

Recuperação de Informação:

a representação do subjetivo

Edberto Ferneda

Como citar: FERNEDA, E. Recuperação de Informação: a representação do subjetivo. *In:* JORENTE, M. J. V.; SEGUNDO, R. S.; MONTOYA, J. A. F.; MARTÍNEZ-ÁVILLA, D.; NAKANO, N. (org.) **Curadoria Digital e Gênero na Ciência da Informação**. Marília: Oficina Universitária; São Paulo: Cultura Acadêmica, 2021. p. 193-209.
DOI: <https://doi.org/10.36311/2021.978-65-5954-142-3.p193-209>



All the contents of this work, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença Creative Commons Atribuição-NãoComercial-SemDerivações 4.0 (CC BY-NC-ND 4.0).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia Creative Commons Reconocimiento-No comercial-Sin derivados 4.0 (CC BY-NC-ND 4.0).

CAPÍTULO 8

Recuperação de Informação: a representação do subjetivo

Edberto Ferneda

RESUMO

Recuperação de informação designa a operação pela qual se seleciona documentos de um acervo em função de uma determinada demanda informacional. Em um sistema computacional, a recuperação se efetivada por meio da comparação entre as representações dos documentos e a representação da necessidade de informação do usuário. Um documento é recuperado se a sua representação coincidir total ou parcialmente com a representação da necessidade do usuário. A correta interpretação de tais representações é fundamental para a eficiência de um sistema de informação. Portanto, recuperar informação implica em operar seletivamente sobre um conjunto de itens de informação, o que envolve processos cuja formalização e automação só são possíveis por meio de simplificações de conceitos tipicamente subjetivos. O objetivo deste texto é apresentar e avaliar as formas de representação computacional dos conceitos e operações inerentemente subjetivos que fazem parte do processo de recuperação de informação. Trata-se de um trabalho de abordagem qualitativa e natureza básica. Quanto aos objetivos, pode ser considerado como explicativo, e bibliográfica em relação aos procedimentos. Como resultado, pode-se afirmar que o processo de recuperação

de informação é essencialmente impreciso devido a conceitos subjetivos utilizados em seu embasamento teórico. O desenvolvimento de sistemas computacionais de recuperação de informação, tais como os mecanismos de busca da Web, exigem simplificações de tais conceitos para que seja possível representá-los por meio de algoritmos. Inexoravelmente essas simplificações afetam direta ou indiretamente a eficiência dos sistemas de informação. Assim, a automação do processo de recuperação de informação viabiliza a operação de grandes quantidades de dados de uma forma rápida e ágil. Porém, essas características não necessariamente fornecem resultados consistentes ou satisfatórios. Informação, tomada por seu significado de senso-comum, está diretamente relacionada ao seu significado, o que implica procedimentos baseados nas capacidades e habilidades humanas de abstração, apreensão e representação de sua significação.

1 INTRODUÇÃO

A busca por informações utilizando algum tipo de recurso tecnológico tornou-se uma atividade comum na sociedade contemporânea. Ao fazermos uma busca na Web estamos procurando por uma informação que venha a satisfazer uma determinada necessidade. Uma informação é considerada relevante se ela trazer o conhecimento que necessitamos em um determinado momento, em uma determinada situação.

A ideia da utilização de dispositivos eletrônicos na busca por informação teve sua gênese com os trabalhos de Paul Otlet (1934, 2018) e posteriormente com o artigo intitulado *"As We May Think"*, de Vannevar Bush (1945). As ideias veiculadas nesses trabalhos abriram caminho para diversas pesquisas nas décadas seguintes. No início dos anos de 1950, o matemático, físico e cientista da computação Calvin Northrup Mooers (1951) criou o termo *"Information Retrieval"*, inaugurando uma área de pesquisa que se consolidou e se fortaleceu ao longo dos anos. A popularização da Internet e o surgimento da Web trouxeram novos desafios e grande interesse na pesquisa e no desenvolvimento de técnicas para auxiliar na busca e recuperação de informação nesse acervo mundial.

Desde as primeiras pesquisas até os dias atuais, o papel dos sistemas de recuperação de informação passou de simples ferramentas experimentais a sistemas de uso cotidiano, úteis a todos que necessitem de informação para suas atividades. Nesse período, o acelerado avanço tecnológico e inúmeras ideias, conceitos e técnicas foram propostos e desenvolvidos. Porém, a busca por informações relevantes e úteis é ainda uma tarefa árdua. Essa dificuldade leva à reflexão sobre os principais elementos envolvidos no processo de recuperação de informação, que aparentemente são alheios aos avanços tecnológicos, ou pelos menos às tecnologias atualmente disponíveis (FERNEDA, 2013).

Recuperação de informação é a operação pela qual se seleciona documentos de um acervo em função de uma determinada demanda informacional. Na sua essência, a recuperação se dá por meio da comparação entre as representações dos documentos e a representação da necessidade de informação do usuário. Um documento é recuperado se a sua representação coincidir total ou parcialmente com a representação da necessidade do usuário. Recuperar informação implica, portanto, em operar seletivamente sobre um conjunto de itens de informação, o que envolve processos cuja formalização e automação só são possíveis por meio de simplificações de conceitos tipicamente subjetivos (FERNEDA, 2003).

O objetivo deste texto é avaliar as formas de representação computacional dos conceitos e operações inerentemente subjetivos que fazem parte do processo de recuperação de informação.

2 A RELEVÂNCIA DA RELEVÂNCIA

O conceito de relevância é crucial na Recuperação de Informação, sendo muitas vezes utilizado na própria enunciação dos objetivos dessa área. É uma questão fundamental e uma preocupação central para o funcionamento e avaliação dos sistemas de recuperação de informação (SARACEVIC, 2017; MIZZARO, 1997; COOPER, 1971).

O termo “relevância” é geralmente utilizado para identificar um elemento que se sobressai ou se destaca em um determinado conjunto. É utilizado também para discriminar um objeto “de grande valor ou interesse”, ou ainda referenciar “aquilo que importa ou é necessário”.¹

O conceito de relevância pode ser expresso por

¹ FERREIRA, Aurélio Buarque de Holanda. Novo Dicionário da Língua Portuguesa.

diferentes termos. Vannevar Bush (1945) utilizou a expressão “item de importância momentânea”; Mooers (1951) referia-se a “informações úteis”. Termos como “pertinente”, “valioso”, “útil”, “significativo” são utilizados em conotações diferentes, mas geralmente com significado subjacente à relevância.

Como a maioria das noções fundamentais, relevância é intuitivamente bem compreendida - ninguém precisa explicá-la para ninguém no mundo. Essa é a sua força. É por isso que os sistemas que visam a recuperação de informações relevantes para os usuários, incluindo mecanismos de busca e uma variedade de aplicativos de pesquisa nas mídias sociais, são tão bem aceitos globalmente - diferenças em culturas, sociedades e costumes não importam. No entanto, relevância é uma noção humana, não técnica. Essa é a sua fraqueza. Como todas as noções humanas, relevância é confusa. A relevância engloba muitas variáveis que são difíceis de controlar e até mesmo compreender formalmente. A relevância sempre, repito sempre, envolve também um contexto. Todos os algoritmos de busca em todos os sistemas do mundo estão tentando aproximar, com vários graus de sucesso, a noção humana de relevância. É para isso que servem, é por isso que existem.² (SARACEVIC, 2015, p. 27, tradução nossa).

Relevância sempre envolve uma relação. Sempre existe um “para” associado à relevância que se refere a um

2 As most fundamental notions, relevance is intuitively well understood – nobody has to explain it to anybody in the world. That is its strength. That is why the systems aiming at retrieval of relevant information to users, including search engines and a variety of search apps in social media, are so well accepted globally – differences in cultures, societies, and mores do not matter. However, relevance is a human, not a technical, notion. That is its weakness. As all human notions, relevance is messy. Relevance encompasses many variables that are hard to control and even fathom formally. Relevance always, repeat always, involves a context as well. All the search algorithms in all the systems in the world are trying to approximate, with various degrees of success, the human notion of relevance. That is what they are all about, that is why they exist.

contexto, um assunto em questão. Algo é relevante *para* alguém ou *para* um determinado contexto. O conceito de relevância não é necessariamente binário, existindo gradações que se modificam à medida que as intenções e os horizontes cognitivos mudam, ou quando o assunto em questão é alterado (SARACEVIC, 2017, p.17). Segundo Sperber e Wilson (2005, p. 224), “intuitivamente, relevância não é uma questão de tudo ou nada, mas uma questão de graus”. A atribuição desses graus de relevância é um processo inerentemente subjetivo.

Saracevic (1975), Swanson (1986) e Harter (1992) distinguem dois tipos de relevância: “relevância objetiva” e “relevância subjetiva”. A *relevância objetiva* está relacionada aos sistemas, enquanto a *relevância subjetiva* se relaciona à operação e uso de tais sistemas por seus usuários. Segundo Swanson (1986), em um sistema de recuperação de informação sempre caberá ao usuário julgar a relevância das informações resultantes de uma busca. Essa arbitragem tem caráter individual, uma “experiência mental” baseada nas características de cada usuário. Saracevic (2017, p. 24) argumenta que os sistemas são criados por diferentes projetistas, que utilizam diferentes abordagens e diferentes métodos de desenvolvimento. Assim, de certa forma os sistemas também são subjetivos. Portanto, segundo o autor, não existe relevância “objetiva”. Toda relevância é subjetiva, mesmo quando formalizada em um algoritmo.

De forma similar à classificação dicotômica objetiva-subjetiva, diversos autores utilizam a terminologia “relevância do sistema” e “relevância do usuário” (MIZZARO, 1997). A *relevância do sistema* é uma relevância potencial presumida, definida e formalizada a partir de hipóteses ou conjecturas relacionadas com a estrutura de representação dos itens de informação, a forma de organização desses itens e o grau de similaridade de cada item em relação à expressão de busca. No ambiente Web, por exemplo, considerando a sua estrutura

formada por um conjunto de páginas conectadas por links, o algoritmo básico do Google (PageRank) parte da ideia que a quantidade de links que uma página Web recebe de outras páginas pode servir como uma medida da sua relevância (BRIN; PAGE, 2012). Sistemas de biblioteca utilizam critérios de relevância adaptados à estrutura de representação dos itens de seu acervo. O sistema Primo³ classifica os resultados de uma busca com base nos seguintes critérios de relevância (EX LIBRIS, 2015):

- 1. O grau de correspondência de um item em relação à consulta.** Um item é considerado mais relevante se os termos da consulta ocorrerem em campos de metadados específicos do registro do item (autor, título, assunto) e se os termos do registro aparecerem na mesma ordem da consulta;
- 2. A importância acadêmica de um item.** A significância acadêmica do item é calculada a partir de fatores não relacionados à consulta. Para calcular a importância acadêmica de um item é levado em consideração se ele foi publicado em um periódico revisado por pares, quantas vezes foi citado, entre outras características;
- 3. A relevância de um item para o tipo de busca.** O sistema infere o tipo de busca que o usuário está conduzindo. Em uma busca por um assunto amplo ou genérico o sistema acrescenta em seus resultados artigos de referência. Em pesquisas por itens mais específicos, o sistema considera autores, títulos ou outras características para colocar alguns itens no topo da lista de resultados.
- 4. A atualidade de um item.** Pressupõe-se que os usuários geralmente preferem materiais recentes.

³ Ex Libris Primo é um conjunto ferramentas desenvolvido e comercializado pela empresa Ex-Libris que implementam recursos de a busca e recuperação em acervos de objetivos digitais.

Um sistema de recuperação atribui relevância seguindo critérios formalizados por seus algoritmos. Esses algoritmos tem a função principal de comparar a representação de cada documento do acervo com a expressão de busca enunciada pelo usuário. O resultado dessa comparação é um valor numérico que representa o grau de relevância de cada documento em relação à busca. Esse grau de relevância é geralmente utilizado para ordenar (ranquear) o conjunto de documentos resultante de uma busca.

A partir do conjunto de documentos resultantes de sua busca, o usuário julga a relevância dos itens recuperados (*relevância do usuário*) utilizando o seu conhecimento sobre o assunto pesquisado. Borlund (2003) argumenta que a relevância é um conceito cognitivo multidimensional cujo significado é amplamente dependente das percepções dos usuários e de suas necessidades. O julgamento de relevância do usuário é balizado inicialmente pela sua necessidade de informação. Porém, a importância dada a certas dimensões de relevância pode mudar dinamicamente, conforme o usuário avance na análise dos resultados.

O conceito de relevância tem desempenhado um papel importante no desenvolvimento dos sistemas de recuperações de informação. Se a eficiência de um sistema está na sua capacidade em recuperar documento relevantes, essa eficiência pode ser medida pela proximidade entre a relevância do sistema e a relevância do usuário. A relevância do sistema pode ser formalizada utilizando característica relacionadas à organização e à representação dos itens de um acervo. Porém, a relevância do usuário escapa a qualquer tipo formalização ou representação.

3 A INCOMPLETUDE DA REPRESENTAÇÃO

Naquele Império, a Arte da Cartografia alcançou tal Perfeição que o Mapa de uma

única Província ocupava toda uma Cidade, e o mapa do Império, toda uma Província. Com o tempo, estes Mapas Desmedidos não foram satisfatórios e os Colégios de Cartógrafos levantaram um Mapa do Império que tinha o tamanho do Império e coincidia pontualmente com ele. Menos Afeitas ao Estudo da Cartografia, as Gerações Seguintes entenderam que esse dilatado Mapa era Inútil e não sem Impiedade o entregaram às Inclemências do Sol e dos Invernos. Nos desertos do Oeste perduram despedaçadas Ruínas do Mapa, habitadas por Animais e por Mendigos; em todo o País não há outra relíquia das Disciplinas Geográficas. (Suárez Miranda, *Viajes de varones prudentes*, livro quarto, cap. XLV, Lérida, 1658)

Jorge Luis Borges – do rigor na ciência

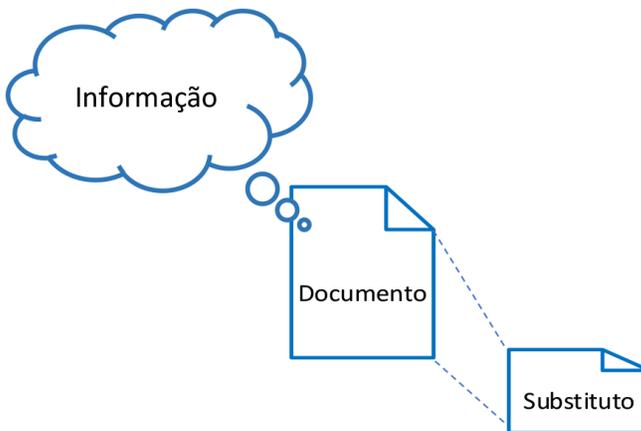
Toda representação é incompleta. Se não fosse incompleta não seria uma representação. Uma representação é comumente mais breve ou menor do que o objeto representado, restringindo-se às características consideradas mais relevantes. Portanto, a criação de uma representação envolve escolhas sobre o que nela será incluído e o que será descartado. Alguma coisa do original é sempre perdida. Uma representação sempre será uma versão distorcida do real, mesmo que apenas por sua incompletude (SARACEVIC, 1991).

O processo de recuperação de informação envolve duas instâncias de representação: a representação de cada item de informação de um determinado acervo e a representação da necessidade de informação do usuário por meio de uma expressão de busca (consulta). Segundo Belkin, Oddy e Brooks (1982a), essas duas representações são de natureza distintas. Um item de informação (documento) é a representação de um “estado de conhecimento coerente”, enquanto uma consulta é a representação de um “estado anômalo de conhecimento”. Existem situações em que o usuário é capaz de especificar exatamente quais informações

são necessárias para resolver um determinado problema. Porém, a situação mais comum é aquela em que o usuário não tem conhecimento prévio da informação que necessita, nem tampouco consegue formalizá-la em uma expressão de busca.

A representação de um documento inclui os elementos descritivos que o identificam e o caracterizam em um acervo, assim como os elementos indicativos de seu conteúdo informativo. A Figura 1 ilustra os elementos do processo de representação documental definidos por Mizzaro (1998).

Figura 1 - Representação do processo de representação de um documento



Fonte: elaborada pelo autor.

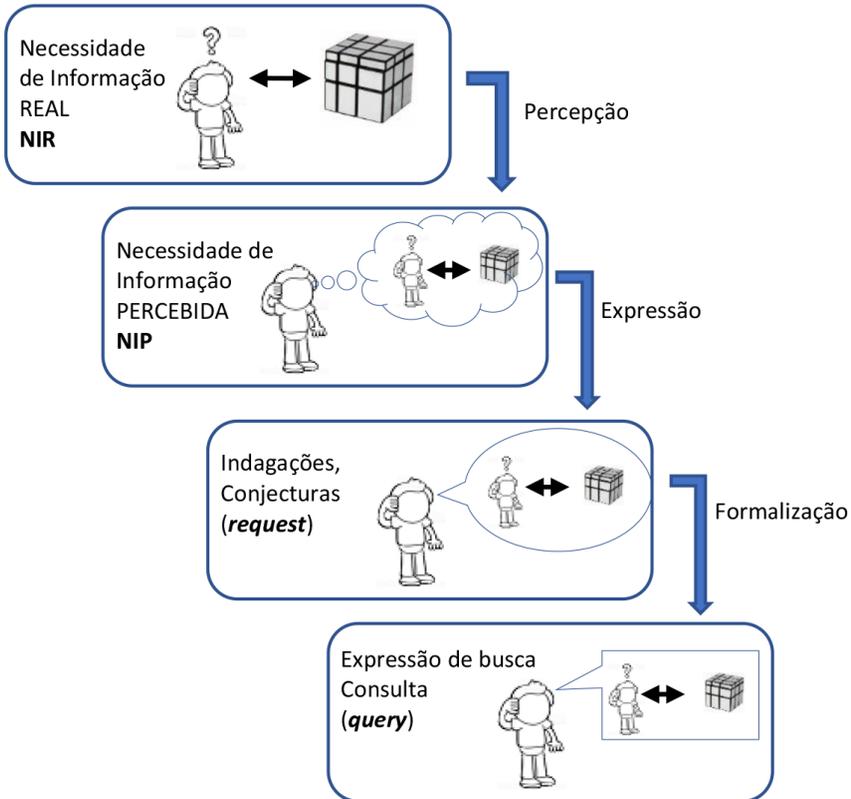
Um *documento* é a representação física de um conhecimento, a materialização de uma *informação*. É a entidade que o usuário de um sistema de recuperação de informação obtém em resposta à sua busca. O “substituto” (*surrogate*) é a representação do documento, consistindo de elementos que o discriminem dentre os demais itens do acervo. É o elemento que será comparado com a expressão de busca, responsável pela recuperação do documento. Mizzaro (1998) ordena esses três elementos da seguinte forma:

substituto < documento < informação

Em um sistema de recuperação de informação, o acervo documental é constituído a priori, sendo passível de ser processado por técnicas automatizadas tais como indexação automática, mineração de textos, entre outras. Por outro lado, a necessidade de informação do usuário só é percebida após a sua enunciação por meio de uma expressão de busca e a sua interpretação é dificultada pelo número reduzido de termos que normalmente é utilizado. Porém, a partir de sua definição, a expressão de busca pode ser utilizada em processos interativos que visam resolver possíveis ambiguidades ou que permitam o seu enriquecimento semântico (PANSANI JUNIOR, 2021).

A Figura 2 ilustra o processo de representação de uma busca definido por Mizzaro (1998). Possui quatro entidades – *Necessidade de Informação Real (NIR)*; *Necessidade de Informação Percebida (NIP)*, *Indagações ou conjecturas (request)* e *Expressão de busca ou consulta (query)* – e três operações: *Percepção*, *Expressão* e *Formalização*.

Figura 2 - Representação do processo de busca por informação



Fonte: adaptada de Mizzaro (1998).

Segundo Belkin, Oddy e Brooks (1982a; 1982b), uma necessidade de informação surge a partir de uma reconhecida anomalia no estado de conhecimento do usuário sobre algum assunto ou situação problemática da qual não consegue especificar com precisão o que é necessário para resolvê-la. Mizzaro (1998) denomina essa necessidade inicial de *Necessidade de Informação Real (NIR)*. O usuário percebe a sua necessidade e constrói uma representação mental, possivelmente incompleta ou incorreta em relação à NIR: a *Necessidade de Informação Percebida (NIP)*. Por meio de conjecturas ou indagações (*request*), o usuário expressa sua necessidade em uma linguagem humana, uma linguagem

natural. Finalmente, o usuário formaliza suas indagações em uma consulta (*query*) utilizando a linguagem fornecida pelo sistema de recuperação de informação. A cada nível de representação ocorre uma perda ou uma distorção em relação ao nível anterior. Os elementos constituintes desse processo podem ser ordenados da seguinte forma (MIZZARO,1998):

query < request < NIP < NIR

Portanto, a consulta (*query*) é a materialização linguística possivelmente incompleta de uma necessidade de informação, após uma sequência de representações mentais.

Em essência, o processo de recuperação de informação se efetiva pela comparação entre representações: a representação da necessidade de informação do usuário e a representação de cada documento de um acervo. O resultado dessa comparação será geralmente um número que representa o grau de relevância do documento em relação à busca e irá posicioná-lo na lista de resultados.

4 IMPRECISÃO MATEMÁTICA

[...]

- Está bem - disse o computador. - A Resposta à Grande Questão...

- Sim...!

- Da Vida, o Universo e Tudo o Mais... - disse Pensador Profundo.

- Sim!

- É... - disse Pensador Profundo, e fez uma pausa.

- Sim...! -É... - Sim...!!!...?

- Quarenta dois - disse Pensador Profundo, com uma majestade e uma tranquilidade infinitas.

Douglas Adams – O Guia do Mochileiro das Galáxias

Os primeiros computadores pesavam várias toneladas e sua programação era feita por meio da conexão direta de seus circuitos. Na década de 1950 a programação era feita através da transmissão de instruções em código binário por meio de cartões ou fitas perfuradas. Com o surgimento das linguagens de programação, o código binário ficou restrito ao núcleo do computador e a comunicação com o mundo externo era feita por uma nova camada de programa. “Aquilo que ontem fora interface torna-se órgão interno” (LÉVY, 1993, p. 101). Atualmente os computadores são constituídos por um conjunto de dispositivos e camadas de programas que se comunicam umas com as outras, permitindo grande distanciamento do seu núcleo binário.

Binária, a informática? Sem dúvida, em um certo nível de funcionamento de seus circuitos, mas faz tempo que a maioria dos usuários não mais tem qualquer relação com esta interface. Em que um programa de hipertexto ou de desenho é “binário”? (LÉVY, 1993, p. 102).

Em resposta à questão colocada por Pierre Lévy, podemos confirmar que atualmente utilizamos um computador sem precisar termos conhecimento do funcionamento de seus circuitos, da mesma forma como utilizamos qualquer outro aparelho eletrônico. Porém, a alma binária de um computador perpassa todas as suas camadas de programas e limita sua capacidade de efetuar tarefas que a maioria dos seres humanos fazem com relativa facilidade.

No processo de recuperação de informação, recursos computacionais viabilizam a operação de grandes acervos documentais, como é o caso da Web. No entanto, a natureza dos computadores exige a matematização de conceitos e processos tipicamente subjetivos. A relevância, agora destituída de sua subjetividade, passa a ser um número. A estratégia primária da automação do processo de representação documental (indexação) é a simples contagem de palavras. As palavras com maior número de ocorrências superfície

textual de um documento são eleitas como representantes de seu conteúdo intelectual. A necessidade de informação é representada por um conjunto de palavras destituídas de seus significados.

A automação do processo de recuperação de informação impõe uma lógica na qual a informação deve ser numericamente definida no interior de um sistema fechado, que desconsidera alguns fatores humanos envolvidos nesse processo.

5 CONSIDERAÇÕES

O termo “subjetivo” é definido como o “que pertence ao sujeito pensante e a seu íntimo”; “pertinente a ou característico de um indivíduo; individual, pessoal, particular”⁴. Subjetivo é tudo aquilo que é próprio do sujeito ou a ele relativo. É algo que está baseado em uma interpretação individual.

Os conceitos envolvidos no processo de recuperação de informação são tipicamente subjetivos. O desenvolvimento de sistemas computacionais exige simplificações de tais conceitos para que seja possível formalizá-los e representá-los por meio de algoritmos e programas. Essas simplificações afetam direta ou indiretamente a eficiência dos sistemas de informação. Observa-se que grande parte das pesquisas em Recuperação de Informação está concentrada na busca formas mais eficiente de representar a subjetividade envolvida nesse processo.

A automação do processo de recuperação de informação viabiliza a operação de grandes quantidades de dados de uma forma rápida e ágil. Porém, não necessariamente fornece resultados consistentes ou satisfatórios. Informação, tomada por seu significado de senso-comum, está diretamente

4 FERREIRA, Aurélio Buarque de Holanda. Novo Dicionário da Língua Portuguesa.

relacionada ao seu significado, o que implica procedimentos baseados nas capacidades e habilidades humanas de abstração, apreensão e representação de sua significação.

REFERÊNCIAS

BELKIN, N. J.; ODDY, R. N.; BROOKS, H. M. ASK for Information Retrieval: part I. Background and Theory. **Journal of Documentation**, [s. l.], v. 38, n. 2, p. 61-71, 1982a.

BELKIN, N. J.; ODDY, R. N.; BROOKS, H. M. ASK for information retrieval: part II. Results of a design study. **Journal of Documentation**, [s. l.], v. 38, n. 3, p. 145-164, 1982b.

BORLUND, P. The concept of relevance in IR. **Journal of the American Society for Information Science and Technology**, [s. l.], v. 54, n. 10, p. 913-925, 2003.

BRIN, S.; PAGE, L. Reprint of: The anatomy of large-scale hypertextual Web search engine. **Computer Networks**, [s. l.], v. 56, n. 18, p. 3825-3833, 2012.

BUSH, V. As We May Think. **Atlantic Monthly**, [s. l.], v. 176, p. 101-108, 1945.

COOPER, W. S. A Definition of Relevance for Information Retrieval. **Information Storage and Retrieval**, [s. l.], v. 7, n. 1, p. 19-37, 1971.

EX LIBRIS. **Primo discovery**: search, ranking, and beyond. [S. l.: s. n.], 2015.

FERNEDA, E. **Recuperação da Informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. 2003. 147 f. Tese (Doutorado em Ciência da Comunicação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2003.

FERNEDA, E. **Ontologia como recurso de padronização terminológica em um Sistema de Recuperação de Informação**. 2013. Relatório (Estágio Pós-Doutoral em Ciência da Informação) – Programa de Pós-Graduação em Ciência da Informação, Universidade Federal da Paraíba, João Pessoa, 2013.

HARTER, S. P. Psychological relevance and information science. **Journal of the American Society for Information Science**, [s. l.], v. 53, n. 4, p. 257-270, 1992.

LÉVY, P. **As tecnologias da Inteligência**: o futuro do pensamento na era da informática. São Paulo: Editora 34, 1993.

MIZZARO, S. Relevance: the whole history. **Journal of the American Society for Information Science**, [s. l.], v. 48, n. 9, p. 810-832, 1997.

MIZZARO, S. How many relevances in Information Retrieval? **Interacting with Computers**, [s. l.], v. 10, n. 3, p. 303-320, 1998.

MOOERS, C. N. Zatocoding applied to mechanical Organization of Knowledge. **American Documentation**, [s. l.], v. 2, n. 1, p. 20-32, 1951.

OTLET, P. **Traité de documentation**: le livre sur le livre: théorie et pratique. Bruxelles: Mundaneum, 1934.

OTLET, P. **Tratado de documentação**: o livro sobre o livro: teoria e prática. Brasília: Briquet de Lemos/Livros, 2018.

PANSANI JUNIOR, E. A. **Contextualização e expansão de consultas em Sistemas de Recuperação de Informação**: um método baseado em ontologias de domínio. 2021. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2021.

SARACEVIC, T. Information science: origin, evolution and relations. In: INTERNATIONAL CONFERENCE ON CONCEPTIONS OF LIBRARY AND INFORMATION SCIENCE: HISTORICAL, EMPIRICAL AND THEORETICAL PERSPECTIVES, 1991, Finland. **Proceedings** [...]. Helsink: [s. n.], 1991.

SARACEVIC, T. Relevance: a review of and framework for the thinking on the notion in information science. **Journal of the American Society for Information Science**, [s. l.], v. 26, n. 6, p. 321-343, 1975.

SARACEVIC, T. **The notion of relevance in Information Science**: everybody knows what relevance is. But, what is it really? Williston: Morgan & Claypool, 2017.

SPERBER, D.; WILSON, D. Teoria da relevância. **Linguagem em (Dis) curso**, Santa Catarina, v. 5, p. 221-268, 2005.

SWANSON, D. R. Subjective versus objective relevance in bibliographic retrieval systems. **Library Quarterly**, [s. l.], v. 56, n. 4, 1986.



61



62



60



59



58



63



64



65



66



68



67

67