

Regina Célia Figuciredo

ESTUDO COMPARATIVO DE JULGAMENTOS
DE RELEVÂNCIA DO USUÁRIO E NÃO-USUÁRIO DE
SERVIÇOS DE DISSEMINAÇÃO SELETIVA DA INFORMAÇÃO

Dissertação apresentada ao Instituto
Brasileiro de Informação em Ciência e
Tecnologia para obtenção do Título
de Mestre em Ciência da Informação.

São Paulo

1978

AGRADECIMENTOS

Ao Dr. Tefko Saracevic, pela orientação.

A Gilda Maria Braga, pelas sugestões, leitura e crítica desta dissertação.

A Terezine Arantes Ferraz, pela dedicação, incentivo e apoio para a realização deste trabalho, pela leitura criteriosa do manuscrito e pelo importante papel que teve em minha formação.

Ao Prof. Romulo Ribeiro Pieroni, Superintendente do Instituto de Energia Atômica, pelas facilidades concedidas na elaboração deste trabalho.

Ao Prof. Dr. Rui Ribeiro Franco, pelas sugestões.

A todos do Centro de Informações Nucleares pela oportunidade de realização deste trabalho.

A todos que, direta ou indiretamente, me incentivaram e colaboraram para a concretização deste trabalho.

Resumo

Os julgamentos de relevância, atribuídos a um grupo de documentos da área de energia nuclear, por usuários de um serviço de disseminação seletiva da informação foram comparados àqueles atribuídos por um não usuário.

A avaliação de relevância foi feita pelos títulos dos documentos. Observou-se também a utilidade dos títulos como determinantes da relevância ou não-relevância dos documentos.

Concluiu-se que os julgamentos dos usuários e do não usuário foram similares, sendo aceitável a porcentagem de discordâncias. Considerou-se que a porcentagem de concordâncias não foi maior devido à diferença de pontos de vista entre usuário e não-usuário: enquanto o primeiro avalia o documento para sua necessidade de informação, o segundo o avalia em função da questão proposta.

Verificou-se que a frequência de descritores do perfil nos títulos dos documentos não discrimina, por si só, os documentos relevantes dos não-relevantes.

Abstract

User and non-user relevance judgments, of a group of documents in the field of nuclear energy, were compared.

Assessments of relevance were done by document titles. The effectiveness of titles in determining relevance or non-relevance of documents was observed.

It was concluded that user and non-user judgements were similar, with an acceptable number of disagreements. It was considered that the level of agreements was not higher because users evaluate documents to their information needs, while non-users evaluate them to the proposed question.

It was stated that the frequency of profile descriptors in document titles does not differentiate relevant documents from non-relevant ones.

SUMÁRIO

1 - INTRODUÇÃO.....	1
1.1 - Hipóteses.....	4
1.2 - Definição dos conceitos.....	5
Notas.....	7
2 - REVISÃO DA LITERATURA.....	8
2.1 - Relevância e pertinência.....	9
2.2 - Experimentos com julgamentos de relevância	20
2.3 - Métodos e medidas.....	30
Notas.....	44
Referências bibliográficas.....	45
3 - METODOLOGIA.....	49
3.1 - Operacionalização dos conceitos.....	49
3.2 - Coleta de dados.....	52
Notas.....	56
4 - COMPARAÇÃO ENTRE JULGAMENTOS DE RELEVÂNCIA.....	57
4.1 - Análise dos dados e resultados.....	57
4.2 - Cálculo das medidas propostas.....	67
4.3 - Considerações sobre os resultados.....	71
Referências bibliográficas.....	75
5 - ANÁLISE SEMÂNTICA.....	76
5.1 - Análise dos documentos.....	76
5.2 - Resultados.....	79
5.3 - Considerações sobre os resultados.....	92
Notas.....	95
Referências bibliográficas.....	97

6 - CONCLUSÕES..... 98

Notas..... 103

7 - LISTA DAS REFERÊNCIAS BIBLIOGRÁFICAS..... 104

ANEXOS

Anexo I - Lista dos perfis do SDI manual introduzidos no SDI automatizado.....

Anexo II - Descritores dos perfis.....

1 INTRODUÇÃO

Um serviço de disseminação seletiva da informação (em inglês, selective dissemination of information, SDI ¹) tem por objetivo a recuperação e disseminação periódica e seletiva de informações, em áreas de interesses específicos dos usuários.

Um serviço de SDI tanto pode ser manual como automatizado. Cabe às instituições patrocinadoras dos serviços analisar, de acordo com seus objetivos, condições físicas e materiais, as vantagens e desvantagens de cada uma dessas modalidades, antes de implantar um serviço de SDI.

A sistemática de execução de um serviço de SDI é basicamente a mesma, seja ele automatizado ou manual. Inicialmente, o usuário define seu perfil. Entende-se por perfil o conjunto de dados que caracte-

terizam o interesse do usuário. Esse conjunto de dados compreende, entre outros, a descrição do assunto, a determinação de áreas de interesse, a seleção de descritores e fontes de referência a serem utilizadas etc. Esses dados variam de um serviço a outro, dependendo das características individuais estabelecidas para cada serviço.

Depois de definido o perfil, são analisadas, manual ou automaticamente, as fontes de referência selecionadas ou disponíveis na instituição, conforme o caso. Toda informação encontrada sobre os assuntos dos perfis é recuperada. As informações são comumente enviadas ao usuário em uma destas três formas: apenas a referência bibliográfica, a referência bibliográfica a acompanhada do resumo e, finalmente, a informação completa, ou seja, cópia do documento original na íntegra.

Ao receber a informação recuperada, o usuário avalia essa informação de acordo com sua importância como resposta à necessidade de informação expressa na questão. O usuário faz nesse momento um julgamento de relevância², isto é, avaliação da informação recuperada em relação à questão por ele proposta.

Na área nuclear existe um sistema de informação, International Nuclear Information System- INIS, patrocinado pela International Atomic Energy Agency

(IAEA). Os países membros da IAEA, participantes do INIS, coletam e promovem a entrada de dados ("input") referentes à literatura nuclear produzida em cada país. A IAEA processa essa informação e a divulga sob as formas de bibliografia impressa - o INIS Atomindex - e fitas magnéticas. As fitas são acessíveis aos países membros que desejem recuperar e disseminar as informações contidas nessas fitas, em bases nacionais.

No Brasil, a Comissão Nacional de Energia Nuclear (CNEN), através de seu Centro de Informações Nucleares (CIN), desenvolveu um serviço de SDI automatizado, utilizando as fitas magnéticas do INIS. As fitas são recebidas e processadas quinzenalmente e a informação recuperada é enviada periodicamente aos usuários do sistema.

Ainda no campo da energia nuclear, foi implantado um serviço manual de SDI pela Divisão de Informação e Documentação Científicas (DIOC), do Instituto de Energia Atômica (IEA), de São Paulo. Esse serviço abrange as mesmas áreas de interesse atendidas pelo SDI do CIN, sendo dirigido, entretanto, somente ao atendimento das pesquisas prioritárias sendo desenvolvidas pela comunidade científica do IEA.

A bibliotecária encarregada do SDI

manual do IEA teve oportunidade, em determinada época, de participar das atividades desenvolvidas pelo serviço de SDI automatizado do CIN.

Em consequência da identidade de propósitos dos dois serviços, os perfis anteriormente atendidos pelo SDI manual do IEA foram incluídos no SDI automatizado do CIN.

Essa inclusão motivou posteriormente a idéia contida no tema central desta dissertação: analisar julgamentos de relevância.

Com esse objetivo em mente, a bibliotecária comparou, durante um período de aproximadamente seis meses, os julgamentos de relevância dos usuários com os seus próprios (não-usuário), atribuídos a um mesmo grupo de documentos.

1.1 Hipóteses

Foram formuladas as seguintes hipóteses para o estudo:

1) Não há diferença entre os julgamentos de relevância emitidos pelos usuários e aqueles

emitidos por um não-usuário, desde que este último tenha conhecimentos básicos dos assuntos cobertos pelo SDI e familiaridade com o enfoque das questões propostas.

2) São relevantes documentos cujos títulos contêm mais de dois descritores mencionados no perfil e, não-relevantes aqueles cujos títulos não contêm nenhum dos descritores do perfil.

1.2 Definição dos conceitos

Os conceitos empregados neste trabalho, assim como foram propostos por diversos autores na literatura sobre o assunto, serão apresentados no item 2.

Para o desenvolvimento desta dissertação, considerou-se necessário precisar esses conceitos de acordo com o contexto em que foram utilizados.

Por relevância entendeu-se a relação entre os documentos e a questão proposta pelo usuário. E por pertinência, a relação entre os documentos e a necessidade de informação do usuário.

Os juizamentos de relevância são

as avaliações dos documentos recuperados sob o ponto de vista do interesse ou não-interesse dos documentos para a questão proposta.

O enfoque dado a esses conceitos nem sempre foi o mesmo encontrado na literatura sobre o assunto. Na revisão da literatura procurar-se-á destacar as divergências entre os conceitos adotados neste trabalho e os apresentados pelos autores em questão.

Notas

1. A sigla SDI será adotada neste trabalho, pois assim aparece na literatura brasileira sobre o assunto.
2. Os conceitos de relevância e de julgamentos de relevância serão apresentados nos itens 1.3 e 2.1.

REVISÃO DA LITERATURA

A revisão da literatura limitou-se a artigos que apresentam aspectos e fatores que influenciam julgamentos de relevância. Algumas experiências sobre comparações entre julgamentos de relevância do usuário e do sistema (ou não-usuário) e estudos sobre fatores que podem determinar os julgamentos de relevância do usuário foram particularmente revisados por estarem diretamente relacionados ao âmbito deste estudo. Devido à dificuldade de medir os julgamentos do usuário, a maior parte dos artigos lidam com medidas de relevância sob o ponto de vista do sistema e tentam aproximá-las a uma medida mais restrita.

A literatura referente aos conceitos de relevância e pertinência também foi analisada.

2.1 Relevância e pertinência

BRADFORD⁽¹⁾ foi o primeiro a usar o termo "relevante", no contexto em que ele é usado hoje em Ciência da Informação, quando falou em artigos "relevantes a um assunto".

Nas décadas de 40 e 50, MOEERS⁽²⁾, TAUBE⁽³⁾ e PERRY⁽⁴⁾ foram os primeiros a falar sobre relevância em relação à recuperação da informação. Eles reconheceram que nem tudo o que seria recuperado pelos sistemas de recuperação da informação seria relevante. Eles se preocuparam mais com a não-relevância, com a recuperação indesejada, do que propriamente com relevância. A não-relevância foi apontada pelos autores como causa de má utilização ou adequação dos sistemas.

Em 1958, durante a "International Conference for Scientific Information" (ICSI)⁽⁵⁾, o conceito de relevância foi discutido e explorado por diversos autores, que concordaram em determinados pontos:

"- relevância não é, exclusivamente, uma propriedade de documentos;

- relevância não é uma propriedade dicotômica;

- existe uma "relevância para o usuário" que deve ser julgada." 1

Várias definições de relevância apareceram na década de 60, principalmente como resultado da afirmativa de que não era muito claro o significado do termo, como afirmou SARACEVIC⁽⁶⁾. Muitas definições eram somente paráfrases de definições anteriores, numa tentativa de apresentar algo mais que hipóteses e fatores envolvidos nas relações entre usuário/documento ou sistema/documentos. As definições estabeleceram-se dentro de um padrão geral: Relevância é o A de um B existente entre um C e um D determinado por um E.² Nessas definições os espaços em branco eram preenchidos com palavras de significado similar, tais como:

A	B	C
medida	correspondência	documento
grau	utilidade	artigo
dimensão	conexão	texto
estimativa	satisfação	referência
relação	significado	fato

D	E
questão	pessoa
pedido de informação	jugador
informação utilizada	usuário
ponto de vista	requisitante
requisição de informação	especialista da informação

No empenho para encontrar uma definição adequada de relevância, e, provavelmente, como consequência desse empenho, vários termos foram propostos por diversos autores para identificar diferentes aspectos de relevância. Pertinência, relevância do sistema, relevância do usuário etc foram termos muitas vezes conceituados com o sentido de relevância.

Ao especificar os fatores, elementos e relações que a noção de relevância compreende, os cientistas da informação desenvolveram muitos pontos de vista sobre relevância, de acordo com as características que eram mais importantes na época em que foram propostos. Assim, de acordo com SARACEVIC⁽⁶⁾, pode-se ter a relevância do ponto de vista do sistema, do destino da informação, da literatura sobre o assunto, do conhecimento do assunto, e por uma visão pragmática. Esses pontos de vista não são mutuamente exclusivos; cada um deles completa alguns aspectos que não haviam sido reconhecidos anteriormente.

O ponto de vista do sistema está relacionado com os aspectos internos do sistema, isto é, indexação, classificação, organização de arquivos etc.

O ponto de vista do destino da informação equiparou relevância com julgamentos de relevância

cia, considerando os fatores que afetam as respostas dos usuários.

O ponto de vista da literatura sobre o assunto baseou-se em considerações sobre a estrutura das literaturas, por exemplo, as redes de citações, aparecimento de artigos nos periódicos, obsolescência, uso da literatura.

O ponto de vista de conhecimento do assunto relacionou-se com a natureza, estrutura e extensão do conhecimento do assunto que estava compreendido na questão. Pode ser encarado em termos da questão ou da bagagem de conhecimentos do usuário.

O ponto de vista pragmático concentrou-se na aplicação imediata ou no problema à mão como justificativa de custo-benefício.

As definições de relevância quase sempre estavam acompanhadas por uma tentativa de distinção entre relevância e pertinência.

POLUSHKIN⁽⁷⁾ afirmou que relevância é usualmente entendida como a característica do grau de correlação entre o conteúdo do documento recuperado e o conteúdo da questão. Relevância pode servir como uma ca

racterísticas da relação semântica entre o documento e a questão.

E, por pertinência, afirma POLUSHKIN, entende-se a característica do grau de correlação entre o conteúdo do documento recuperado e o pedido de informação, em conjunto com as características puramente subjetivas do usuário específico.

SARACEVIC⁽⁶⁾ definiu relevância como uma medida da eficácia do contato entre uma fonte e um destinatário num processo de comunicação. Salientou que o conceito de necessidade de informação deu origem à noção de pertinência. O processo de formulação da questão e resposta obtida à questão foi representado por REES e SARACEVIC⁽⁸⁾ como se vê na Fig. 1.

FOSKETT⁽⁹⁾ sugeriu que a verdadeira distinção é que relevância poderia ser tomada no sentido de pertencente ao campo/assunto/universo do discurso delimitado pelos termos da questão, tal como estabelecido pela opinião geral dos pesquisadores naquele campo; enquanto pertinência poderia ser definida como a adição de novas informações às já existentes na mente do usuário, as quais são úteis para ele no trabalho que o levou a formular a questão.

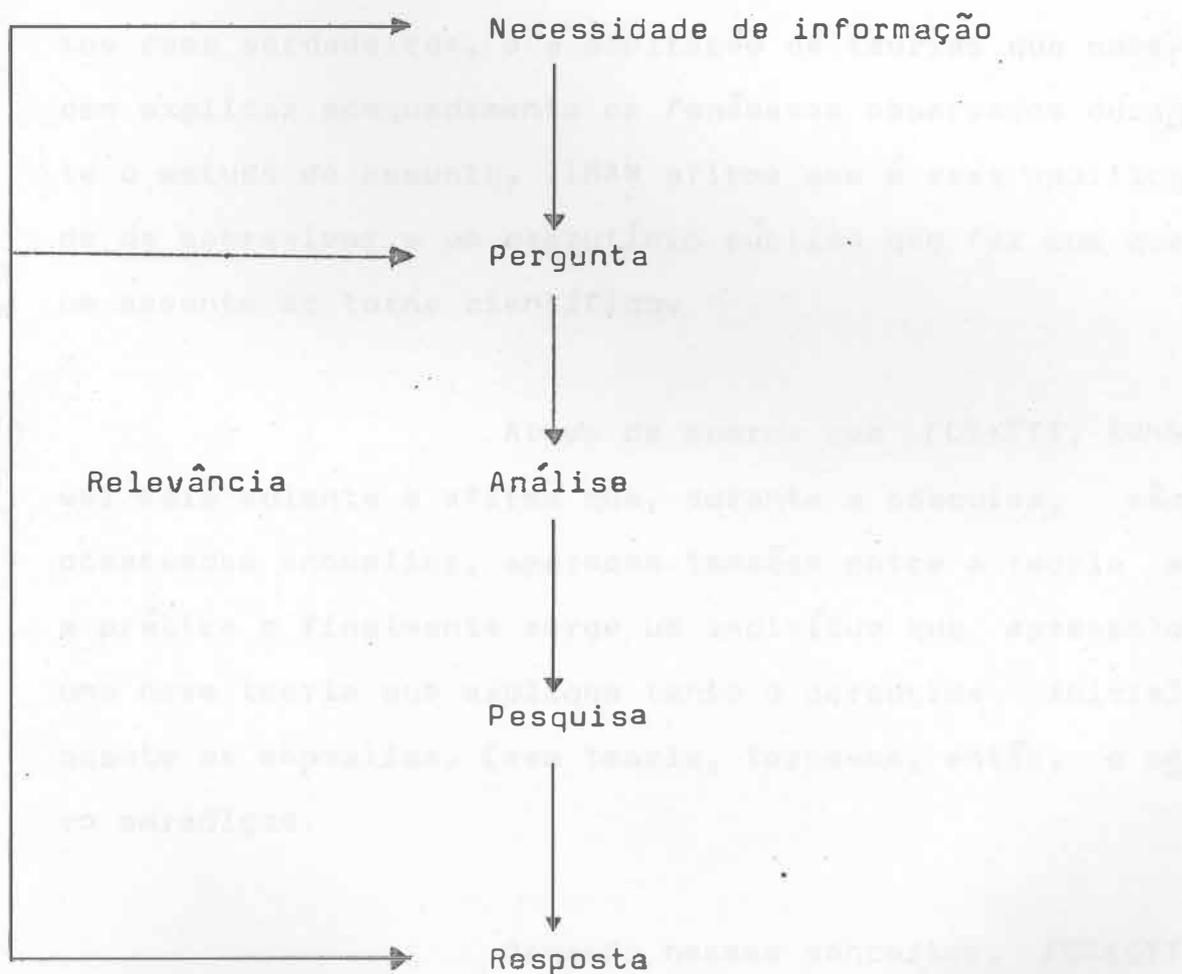


Fig. 1 - Processo de formulação da questão e resposta

FOSKETT baseou suas definições no conceito de "paradigma" de KUHN⁽¹⁰⁾, isto é, padrão de pensamento aceito num determinado campo do conhecimento, e na visão da ciência de ZIMAN⁽¹¹⁾. Segundo apresenta FOSKETT³, tanto KUHN como ZIMAN alegam que em qualquer campo de assunto há um conjunto de dados comumente aceitos como verdadeiros, e a aceitação de teorias que parecem explicar adequadamente os fenômenos observados durante o estudo do assunto. ZIMAN afirma que é essa habilidade de sobreviver a um escrutínio público que faz com que um assunto se torne científico.

Ainda de acordo com FOSKETT, KUHN vai mais adiante e afirma que, durante a pesquisa, são observadas anomalias, aparecem tensões entre a teoria e a prática e finalmente surge um indivíduo que apresenta uma nova teoria que explique tanto o paradigma inicial quanto as anomalias. Essa teoria, torna-se, então, o novo paradigma.

Baseado nesses conceitos, FOSKETT redefiniu relevância como sendo uma parte do paradigma, ou conhecimento público ou consenso em um campo; e pertinência como sendo relacionada com o padrão de pensamento específico na mente de um usuário específico.

LANCASTER⁽¹²⁾ sugeriu também uma

distinção entre relevância e pertinência. Ele aconselha que, em termos de conveniência, se distinguísse pertinência como a relação entre um documento e uma necessidade de informação e relevância como a relação entre o documento e a questão.

Afirma LANCASTER que os julgamentos de pertinência são muito mais transitórios que os julgamentos de relevância. O julgamento de pertinência reflete o valor de um documento, num determinado tempo, para satisfazer a uma determinada necessidade de informação que, ao contrário da formulação da questão, pode mudar facilmente. Os julgamentos de pertinência são influenciados tanto pela passagem do tempo como pela seqüência na qual as decisões são tomadas durante o desenvolvimento do estudo.

Por sua vez, salienta LANCASTER que os julgamentos de relevância podem não refletir o grau de êxito alcançado para satisfazer os usuários. É possível que um documento seja considerado relevante por todos os membros de um júri mas que o usuário decida que ele não tem valor para satisfazer sua necessidade de informação.

HARMON(13) distinguiu relevância do usuário de relevância do sistema. Relevância do usuá-

rio pode ser vista como uma avaliação feita pelo usuário da extensão da relação entre a resposta do sistema e a sua necessidade de informação. A relevância do sistema pode ser vista como a avaliação feita pelo sistema do grau de relação entre a informação contida no sistema e a questão proposta pelo usuário.

A distinção feita por HARMON entre relevância do usuário e relevância do sistema se aproxima dos conceitos de relevância e pertinência expostos nesta dissertação. O que HARMON chamou de relevância do usuário equivale à pertinência, pois se refere à relação entre documento e necessidade de informação. Relevância do sistema equivale a relevância propriamente dita, pois está relacionada à questão proposta pelo usuário.

WILSON⁽¹⁴⁾ tentou explicar relevância de uma maneira diferente. Afirmou que relevância é uma noção altamente geral e vaga que pode se tornar específica e precisa de inúmeras maneiras.

Admitiu que relevância não é sempre a noção correta a ser empregada na avaliação de desempenho dos sistemas de informação. Segundo afirma o autor, relevância não é uma noção única, mas muitas.

Para definir o que chamou de "rele

vância situacional", WILSON distinguiu anteriormente relevância lógica de relevância psicológica. Relevância psicológica está relacionada com os usos e efeitos reais da informação: como as pessoas utilizam a informação, como seus pontos de vista realmente mudam ou deixam de mudar em consequência do recebimento da informação. A relevância lógica está relacionada com os usos ou efeitos somente na medida em que esses se adaptam aos padrões de crítica, somente quando correspondem ao que deveria acontecer ou ser feito. A noção de relevância situacional apresentada por WILSON incorpora duas outras noções distintas de relevância: a relevância lógica de Cooper e a relevância evidencial baseada na lógica indutiva. Relevância situacional é relevância para uma situação individual específica, para a situação como o indivíduo a vê, não como outros a vêem ou como ela "realmente" é. É também relevância em relação ao cabedal de conhecimentos de um indivíduo em particular e muda na medida em que esse cabedal se altera. A relevância situacional foi considerada por WILSON como o produto das relações entre o cabedal de conhecimentos de uma pessoa, suas preferências e seus interesses.

CUADRA e KATTER⁽¹⁵⁾ salientam também o aspecto "situacional" da relevância. Afirmam que há duas maneiras principais nas quais o termo relevância foi usado na literatura: 1) Para indicar a relação

de pontos e que a relevância pode ser influenciada por um sem número de variáveis.

O estudo mostrou que os julgamentos feitos pelos julgadores sem nenhuma orientação e os feitos de acordo com instruções de uso da informação diferiram sensivelmente. Essas instruções de uso representaram fatores e situações que, segundo os autores, afetam os julgamentos de relevância.

Assim como esses pontos de vista, muitos outros surgiram durante o desenvolvimento dos sistemas de informação numa tentativa de melhorar o desempenho e avaliar a eficiência de vários sistemas.

2.1 Experimentos com julgamentos de relevância

Os experimentos sobre relevância foram desenvolvidos de acordo com cinco classes gerais de variáveis (como foi revisto por SARACEVIC⁽⁶⁾): 1) documentos e representações de documentos; 2) questões ("queries"); 3) situações e condições de julgamento; 4) formas de expressão; e 5) características humanas (pessoas).

Os autores dos estudos revisados neste capítulo fixaram sua atenção principalmente em documentos e suas representações e na consistência de julgamentos de relevância, assim como na concordância desses julgamentos na medida em que são afetados por certas características humanas.

O trabalho mais representativo que se teve conhecimento sobre a compreensão dos critérios de escolha empregados pelo usuário foi o de GIFFORD e BAUMANIS⁽¹⁶⁾, onde é descrita uma investigação empírica sobre o papel dos documentos nos julgamentos de relevância. O estudo foi baseado na pergunta: "O que em um documento provoca o julgamento do usuário?" A questão foi diretamente explorada, principalmente pela determinação de diferenças textuais significativas que pudessem ser encontradas entre documentos relevantes e não-relevantes. Os resultados foram encorajadores. Ao examinar as avaliações de 61 questões, os autores encontraram diferenças textuais correspondentes às diferenças de avaliação, as quais eram geralmente diferenças em ocorrência de palavras. Os autores afirmaram que uma declaração dessas diferenças, como propriedades a serem exigidas para resumos relevantes (ou para resumos parcialmente relevantes), sempre constitui uma interpretação significativa e relativamente não-ambígua da questão.

Na amostra examinada pelos autores os documentos relevantes distinguiram-se com regularidade dos não-relevantes. Foram encontradas diferenças textuais, que estavam diretamente ligadas a ocorrências de palavras, que correspondiam a diferenças na escala de avaliação. Essa observação levou os autores a sugerirem que essas diferenças de ocorrências de palavras nos textos provocam os julgamentos dos usuários. Afirmam ainda que os usuários julgam os documentos que respondem a uma questão mediante um padrão relativamente consistente.

Contudo, em relação à noção geral de relevância, ou melhor, ao que torna um documento relevante, nenhuma evidência de atributos gerais foi encontrada. Afirmaram os autores que os documentos relevantes para uma questão específica geralmente têm propriedades significativas em comum, mas isto não se estende de uma questão à outra.

Também relacionado aos documentos foi o experimento realizado por TAGUE⁽¹⁷⁾, que tentou classificar as relações entre as palavras da questão e as palavras usadas para indexar documentos (respostas) relevantes, relativamente relevantes e não-relevantes. No estudo, a avaliação sobre a relevância dos documentos foi feita pelo próprio usuário que formulou a questão.

Especificamente, as seguintes perguntas foram feitas:

1) Os documentos considerados relevantes pelos usuários são mais prováveis de terem sido indexadas pelos termos usados na proposição da questão do que os considerados parcialmente relevantes ou não-relevantes?

2) Os documentos relevantes são mais prováveis de serem indexados por palavras do thesaurus relacionadas às palavras da questão pelo código semântico⁵ do que os documentos parcialmente relevantes ou não-relevantes?

3) Os documentos relevantes são mais prováveis de conterem palavras acrescentadas pelo analista do sistema às palavras da questão que os parcialmente relevantes ou não-relevantes?

4) Quais as diferenças nas distribuições dos três tipos de origem de palavras⁶ e de todas as outras palavras incluídas no programa de busca (isto é, forma da estratégia de busca usada na pesquisa do arquivo) em documentos relevantes, parcialmente relevantes e não-relevantes?

5) Quais são as diferenças na distribuição das palavras da questão em títulos, resumos e índices feitos pela máquina (resumos telegráficos) para respostas relevantes, parcialmente relevantes ou não-relevantes?

Algumas conclusões, dentre elas as seguintes, foram apresentadas pelo autor: 1) As palavras do título são mais efetivas como termos de avaliação; 2) As palavras da questão são mais úteis para localizar respostas relevantes; 3) A frequência de palavras da questão é maior em documentos relevantes que nos parcialmente relevantes ou não-relevantes; 4) O número de palavras, mais do que a relação entre as palavras da questão original, é que caracteriza as respostas relevantes.

Conquanto o principal objetivo do estudo tenha sido o de relacionar propriedades de documentos relevantes e não-relevantes à eficácia de indexação e de diferentes tipos de representação de documentos (títulos, resumos convencionais e resumos telegráficos), as conclusões apresentadas por TAGUE podem ser úteis ao entendimento do "comportamento" dos documentos sob o ponto de vista do usuário.

SARACEVIC, em outro estudo⁽¹⁸⁾, examinou uma variedade de aspectos relacionados com tes-

tes de sistemas de recuperação, entre eles: usuários e proposições de questões, e julgamentos de relevância baseados em diferentes formatos de saída. Cada representação de documentos (títulos, resumos e texto completo) foi julgada em uma escala de três pontos: relevantes, parcialmente relevantes e não-relevantes. Os julgamentos feitos sobre texto completo foram tomados como padrão. O autor mediu relevância em termos de sensibilidade ("sensitivity"), especificidade ("specificity") e efetividade ("effectiveness"). As medidas de avaliação revelaram que há maior concordância sobre o que não é relevante do que sobre o que é relevante.

Salienta o autor que uma comparação entre as indicações de relevância dadas pelo usuário e pelo sistema mostrou que os usuários se saíram consideravelmente melhor do que os sistemas ao reconhecerem respostas não-relevantes (especificidade): os sistemas nem se aproximaram da habilidade do usuário em reconhecer a não-relevância. SARACEVIC declarou que não se deve esperar que um sistema atribua relevância melhor do que um usuário; afirma ser desejável, contudo, que um sistema tenha um desempenho tão bom quanto o do usuário.

Este estudo conduziu SARACEVIC à conclusão de que diferentes formatos têm um efeito signi

ficativo sobre os julgamentos de relevância do usuário.

RATH et alii⁽¹⁹⁾ realizaram um estudo comparativo entre quatro tipos de indicadores léxicos de conteúdo (títulos, dois tipos de resumos e textos).

Tomaram por base que as duas funções de um indicador léxico de conteúdo são: 1) Determinar se um documento é relevante para um propósito específico; e 2) Evidenciar alguma informação do documento sem que seja necessário ler o texto completo.

Os resultados indicaram aos autores que as pessoas são capazes de determinar a "utilidade" de documentos e respostas à questões baseadas neles mais rapidamente do que poderia ser predito com base na velocidade de leitura das pessoas. Concluíram os autores, ainda, que o uso de títulos na busca de documentos sem resumo adicional parecia levar a um alto número de documentos aceitos que poderiam ter sido rejeitados caso houvesse maior informação disponível. Para os autores, não houve diferença significativa entre as avaliações usando textos completos e resumos.

O'CONNOR^(20, 21) conduziu diversos

estudos para explicar discordâncias nos julgamentos de relevância e os fatores que estavam relacionados e/ou causando essas discordâncias. No primeiro estudo, analisou discordâncias em relação aos formulários de requisição de pesquisas bibliográficas. Partiu do princípio que as discordâncias sobre relevância ocorrem porque os julgadores interpretam de forma diferente as requisições de pesquisas, principalmente porque elas são pouco claras. O'CONNOR concorda com a afirmação de MODERS(22) que diz que "qualquer pergunta feita pelo usuário pode ser interpretada de várias maneiras. Dependendo da interpretação, a relevância dos documentos recuperados vai mudar."⁷

Várias espécies de requisições de pesquisas bibliográficas que não estavam claras foram examinadas pelo autor como possíveis causas de discordâncias. Sugeriu ele que os julgadores podem concordar quanto ao significado de uma requisição de pesquisa bibliográfica, mas discordam sobre o fato de um determinado documento ter as características especificadas na requisição (conclusivo, inédito, novo, útil para o estudo etc). O'CONNOR decidiu que as causas básicas de discordâncias sobre relevância são as diferenças de interpretação das requisições ou documentos, e não fatores tais como educação, experiência, função etc dos julgadores, e o que eles consideram como o objetivo, ambiente e momento da propo-

sição da questão.

Em outro estudo⁽¹¹⁾, concordâncias e discordâncias sobre documentos-resposta foram comparadas. Dois julgadores, trabalhando independentemente, com pararam oitenta e dois documentos e trinta questões, em documentação e áreas relacionadas, para encontrar pares de questões e documentos-resposta. Eles discutiram suas discordâncias tentando resolvê-las. Várias espécies de concordâncias foram descritas numa tentativa de encontrar procedimentos racionais para resolver discordâncias.

BARHYDT⁽²³⁾ realizou um estudo para testar a eficácia de julgamentos de relevância de não-usuários sobre respostas a questões apresentadas a um centro piloto. Cada conjunto de respostas foi avaliado pelo usuário, por um especialista no assunto e por um especialista do sistema. A relevância foi julgada de duas maneiras: relevante ou não-relevante. Os resultados indicaram que o especialista no assunto não é melhor (ou pior) do que o especialista do sistema ao avaliar a relevância do documento. E que ambos não tiveram muito êxito em igualar sua avaliação com a avaliação ideal do usuá-rio.

O autor concluiu que o usuário é o melhor árbitro da relevância do documento, e que este de

ve ser considerado o único árbitro na operação de qualquer sistema.

RESNICK e SAVAGE⁽²⁴⁾ desenvolveram um experimento controlado para comparar a relativa consistência humana inter- e intra-assunto ao julgar a relevância de documentos em relação aos seus interesses gerais, com base em citações, resumos, termos de indexação, e textos completos. Os julgamentos humanos foram testados em relação a: 1) A habilidade de julgar consistentemente a relevância dos itens; 2) O efeito dos diferentes materiais léxicos (citações, resumos, termos de indexação e textos completos) nos julgamentos; 3) A habilidade de confirmar julgamentos anteriores de relevância.

Os resultados mostraram aos autores que, sob as condições do experimento, os julgamentos de relevância humanos foram consistentes, independentemente do tipo de material no qual o julgamento fôra baseado, exceto no caso dos resumos.

Salientam os autores que o que permanece como um problema é definir um procedimento para a quantificação dos resultados dos julgamentos humanos.

2.3 Métodos e medidas

Os autores empregaram vários métodos e medidas para conceituar julgamentos de relevância e para determinar conexões semânticas entre questões e textos relevantes. Serão apresentados alguns desses métodos e medidas, os quais, em sua maioria, foram descritos e utilizados nos estudos mencionados anteriormente (item 2.2).

Como mencionou SARACEVIC⁽⁶⁾, as primeiras unidades quantitativas de medida foram: revocação ("recall") e precisão ("precision"), definidas por KENT et alii⁽²⁵⁾. Revocação é a razão das respostas relevantes recuperadas sobre o total de respostas relevantes do sistema. Precisão é a razão das respostas relevantes recuperadas sobre o total de respostas recuperadas.

Algumas disciplinas e teorias, como, por exemplo, a Bibliometria e a teoria epidêmica de GOFFMAN⁽²⁶⁾ em relação à relevância, foram apresentadas na literatura mas não serão revisadas neste trabalho.

BARHYDT⁽²³⁾ e SARACEVIC⁽¹⁸⁾ apresentaram algumas medidas de avaliação do desempenho de sistemas: sensibilidade ("sensitivity"), especificidade

("specificity") e efetividade ("effectiveness"). A sensibilidade (S_e) mede a habilidade do sistema em fornecer ao usuário respostas relevantes do arquivo. A especificidade (S_p) mede a habilidade dos sistema em não fornecer ao usuário respostas não-relevantes. A efetividade (E_s) é uma adição de S_e e S_p menos uma constante 1.

$$S_e = \frac{\text{número de documentos recuperados}}{\text{total de documentos relevantes no arquivo}}$$

$$S_p = \frac{\text{número de documentos não-relevantes não recuperados}}{\text{total de documentos não-relevantes no arquivo}}$$

Segundo SARACEVIC, a sensibilidade pode ser considerada como equivalente à medida de revocação, se algumas aproximações forem feitas. Mas especificidade não é o mesmo que precisão que é o número de documentos relevantes recuperados sobre o total de documentos recuperados.

GEBHARDT⁽²⁷⁾ propôs um modelo probabilístico para a avaliação de relevância dos documentos. O autor acreditou que frequentemente os julgadores não concordam completamente e, então, considerou relevância

como uma variável aleatória. Definiu como "qualidade de seleção" a medida para a habilidade de um sistema de recuperação em encontrar os documentos desejados pelo usuário. GEBHARDT considerou que uma simples separação entre relevante e irrelevante é o pior método para avaliar relevância. Declarou que usualmente os documentos são dividitos em relevantes e não-relevantes (com relação a um determinado problema) e em recuperados e não-recuperados pelo sistema de recuperação (Tab. I).

Uma generalização das medidas de revocação e precisão foi definida para que essas se adequassem mais a uma situação real. O autor considerou que uma divisão em duas partes existe do lado do sistema de recuperação (documentos recuperados e não-recuperados) em vez de ser do lado do usuário (ou julgador). Como passo seguinte na generalização, considerou a possibilidade de serem assinalados não só os valores "relevante" ou "irrelevante" para cada documento, mas também uma medida de relevância que pode ser padronizada para o intervalo 0-1, onde 1 significa relevante e 0, irrelevante. Ficou em aberto, entretanto, a apresentação de uma escala pa-drão para a atribuição de julgamentos de relevância a do-cumentos.

Os resultados apresentados por GEBHARDT indicaram que mesmo um sistema ideal não conse-

Tab. I - Avaliação de relevância

	Relevantes	Não-relevantes
Documentos Recuperados	a	b
Documentos não Recuperados	c	d

Então:

Revocação: $a/(a+c)$ → Relevante Recuperado/Relevante

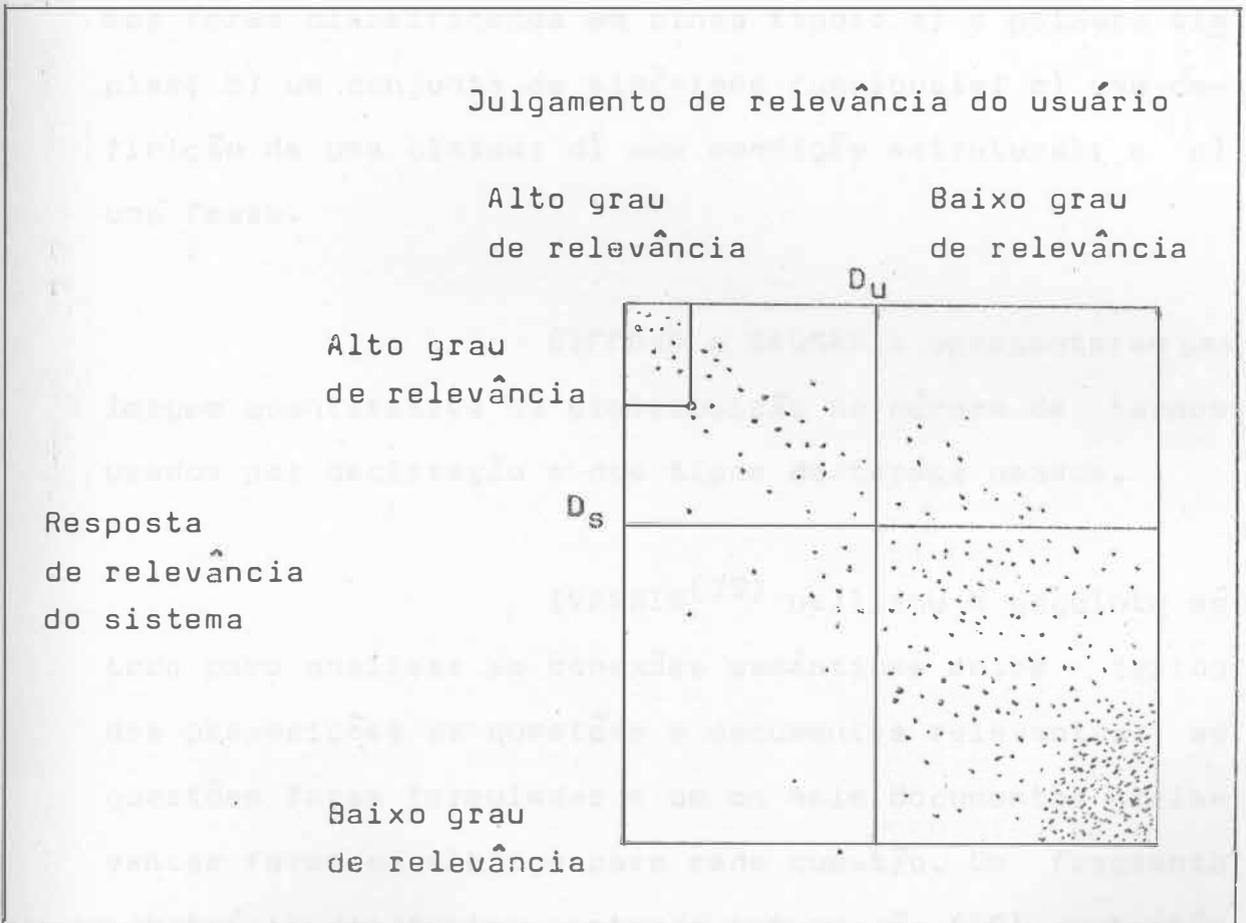
Precisão : $a/(a+b)$ → Relevante Recuperado/Recuperado

que encontrar todos os documentos relevantes e apenas aqueles.

KING⁽²⁸⁾, em uma revisão das medidas de desempenho dos sistemas, sugeriu uma representação da relação entre julgamentos de relevância do sistema e do usuário, na qual os valores de julgamento do usuário foram colocados num gráfico contra os valores correspondentes de relevância do sistema para cada documento de uma determinada questão (Graf. 1).

GIFFORD e BAUMANIS⁽¹⁶⁾ analisaram sessenta e uma questões e suas respostas para determinar diferenças textuais entre documentos relevantes e não-relevantes. As respostas para cada questão foram classificadas em relevante, parcialmente relevante e não-relevante. Foram identificadas ocorrências comuns de palavras, frases e outras propriedades textuais. A tabulação das classificações e da ocorrência de propriedades significativas em cada documento foi feita para cada questão.

Uma declaração da combinação das propriedades foi demonstrada pela tabulação dessas propriedades para cada questão. A correlação entre a ocorrência de propriedades textuais e a classificação para cada questão, e, cumulativamente, para grupos de ques-



Graf. 1 - Julgamentos de relevância dos usuários contra a resposta de relevância do sistema para documentos individuais e uma determinada questão

tões, foi também tabulada (Tab. II).

Seis níveis de êxito foram determinados pelos autores; o êxito foi medido pela porcentagem de todas as respostas à questão, reunidas nos grupos de classificação, de acordo com critérios textuais. Os termos foram classificados em cinco tipos: a) a palavra simples; b) um conjunto de sinônimos funcionais; c) uma definição de uma classe; d) uma condição estrutural; e e) uma frase.

GIFFORD e BAUMANIS apresentaram uma imagem quantitativa da distribuição do número de termos usados por declaração e dos tipos de termos usados.

IVANKIN⁽²⁹⁾ utilizou o seguinte método para analisar as conexões semânticas entre textos das proposições de questões e documentos relevantes: as questões foram formuladas e um ou mais documentos relevantes foram escolhidos para cada questão. Um fragmento arbitrário dos textos contendo informação (IF) também foi identificado. Durante a investigação, as correspondências entre as questões e cada IF de um documento relevante foram estudadas. As palavras-chave das questões e dos ifs foram identificadas.

O autor estudou:

Tab. II - Classificação hipotética dos resumos por propriedades textuais relacionadas às avaliações

Tabulação das propriedades							
Número do resumo	Avaliação	Propriedades*					
		A	B	C	D	E	F
10	R	X		X	X		
17	R		X	X		X	
14	R	X	X	X	X		X
23	P			X	X	X	
11	P			X		X	
27	N			X			
20	N			X			X
16	N	X			X		X

Declarações dos requisitos

declaração r: (A ou B) e C e (D ou E)

declaração p: C e (D ou E)

declaração n: nenhuma

Tabulação da correlação**

Nível da declaração	Avaliação			Total
	N	P	R	
r	0	0	3	3
p	0	2	0	2
n	3	0	0	3
Total	3	2	3	8

* X denota a presença da propriedade naquele resumo.

** Esse exemplo hipotético mostra perfeita correlação. Cem por cento dos resumos estão propriamente colocados através dos critérios textuais.

- a) as características quantitativas dos conjuntos de palavras-chave nos IF e nas questões;
- b) as palavras-chave específicas e gerais dos conjuntos individuais (IFs e questões);
- c) os diferentes tipos de relações e suas frequências de ocorrências;
- d) a distribuição dos IFs nos documentos relevantes;
- e) a produtividade (número de palavras-chave selecionadas) dos vários elementos estruturais de um documento.

IVANKIN fez ainda uma análise estatística das relações entre palavras-chave, suas tendências e as combinações entre todos os conjuntos de palavras-chave.

Para medir a consistência dos julgamentos de relevância humanos, RESNICK e SAVAGE⁽²⁴⁾ selecionaram um grupo de técnicos e um grupo de documentos. Para cada um dos documentos (D), foram preparados um cartão contendo a referência bibliográfica (C), um cartão contendo um resumo (A) - que incluía a referência bibliográfica - e outro cartão contendo termos de indexação (T). Cada indivíduo foi instruído para indicar os documentos relevantes e não-relevantes para seus interesses.

As tarefas dos quatro grupos foram executadas independentemente.

Aproximadamente um mês depois, os autores refizeram a experiência. Os técnicos tiveram que completar todas as respostas novamente e indicar quais os ítems que eles sentiram que haviam respondido da mesma maneira na primeira experiência.

Como já foi mencionado na revisão dos estudos sobre julgamentos de relevância (item 2.2), os resultados mostraram que, nesse experimento, os julgamentos de relevância humanos foram consistentes.

HARMON⁽¹³⁾ propôs um modelo de necessidade de informação do usuário que pode ser útil para interpretar a relevância do usuário. O modelo envolve conjuntos cognitivos dos usuários, que são organizados e ordenados de acordo com critérios específicos.

O autor identificou durante a pesquisa quatro estádios fundamentais de transformação de necessidades dos usuários:

I - Informação insuficiente e desorganizada;

II - Informação insuficiente mas organizada;

III - Informação suficiente mas desorganizada;

IV - Informação suficiente e organizada (Tab. III).

Segundo HARMON, a pesquisa deve ser vista como um processo cíclico ou modular envolvendo tanto a aquisição como a ordenação da informação. Mediante esse processo, a mudança da configuração da necessidade da informação influencia a natureza das perguntas geradas. Por sua vez, as respostas às perguntas modificam a necessidade de informação (Fig. 2).

Como se verificou através dessa revisão, a relevância ainda está longe de ser bem compreendida e conceituada. Diferentes sistemas desenvolveram pontos de vista diferentes sobre relevância. Cada um dos experimentos realizados, ou sendo realizados, dá uma nova dimensão para um enfoque das propriedades gerais, efeitos e fatores que afetam relevância. E os autores concluem que há muita pesquisa a ser efetuada nessa área, que experimentos idênticos devem ser realizados em outras áreas do conhecimento, com amostras maiores ou com mais controle das variáveis.

Não há muito o que comparar entre os experimentos descritos na literatura pois a popula-

ção analisada, a coleta de dados, os métodos, as variáveis externas, os tipos de representação dos documentos etc variam bastante de um experimento para outro. Há, contudo, um objetivo comum: esclarecer relevância como uma medida de qualidade da informação recuperada e interpretar fatores que afetam julgamentos de relevância.

	Informação Desorganizada	Informação Organizada
Informação Insuficiente	I	II
Informação Suficiente	III	IV

Tab. III - Configuração das necessidades de informação durante o período de estudo

	Informação Desorganizada	Informação Organizada
Informação Insuficiente	I	II
Informação Suficiente	III	IV

Fig. 2 - Configuração das necessidades de informação

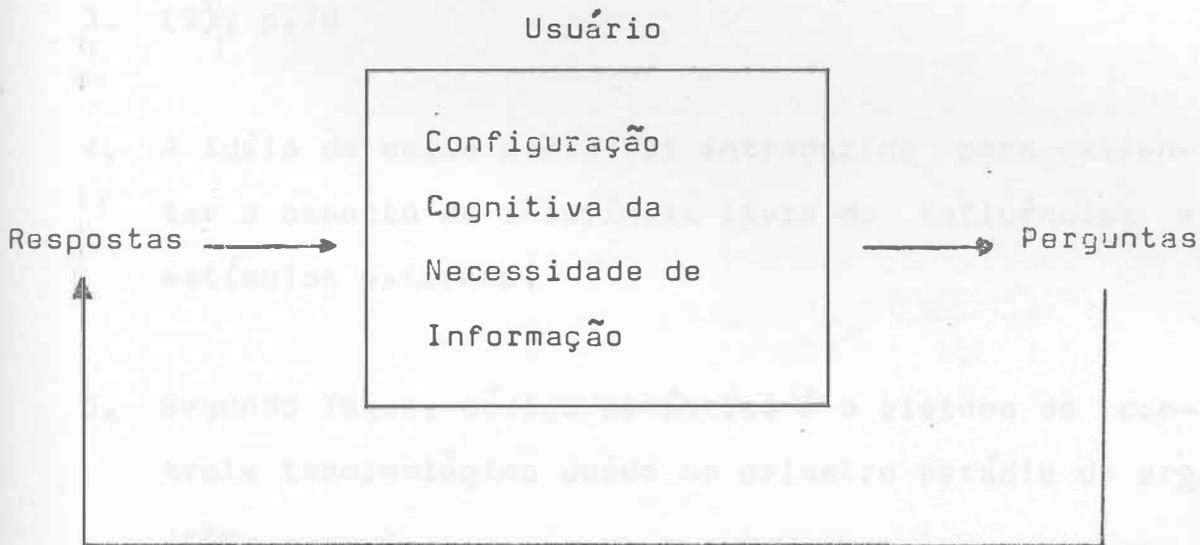


Fig. 2 - Configuração das necessidades de informação

Notas

1. Citado à página 76 de:

FIGUEIREDO, L.M. de. O conceito de relevâncias e suas implicações. Ci. Inf., 6(2):75-8, 1977.

2. (6), p.328

3. (9), p.78

4. A idéia da caixa preta foi introduzida para salientar o aspecto da relevância livre de influências e estímulos externos.

5. Segundo Tague, código semântico é o sistema de controle terminológico usado no primeiro estágio do projeto.

6. As palavras codificadas no programa incluem um ou mais dos seguintes tipos:

1) Palavras significativas da questão;

2) Palavras relacionadas no thesaurus às palavras da questão pelo código semântico, isto é, palavras que incluem ao menos um dos fatores semânticos relacionados às palavras da questão;

3) Palavras não relacionadas às palavras da ques-

tão pelo código semântico, mas mencionadas pelo analista da questão, por intuição ou conhecimento do arquivista, como sendo relacionadas às palavras da questão.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BRADFORD, S.C. Sources of information on specific subjects. Engineering: 85-6, 1934. apud FIGUEIREDO, L.M. de. O conceito de relevância e suas implicações. Ci. Inf., 6(2):75, 1977.
2. MOOERS, C.S. Coding, information retrieval, and the rapid selector. American Documentation, 1(4): 225-229 (1950) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975.
3. TAUBE, M. et alii. Storage and retrieval of information by means of the association of ideas. American Documentation, 6(1):1-17 (1955) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975.
4. PERRY, J.W. Superimposed punching of numerical codes on handsorted, punch cards. American Documentation, 2(4):205-212 (1951) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975.

5. NATIONAL ACADEMY OF SCIENCES. Proceedings of the International Conference on Science Information, 2 Vols. National Academy of Sciences, Washington, D.C. (1959) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975.
6. SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):321-43, Nov.-Dec. 1975.
7. POLUSHKIN, V.A. Relevance and pertinence. Autom. Docum. Math. Linguistics, 7(1):52-4, 1973.
8. REES, A.M. & SARACEVIC, T. Conceptual analysis of questions in information retrieval systems. Proc. Am. Docum. Inst., Part 2:175-7, 1963.
9. FOSKETT, D.J. A note on the concept of "relevance". Inf. Storage Retrieval, 8(2):77-8, Apr. 1972.
10. KUHN, T.S. The structure of scientific revolutions. University of Chicago Press (1962) apud FOSKETT, D.J. A note on the concept of "relevance". Inf. Storage Retrieval, 8(2):77-8, Apr. 1972.
11. ZIMAN, J. Public Knowledge: the social dimension of science. Cambridge University Press, (1968) apud FOSKETT, D.J. A note on the concept of "relevance". Inf. Storage Retrieval, 8(2):77-8, Apr. 1972.
12. LANCASTER, F.W. Some notes on the distinction between pertinence and relevance. In: _____. Guidelines for the evaluation of information systems and services. [s.l.] [s.ed.], Jan. 1977. [Prepared for Unesco under contract].

13. HARMON, G. Information need transformation during inquiry: a reinterpretation of user relevance. Proc. ASIS, 7:41-3, 1970.
14. WILSON, P. Situational relevance. Inf. Storage Retrieval, 9(8):457-71, 1973.
15. CUADRA, C.A. & KATTER, R.V. Opening the black box of "relevance". J. Docum., 23(4):291-303, Dec. 1967.
16. GIFFORD, C. & BAUMANIS, G.J. On understanding user choices: textual correlates of relevance judgements. Am. Docum., 20(1):21-6, Jan. 1969.
17. TAGUE, J. Matching of question and answer terminology in an educational research file. Am. Docum., 16(1):26-32, Jan. 1965.
18. SARACEVIC, T. Selected results from an inquiry into testing of IR systems. J. Am. Soc. Inf. Sci., 22(2):126-39, Mar.-Apr. 1971.
19. RATH, G.J. et alii. Comparisons of four types of lexical indicators of content. Am. Docum., 12(2):126-30, 1961.
20. O'CONNOR, J. Relevance disagreements and unclear request forms. Am. Docum., 18(3):165-77, 1967.
21. O'CONNOR, J. Some independent agreements and resolved disagreements about answer-providing documents. Am. Docum., 20(4):311-9, 1969.
22. MODERS, C. The intensive sample test, Zator Co., Cambridge, Mass. (August 1959) apud O'CONNOR, J. Relevance disagreements and unclear request forms. Am. Docum., 18(3):165-77, 1967.

23. BARHYDT, G.C. The effectiveness of non-user relevance assessments. J. Docum., 23(2):146-9, e (3):251, 1967.
24. RESNICK, A. & SAVAGE, T.R. The consistency of human judgements of relevance. Am. Docum., 15(2):93-5, 1964.
25. KENT, A. et alii. Machine literature searching. VIII. Operational criteria for designing information retrieval systems. American Documentation, 6(2): 93-101, 1955. apud SARACEVIC, T. Relevance: a review and framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(5):321-43, Nov.-Dec. 1975.
26. GOFFMAN, W. An indirect method of information retrieval. In: SARACEVIC, T., ed. Introduction to information science. New York, Bowker, 1970. p.485-92.
27. GEBHARDT, F. A simple probabilistic model for the relevance assessment of documents. Inf. Process. Mgmt., 11(1/2):59-65, Jun. 1975.
28. KING, D.W. Design and evaluation of information systems. A. Rev. Inf. Sci. Technol., 3:70-4, 1968.
29. IVANKIN, V.I. A study of the semantic relationships between the texts of a query and of a relevant document. Autom. Docum. math. Linguistics, 9(1): 20-7, 1975.

3 METODOLOGIA

3.1 Operacionalização dos conceitos

As medidas de relevância que serão utilizadas neste estudo são: sensibilidade e especificidade.

Sensibilidade é o número de documentos recuperados sobre o total de documentos relevantes no arquivo. Especificidade é o número de documentos não-relevantes não recuperados sobre o total de documentos não-relevantes do arquivo.

Quando os perfis que serviram para o estudo estavam sendo trabalhados pelo SDI manual, a avaliação dos documentos a serem recuperados era feita pela bibliotecária, comparando-se o documento original com os

descritores escolhidos pelos usuários. Eram recuperados os documentos que tratassem do assunto específico a ser pesquisado, ou seja, que contivessem alguns ou todos os descritores selecionados para definir e delimitar o assunto, de acordo com as relações entre os descritores mencionadas na descrição do assunto.

A informação recuperada era enviada ao usuário em forma de fichas contendo as referências bibliográficas. A cada recebimento de referências, o usuário analisava o conteúdo da informação recebida e respondia numa escala de dois pontos: de interesse ou não-interesse.

No caso do SDI automatizado, o procedimento é semelhante: o computador seleciona documentos de interesse para os perfis comparando os descritores selecionados no perfil com os descritores atribuídos ao documento na indexação. Na ocasião do estabelecimento do perfil, o usuário deve dar a cada descritor um peso numérico¹ que indica o grau de interesse da palavra para sua pesquisa. O computador, ao comparar os descritores do documento com os do perfil, soma os pesos das palavras encontradas através de uma fórmula específica. Se a soma for maior ou igual a um patamar² pré-estabelecido o documento é recuperado.

A cada quinzena, quando é feita a recuperação dos documentos, são enviadas as referências selecionadas ao usuário em forma de cartões. A cada bloco de cartões corresponde um cartão de interesse, onde vêm registrados os números dos documentos selecionados. Esse cartão deve ser devolvido pelo usuário com os julgamentos de relevância atribuídos a cada documento: "muito interesse" (M), "relativo interesse" (R) e "nenhum interesse" (N).

Neste estudo, decidiu-se comparar os julgamentos dos usuários com o do não-usuário (bibliotecária), atribuídos aos mesmos documentos.

Com os dados coletados pensava-se obter os seguintes indicadores:

a) Qual o grau de relevância alcançado, considerando-se o total de documentos recuperados pelo SDI automatizado em dois níveis: as referências M, isoladamente, e, em outro grupo, as referências M e R em conjunto;

b) Qual o índice de concordâncias e discordâncias entre as avaliações do sistema, do usuário ou do não-usuário.

nº do perfil no SDI/IEA	nº do perfil no CIN	áreas de interesse
P5	SDI/CIN 0278-1	E36 (5)
INIS <u>6</u> (20) 207174-01-2(R)(R) avaliação feita pelo não-usuário	→ avaliação feita pelo usuário	

Fig. 3 - Modelo de ficha para coleta de dados

Depois de definidas as medidas de relevância e demonstrados graficamente os índices de concordâncias e discordâncias das avaliações, passou-se à análise dos dados propriamente dita.

Foi escolhida para a análise uma amostra dos documentos onde houve concordância de julgamentos e os documentos que causaram divergências extremas (M e N).

Os documentos avaliados em R não foram considerados para o estudo pois não apresentaram uma característica significativa, isto é, o índice R apresentou-se em número quase idêntico tanto em relação ao índice M como ao índice N, sem mostrar uma tendência a se associar a um ou a outro índice.

Para os documentos divergentes (127) trabalhou-se com a população total. Como o número de documentos onde houve concordância de julgamentos era maior (880), foi extraída uma amostra para o estudo (201 documentos que usuário e não-usuário avaliaram N e 63 documentos que ambos avaliaram M).

O cálculo da amostra foi feito tendo-se por base um nível de 30% de confiabilidade. Esse nível foi considerado suficiente pois pareceu represen-

tar o comportamento da população, além de permitir que se trabalhasse com um número de documentos semelhante aos dos documentos divergentes.

A amostra foi extraída utilizando-se uma tabela de números aleatórios⁴ para que todos os documentos tivessem a mesma probabilidade de serem escolhidos.

Cada referência foi analisada no sentido de detectar as palavras do título que fossem iguais ou semanticamente relacionadas aos descritores do perfil. Tentou-se relacionar essa ocorrência comum de palavras com alguma propriedade que estivesse relacionada com a avaliação M ou N.

4. Table of random numbers. - JOE KLUGH, Ed. - McGraw-Hill, New York, 1970, p. 228-32.

Notas

1. No sistema manual o peso numérico não era dado, mas o "peso", ou seja, o grau de interesse do descritor era definido pelo usuário na hora da entrevista para a formulação do perfil.
2. Patamar é um limite numérico estabelecido para cada perfil, abaixo do qual os documentos não devem ser recuperados, isto é, a soma dos pesos dos descritores do documento não deve ser inferior ao patamar.
3. No período final da avaliação, pediu-se ao CIN uma listagem de todos os cartões-interesse recebidos, referentes ao período em que foi feita a avaliação, para completar os dados das fichas. Essa listagem foi necessária pois muitos usuários responderam bem depois que haviam recebido as referências e depois que se havia encerrado a coleta de dados.
4. TABLE of random numbers. In: KLUGH, H.E. Statistics: the essentials for research. New York, Wiley, 1970. p.338-9.

4 COMPARAÇÃO ENTRE JULGAMENTOS DE RELEVÂNCIA

A primeira hipótese aventada nesta dissertação propõe uma comparação entre os julgamentos de relevância do usuário e do não-usuário para verificar a compatibilidade desses julgamentos.

4.1 Análise dos dados e resultados

O total de documentos avaliados (2033 documentos) foi distribuído em uma tabela (Tab.IV) de acordo com o grau de avaliação que lhes foi conferido tanto pelo usuário como pelo não-usuário. Foi feita uma representação esquemática dessa tabela para melhor compreensão dos dados anotados e da terminologia usada para se referir a eles.

		Julgamentos de relevância do não-usuário		
		M	R	N
J u l g a m e n t o s d o u s u á r i o	M	MM	MR	MN
	R	RM	RR	RN
	N	NM	NR	NN
				Total

Nessa tabela as casas representadas em negrito evidenciam o número de documentos onde houve concordâncias de julgamentos.

Para se referir aos vários quadros da tabela usou-se sempre uma notação dupla (MM, MR, MN etc) onde, a primeira letra sempre se refere ao julgamento do usuário e a segunda, ao do não usuário.

Tab. IV - Tabela das avaliações

		NÃO-USUÁRIO		
		M	R	N
USUÁRIO	M	210	263	110
	R	91	281	260
	N	17	131	670
total				2033

Assim, como se pode observar na Tabela IV, 210 documentos foram considerados M pelo usuário e também pelo não-usuário; 263 foram considerados M pelo usuário e R pelo não-usuário; 91 foram considerados R pelo usuário e M pelo não-usuário; e assim por diante.

Houve maior divergência entre a avaliação do usuário e do não-usuário em MR (263) do que em RM (91); da mesma forma, em MN (110) a divergência foi maior do que em NM (17). E assim aconteceu, regularmente, que a divergência entre dois níveis (M e R, por exemplo) foi maior no sentido usuário - não-usuário que no inverso.

Esse fato parece indicar que a tendência do não-usuário foi sempre a de avaliar as referências abaixo do interesse que os documentos pudessem ter para a questão do ponto de vista do usuário. Uma avaliação de "interesse" e "não-interesse" simplesmente, talvez aproximasse mais os julgamentos.

Através da tabela percentual pode-se ter uma visão mais clara dos totais apresentados acima e do que eles representam em relação à população total de documentos (Tab. V).

Apenas 6,25% (5,41 + 0,84) dos do-

Tab. V - Tabela percentual das avaliações

		NÃO USUÁRIO		
		M	R	N
USUÁRIO	M	10,33%	12,94%	5,41%
	R	4,48%	13,82%	12,79%
	N	0,84%	6,44%	32,95%
				100%

cumentos analisados mostraram diferenças extremas de avaliação (MN e NM).

Em 57,10% (10,33 + 13,82 + 32,95) dos casos houve concordância de julgamentos entre os usuários e o não-usuário.

Para efeito deste estudo, essa tabela poderia ser resumida (Tab. VI).

Esses totais representam somente 49,53% da população total de documentos analisados. Mas, por outro lado, representam a população total de documentos que interessam para que se possa testar as hipóteses.

Mais uma vez, pode-se notar que o comportamento do índice R nas avaliações foi quase igual para os usuários e o não-usuário (isto é, 632 e 675 respectivamente).

Houve concordância, contudo, quanto aos índices M e N. Percentualmente, isso significaria que 28,68% dos documentos foram considerados M pelos usuários, enquanto 15,65% foram considerados M pelo não-usuário.

Da mesma forma, temos que 40,21%

Tab. VI - Tabela percentual das avaliações considerando-se só os índices M e N

		NÃO-USUÁRIO	
		M	N
USUÁRIO	M	10,33%	5,41%
	N	0,84%	32,95%

Tab. VII - Tabela comparativa das avaliações

	USUÁRIO	NÃO-USUÁRIO
M	583	318
R	632	675
N	818	1040
Total	2033	2033

dos documentos foram considerados N pelos usuários, para 51,15% avaliados N pelo não-usuário.

Como já foi esclarecido anteriormente, nem sempre os documentos considerados M, R ou N pelos usuários foram os mesmos considerados M, R ou N, respectivamente, pelo não-usuário. Assim, se se reportar à tabela percentual das avaliações (Tab. V), pode-se notar que dos 28,68% dos documentos considerados M pelos usuários, 12,94% foram considerados R pelo não-usuário.

Foram analisados também os totais das avaliações M, R e N feitas pelo usuário (583, 632, 818 respectivamente) e pelo não-usuário (318, 675, 1040 respectivamente), sem implicar as diferenças individuais de avaliação para cada documento, isto é, sem que os documentos considerados M pelo não-usuário (318), por exemplo, façam parte dos 583 documentos julgados M pelos usuários.

A Tab. VII é uma representação comparativa entre os julgamentos do usuário e do não-usuário em relação ao total de documentos analisados.

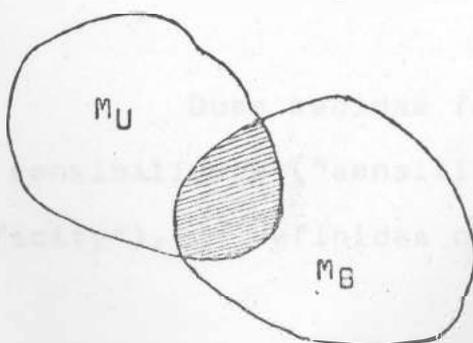
A fim de proporcionar maior clareza para a comparação dos dados apresentados, pode-se resumir as tabelas anteriores em uma só tabela, cuja repre

sentação esquemática é dada a seguir:

		NÃO-USUÁRIO		
		$\sum M_B$	$\sum R_B$	$\sum N_B$
USUÁRIO	$\sum M_U$	M_{U*B}		
	$\sum R_U$		R_{U*B}	
	$\sum N_U$			N_{U*B}

onde, por exemplo, M_{U*B} significa:

$M_U \cap M_B$, e U = usuário
B = não-usuário



Na Tabela VIII, os números das colunas em destaque (canto superior e esquerdo, respectivamente) representam a somatória dos documentos considerados M, R ou N pelos usuários (583, 632, 818 respectivamente) e pelo não-usuário (318, 675, 1040 respectivamente). Os quadros internos preenchidos mostram o total de documentos M (210), R (281) ou N (670) em que houve concordância de julgamentos.

Pode-se visualizar assim o total de documentos considerados M, R ou N pelos usuários e pelo não usuário, respectivamente, em relação ao total de documentos avaliados, por ambos, da mesma forma.

Nota-se que dos 583 documentos julgados M pelos usuários e dos 318 julgados M pelo não-usuário, 210 foram julgados M por ambos e assim por diante.

4.2 Cálculo das medidas propostas

Duas medidas foram propostas para este estudo: sensibilidade ("sensitivity") e especificidade ("specificity"), já definidas nos itens 2.3 e 3.1.

Essas medidas forneceram dados com

Tab. VIII - Documentos em que houve concordâncias em relação ao total das avaliações

		NÃO-USUÁRIO		
		M	R	N
		318	675	1040
USUÁRIO	M	583	210	
	R	632		281
	N	818		670

os quais se poderá avaliar a habilidade do não-usuário em recuperar respostas relevantes para as questões propostas pelos usuários, em relação à habilidade do usuário em determinar os documentos relevantes para sua própria pesquisa.

Para efeito deste estudo, as medidas definidas anteriormente foram alteradas para:

$$Se = \frac{\text{n}^{\circ} \text{ de documentos julgados } \underline{M} \text{ por ambos}}{\text{n}^{\circ} \text{ de documentos julgados } \underline{M} \text{ pelos usuários}}$$

$$Sp = \frac{\text{n}^{\circ} \text{ de documentos julgados } \underline{N} \text{ por ambos}}{\text{n}^{\circ} \text{ de documentos julgados } \underline{N} \text{ pelos usuários}}$$

Desta forma, o total de documentos julgados M ou N pelos usuários representa, respectivamente, o total de documentos relevantes ou irrelevantes do arquivo analisado.

O cálculo das medidas forneceu os seguintes resultados:

$$Se = \frac{210}{583} = 0,36$$

$$S_p = \frac{670}{818} = 0,81$$

Normalmente, o nível de especificidade é bem maior do que o de sensibilidade. Os resultados obtidos acima, vêm confirmar esse fato.

Em geral, é mais fácil para um substituto do usuário concordar sobre o que não é de interesse para a questão proposta pelo usuário do que sobre o que é de muito interesse. A medida do grau de interesse (muito, relativo) é subjetiva e varia de acordo com necessidades e pontos de vista do usuário no momento da avaliação.

A princípio, pareceu que se a avaliação tivesse sido feita só em termos de "interesse" e "não-interesse" chegar-se-ia a resultados bem diferentes. Não foi possível, contudo, testar essa hipótese pois tanto os usuários como o não-usuário haviam feito a avaliação numa escala de três pontos: M, R e N.

Fez-se uma aproximação, associando-se os documentos R aos M para o cálculo de Se (210 + 281 e 583 + 632). Esse dado, contudo, não parece representar a realidade pois o que foi considerado R poderia, muitas vezes, estar mais associado aos documentos N que

aos M.

De qualquer forma, o resultado não foi muito diferente do encontrado anteriormente, considerando-se só os documentos de muito interesse.

$$\text{Assim, } Se_{M+R} = \frac{491}{1215} = 0,40$$

Para o cálculo de S_p também foi feita uma aproximação, associando-se os documentos R aos N (281 + 670 e 818 + 632).

$$\text{Assim, } Sp_{N+R} = \frac{951}{1450} = 0,65$$

O grau de especificidade diminuiu quando se associou os documentos R aos N. Mas, mesmo assim, foi maior que o grau de sensibilidade.

4.3 Considerações sobre os resultados

Através da análise dos julgamentos de rele-

vância verificou-se que há diferença entre os julgamentos do usuário e do não-usuário. No entanto, a diferença não foi significativa a ponto de comprometer o papel de um não-usuário ao avaliar documentos para as questões propostas.

Como já foi demonstrado anteriormente, apenas 6,25% dos documentos analisados causaram divergências extremas de julgamentos entre usuários e o não-usuário. Por outro lado, para 57,10% dos documentos houve concordância de julgamentos.

Houve maior concordância sobre o que não é relevante do que sobre o que é relevante. Assim, 28,68% dos documentos foram julgados M pelos usuários e só 10,33% foram documentos em que houve concordância de julgamentos. Mas, de 40,23% dos documentos considerados N pelos usuários, 32,95% eram documentos em que houve concordância de julgamentos.

Esse resultado foi também encontrado na literatura. SARACEVIC⁽¹⁾, por exemplo, concluiu que, de acordo com as medidas de avaliação utilizadas em seu experimento (sensibilidade e especificidade), a concordância sobre o que é relevante foi consideravelmente menor que a concordância sobre o que é não-relevante.

As conclusões apresentadas por SARACEVIC⁽²⁾, em outro trabalho, mostram que quanto menos um não-usuário sabe sobre uma questão, maior a tendência de julgar documentos relevantes. Afirma SARACEVIC que os grupos de pessoas não conhecedoras do assunto (por exemplo, especialistas de sistemas de recuperação da informação e bibliotecários) tendem a fazer julgamentos de relevância relativamente altos. No entanto, o autor afirma que se espera que a avaliação de relevância de documentos seja similar para fornecedores de informação (isto é, substitutos do usuário) com algum conhecimento do assunto e para os usuários (isto é, pessoas com extenso conhecimento do assunto).

De fato, essa afirmação foi comprovada neste estudo. Embora a avaliação tenha sido feita por um bibliotecário, a tendência foi emitir julgamentos de relevância baixos. Isto porque, no presente caso, o bibliotecário, pelo fato de ter tido um treinamento sobre as disciplinas relacionadas à energia nuclear, se enquadra na segunda observação de SARACEVIC. Isto é, o fato de ter conhecimento do assunto fez com que os julgamentos de relevância do não-usuário fossem similares aos dos usuários.

Considera-se que as discordâncias de julgamentos muitas vezes ocorreram por falta de uma

definição da escala de avaliação adotada.

Em alguns experimentos(1,3,4) foram definidas as categorias, ou melhor, os níveis de avaliação, inclusive acrescentando que o fato de um documento já ser conhecido pelo usuário, ou estar escrito em língua desconhecida etc não deveria ser considerado na avaliação. Os resultados indicaram que os julgamentos de relevância são melhor manipulados quando os julgadores são instruídos adequadamente.

1. SPALKEVIC, F. ...
2. GIFFORD, L. & BRUNNEN, G.J. On understanding user objectives (textual correlates of relevance judgments). *Inf. Secur.*, 20(1):21-6, Jan. 1967.

REFERÊNCIAS BIBLIOGRÁFICAS

1. SARACEVIC, T. Selected results from an inquiry into testing of IR systems. J. Am. Soc. Inf. Sci., 22(2):126-39, Mar.-Apr. 1971.
2. SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):340-1, Nov.-Dec. 1975.
3. CUADRA, C.A. & KATTER, R.V. Opening the black box of "relevance". J. Docum., 23(4):291-303, Dec. 1967.
4. GIFFORD, C. & BAUMANIS, G.J. On understanding user choices: textual correlates of relevance judgements. Am. Docum., 20(1):21-6, Jan. 1969.

5 ANÁLISE SEMÂNTICA

A segunda hipótese proposta para este estudo está relacionada ao conteúdo dos títulos de documentos como elemento influenciador dos julgamentos de relevância. A referência bibliográfica do documento era o único elemento disponível na ocasião da avaliação e, assim, a avaliação foi feita com base nos títulos.

5.1 Análise dos documentos

A determinação de propriedades textuais nos documentos analisados, que pudessem testar a hipótese mencionada acima, foi feita da seguinte forma:

- 1) Dois mil e trinta e três documentos foram analisados pelos usuários de 25 perfis e pelo não-usuário.

2) As respostas dadas por ambos foram comparadas e, posteriormente, agrupadas de acordo com concordâncias e discordâncias de julgamentos.

3) Foram escolhidos para a análise os grupos de documentos onde houve concordância de julgamento (MM, 210 e NN, 670) e onde houve discordância extrema (MN, 110 e NM, 17). Como já foi mencionado anteriormente, o índice R não foi considerado.

Convém salientar que para os grupos MN (110) e NM (17) trabalhou-se com a população total de documentos. Para os grupos MM (210) e NN (670) foi selecionada uma amostra de documentos, cujo índice de confiabilidade pareceu satisfatório para representar as propriedades dos documentos que iriam ser verificadas.

4) Analisou-se primeiramente os grupos MN e NM, cujo número de documentos (110 e 17, respectivamente) era menor. Além desse fato, esperava-se que esses documentos apresentassem algumas características que os diferenciavam dos documentos onde houve concordância de julgamentos.

5) Iniciou-se a análise dos documentos do grupo NN (670) extraíndo-se uma amostra com 30%

dos documentos da população total (201 documentos). Analisados os documentos dessa amostra, verificou-se que eles já apresentavam características diferentes daquelas dos grupos anteriores. Como o número de documentos da amostra era proporcional ao dos grupos MN (110) e NM (17), decidiu-se que 30% do total de documentos daria um resultado confiável em termos deste estudo.

A análise do grupo MM foi feita da mesma forma que para o grupo NN.

6) Os títulos dos documentos de cada grupo foram comparados com os descritores dos perfis para os quais haviam sido recuperados.

Convém explicar que foram considerados termos equivalentes aos descritores dos perfis os de igual significado¹ desses descritores.

7) Fez-se uma tabulação dos dados de cada grupo de documentos de acordo com o número de descritores do perfil que apareceram nos títulos.

Quatro grupos foram encontrados:

- a) nenhum descritor do perfil no título
- b) um descritor do perfil no título

c) dois descritores do perfil no título

d) três descritores do perfil no título.

8) Comparou-se os resultados para cada grupo de documentos, definindo-se, então, as características de cada grupo e a ocorrência das propriedades acima mencionadas (a, b, c, d) em relação às avaliações (M, R e N)

5.2 Resultados

Os resultados da análise mostraram que o número de descritores do perfil que aparecem nos títulos dos documentos recuperados não é um fator decisivo para a avaliação de um documento como sendo relevante ou não-relevante.

Verificou-se, contudo, que os documentos considerados relevantes tinham mais descritores do perfil que os não-relevantes.

Observou-se que 65,67% dos títulos dos documentos considerados como não-relevantes, tanto

pelo usuário como pelo não-usuário, não continham nenhum descritor selecionado pelo perfil.

Por outro lado, 84,12% dos documentos do grupo MM tinham um ou mais descritores do perfil no título, sendo que 38,09% dos documentos tinham dois ou mais descritores.

A Tabela IX mostra o número de documentos analisados para cada grupo e sua distribuição nas quatro classes de ocorrência de descritores, definidas no item 5.1.

Pode-se observar, com relação aos grupos NM e MN, que para o julgamento do não-usuário, o aparecimento de descritores do perfil nos títulos dos documentos influencia o julgamento. Assim é que 14 dos 17 documentos (76,47% dos documentos do grupo) considerados N pelo usuário e M pelo não-usuário continham ao menos um descritor do perfil.

Analisando-se esses descritores, verificou-se que eram descritores importantes para o perfil. Por exemplo, aos descritores "HEAT TRANSFER" e "FLUID FLOW" apareceram em documentos selecionados para o perfil 244-1 - Medida de parâmetros térmicos, barreira térmica e teste de materiais -, que não foram considera-

Tab. IX - Distribuição dos documentos de acordo com ocorrência de descritores

		Ocorrências de descritores				Total
		a	b	c	d	
G R U P O S d e D O C U M E N T O S	N M	3	13	1	0	17
	M N	58	45	7	0	110
	M M	10	29	21	3	63
	N N	132	67	2	0	201
						391

dos relevantes para o usuário. No entanto, "FLUID FLOW" tem peso 4 e "HEAT TRANSFER", 5 na definição do perfil. (Ver Anexos 1 e 2)

Da mesma forma, 52,72% dos documentos julgados M pelo usuário e N pelo não-usuário não continham nenhum descritor no título.

Nesse grupo MN, contudo, 40,99% dos documentos continham no título um descritor do perfil. Verificando-se esses documentos, constatou-se que não haviam sido atribuídos pesos muito altos no perfil à maior parte desses descritores, o que significa que eram descritores apontados pelo usuário como secundários.

Assim, aconteceu terem sido considerados relevantes documentos não muito relacionados ao âmbito da pesquisa proposto inicialmente pelo usuário.

Na pesquisa sobre: "Determinação de impurezas em materiais combustíveis nucleares, por espectrometria de raios gama de captura de neutrons térmicos" (Perfil 233-1), por exemplo, não havia sido definido que as aplicações das técnicas mencionadas na medicina pudessem ser de interesse para o usuário. Porém, títulos como:

- "Application of X-ray fluorescence

analysis in medicine"

- "Role of neutron activation analysis in biomedical and environmental studies"

- "Survey of 43 constituent elements in kidney samples by instrumental activation analysis"

foram considerados de muito interesse para o usuário.

Da mesma forma, documentos cujos títulos não continham descritores dos perfis causaram divergência entre julgamentos de relevância do usuário e do não-usuário.

Para o perfil 54-1 sobre: "Características de pós de UO_2 de grau cerâmico" foram selecionados os seguintes documentos:

- "Composition of the cubic nitride in the system nickel-tantalum-nitrogen"

- "Solidus temperature of zirconium carbide in the homogeneity portion"

- "The creep of Si_3N_4 under oxidizing and non-oxidizing conditions"

Nenhum desses documentos continha descritores do perfil (Ver Anexo 2) e estes foram considerados de muito interesse pelo usuário.

Os documentos: "The reliability of mercury analysis in environmental materials", e "Determination of uranium in the red blood cells of the workers in the chemical processing of uranium ore" foram selecionados para o perfil 240-1 - "Utilização de fontes de Cf^{252} para métodos analíticos (análise por ativação)" cujos descritores são: "Californium 252", "Radioactivation" e "Activation analysis", tendo sido considerados M pelo usuário.

Igualmente, o usuário do perfil 240-1 considerou M um documento cujo título era: "Geo- and cosmochemistry and archaeology" e que tratava principalmente de análise de amostras lunares. Nenhum descritor do perfil aparecia no título do documento.

O fato dos títulos dos documentos analisados não conterem qualquer descritor do perfil, ou conterem descritores de menor interesse (de peso menor), relacionados a elementos ou aspectos não apontados como descritores ou mencionados na descrição do assunto, levou provavelmente o não-usuário a não considerar os documentos como relevantes.

A partir desses resultados pôde-se verificar que o aparecimento ou não de descritores do perfil no título justifica e autoriza o não-usuário a

classificar os documentos em relevantes e não-relevantes.

Essa observação, que foi feita a partir dos resultados apresentados pelos grupos MN e NM, pode ser estendida aos documentos dos grupos MM e NN.

Pela tabela de distribuição dos documentos dos grupos de concordâncias e discordâncias de julgamentos segundo a ocorrência de descritores dos perfis (Tab. IX), observa-se que apenas 10 documentos do grupo MM (15,87% do total) não continham descritores dos perfis e 69 documentos NN (34,32% do total) continham ao menos um descritor dos perfis.

Quanto à relação entre os descritores contidos nos títulos, verificou-se que a maior parte dos documentos que causaram divergências de julgamentos, e que continham ao menos um descritor do perfil, não mostraram o relacionamento entre os descritores indicado na questão.

Por exemplo, para o perfil 240-1 - "Utilização de fontes de Cf²⁵² para métodos analíticos (análise por ativação)" - vários documentos que foram considerados de muito interesse pelos usuários continham só o descritor "activation analysis" nos títulos.

Porém, na formulação do perfil do SDI automatizado, apareciam só os descritores "Activation analysis", "Radioactivation" e "Californium 252", sendo que Cf²⁵² tinha peso 10 em relação aos outros dois descritores de peso 5.

De acordo com o que foi exposto, títulos como os abaixo mencionados não deveriam ter sido considerados de muito interesse:

- "Comparison of atomic absorption and neutron activation analysis for the determination of silver, chromium, and zinc in various marine organisms";

- "Determination of gold in platinum by neutron activation analysis";

- "Monostandard activation analysis of ancient pigments of the murals in the Takamatsuzuka tumulus";

- "Cadmium by in vivo neutron activation analysis".

Outro exemplo poderia ser o perfil 54-1 - Características de pós de UO₂ de grau cerâmico - onde "Uranium dioxide" (UO₂) teve maior peso que as outras palavras (Ver Anexo 2) e, no entanto, vários documentos que mencionavam outros elementos foram considerados de muito interesse. Por exemplo,

- "Microstructure and properties of

fused-cast $\text{NaNbO}_3\text{-BaTiO}_3$ ceramics";

- "Influence of calcia addition on sintering behaviour of zirconia";

- "Effect of sintering temperature on the structure and compacting of dispersion hardened nickel";

- "On the effect of tungsten carbide grain size on the alloy WC-Co strenght".

Durante o estudo, alguns usuários foram contatados nesse sentido, pois passaram a considerar de interesse elementos e métodos não definidos no perfil. Assim, por exemplo, o usuário do perfil 251-1 - Preparação de compostos marcados e radiofármacos - que haviam selecionado os seguintes elementos para o perfil: Tc^{99m} , In^{113m} , I^{131} , Cr^{51} , Hg^{203} , Hg^{197} , I^{125} , H^3 , C^{14} - considerou, por exemplo, esses documentos como de muito interesse:

- "Method for the preparation and quality control of ^{68}Ga radiopharmaceuticals" e

- "Fluorine-18 labelled L-p fluoro-phenyl-alanine".

Segundo informação do usuário, os radioisótopos não indicados no perfil seriam considerados também como de interesse relativo se ligados à marcação de compostos. No entanto, os radioisótopos menciona-

dos nos documentos a acima foram considerados como de muito interesse pelo usuário.

Enquanto se fazia a análise dos resultados, notou-se que determinados perfis provocaram maiores discordâncias ou concordâncias de julgamentos que outros. Essa observação poderia conduzir a um ponto de vista mais crítico na interpretação dos resultados. Esse fato parece confirmar que não só as características textuais dos documentos devem ser levadas em conta para análise dos documentos do usuário, como também as flutuações de pontos de vista que ocorrem no momento da avaliação.

Com relação às discordâncias de julgamentos, alguns perfis se destacaram por terem causado divergências em quase todos os envios de referências, enquanto outros se caracterizaram por não terem apresentado julgamentos discordantes (Tab. X).

Verificando-se as tabelas de concordâncias de julgamentos (Tabs. XI e XII) vê-se que os perfis que não haviam provocado discordâncias de julgamento estavam melhor representados na tabela de concordâncias de julgamentos do grupo NN. Esses perfis apresentaram durante todo o período do estudo um índice de relevância baixo sendo que as avaliações oscilaram sempre en

Tab. X - Discordâncias MN e NM por perfil

	6(18)	6(19)	6(20)	6(21)	6(22)	6(23)	6(24)	7(1)	7(2)	7(3)	7(4)	TOTAL
54-1	1	-	-	-	-	1	2	-	4	4	3	15
81-1	-	-	-	-	-	-	-	-	-	-	-	-
85-1	1	-	2	-	2	-	4	-	3	-	-	12
229-1	-	-	-	3	-	-	1	-	-	-	-	4
230-1	2	-	-	-	-	-	-	-	-	-	-	2
231-1	-	-	-	-	-	-	-	-	-	-	-	-
232-1	-	-	-	-	-	-	-	-	-	-	-	-
233-1	2	2	1	2	1	2	10	1	-	4	4	29
233-2	-	-	-	-	-	-	-	-	-	-	-	-
234-1	-	-	-	-	-	-	-	-	1	-	-	1
235-1	2	-	-	-	2	-	-	-	-	-	-	4
239-1	2	-	2	-	-	2	-	-	-	-	-	6
240-1	3	-	4	3	2	5	3	-	-	-	-	20
241-1	-	-	-	-	-	-	-	-	-	-	-	-
242-1	1	-	-	-	-	-	-	-	-	-	-	1
243-1	-	-	-	-	-	-	-	-	-	-	-	-
244-1	2	-	1	-	-	-	2	14	2	-	-	21
245-1	-	-	-	1	3	1	-	-	1	-	-	6
246-1	-	-	-	-	-	-	-	-	-	-	-	-
251-1	-	-	-	-	-	1	-	1	-	3	-	5
252-1	-	-	-	-	-	-	-	-	-	-	-	-
278-1	-	-	-	1	-	-	-	-	-	-	-	1
278-2	-	-	-	-	-	-	-	-	-	-	-	-
279-1	-	-	-	-	-	-	-	-	-	-	-	-
TOTAL	16	2	10	10	10	12	22	16	11	11	7	127

Tab. XI - Concordâncias MM por perfil

	6(18)	6(19)	6(20)	6(21)	6(22)	6(23)	6(24)	7(1)	7(2)	7(3)	7(4)	TOTAL
54-1	2	4	-	3	6	1	2	1	-	1	-	20
81-1	-	-	-	1	5	-	1	1	-	1	1	10
85-1	11	9	4	3	3	-	6	-	-	1	1	38
229-1	1	-	2	3	-	-	3	1	-	1	2	13
230-1	-	1	-	-	2	3	-	1	-	2	1	10
231-1	-	-	-	-	-	-	-	-	-	-	-	-
232-1	-	-	-	-	-	-	-	-	-	-	-	-
233-1	9	1	2	1	2	1	2	-	2	1	2	23
233-2	4	3	2	5	2	1	2	-	-	-	-	19
234-1	1	6	-	1	-	-	1	1	-	-	-	10
235-1	2	-	-	-	-	-	1	-	-	-	-	3
239-1	-	-	-	-	-	-	-	-	-	-	2	2
240-1	-	-	-	-	-	-	-	1	1	-	1	3
241-1	-	-	-	-	-	-	-	-	-	-	-	-
242-1	1	-	-	-	1	1	2	1	-	-	-	6
243-1	1	2	-	2	1	2	3	-	-	-	-	11
244-1	4	3	1	2	-	-	2	4	-	1	-	17
245-1	-	3	-	1	1	-	-	-	-	-	1	6
246-1	-	-	-	-	-	-	-	-	-	-	-	-
251-1	-	1	3	-	-	1	-	-	-	5	1	11
252-1	-	-	-	-	-	-	-	1	-	1	-	2
278-1	-	-	-	1	1	-	-	-	-	-	2	4
278-2	-	-	-	-	-	-	-	-	-	-	-	-
279-1	-	-	-	-	1	-	-	-	-	-	1	2
TOTAL	36	33	14	23	25	10	25	12	3	14	15	210

Tab. XII - Concordâncias NN por perfil

	6(18)	6(19)	6(20)	6(21)	6(22)	6(23)	6(24)	7(1)	7(2)	7(3)	7(4)	TOTAL
54-1	-	-	-	-	-	-	-	-	-	-	-	-
81-1	1	7	1	1	2	1	1	-	4	1	4	23
85-1	15	5	2	4	4	5	3	1	1	-	1	41
229-1	4	-	6	-	-	3	1	2	1	4	1	22
230-1	-	-	2	-	-	-	-	-	-	-	-	2
231-1	-	-	-	-	-	-	-	-	-	-	-	-
232-1	4	4	5	3	1	2	2	2	1	3	5	32
233-1	3	6	3	11	5	4	12	3	7	2	6	62
233-2	2	-	-	3	2	1	-	-	1	1	1	11
234-1	3	3	1	-	1	-	12	-	4	-	-	24
235-1	24	22	6	12	10	-	1	1	2	-	1	79
239-1	13	4	9	5	-	11	3	1	-	1	-	47
240-1	43	-	10	10	12	3	7	2	10	3	5	105
241-1	1	2	-	1	1	2	18	-	7	2	-	34
242-1	-	-	-	-	1	-	-	-	-	-	-	1
243-1	1	3	1	2	3	4	1	-	-	-	1	16
244-1	1	2	-	3	-	1	5	-	-	-	-	12
245-1	5	7	1	16	9	8	22	-	31	2	3	104
246-1	-	-	-	-	-	-	-	-	-	-	-	-
251-1	-	-	-	-	-	-	1	1	-	1	-	3
252-1	-	-	-	-	1	1	-	-	1	-	-	3
278-1	-	-	-	12	1	21	2	-	2	-	-	38
278-2	-	-	-	-	-	4	-	1	1	-	-	6
279-1	-	-	-	2	-	2	-	-	-	1	-	5
TOTAL	120	65	47	85	53	73	89	11	73	21	28	665

tre N e R.

Os perfis que não tiveram nenhuma representação nas três tabelas foram analisados. Verificou-se que as discordâncias e concordâncias estavam relacionadas ao índice R, que não foi considerado para este estudo.

5.3 Considerações sobre os resultados

Os resultados alcançados indicaram que a frequência de descritores do perfil no título, não discrimina, por si só, os documentos relevantes dos não relevantes.

TAGUE⁽¹⁾, contudo, concluiu que o número de palavras, mais do que a relação entre as palavras da questão, é que caracteriza as respostas relevantes. Afirmou, ainda, que a frequência de palavras da questão é maior em documentos relevantes que nos parcialmente relevantes e não relevantes.

Verificou-se, realmente, que os documentos relevantes tinham mais descritores do perfil que os não-relevantes. Mas essa verificação não deu con-

dições para que se estabelecesse um padrão de frequência de descritores nos documentos relevantes ou nos não-relevantes, que designasse um padrão relativo de julgamentos de relevância.

GIFFORD e BAUMANIS⁽²⁾ concluíram em seu estudo que há diferenças textuais correspondentes a julgamentos de relevância e que a determinação dessas propriedades poderia conduzir ao estabelecimento do padrão de julgamentos do usuário. Contudo, os próprios autores dizem que sua amostra poderia não ser considerada significativa pois: a) foi dada explicação aos usuários sobre os níveis de avaliação; b) as questões foram dirigidas a um tópico específico; c) resumos foram colocados à disposição dos usuários.

IVANKIN⁽³⁾ concluiu que os títulos quase sempre contêm os descritores que o thesaurus do sistema de recuperação da informação incluiu. Declarou que a ocorrência de palavras no título deve ser o melhor critério para seleção das palavras como descritores do documento na hora da indexação.

TAGUE⁽¹⁾ apontou que as palavras da questão são mais úteis para localizar respostas relevantes, mas não devem ser usadas exclusivamente pois 14% das respostas relevantes não continham nenhuma palavra

da questão.

O que se verificou neste estudo foi que 14,28% das respostas consideradas relevantes tanto pelos usuários como pelo não-usuário não continham descritores do perfil. Esse resultado está de acordo com a conclusão apontada por TAGUE.

Embora as condições e elementos de estudo não tenham sido os mesmos nos experimentos revisados e neste experimento, as conclusões foram compatíveis.

Notas

1. Os termos de "igual significado" foram determinados por:
 - sinônimos que haviam sido mencionados pelos usuários no perfil do SDI manual. Ex: "microspheres" e "coated fuel particles"; os dois termos foram apresentados pelo usuário como equivalentes, mas "microsphere", conquanto não apareça no Thesaurus do INIS², é usado na literatura;
 - formas diferentes de apresentação do termo, isto é, como verbo, substantivo, adjetivo. Ex: "preparation of enriched uranium" foi considerado equivalente ao descritor "uranium enrichment"
 - "narrower terms" que aparecem no thesaurus. Ex: Os "narrower terms" de descritores gerais como "reactor instrumentation", "rare earth complexes" etc, que aparecem nos títulos de documentos, foram considerados equivalentes ao descritor mais geral do perfil.

2. O Thesaurus do INIS é o vocabulário controlado usado para indexação e recuperação da informação no sistema INIS. Sua estrutura é a mesma de qualquer thesaurus:
 - BT (broader terms) - termos gerais
 - Nt (narrower terms) - termos específicos

REFERENCIAS BIBLIOGRÁFICAS

RT (related terms) - termos relacionados

UF (used for) - termos proibidos

1. TAGUE, J. Matching of question and answer
terminology in an educational research file. Am.
J. Libr., 19(1):36-38, Jan. 1965.

2. GIBSON, C. & SHANNON, S.J. An experimental study
of the effect of subject matter on the
Am. J. Libr., 12(1):71-6, Jan. 1969.

3. ISAKI, M. A study of the semantic relationships
between the texts of a query and a relevant
document. Am. J. Libr., 11(1):22-31, 1967.

REFERÊNCIAS BIBLIOGRÁFICAS

1. TAGUE, J. Matching of question and answer terminology in an educational research file. Am. Docum., 16(1):26-32, Jan. 1965.
2. GIFFORD, C. & BAUMANIS, G.J. On understanding user choices: textual correlates of relevance judgements. Am. Docum., 20(1):21-6, Jan. 1969.
3. IVANKIN, V.I. A study of the semantic relationships between the texts of a query and of a relevant document. Autom. Docum. math. Linguistics, 9(1): 20-7, 1975.

CONCLUSÕES

De acordo com os resultados obtidos para o teste da primeira hipótese proposta nesta dissertação, concluiu-se que há discordâncias entre os julgamentos do usuário e do não-usuário. Essa diferença, contudo, não foi significativa a ponto de comprometer o papel de um não-usuário ao avaliar documentos, em substituição ao usuário.

Considera-se que o fato de não haver ocorrido um índice de concordância mais alto se deve ao fato de que, enquanto o usuário avalia os documentos para sua necessidade de informação, o não-usuário os avalia para a questão proposta. Essa conclusão poderia ser testada se se instrísse os usuários no sentido de formular a questão o mais próximo possível de sua necessidade de informação e de avaliar os documentos sob o ponto de vista exclusivo da questão proposta.

Em relação aos serviços de SDI podemos concluir que:

1) O não-usuário pode avaliar documentos ou questões em substituição ao usuário.

Considerando-se que os serviços de SDI foram projetados para atender às necessidades de informação do usuário, a melhor forma de aproximar as avaliações do sistema ou não-usuário às do usuário é fazer com que este esclareça de forma adequada a questão.

Deve haver no formulário de proposição da questão espaço reservado à descrição do assunto, ou explicação detalhada da questão, para que o especialista do sistema possa entender com clareza o assunto envolvido na questão, a finalidade e as circunstâncias determinantes do estudo e o tipo de informação desejada.

2) Como os julgamentos de relevância do usuário mudam com o tempo, deve-se criar condições que façam o usuário interagir com o sistema.

É aconselhável que, periodicamente, se renove o contato com o usuário para reformular o perfil, se for o caso, para que a relevância alcançada pelo sistema possa se comparar à relevância esperada pelo usu

ário.

3) Deve-se definir para o usuário a escala de avaliação adotada para que essa possa entendida da mesma forma que o sistema a interpreta.

A segunda hipótese apresentada neste estudo relaciona a frequência de aparecimento de descritores do perfil nos títulos de documentos aos julgamentos de relevância. Concluiu-se que o número de descritores do perfil nos títulos não é um fator determinante de relevância ou não-relevância de documentos.

Os resultados da análise feita nesta dissertação levam a algumas conclusões sobre os títulos de documentos e descritores do perfil em relação aos julgamentos de relevância:

1) O aparecimento ou não de descritores do perfil nos títulos dos documentos é para o não-usuário um critério de julgamento.

2) Deve-se procurar estabelecer, tanto quanto possível, relações semânticas entre descritores, desde que essas relações não signifiquem um conceito já adotado no thesaurus (ex., "uranium enrichment", "reactor control systems", "biological radiation effects").

A descrição do assunto, que foi sugerida nestas conclusões como elemento necessário à formulação da questão, torna evidente as relações entre os descritores para o analista da questão.

A análise da questão é necessária não só às operações com lógica booleana, mas, também, para recuperação através do sistema de pesos. O relacionamento entre os descritores no sistema de pesos deve ser utilizado para estabelecimento do patamar.

Para maior esclarecimento da conexão semântica entre documento e questão e sobre a utilidade dos títulos para determinação da relevância de documentos, pode-se sugerir outro estudo nas mesmas bases em que este foi feito, utilizando-se também o resumo¹.

Uma idéia sugerida, mas não testada, durante este estudo foi que quanto mais detalhada é uma escala de avaliação maior o índice de discordâncias de julgamentos entre usuário e não-usuário. Pareceu relativamente fácil para o não-usuário concordar que um documento seja relevante ou não-relevante face à questão. O mesmo não aconteceu, entretanto, quanto à determinação do grau de relevância (muito, relativo etc) do documento em relação à necessidade de informação do usuário.

Além disso, do ponto de vista do sistema existe uma avaliação dicotômica: relevante e não-relevante; do ponto de vista do usuário existem diferentes graus de relevância dos documentos.

Todas essas conclusões conduzem a uma só afirmação: não se consegue determinar uma medida precisa de relevância, pois os usuários avaliam a "relevância" em relação à sua necessidade de informação (pertinência), enquanto o não-usuário, ou o sistema, avalia a relevância em relação à questão.

Todos os trabalhos revisados na literatura e inclusive este, constituem uma tentativa de aproximação entre relevância e pertinência para que se possa quantificar ou padronizar uma medida de relevância.

Notas

1. O CIN, na época em que foi feito este estudo, fornecia aos usuários somente a referência dos documentos recuperados. Posteriormente, passou a fornecer também os resumos.

BRADY, G. L. The effectiveness of two-year relevance assessments. J. Doc., 22(4):214-9, 4 (1971), 1971.

BRADY, G. L. Sources of information on available subjects. Information, 19(4), 1964. Information, L.S. de. O conceito de relevância e sua medição. EL. Inf., 1(2):175-8, 1972.

BRADY, G. L. & PETER, R. E. Opening the black box of "relevance". J. Doc., 22(4):214-225, Mar. 1971.

BRADY, G. L. A note on the concept of "relevance". J. Doc., 22(4):214-225, Mar. 1971.

BRADY, G. L. A simple probabilistic model for the relevance assessment of documents. Inf. Systems, 2(1):111-120, Jan. 1975.

BRADY, G. L. & PETER, R. E. On understanding user-provided relevance criteria. J. Doc., 22(4):214-225, Mar. 1971.

BRADY, G. L. An analysis of the concept of "relevance". Information, 19(4), 1964. Information, L.S. de. Information, New York, 1970. p. 111-120.

7 LISTA DAS REFERÊNCIAS BIBLIOGRÁFICAS

- BARHYDT, G. C. The effectiveness of non-user relevance assessments. J. Docum., 23(2):146-9, e (3):251, 1967.
- BRADFORD, S.C. Sources of information on specific subjects. Engineering: 85-6, 1934. apud FIGUEIREDO, L.M. de. O conceito de relevância e suas implicações. Ci. Inf., 6(2):75-8, 1977.
- CUADRA, C.A. & KATTER, R.V. Opening the black box of "relevance". J. Docum., 23(4):291-303, Dec. 1967.
- FOSKETT, D.J. A note on the concept of "relevance". Inf. Storage Retrieval, 8(2):77-8, Apr. 1972.
- GEBHARDT, F. A simple probabilistic model for the relevance assessment of documents. Inf. Process. Mgmt., 11(1/2):59-65, Jun. 1975.
- GIFFORD, C. & BAUMANIS, G.J. On understanding user choices: textual correlates of relevance judgements. Am. Docum., 20(1):21-6, Jan. 1969.
- GOFFMAN, W. An indirect method of information retrieval. In: SARACEVIC, T., ed. Introduction to information science. New York, Bowker, 1970. p.485-92.

- HARMON, G. Information need transformation during inquiry: a reinterpretation of user relevance. Proc. ASIS, 7:41-3, 1970.
- IVANKIN, V.I. A study of the semantic relationships between the texts of a query and of a relevant document. Autom. Docum. math. Linguistics, 9(1): 20-7, 1975.
- KENT, A. et alii. Machine literature searching. VIII. Operational criteria for designing information retrieval systems. American Documentation, 6(2): 93-101, 1955. apud SARACEVIC, T. Relevance: a review and framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):321-43, Nov.-Dec. 1975.
- KING, D.W. Design and evaluation of information systems. A. Rev. Inf. Sci. Technol., 3:70-4, 1968.
- KUHN, T.S. The structure of scientific revolutions. University of Chicago Press (1962) apud FOSKETT, D.J. A note on the concept of "relevance". Inf. Storage Retrieval, 8(2):77-8, Apr. 1972.
- LANCASTER, F.W. Some notes on the distinction between pertinence and relevance. In: _____. Guidelines for the evaluation of information systems and services. [s.l.], [s.ed.], Jan. 1977. [Prepared for Unesco under contract]
- MOODERS, C.S. Coding, information retrieval, and the rapid selector. American Documentation, 1(4): 225-229 (1950) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975.

MOEDERS, C.S. The intensive sample test, Zator Co., Cambridge, Mass. (August 1959) apud O'CONNOR, J. Relevance disagreements and unclear request forms. Am. Docum., 18(3):166, 1967.

NATIONAL ACADEMY OF SCIENCES. Proceedings of the International Conference on Science Information, 2 Vols. National Academy of Sciences, Washington, D.C. (1959) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975.

O'CONNOR, J. Relevance disagreements and unclear request forms. Am. Docum., 18(3):165-77, 1967.

_____. Some independent agreements and resolved disagreements about answer-providing documents. Am. Docum., 20(4):311-9, 1969.

PERRY, J.W. Superimposed punching of numerical codes on handsorted, punch cards. American Documentation, 2(4):205-212 (1951) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975.

POLUSHKIN, V.A. Relevance and pertinence. Autom. Docum. math. Linguistics, 7(1):52-4, 1973.

RATH, G.J. et alii. Comparisons of four types of lexical indicators of content. Am. Docum., 12(2): 126-30, 1961.

REES, A.M. & SARACEVIC, T. Conceptual analysis of questions in information retrieval systems. Proc. Am. Docum. Inst., Part 2:175-7, 1963.

- RESNICK, A. & SAVAGE, T.R. The consistency of human judgements of relevance. Am. Docum., 15(2):93-5, 1964.
- SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):321-43, Nov.-Dec. 1975.
- _____. Selected results from an inquiry into testing of IR systems. J. Am. Soc. Inf. Sci., 22(2):126-39, Mar.-Apr. 1971.
- TAGUE, J. Matching of question and answer terminology in an educational research file. Am. Docum., 16(1):26-32, Jan. 1965.
- TAUBE, M. et alii. Storage and retrieval of information by means of the association of ideas. American Documentation, 6(1):1-17 (1955) apud SARACEVIC, T. Relevance: a review of and a framework for the thinking on the notion in information science. J. Am. Soc. Inf. Sci., 26(6):327, Nov.-Dec. 1975
- WILSON, P. Situational relevance. Inf. Storage Retrieval, 9(8):457-71, 1973.
- ZIMAN, J. Public Knowledge: the social dimension of science. Cambridge University Press, (1968) apud FOSKETT, D.J. A note on the concept of "relevance". Inf. Storage Retrieval, 8(2):77-8, Apr. 1972.

Anexo I - Lista dos perfis do SDI manual do IEA introduzidos
no SDI automatizado do CIN

Nº Perfil SDI/IEA	Título	Nº Perfil SDI/CIN
P1	- Características de pós UO_2 de grau cerâmico	54-1
P2	- Fabricação e controle da qualidade de pastilhas de UO_2	81-1
P3	- Utilização de um espectrômetro Ge(Li) na de terminação de impurezas em materiais combus tíveis nucleares, por espectrometria de rai os gama de captura de neutrons térmicos	233-1
P4	- Determinação do nº médio de neutrons emiti dos na fotofissão do ^{238}U	232-1
P5	- Unidade crítica	278-1 278-2 279-1
P7	- Fabricação de microesferas de U e Th	230-1
P8	- Termoluminescência como efeito das radia ções em cristais iônicos e sua aplicação em dosimetria da radiação	234-1
P9	- Efeito de radiações nas várias propriedades de cristais iônicos naturais brasileiros e sua aplicação em dosimetria	
P10	- Medidas da corrente termoiônica nos cris tais iônicos irradiados	231-1
P11	- Controle de radiocontaminação interna	229-1
P12	- Tratamento de resíduos radioativos líquidos	239-1
P13	- Preparação de compostos marcados e radiofár macos	251-1
14	- Difusão em cristais iônicos	241-1
P15	- Optimal reactor strategies for Brazil	242-1
P17	- Uso de geradores na produção de radioisóto pos	252-1
P18	- Controle de reatores e instrumentação nucle ar	85-1

P19	- Atrito interno	-	235-1
P20	- Constantes de estabilidade dos complexos	-	245-1
P21	- Desenvolvimento e aperfeiçoamento de fontes de radiação e dispositivos de irradiação	-	246-1
P22	- Utilização de fontes de californio-252 para métodos analíticos	-	240-1
P23	- Enriquecimento de urânio: Processos aerodinâmicos		
P24	- Medida de parâmetros térmicos, barreira térmica e teste de materiais	-	244-1
P25	- Tecnologia da purificação e produção da grafita nuclear	-	243-1

Anexo II - Descritores dos perfis

0054-1 (P1)

5	Uranium dioxide	1	Sintered materials
3	Powder metallurgy	1	Powders
1	Compacting	1	Density
1	Impurities	1	Texture
1	Grain size	1	Microstructure
1	Growth	1	Sintering

0081-1 (P2)

5	Uranium dioxide	1	Ceramics
1	Cermets	1	Fuel pellets
1	Fabrication	1	Quality control
1	Performance testing	1	Sintering
1	Fuel elements	1	Powder metallurgy
3	Plutonium oxides	3	Thorium oxides

0085-1 (P18)

1	Analog systems	1	CAMAC system
2	Data acquisition systems	2	Electronic circuits
1	Pulse circuits	2	Reactor control systems
1	Temperature measurement	1	Transducers
1	Radiation detection	1	Simulators
1	Amplifiers	1	Charged particle detection
2	Digital systems	1	On-line systems
1	Reactor instrumentation	1	Pulse analyzers

0229-1 (P11)

1	Metabolism	1	Biological radiation effects
2	Urine	2	Uranium
2	Thorium	2	Fluorine
2	Plutonium	1	Toxicity
1	Biological materials	2	Radionuclide kinetics
2	Tritium	2	Iodine-131

0230-1 (P7)

3	Uranium	3	Uranium dioxide
5	Coated fuel particles	1	Coatings
1	Nuclear fuels	3	Thorium

2 Pyrolytic carbon
1 Surface coating

0231-1 (P10)

3 Thermionic conversion
1 Dipoles
1 Lithium fluorides
1 Calcium fluorides
3 Ionic conductivity
1 Relaxation losses
1 Potassium bromides
1 Rare earths

0232-1 (P4)

5 Uranium-238
1 Thorium-232
5 Fission spectra

0233-1 (P3)

2 Activation analysis
2 Gamma spectrometers
2 Si semiconductor detectors
2 Trace amounts
1 Gamma spectra
1 X-ray spectra

0233-2 (P23)

5 Uranium
6 Isotope separation
1 Centrifugation

0240-1 (P22)

10 Californium-252
5 Activation analysis

0241-1 (P14)

5 Diffusion
2 Films
2 Pressure dependence
1 Potassium chlorides
1 Potassium bromides
1 Cadmium fluorides

1 HTGR type reactors

2 Ionic crystals
1 Electric currents
1 Potassium chlorides
1 Sodium chlorides
1 Irradiation
1 Doped materials
1 Cadmium fluorides

3 Fission neutrons
5 Photofission
1 Multiple production

1 Californium-252
2 X-ray fluorescence analysis
2 Semiconductor detectors
2 Capture
2 Gamma spectroscopy
1 Impurities

1 Uranium hexafluoride
6 Separation nozzle method
1 Ultracentrifuges

5 Radioactivation

1 Calcium fluorides
3 Ionic crystals
2 Temperature dependence
1 Sodium chlorides
1 Lithium fluorides

0242-1 (P15)

6 Fuel cycle	5 Thorium
1 PWR type reactors	1 Cost benefit analysis
1 HTGR type reactors	3 Uranium
1 Economic policy	1 Nuclear fuels
1 BWR type reactors	

0243-1 (P25)

8 Graphite	1 Physical properties
1 Impurities	2 Production
1 Machining	1 Annealing
1 Fabrication	2 Purification
1 Mechanical properties	1 Graphite moderated reactors
1 Mining	

0244-1 (P24)

5 Heat transfer	1 Pressure gages
1 Impurities	1 Flow rate
4 Fluid flow	1 Water
1 Thermocouples	1 Helium
1 Temperature measurement	1 Gas chromatography
1 Data acquisition	1 Pressure measurement
1 Thermal insulation	

0245-1 (P20)

5 Tetracyclines	5 Rare earth complexes
5 Stability	2 Structural chemical analysis
3 Ligands	

0246-1 (P21)

1 Gamma radiography	3 Sealed sources
1 Iridium	1 Cobalt
1 Calibration	1 Irradiation capsules
1 Irradiation devices	5 Gamma sources
2 Leak testing	1 Terbium
1 Thulium	1 Rabbit tubes
1 Irradiation procedures	

0252-1 (P17)

6 Radioisotope generators	4 Isotope production
---------------------------	----------------------

1 Cesium-137
1 Molybdenum-99
1 Daughter products
1 Barium-137
1 Technetium-99

1 Tin-113
1 Tellurium-132
1 Minutes living radioisotopes
1 Indium-113
1 Iodine-132

0278-1 (P5A)

1 Containment buildings
2 Shields
3 GCR type reactors
2 Construction
1 Electrical equipment

3 Reactor components
1 Reactor cores
5 Zero power reactors
2 Planning

0278-2 (P5B)

1 Control rod drives
2 Reactor instrumentation
3 GCR type reactors
1 Control elements
1 Neutron flux
1 Radiation monitors

3 Reactor control systems
1 Control rod worths
1 Temperature measurement
5 Zero power reactors
1 Pressure measurement

0279-1 (P5C)

5 Zero power reactors
2 Fault tree analysis
1 Reactivity coefficients
2 Reactor safety
1 Control
3 GCR type reactors

1 Reactor instrumentation
1 Control rod worths
1 Reactor start-up
2 Criticality
1 Containment buildings

0234-1 (PB/9)

6 Thermoluminescence
1 Glow curve
1 Calcium fluorides
1 Fluorite
1 Annealing
1 Potassium bromides

3 Lithium fluorides
1 Potassium chlorides
1 Cadmium fluorides
2 Point defects
1 Sodium chlorides
1 Electron spin resonance