preencha a FOLHA DE AVALIAÇÃO DO SISTEMA.
7. TODOS OS ITENS DA FOLHA DE PESQUISA E DA FOLHA DE AVALIAÇÃO DO SISTEMA devem ser respondidos.

COMO CONSULTAR O ÍNDICE KWIC

Este índice é constituído de duas partes:

- 1) uma relação numérica dos documentos da coleção;
- 2) um índice alfabético pelas palavras significati

vas do título do documento, seguido o número do documento, como aparece na relação numérica:

Há entradas por todas as palavras significativas do título de cada documento indexado. Um documento cujo título seja Uma Política para Exportação de Café terá entradas começando por POLÍTICA, EXPORTAÇÃO e CAFÉ.

Para consultar este índice procure pela palavra que representa o assunto desejado, bem como por seus sinônimos, pois o autor pode ter usado qualquer um deles no título do seu trabalho. Se deseja um documento sobre TANGERINA, procure também por BERGAMOTA, LARANJA CRAVO e MEXERICA.

Lembre-se de que uma obra mais geral pode trazer a informação que procura. Um trabalho sobre LARANJAS, por exemplo, poderá tratar de TANGERINA.

Note que a palavra do título que determina a ordem alfabética fica no meio da página e as palavras restantes do título vêm antes ou depois dela.

Localizando um documento que o interesse anote o seu número e procure-o na relação numérica, onde encontrará a sua descrição bibliográfica completa.

17	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

## ANEXO 58

## COMO CONSULTAR UM ÍNDICE UNITERMO

Este índice é dividido em duas partes:

1) uma relação numérica de todos os artigos incluídos no índice (cada artigo tem o seu número próprio).

2) um índice em fichas alfabetado pelos nomes dos assuntos que encabeçam as fichas.

Nessas fichas são anotados os números de todos os artigos que tratam do assunto que encabeça as fichas. Esses números são distribuídos na ficha unitermo pela unidade.

Há uma ficha para cada uma das palavras que compõem o assunto dos documentos. Um documento sobre a POLÍTICA DO BRASIL PARA COMÉRCIO DO CAFÉ aparecerá em quatro fichas - POLÍTICA, BRASIL, COMÉRCIO, CAFÉ.

Para consultar o índice, retiram-se as fichas de todas as palavras que compõem os assuntos procurados e verificam-se quais os números dos documentos que aparecem em todas elas. Só esses documentos tratam do assunto procurado. Ex.:

POLÍTICA										COMÉRCIO / COMERCIALIZAÇÃO/COMERCIAL		8	9		
CAFÉ										8	9	(28)	19		
BRASIL / BRASILEIRO										8	9	(28)	19		
0	1	2	3	4	5	6	7	8	9	(28)	19		29		
20	11	12	23		5	36	37	(28)			49				39
	71						77								

No exemplo só o doc. 28 trata de POLÍTICA do BRASIL pa

ra COMÉRCIO DO CAFÉ.

Repare, também, que foram reunidas na mesma ficha as várias formas de uma mesma palavra. Ex.: COMÉRCIO/COMERCIALIZAÇÃO/COMERCIAL.

Se por acaso, não encontrar determinada palavra procure um sinônimo ou um assunto mais geral que possa incluir o assunto procurado.



COMO CONSULTAR O CATÁLOGO ALFABÉTICO DE ASSUNTOS

Neste catálogo as fichas que descrevem os documentos são encabeçadas pelo nome dos assuntos de que tratam e estão ordenadas alfabeticamente.

Procure pelo assunto específico. Se necessita de um documento sobre ROSAS, localize a palavra ROSAS e não FLORES ou BOTÂNICA, mas, se não houver documentos sobre o assunto que procura, lembre-se de que uma obra mais geral, como uma obra sobre FLORES, pode trazer a informação que deseja.

Caso não encontre o assunto que procura, tente um sinônimo. O catálogo tentará guiá-lo remetendo-o dos sinônimos para os nomes escolhidos para as entradas ou levando-o dos assuntos gerais aos específicos.

## COMO CONSULTAR O CATÁLOGO SISTEMÁTICO

Neste catálogo os assuntos são representados por símbolos de um sistema de classificação, que aparecem ao alto de cada ficha, no canto direito.

O catálogo está dividido em duas partes:

1) a primeira parte é um índice no qual os nomes dos assuntos aparecem em ordem alfabética, trazendo ao lado os símbolos de classificação que lhes correspondem;

2) a segunda parte é uma relação, em fichas, de todos os documentos existentes na coleção, tendo ao alto, no canto direito, o símbolo de classificação, que representa o assunto tratado em cada documento e serve, também, como meio de ordenação.

Os símbolos de classificação, nas classificações decimais, são números decimais e como tais são ordenados.

Os símbolos da classificação de biblioteconomia do Classification Research Group (CRG) são constituídos de letras maiúsculas e minúsculas e são ordenados alfabeticamente pelas letras maiúsculas e sub-ordenados pelas letras minúsculas.

A consulta é feita localizando no índice o nome do assunto procurado, a fim de descobrir o seu símbolo de classificação, e, depois, consultando a segunda parte, aquela em que aparece a relação de todos os documentos da coleção, procurando pelo símbolo de classificação.

Sob cada símbolo de classificação estão relacionados os documentos que tratam do assunto que representam.

Não localizando o que deseja, convém procurar também nas fichas anteriores e posteriores, pois obras sobre assuntos mais gerais ou mais específicos podem trazer a informação necessária.

## LINGUAGEM DE INDEXAÇÃO:

## FOLHA DE PESQUISA

TODOS OS ITENS ABAIXO DEVEM SER RESPONDIDOS.

1. Hora de início desta pesquisa:

2. Pergunta: (Nº 1)

Para que tipo de leitor o índice em cadeia é indicado ?

3. Relacione aqui todas as entradas procuradas no índice alfabético, na ordem em que as pesquisou.

4. Indique aqui o Número de Ordem dos documentos localizados que parecem responder à pergunta (aqueles documentos que você irá examinar para certificar-se de que respondem a pergunta). Faça esta anotação antes de examinar os documentos.

5. Indique aqui

a) o Número de Ordem dos documentos que respondem à pergunta;

ou

b) o fato de não ter encontrado a resposta

6. Hora em que terminou esta pesquisa:

Assinatura: \_\_\_\_\_

## LINGUAGEM DE INDEXAÇÃO:

## FOLHA DE AVALIAÇÃO DO SISTEMA

Data:

Assinatura:

Período escolar:

- |   |     |     |
|---|-----|-----|
| 1. Você já tinha consultado um catálogo deste tipo ?  | SIM | NÃO |
| 2. As instruções sobre o modo de utilizar este catálogo foram suficientes ?                         | SIM | NÃO |
| 3. Você achou fácil usar este tipo de catálogo ?  | SIM | NÃO |
| 4. O que foi mais difícil:  |     |     |
| localizar a entrada no índice alfabético ?  | ( ) |     |
| localizar a ficha do documento ?  | ( ) |     |
| 5. Achou difícil escolher os documentos que poderiam responder as perguntas ?                       | SIM | NÃO |
| 6. Você ficou cansada de procurar estas 10 perguntas ?  | SIM | NÃO |
| 7. Se ficou cansada foi:  |     |     |
| por causa do catálogo ?   | ( ) |     |
| por ter que preencher a folha de pesquisa ?   | ( ) |     |
| 8. Você teve vontade de continuar a pesquisar depois de ter terminado o trabalho ?                  | SIM | NÃO |
| 9. Você acha que um leitor comum terá dificuldade em utilizar este tipo de catálogo ?               | SIM | NÃO |
| 10. Você preferiria outro tipo de catálogo ?  | SIM | NÃO |
| 11. Que outro tipo preferiria ?   |     |     |
| <hr/>   |     |     |
| <hr/>   |     |     |
| 12. Há alguma coisa que você gostaria de acrescentar sobre esta experiência ? Faça-o, se o desejar. |     |     |
| <hr/>   |     |     |
| <hr/>   |     |     |

## SYNOPSIS

PIEADADE, M.A.R. - Estudo comparativo de algumas linguagens de indexação: eficácia e tempo de pesquisa. Rio de Janeiro, 1975. p. Thesis

Reports a comparative test on information retrieval by KWIC, Uniterm, Subject Headings, Dewey Decimal Classification, Universal Decimal Classification and the Classification of Librarianship and Information Sciences (a faceted classification developed by the Classification Research Group).

One hundred Brazilian documents on Library and Information Sciences were selected for the test and every document was indexed according to the six systems covered by the evaluation.

Eight librarians prepared ten questions that could be answered completely by one or more of the documents in the test collection.

Sixty students of the Faculdade de Biblioteconomia e Documentação of the Universidade Santa Ursula, divided randomly into groups of ten, searched in the six indexes to find the documents that would answer the ten questions.

Each searcher was instructed to record time spent in search, the entries searched, documents selected for examination, and documents considered to answer the questions posed. To these data were added the number of documents assigned to each term consulted by the searchers.

From the data collected were determined the performance figures for each system.

Analysis of the data showed that alphabetical index systems (KWIC, Uniterm and Subject Headings) performed better than the classified indexing languages. Among the latter, the faceted classification had the best performance.

Considering degree of success in answering the test questions, the systems were ranked as follows: Subject Headings (96%), Uniterm (94%), KWIC (85%), Classification of Librarianship and Information Sciences (84%), Universal Decimal Classification (75%) and Dewey Decimal Classification (73%).

Analysis of failures showed that most were due to human failures in indexing and searching rather than to any intrinsic defects in the languages themselves.

Searchers preferences favoured Subject Headings.

Final results showed that the Subject Headings were ranked first according to five of the six variables studied, and were ranked second on the other variable.

ERRATA

nas páginas

4, 6, 7, 8, 9, 10, 13, 16, 18, 19, 22, 23, 31, 32  
33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46,  
47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60,  
61, 62, 63, 64, 65, 66, 67, 68, 69, 70

onde está pesquisa leia-se busca

onde está pesquisador leia-se encarregado da recuperação