

Linguagens de Indexação: uma experiência de análise e avaliação

MARIA MARTHA DE CARVALHO *

TANIA MARA BOTELHO **

WANDA MARIA MAIA DA ROCHA PARANHOS ***

Descrição de uma experiência que compara e avalia cinco diferentes linguagens de indexação, adotando-se as medidas de sensibilidade e especificidade para se avaliar a eficácia de cada uma.

1. OBJETIVOS

A finalidade do presente artigo é descrever uma experiência que consistiu em comparar e avaliar cinco diferentes linguagens de indexação. A experiência visou chegar a uma avaliação da *eficácia* de cada uma das linguagens de indexação, usando para tal um conjunto de medidas.

As linguagens de indexação utilizadas foram:

- a) Unitermo, utilizando termos retirados dos textos dos documentos. — Unitermo.

* Professora da Escola de Biblioteconomia da UFMG.

** Cientista de Informação do Departamento de Suporte de Sistemas, SERPRO.

*** Professora do Departamento de Biblioteconomia da UFPR.

- b) EJC Thesaurus, 1964 (Do "Engeneering Joint Council") — EJC.
- c) ERIC Thesaurus, 1970 (Do "Educational Resources Information Center") — ERIC.
- d) Library Literature Subject Headings List, 1963. — LLSHL.
- e) Schultz, G. Thesaurus of Information Science. 1968. — TIS.

2. DESCRIÇÃO DA EXPERIÊNCIA

A experiência aqui descrita, seguirá uma metodologia específica, cujas etapas estão representadas no fluxograma anexo (figura 1).

O arquivo sujeito à experiência de avaliação consistiu nos 65 artigos do livro do Dr. Tefko Saracevic (1). Cada um dos artigos foi, em conjunto, indexado por 6 diferentes modos em cada uma das cinco linguagens de indexação utilizadas (portanto, cada artigo recebeu o equivalente a seis "cabecinhos de assunto" em cada linguagem).

O conjunto das informações referentes ao artigo, isto é, a identificação dos 65 artigos e os respectivos seis termos de indexação, alocados a cada um deles nas cinco linguagens foram armazenados em computador, obedecendo ao formato FAMULUS (figura 2), sistema usado na School of Library Science da University of Missouri.

Segundo o referido formato, os campos se dividem da maneira especificada a seguir:

NOME DO CAMPO	ESPECIFICAÇÃO
DOC =	Nº do documento no livro-base.
A =	Autor(es), usando o sobrenome seguido do prenome ou iniciais. Se mais de um autor, foi usado "1", "2" para acrescentar os demais.
T =	Título do artigo, título do periódico, ou se este último não constava, o título do livro-base.
D =	Data de publicação do artigo, usando-se dois dígitos para o ano (00 a 99) e dois dígitos para o mês (01 a 12).
AB =	Resumo do artigo.
IS1	Conjunto de unitermos usados do texto do documento.
IS2	Conjunto de termos selecionados do EJC Thesaurus.
IS3	Conjunto de termos selecionados do ERIC Thesaurus.
IS4	Conjunto de termos selecionados do Library Literature Subject Headings List.
IS5	Conjunto de termos selecionados do Thesaurus of Information Science, de G. Schultz.

A primeira listagem produzida pelo computador foi uma relação dos termos utilizados para indexação dos 65 artigos em cada linguagem. A lista foi organizada em duas diferentes ordens (ver figura 3):

- a) ordem alfabética dos termos;
- b) ordem numérica decrescente de frequência de uso do termo.

Concluída essa etapa, estava o arquivo pronto para receber perguntas. O resultado obtido para cada uma das questões propostas mediante o uso das cinco linguagens permitiria a sua comparação e avaliação.

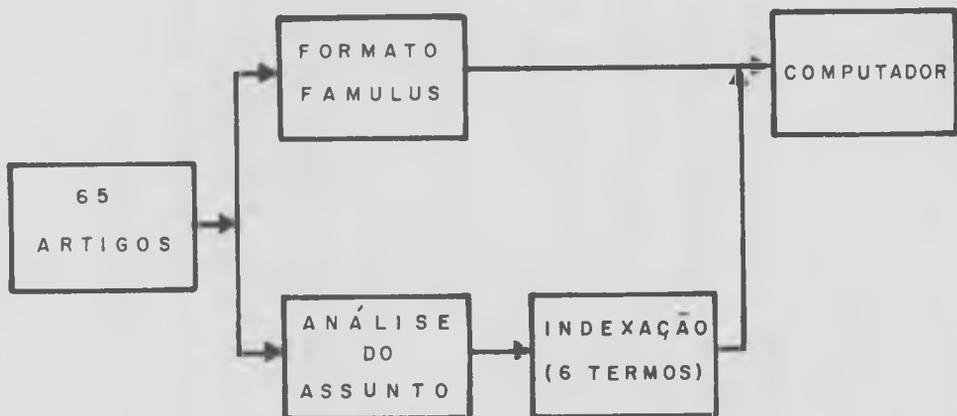
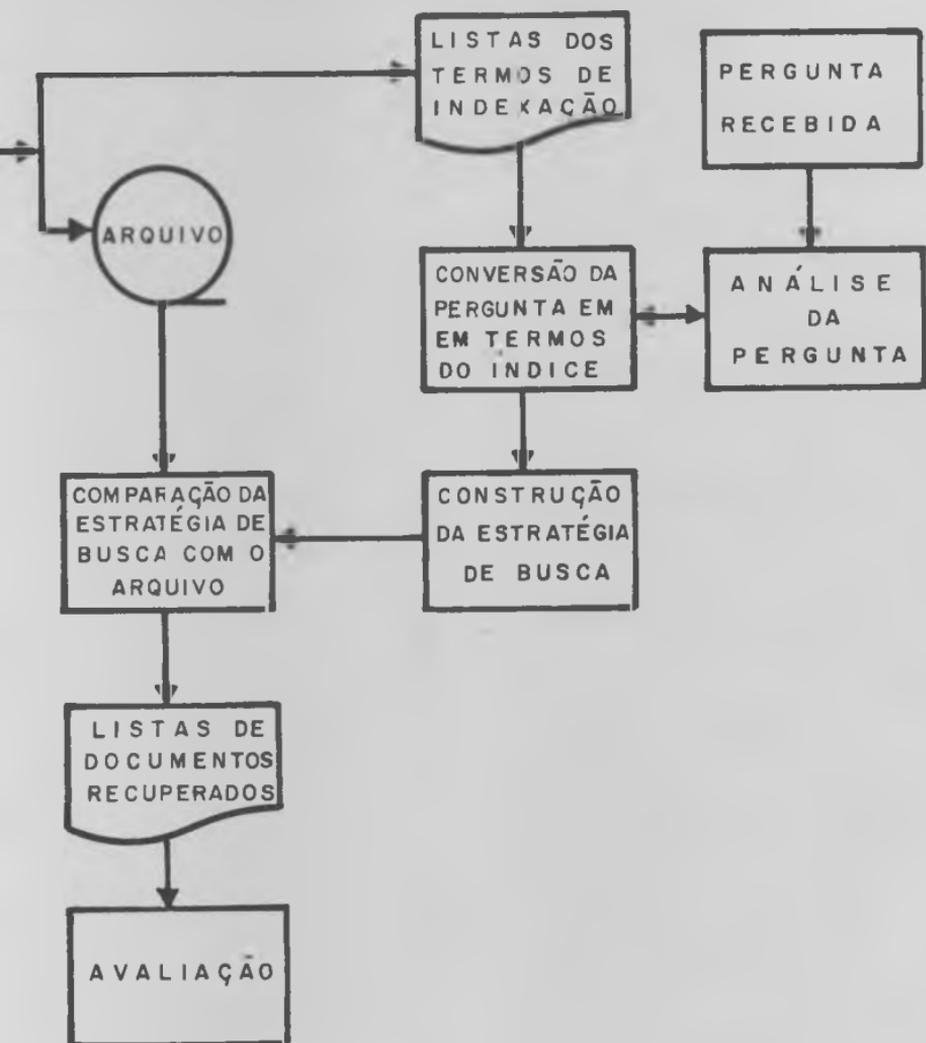


Fig. 1 - FLUXOGRAMA



DA METODOLOGIA

DOC

001

A

RAPOPORT ANATOL

T

WHAT IS INFORMATION?, ETC: A REVIEW OR GENERAL SEMANTICS

D

5304

AB

THE MORE DISORDERED SOMETHING IS, THE MORE INFORMATION IS NEEDED TO MAKE IT KNOWN. THUS, INFORMATION IS THE CARRIER OF ORDER AND IT CAN BE QUANTITATIVELY EQUATED TO THE PROCESSES INVOLVED IN ORDERING PORTIONS OF THE WORLD. THE WORTH OF INFORMATION CAN BE DESCRIBED THROUGH QUANTITATIVE AND PRECISE MATHEMATICAL METHODS. THE FORMULA DEVELOPED TO SHOW AMOUNT OF INFORMATION CONVEYED IS THE SAME FORMULA FOR ENTROPY WHICH MEASURES THE EQUALIZATION OF TEMPERATURES THROUGHOUT A SYSTEM. STUDY OF THIS FORMULA SHOWS THAT ENTROPY AND INFORMATION DESCRIBE SIMILARLY STRUCTURED EVENTS. LIKEWISE, LIFE AND EVOLUTION DEPEND ON THIS SAME ORDERING PROCESS. INFORMATION THEORY HOLDS THE PROMISE OF A PRECISE LANGUAGE TO DESCRIBE THESE AND OTHER PROCESSES INVOLVING THE EMERGENCE OF ORDER FROM CHAOS.

IS1

INFORMATION, SELECTION, MEASURE, COMMUNICATION, MATHEMATICAL, THEORY.

IS2

INFORMATION THEORY8, PROBABILITY THEORY10, MATHEMATICAL ANALYSIS10, COMMUNICATION THEORY8, REDUNDANCY7 LINGUISTICS10.

IS3

INFORMATION THEORY, COMMUNICATION PROBLEMS, COMPUTATIONAL LINGUISTICS, PROBABILITY THEORY, REDUNDANCY, SEMANTICS.

IS4

RESEARCH TECHNIQUES, COMMUNICATION, CODING SYSTEMS, INFORMATION THEORY, DOCUMENTS-SELECTION, CONTENT ANALYSES.

IS5

INFORMATION SCIENCE, COMMUNICATION, STORAGE AND RETRIEVAL, NOISE, REDUNDANCY, CODING.

FREQUENCY WORD/PHRASE	FREQUENCY WORD/PHRASE
2 ABSTRACT	5 SEMANTIC
2 ABSTRACTING	5 THEORY
1 ACQUISITION	4 AUTOMATION
1 ALGEBRA	4 CITATION
2 ANALYSIS	4 EFFECTIVENESS
1 ANSWERING	4 LITERATURE
1 APPLICATION	4 PAPER
1 ARRANGEMENT	4 SCIENTIFIC
2 ASSOCIATION	4 STATISTICAL
1 ASSOCIATIVE	4 SYSTEM
1 AUTHORS	4 SYSTEMS
4 AUTOMATION	3 CLASSIFICATION
1 AVAILABILITY	3 CONSISTENCY
1 AVERAGE	3 EFFICIENCY
1 BEHAVIOUR	3 FLOW
1 BIBLIOGRAPHIC	3 LIBRARY
1 BIBLIOGRAPHIC CONTROL	3 MATHEMATICAL
1 BIBLIOGRAPHIE	3 PERFORMANCE
1 BIBLIOGRAPHY	3 PRECISION
1 BIBLIOMETRIC	2 RECALL
1 BIBLIOMETRICS	3 RELEVANCE
2 BOOK	3 RESEARCH
1 BOOLEAN	3 SEARCH
2 BRADFORD	3 TEST
1 BRONCHING	3 USER
1 CALCULUS	2 ABSTRACT
1 CHAINING	2 ABSTRACTING
1 CHANNEL	2 ANALYSIS
4 CITATION	2 ASSOCIATION
1 CITATIONS	2 BOOK
3 CLASSIFICATION	2 BRADFORD
2 CLUSTER	2 CLUSTER
1 CODING	2 CONTROL
1 COGNIZANCE	2 DECISION
1 COLLECTION	2 FILE
1 COLLECTIVE	2 FREQUENCY
5 COMMUNICATION	2 FUNCTION
1 COMPILATION	2 GROWTH
8 COMPUTER	2 DECISION
3 CONSISTENCY	1 DECISION-MAKING
1 CONTEXT	1 DESCRIPTIONS
2 CONTROL	1 DESCRIPTOR
1 COORDINATE	1 DESCRIPTORS
1 COORDINATION	1 DIMENSIONS
1 COST	1 DISCREPANCY
1 CRANFIELD	1 DISPERSION
1 CURRENT AWARENESS	5 DISTRIBUTION
1 CRITERIA	1 DISTRIBUTIONS
15 INDEXING	2 INDEX
9 INFORMATION	2 JOURNAL
9 RETRIEVAL	2 LAW
9 COMPUTER	2 METHOD
8 MEASURE	2 METHODOLOGY
6 EVALUATION	2 NETWORK
6 LANGUAGE	2 NUCLEUS
6 MODEL	2 PROCESSING
5 COMMUNICATION	2 PRODUCTIVITY
5 DISTRIBUTION	2 QUESTION

FIG. 3 — Lista dos termos utilizados na indexação: à esquerda ordem alfabética e à direita ordem numérica.

Foram elaboradas doze perguntas, tendo-se em vista que o livro-base está dividido em doze capítulos englobando textos de assuntos afins. O conjunto ideal de respostas para cada pergunta correspondia aos artigos pertencentes ao capítulo para o qual a pergunta havia sido formulada. As perguntas usadas para o teste são mostradas na figura 4.

O conjunto ideal para cada pergunta é o seguinte:

<i>Pergunta</i>	<i>Artigos que formam o conjunto ideal</i>	<i>Total</i>
1	1, 2, 3, 4, 5	5
2	6, 7, 8, 9, 10, 11, 12, 13,	8
3	14	1
4	15, 16, 17,	3
5	18, 19,	2
6	20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39,	20
7	40, 41, 42,	3
8	43, 44, 45, 46,	4
9	47, 48, 49,	3
10	50, 51, 52, 53,	4
11	54, 55, 56, 57, 58, 59, 60, 61, 62,	9
12	63, 64, 65,	3

Twelve questions for negotiation and analysis

1. I notice that people from several disciplines call themselves information scientists. What do people in this area mean by information?
2. What exactly have information scientists meant when they spoke of the behavior, or growth, of a literature as representative of the growth of knowledge? Isn't knowledge a psychological phenomena in the users of information and not necessarily connected to literature at all?
3. What has been meant by the term relevance in information science?
4. Can systems analysis or simulation tell us anything about the models or structural elements which are common to information retrieval systems?
5. Have information scientists shown concern for acquisition and selection processes in libraries and special collections?
6. What do you have on methods, theories, approaches to the representation of information for retrieval systems? I am not interested in problems arising from human inefficiency.
7. What properties of bibliographic records should I consider in organizing a file for efficient searching?
8. I need samples of the sort of work information scientists are doing in question analysis, search strategy, search procedures, and other such problems.
9. Are the hyperbolic distributions which characterize such empirical formulations as Bradford's law of scatter a fruitful approach to the study of existing media of knowledge dissemination?
10. What effect do human operating procedures have on the indexing process?
11. I need everything you have on testing and evaluation of information storage and retrieval systems.
12. What does it cost to operate I.R. systems? How can I study the effect of the growth of my file on costs?

FIG. 4 — Lista das 12 perguntas usadas para o teste.

O passo seguinte na metodologia, após a indexação, foi a análise da pergunta. O processo de discussão da pergunta consistiu em se destacar todos os conceitos-chave relacionados com as necessidades do usuário. Assim, aqueles conceitos foram traduzidos em termos de indexação (selecionando-se os termos de indexação utilizados em cada linguagem), a fim de se poder construir uma estratégia de busca para cada pergunta. Para cada questão foi feita uma abordagem genérica e outra específica. Cada estratégia de busca foi elaborada levando-se em consideração dois fatores:

- a) a frequência de uso do termo;
- b) o conjunto ideal de documentos a serem recuperados.

Os conceitos-chave revelados na discussão não foram simplesmente traduzidos um a um para o vocabulário utilizado, correspondente às cinco linguagens de indexação. Os termos que comporiam a estratégia de busca genérica ou específica correspondente a cada questão, em cada linguagem, foram escolhidos levando-se em consideração os dois fatores acima mencionados. As estratégias de busca foram elaboradas com operadores Booleanos (figura 5):

Operador "OU" na listagem o símbolo usado foi /

Operador "E" na listagem o símbolo usado foi &

Operador "NÃO" na listagem o símbolo usado foi ¬

A estratégia de busca para cada linguagem de indexação foi comparada com o campo correspondente armazenado no computador. Foram recuperados como respostas pertinentes a cada questão os artigos que, pelo seu conjunto de seis termos de indexação em cada

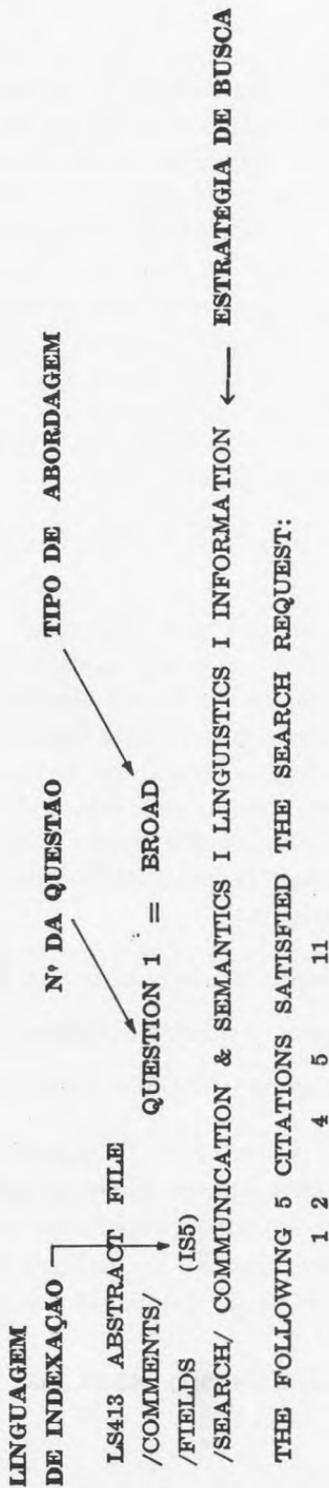


FIG. 5 — Exemplo de estratégia de busca para a pergunta 1 (genérica)

linguagem, satisfizessem a estratégia de busca elaborada. Feita essa comparação, o computador produziu listagens correspondentes aos resultados das buscas relativas às doze questões propostas, nas cinco linguagens em estudo. Com base nessas listagens, pudemos identificar o conjunto de documentos relevantes recuperados ou não pelo sistema, bem como aqueles documentos não relevantes recuperados ou não.

Em seguida, foram aplicadas a esses dados as medidas de sensibilidade e especificidade, para determinar, então, a eficácia de cada linguagem. Para calcular esses dados usamos a seguinte tabela 2 x 2.

	RELEVANTE	NAO RELEVANTE
recuperado	a) relevante encontrado.	b) não relevante encontrado.
não recuperado	c) relevante não encontrado.	d) não relevante não encontrado.

Eficácia (Ef) é a medida que permite ao sistema desempenhar a tarefa para a qual foi elaborado. Pode ainda ser definida como a soma da sensibilidade (Se) e da especificidade (Sp).

Sensibilidade mede a capacidade do sistema em fornecer ao usuário documentos relevantes.

Especificidade mede a capacidade do sistema em não fornecer ao usuário documentos irrelevantes.

As medidas de "Se" e "Sp" são definidas como funções de quatro variáveis apresentadas na tabela

2 x 2, onde a soma das variáveis — “a”, “b”, “c”, “d” — equivalem ao número total de documentos no arquivo.

Assim,

$$\text{Se} = \frac{a}{a + c} \quad \begin{array}{l} \text{a extensão de Se é de 0 a 1:} \\ \text{nada relevante} = 0 \\ \text{todos relevantes} = 1 \end{array}$$

$$\text{Sp} = \frac{d}{b + d} \quad \begin{array}{l} \text{a extensão de Sp é de 0 a 1:} \\ \text{nada irrelevante} = 1 \\ \text{todos irrelevantes} = 0 \end{array}$$

$$\text{Ef} = \text{Se} + \text{Sp} \quad \begin{array}{l} -1 \text{ a extensão de Ef de } -1 \text{ a } +1: \\ \text{todos irrelevantes e somente irrelevantes} = -1 \\ \text{todos relevantes e somente relevantes} = +1 \end{array}$$

Por causa de problemas técnicos no Centro onde os dados foram processados, não puderam ser obtidos os resultados referentes às perguntas 11 e 12 usando o LLSHL e suas estratégias de busca. Desse modo, foram levados em consideração apenas os resultados concernentes às perguntas de 1 a 10, em todas as linguagens estudadas. Portanto, os resultados das perguntas 11 e 12 nas outras quatro linguagens não influenciaram a avaliação total.

3. ANÁLISE DOS RESULTADOS

Unitermo é uma linguagem de indexação onde a escolha dos termos é livre. Os termos são escolhidos diretamente dos textos dos documentos. O LJC, o

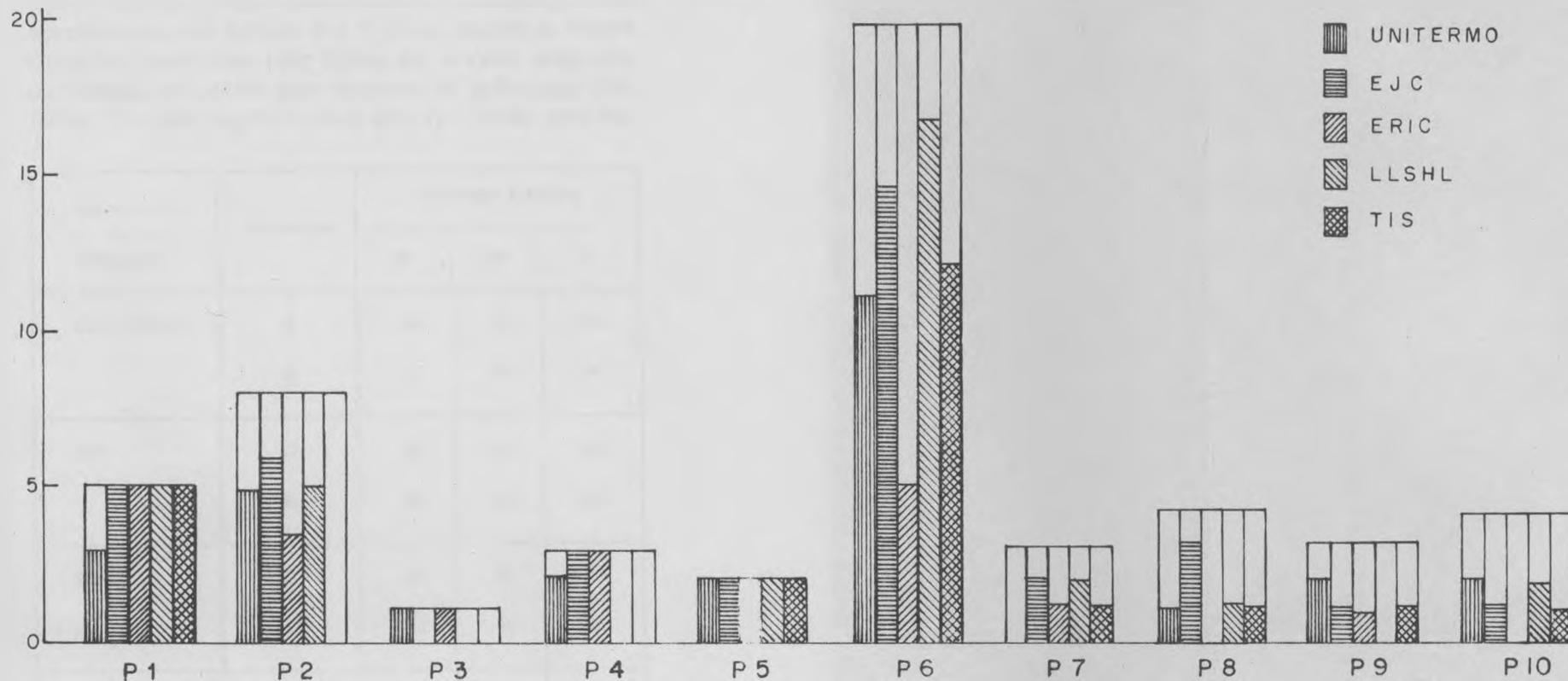


Fig. 6 - GRÁFICO ILUSTRANDO NÚMERO DE DOCUMENTOS RELEVANTES RECUPERADOS PELAS LINGUAGENS DE INDEXAÇÃO

ERIC, o LLSHL e TIS são vocabulários controlados, que permitem ao indexador usar somente os termos neles incluídos. Os resultados da mensuração são apresentados nas figuras 6 e 7. Como mostra a Tabela Geral de Resultados (ver figura 8), o valor mais alto de *eficácia* foi obtido pelo Sistema de Indexação Unitermo. O valor seguinte mais alto foi obtido pela lin-

Linguagem de Indexação	Abordagem	VALORES MEDIOS		
		Se	Sp	Ef
UNITERMO	G	.63	.91	.54
	E	.39	.97	.37
EJC	G	.57	.86	.50
	E	.46	.93	.40
ERIC	G	.41	.91	.37
	E	.41	.96	.38
LLSHL	G	.46	.90	.36
	E	.40	.94	.34
TIS	G	.35	.93	.29
	E	.26	.96	.22

Fig. 7 — Tabela de medidas de acordo com a abordagem da questão (genérica e específica)

Linguagem de Indexação	VALORES MÉDIOS		
	Se	Sp	E1
UNITERMO	.51	.94	.46
EJC	.52	.89	.45
ERIC	.41	.94	.37
LLSHL	.43	.92	.35
TIS	.30	.95	.26

Fig. 8 — Tabela Geral de Resultados por Linguagem de Indexação

guagem de recuperação EJC, com pequena diferença entre as duas linguagens.

Comparando-se esses resultados, podemos sugerir como conclusão a mesma de Keen (2) numa série de testes em que dizia: "As linguagens de indexação de vocabulário não controlado têm um desempenho tão bom quanto as linguagens de indexação de vocabulário controlado, provendo em ambos os casos boa consistência na recuperação". Podemos justificar o bom desempenho do Unitermo e do EJC considerando que a terminologia utilizada na maioria dos artigos vem de diferentes áreas do conhecimento, entre elas a área de engenharia. Assim, termos como teoria da informação, computadores, modelos, simulação, eficiência etc., tornaram-se comumente utilizados na literatura

de Ciência da Informação — e portanto representados pelos Unitermos — e na área de engenharia. Esta é a provável razão pela qual o desempenho do Unitermo e do EJC foram tão similares.

Na avaliação total da experiência podemos salientiar alguns fatores principais que afetaram o desempenho das linguagens de indexação. Esses fatores podem ocorrer em qualquer ponto do processo de recuperação da informação: na indexação, na análise da pergunta, na elaboração da estratégia de busca, no processamento eletrônico.

Podemos identificar alguns erros através da análise das perguntas. A pergunta 1, abordada no sentido genérico, por exemplo, é uma ilustração patente do equilíbrio existente, em termos de desempenho, nas cinco linguagens de indexação (ver figura 9). Exatamente nessa pergunta, o Unitermo teve o pior desempenho devido a uma estratégia de busca, utilizando operação de intersecção (&) o que ocasionou alguma restrição no processo de recuperação. O mesmo fenômeno ocorreu com a LLSHL, que não recuperou o documento de número 4, e com TIS em relação ao documento de número 2. O mesmo pode ser dito sobre a pergunta 4, em relação a LLSHL e TIS. Podemos concluir portanto, que o uso de estratégias utilizando operações de intersecção (&) algumas vezes restringe a sensibilidade e aumenta a especificidade.

Um exemplo de como o homem, através do processo decisório, afeta a recuperação e, conseqüentemente, o desempenho do sistema, pode ser ilustrado pela análise tomada na abordagem genérica da pergunta 2. A estratégia de busca elaborada para recuperação, utilizando TIS, foi:

(Communication/Information) & (Theory/Books/Periodicals / Growth / Behaviour). Um total de oito

documentos formam o conjunto ideal dessa pergunta. A formulação da pergunta exigia duas condições iniciais:

- a) para ser recuperado, o documento deveria ter um dos seguintes termos: "Communication" ou "Information";
- b) satisfeita a condição (a), o documento deveria incluir "Theory" ou "Books" ou "Periodicals" ou "Growth" ou "Behaviour".

Somente três dos oito documentos do conjunto ideal satisfizeram a condição (a), mas nenhum dos três satisfez a condição (b). Portanto, nenhum documento foi recuperado na pergunta 2, na abordagem genérica. Para atingir uma recuperação ótima, deveria ter sido adotada uma das seguintes alternativas:

- a) na fase da indexação as palavras "Communication" ou "Information" deveriam ter sido usadas;
- b) na fase de elaboração da estratégia de busca as palavras "Communication" ou "Information" não deveriam ter sido usadas, ou a formulação feita com os operadores Booleanos deveria ter sido menos restritiva.

Ambos os casos supra-citados envolviam decisão humana, o que nos leva a crer que seja este um dos maiores fatores a influir no desempenho de qualquer sistema de recuperação de informação.

Uma outra importante consideração que gostaríamos de fazer diz respeito à pergunta 3, principalmente, na medida de sensibilidade. O conjunto ideal era composto de somente um documento. Portanto, recuperá-lo em cada uma das linguagens, significava ter a sensibilidade igual a 1; não recuperá-lo significava ter a sensibilidade igual a zero. Nenhuma variação

TERMOS DE INDEXAÇÃO



Nº DOC.	LINGUAGEM DE INDEXAÇÃO	ESTRATÉGIA DE BUSCA (Genérica)	AUTOMATION	CALCULUS	CHANNEL	CODIFICATION	CODING	CODING SYSTEMS	COMPUTATIONAL LINGUISTICS	COMPUTER	COMMUNICATION	COMMUNICATION-EVALUATION	COMMUNICATION-TERMINOLOGY	COMMUNICATION-PROBLEMS	COMMUNICATION THEORY	CONTENT ANALYSIS	DATA	DATA PROCESSING	STATISTICAL MECHANICS
1	IS 1	(INFORMATION COMMUNICATION THEORY) & (SEMANTIC MATH)									•								
	IS 2	INFORMATION THEORY													•				
	IS 3	INFORMATION THEORY (SEMANTICS & LINGUISTICS) COMMUNICATION							•					•					
	IS 4	INFORMATION THEORY & COMMUNICATION						•			•				•				
	IS 5	COMMUNICATION & (SEMANTICS LINGUISTICS INFORMATION)					•				•								
2	IS 1	(INFORMATION COMMUNICATION THEORY) & (SEMANTIC MATH)			•						•								
	IS 2	INFORMATION THEORY																	
	IS 3	INFORMATION THEORY (SEMANTICS & LINGUISTICS) COMMUNICATION																	
	IS 4	INFORMATION THEORY & COMMUNICATION									•					•			
	IS 5	COMMUNICATION & (SEMANTICS LINGUISTICS INFORMATION)									•			•			•		
3	IS 1	(INFORMATION COMMUNICATION THEORY) & (SEMANTIC MATH)																	
	IS 2	INFORMATION THEORY													•				
	IS 3	INFORMATION THEORY (SEMANTICS & LINGUISTICS) COMMUNICATION																	
	IS 4	INFORMATION THEORY & COMMUNICATION									•					•			
	IS 5	COMMUNICATION & (SEMANTICS LINGUISTICS INFORMATION)									•								
4	IS 1	(INFORMATION COMMUNICATION THEORY) & (SEMANTIC MATH)									•								
	IS 2	INFORMATION THEORY			•						•								
	IS 3	INFORMATION THEORY (SEMANTICS & LINGUISTICS) COMMUNICATION													•				
	IS 4	INFORMATION THEORY & COMMUNICATION									•	•	•						
	IS 5	COMMUNICATION & (SEMANTICS LINGUISTICS INFORMATION)									•								
5	IS 1	(INFORMATION COMMUNICATION THEORY) & (SEMANTIC MATH)																	
	IS 2	INFORMATION THEORY	•																
	IS 3	INFORMATION THEORY (SEMANTICS & LINGUISTICS) COMMUNICATION	•			•													
	IS 4	INFORMATION THEORY & COMMUNICATION							•										
	IS 5	COMMUNICATION & (SEMANTICS LINGUISTICS INFORMATION)									•		•						•

Fig. 9 ANÁLIS

na sensibilidade era possível. O conceito-chave envolvido nesta pergunta era "relevance", e somente o Unitermo e o ERIC recuperaram este documento. É interessante analisar-se o que ocorreu com o EJC, o LLSHL e o TIS.

No EJC o termo "relevance" foi utilizado na fase de indexação e omitido na estratégia de busca: a causa do mau desempenho foi, portanto, uma falha no processo decisório de elaboração da estratégia de busca.

Na LLSHL o documento foi corretamente indexado como "Information Retrieval-Evaluation". A estratégia de busca foi "Information Retrieval Evaluation", tendo se omitido o "-" por erro de processamento, o que prejudicou a recuperação.

No TIS a estratégia de busca utilizada foi "Evaluation & Effectiveness". Se o operador Booleano "ou" tivesse sido usado em lugar do "e", o resultado teria sido diferente.

Em relação à pergunta 5, o desempenho do ERIC foi grandemente afetado por um erro de processamento: os termos relacionados no campo "IS3" foram misturados com os termos do campo "IS2". Assim, o computador não pôde identificar o campo "IS3" e, portanto, não fez a comparação. Um fato curioso é que este problema ocorreu somente com os documentos de número 18 e 19, que correspondiam, exatamente, ao conjunto ideal de documentos para recuperação. Se esse problema não tivesse ocorrido, os dois documentos teriam sido recuperados, o que elevaria os valores médios da sensibilidade e da *eficácia* para o ERIC.

4. CONCLUSÕES

Os resultados obtidos através da análise e avaliação da experiência nos levaram às seguintes conclusões:

- a) a linguagem de indexação não controlada tem um desempenho tão bom quanto a controlada;
- b) a medida de sensibilidade parece ter melhor desempenho quando se usa abordagem genérica na elaboração da estratégia de busca. Era de se esperar esse resultado, uma vez que ao se usar estratégia genérica o sistema recupera maior número de documentos, existindo, portanto, maior probabilidade de se recuperar documentos relevantes;
- c) a medida de especificidade parece ter melhor desempenho quando aplicada à estratégia de busca na abordagem específica: isto é compreensível uma vez que, usando a abordagem específica na estratégia de busca, tenta-se recuperar somente os documentos relevantes e, portanto, há menor probabilidade de que sejam recuperados documentos irrelevantes;
- d) esperava-se que o TIS, devotado à Ciência da Informação, e o LLSHL devotado à Biblioteconomia, tivessem os melhores desempenhos. Entretanto, isso não aconteceu e talvez possa ser explicado pelo fato de que o LLSHL foi publicado em 1963, quando os conceitos-chave envolvidos nos 65 artigos do livro base não eram ainda usualmente utilizados em Biblioteconomia. O TIS foi publicado em 1968 e igualmente não teve bom desempenho, provavelmente, porque os conceitos-chave que possui não são tão específicos como seria desejável;
- e) podemos atribuir alguns erros de recuperação a falhas humanas nas fases de indexação e elaboração da estratégia de busca;

- f) outra consideração a ser feita diz respeito ao uso de “e” (intersecção) na estratégia de busca, o que afetou a sensibilidade, embora tivesse sido necessário para a especificidade;
- g) os resultados também podem ter sido influenciados pela falta de padronização (gênero, número etc.) dos termos no momento da conversão em linguagem de máquina;
- h) podemos ainda salientar a dificuldade que existe em se elaborar estratégia de busca com os termos de uma simples listagem, pois nesta o pesquisador não tem acesso às relações sintáticas e de sinonímia que os “thesauri” possuem. Essas relações orientam o indexador e, provavelmente, seriam úteis para ajudar o pesquisador a elaborar melhor sua estratégia de busca;
- i) uma vez que o sistema Unitermo conseguiu o maior valor de *eficácia* pode ser considerado como a melhor linguagem de indexação, dentre as analisadas na presente experiência.

SUMMARY

Practical experience using sensibility and specificity measures in order to compare five different indexing languages and evaluate their effectiveness.

REFERÊNCIAS BIBLIOGRÁFICAS

1. SARACEVIC, Tefko, ed. *Introduction to information science*. New York, R. R. Bowker, 1970, 751 p.
2. KEEN, M. & DIGGER, J. *Report on an information science index language*. Aberystwyth, Wales, College of Librarianship, 1972.