

O CONCEITO DA RELEVÂNCIA E AS TÉCNICAS DE SEARCH ENGINE OPTIMIZATION: Suas implicações no ranqueamento nos mecanismos de busca

LARISSA PAVARINI DA LUZ

GUSTAVO CAMOSSI

JOSÉ APARECIDO LIMA

Resumo

Mecanismos de busca como o Google tornaram-se as principais ferramentas usadas para localizar informações na Web. Além disso, são cruciais para atrair tráfego para um site e representam o principal impulso dos esforços nos mecanismos de busca. Sendo assim, os objetivos deste artigo são de estudar, no âmbito da Ciência da Informação, como o conceito de relevância vem sendo aplicado na otimização de mecanismos de busca e identificar como as organizações podem reagir a essa mudança, adaptando e implementando novas estratégias para suas táticas de *Search Engine Optimization*. Realizou-se uma revisão bibliográfica sobre os conceitos de relevância dentro da Ciência da Informação e *Search Engine Optimization*. Apresentou-se os avanços tecnológicos implementados no mecanismo de busca (Google), destacando-se o funcionamento do seu algoritmo *Hummingbird*, cujo objetivo é aumentar a compreensão da semântica nas buscas e a relevância do conteúdo nos ambientes informacionais digitais. O artigo mostra como as alterações nos algoritmos do Google, especialmente no algoritmo *Hummingbird*, melhoraram a compreensão do contexto, da relevância e do significado das pesquisas feitas no Google e, dessa forma, contribuíram para apresentar resultados mais precisos referente ao tema buscado, como é o caso do elemento informacional conhecido como *Knowledge Graph*, por exemplo.

Palavras-chave Mecanismos de busca; *Search engine optimization*; Relevância; Otimização de motor de busca

THE CONCEPT OF RELEVANCE AND SEARCH ENGINE OPTIMIZATION TECHNIQUES: Its implications for search engine ranking

Abstract

Search engines like Google have become the primary tools used to locate information on the Web. Moreover, they are crucial to attract traffic to a site and represent the main boost of search engine efforts. Thus, the objectives of this article are to study, within the scope of Information Science, how the concept of relevance has been applied in search engine optimization and to identify how organizations can react to this change by adapting and implementing new strategies for their Search Engine Optimization tactics. A literature review was carried out about the concepts of relevance within Information Science and Search Engine Optimization. The technological advances implemented in search engines were presented, highlighting Google and the functioning of its Hummingbird algorithm, whose objective is to increase the understanding of semantics in searches and the relevance of content in sites. The article shows how the changes in Google's algorithms, especially the Hummingbird algorithm, have improved the understanding of context, relevance and meaning of searches made on Google, and, in this way, have contributed to present more precise results regarding the searched subject, as is the case of the informational element known as Knowledge Graph, for example.

Keywords: *Search engines; Search engine optimization; Relevance; Hummingbird*

1 INTRODUÇÃO

A Recuperação da Informação (RI) existe como um processo desde a Antiguidade, porém, o que caracteriza o seu marco moderno é a necessidade de armazenar e favorecer acesso rápido e preciso ao grande volume de documentos. Mooers (1951, p. 51, tradução nossa) criou o termo, enfatizando que ele "[...] engloba os aspectos intelectuais da descrição de informações

e suas especificidades para a busca, além de quaisquer sistemas, técnicas ou máquinas empregadas para o desempenho da operação”. Para (SARACEVIC, 1996), a RI pode ser considerada a vertente tecnológica da Ciência da Informação (CI) e é resultado da relação desta com a Ciência da Computação (CC).

De acordo com(BAEZA-YATES; RIBEIRO-NETO, 2013):

A Recuperação da Informação trata da representação, armazenamento, organização e acesso a itens de informação, como documentos, páginas Web, catálogos online, registros estruturados e semiestruturados, objetos multimídia etc. A representação e a organização dos itens de informação devem fornecer aos usuários facilidade de acesso às informações de seu interesse”(BAEZA-YATES; RIBEIRO-NETO, 2013, p. 1).

De acordo com Fachin (2009), o processo de RI está amparado em coleta, representação, armazenamento, organização e acesso por parte dos usuários. Recuperar informação é “[...] tornar possível e concreto o encontro entre uma pergunta formulada, a informação armazenada e o retorno positivo ao usuário solicitante, quer de forma manual ou automatizada/digital”(FACHIN, 2009, p. 3). De tal modo, para recuperar as informações que verdadeiramente são relevantes, a partir do conjunto de materiais disponíveis, determinadas ferramentas se tornam imprescindíveis.

Neste contexto, são necessárias estruturas de representação e mecanismos de acesso que forneçam tratamento desse novo ambiente e tipo de recurso informacional (LIMA, 2020). São vários os recursos disponíveis na Web que proporcionam liberdade de uso e seleção de tarefas a serem realizadas. Ambientes para diferentes tarefas como *e-commerce*, *e-learning*, *e-science* e *e-government* vêm crescendo gradativamente, tal que as informações ministradas pelos ambientes digitais crescem exponencialmente, e a preocupação de como encontrá-las no momento oportuno tem sido constante (CONEGLIAN et al., 2017).

Portanto, as teorias e técnicas no campo da Organização e Representação e da Informação e a Recuperação da Informação procuram possíveis soluções por meio de elementos teóricos em áreas correlatas que abrangem a esfera interdisciplinar desse campo de estudo, sem perder seu objetivo principal que é de “[...] tratar a informação para o usuário ter a uma recuperação eficiente”(LIMA, 2020, p. 59).

A natureza interdisciplinar da CI é um fato amplamente reconhecido por todos aqueles que estudam a área desde o início de sua concepção (RAYWARD, 1996, p. 4; SARACEVIC, 1996, p. 42). A Ciência da Informação pode ser definida como:

a disciplina que investiga as propriedades e o comportamento da informação, as forças que governam seu fluxo, e os meios de processá-la para otimizar sua acessibilidade e uso. A CI está ligada ao corpo de conhecimentos relativos à origem, coleta, organização, estocagem, recuperação, interpretação, transmissão, transformação e uso de informação. [...] Ela tem tanto um componente de ciência pura, através da pesquisa dos fundamentos, sem atentar para sua aplicação, quanto um componente de ciência aplicada, ao desenvolver produtos e serviços (BORKO, 1968, p. 3, tradução nossa).

Baseando-se na definição de Borko (1968), é possível reconhecer que a Ciência da Informação se atenta às várias etapas relacionadas ao tratamento da informação, desde o armazenamento, recuperação e disseminação. Portanto, na área da CI, a Recuperação da Informação é capaz de auxiliar diversos processos, como a representação e organização, permitindo que a informação seja encontrada pelos usuários.

Saracevic (1996, p. 61–62), também evidencia que, entre os problemas correntes surgidos na CI, está a questão da ecologia informacional surgida com a sociedade da informação e suas Tecnologias da Informação e Comunicação (TIC). Portanto, o isolamento e a incompatibilidade entre esses atores devem ser reduzidos, o que constitui um problema mais social do que técnico, mas que deve ser enfrentado. Uma forma possível de amenizar esses problemas é encontrar técnicas e abordagens como as técnicas de marketing, abrindo assim mais uma fronteira para a pesquisa interdisciplinar (BONIFACIO, 2015).

Considera-se marketing um conjunto de atividades e processos para criar, comunicar, entregar e trocar ofertas que tenham valor para clientes, parceiros e a sociedade em geral (AMA, 2017). Com o marketing surgem as técnicas de *Search Engine Optimization* (SEO) que constituem um conjunto de atividades necessárias para gerar um alto volume de referências bem-sucedidas, originárias de mecanismos de busca e diretórios Web e que têm como objetivo propagar um ambiente informacional digital por meio da análise interna e externa de suas páginas, conteúdos e da quantidade de hiperlinks externos que apontem para essas páginas (JERKOVIC, 2009).

Para atender à necessidade de recuperar informação e, devido a quantidade de informações disponíveis na Web, surgiram os mecanismos de busca (CENDÓN, 2001). Para Gabriel; Kiso, (2020), os mecanismos de busca constituem um sistema de recuperação da informação que tem como finalidade específica ajudar na busca de informações mantidas em ambientes computacionais. São sistemas encarregados de buscar a informação almejada e, portanto, quanto mais adequado for o resultado proporcionado pelo mecanismo e quanto mais rápido ele apresentar esse resultado, mais útil se tornará para quem faz a busca.

A questão é que mecanismos de busca como o Google ou o Bing ajudam os usuários a classificar o conteúdo que aparece nos resultados de busca mais relevante (MACHILL; NEUBERGER; SCHINDLER, 2003; RICHMOND, 2008) em resposta às suas pesquisas, e é por isso que a competição por informações está aumentando. Essa escolha é guiada pelo uso de diferentes algoritmos e mecanismos de indexação da informação. Especificamente, são utilizadas estratégias de disponibilização e recuperação de informação com o objetivo de melhorar o posicionamento dos diferentes sites nos mecanismos de busca, atualmente a principal forma de recuperação de informação disponível na internet. É aqui que o *Search Engine Optimization* (SEO) se destaca (NEVES et al., 2020).

O SEO é um conjunto de técnicas que auxilia o conteúdo da web a aumentar suas chances de aparecer nos primeiros resultados dos *rankings* das ferramentas de busca como Google e Bing (DAVIS, 2006; ENGE et al., 2012; ORDUNA-MALEA; ALONSO-ARROYO, 2017). Uma vez aplicadas adequadamente, tornarão um Website mais visível e melhor posicionado nos mecanismos de busca (LEDFOORD, 2015). A aparição nos primeiros resultados deste ranking leva, portanto, a um maior número de usuários acessando o site (IGLESIAS-GARCÍA; CODINA, 2016). Esse fato, também conhecido como visibilidade na Web, torna necessário o uso de estratégias de posicionamento em mecanismos de busca (DOMENE, 2014; GIOMELAKIS; VEGLIS, 2015).

O algoritmo de classificação (ranqueamento) do mecanismo de busca é uma fusão de procedimentos que definem a localização de um determinado resultado no conjunto total de resultados de uma busca. Sua função é hierarquizar as páginas para entregar os resultados em ordem de relevância para cada usuário. O conceito de relevância é uma forte presença não só nos atuais mecanismos de busca Web, mas de certa forma resume o objetivo das pesquisas em Recuperação de Informação.

Nesse contexto, os objetivos deste estudo são analisar, no âmbito da CI, como as técnicas de SEO podem contribuir para proporcionar uma maior relevância aos ambientes informacionais no processo de Recuperação da Informação.

Para tanto, o trabalho realizou uma revisão de literatura sobre o conceito de “relevância” no contexto da CI, em seguida a sua aproximação com as técnicas de SEO e os estudos Recuperação da Informação.

Tratou-se de pesquisa bibliográfica do tipo revisão narrativa de literatura, considerando seu uso como adequado “[...] para descrever e discutir o desenvolvimento ou o ‘estado da arte’ de um determinado assunto, sob ponto de vista teórico ou contextual”(ROTHER, 2007, p. 5).

Foram analisados dados de fontes bibliográficas(GIL, 2002), coletados em fontes de informações como livros de referência e artigos científicos, buscando-se por definições, sobre o conceito de “relevância” no contexto da CI, em seguida a sua aproximação com as técnicas de SEO, uma vez que o uso das técnicas de SEO estão associadas a processos e atividades do marketing digital. A revisão foi conduzida a partir de experiências e aprendizados prévios(MINAYO, 1994).

2 O CONCEITO DE RELEVÂNCIA

A relevância é um dos elementos fundamentais a serem considerados no desenvolvimento de um Sistema de Recuperação da Informação (SRI) a fim de determinar como processar a informação, solicitá-la e permitir o acesso aos usuários. Neste trabalho, o conceito de relevância percorre todo o caminho de definição do processo de busca e visa ampliar as expressões ou questionamentos levantados pelos usuários, tornando o percurso de recuperação da informação mais satisfatório.

Figueiredo (1977), aponta que no âmbito da CI o conceito de relevância surgiu em meados de 1934, em torno da Lei de Bradford, que consistiu em uma série de estudos que mais tarde foram reunidos sob o nome “Bibliometria”, que procuram medir a produtividade das fontes de informação, relacionando um conjunto de “produtores” (periódicos, autores, termos) a um conjunto de “produtos” (documentos, autoria, frequência). Nesse sentido, Saracevic, em estudos sobre relevância, advertiu “para o fato de que distribuições bibliométricas são, de fato, “distribuições associadas à relevância” (SARACEVIC, 1975, p. 330).

Segundo Figueiredo (1977), Mooers (1950); Perry (1951) e Taube (1955), são os primeiros a fazer uma abordagem sobre a relevância em meados de 1950-1955, mas a primeira indagação mais abrangente do conceito de relevância surgiu em 1958, no decorrer da ‘*International Conference for Scientific Information*’ (ICSI) com Fairthorne, Vickery, Bar-Hillel e outros que acordaram em determinados pontos: (1) relevância não é, exclusivamente, uma propriedade de documentos; (2) relevância não é uma propriedade dicotômica; (3) existe uma ‘relevância para o usuário’ que deve ser julgada.

Conforme Figueiredo (1977, p. 76), por volta de 1960, com Maron e Kuhns, nasceu a primeira teoria de relevância como noção quantitativa, abordando a operação “medida” do ponto de vista probabilístico: “[...] a derivação de uma medida indicando a probabilidade de um documento satisfazer a um determinado pedido”. Segundo a autora, em 1970, Goffman apresentou o conceito de relevância através de um ponto de vista matemático, onde a relevância foi estabelecida como uma dimensão de informação expressa por um documento em relação a uma pergunta.

De acordo com Figueiredo (1977), Goffman elaborou um teorema demonstrando que a relevância determinada simplesmente na base de relações entre a pergunta e o documento não satisfaz o conceito de medida. De acordo com Goffman, (1970 *apud* FIGUEIREDO, 1977), o desenvolvimento contém a exposição ordenada e pormenorizada do assunto pesquisado. Divide-se em seções e subseções que variam em função da abordagem do tema e do método.

Inicia-se com uma revisão teórica acerca do assunto pesquisado, contempla os procedimentos metodológicos utilizados na pesquisa e apresenta os resultados e as discussões.

[...] relevância é determinada não somente pela relação pergunta-documento, mas, também, em termos de relações entre documentos. Em ambos os casos relevância tem as propriedades requeridas por uma medida. Como uma medida, relevância, requer relações entre a pergunta e documentos, e entre os próprios documentos — relevância é uma relação comparativa [...].” (GOFFMAN, 1970 apud FIGUEIREDO 1977, p. 76)

Um outro conceito de relevância é discutido com profundidade por Hjørland. O autor aplica alguns princípios da CI, analisando o domínio a respeito da relevância. Dessa forma, aponta que:

[...] a relevância nunca é ‘de um sistema’, mas sempre ‘humana’ e, portanto, a dicotomia está errada. O determinar quais itens são relevantes em relação a uma determinada meta/tarefa, requer conhecimento do sujeito e é dependente de diferentes teorias/visões. Por conseguinte, os utilizadores dos sistemas de informação não são automaticamente competentes para julgar a pertinência (HJØRLAND, 2010, p. 231 tradução nossa).

De acordo com trabalhos efetuados sobre relevância (SARACEVIC, 1970, 1975), há diversos pontos de vista que se destacam sobre o conceito, que serão apresentados em seguida.

Segundo Saracevic (1975), existem pontos de vista diferentes sobre a relevância, como: ponto de vista do conhecimento do assunto; ponto de vista lógico; ponto de vista do sistema; ponto de vista do destinatário; ponto de vista do conhecimento do destinatário – pertinência e ponto de vista pragmático. (SARACEVIC, 1970) demonstra, inclusive, que os aspectos dos estudos de relevância podem incluir: documentos e representações; questões (*queries*); situações e condições de julgamento; formas de expressão; características humanas (pessoas) e usuários.

Na CI, quando a relevância da prática e da teoria muda para a RI, em vez da organização da informação, a relevância se torna crucial. O gerenciamento da biblioteca está relacionado ao tema e à organização das informações, enquanto na CI a relevância está relacionada à recuperação (SARACEVIC, 2016).

Para avaliar a qualidade dos resultados, os SRI normalmente usam julgamentos humanos que indicam quais resultados são relevantes para uma consulta, ou alguma aproximação de uma ‘informação de referência’ inferida dos cliques dos usuários, ou uma combinação de ambos.

O objetivo básico de todo SRI é fornecer informações relevantes aos usuários em resposta às suas consultas. Desta forma, sistemas são projetados principalmente para responder com informações ou objetos de informação que são potencialmente relevantes para as pessoas. Com o passar dos anos, os SRI fizeram grandes progressos. O primeiro grande avanço prático é o do algoritmo *Hummingbird* desenvolvido pelo Google, que torna os resultados de buscas para os usuários mais precisos. Esse processo é realizado indo além das palavras-chave e considerando não somente termos buscados, mas sinônimos e contextos, como, por exemplo, a localização do usuário e pesquisas anteriores realizadas por ele.

Atualmente o algoritmo *Bidirectional Encoder Representations from Transformers* (BERT) é a principal atualização do mecanismo de busca Google. Ele é baseado em inteligência artificial, capaz de traduzir a linguagem humana para os computadores e permite que o Google entenda o que os usuários estão buscando. De modo geral, é uma atualização criada pelo Google para otimizar e tornar o sistema de buscas mais assertivo. O algoritmo usa a Inteligência Artificial para refinar os resultados de pesquisas e entregar as melhores respostas possíveis para

seus usuários. Assim, cada palavra e frase do que foi buscado é levada em consideração, porque o BERT utiliza uma tecnologia neural para aperfeiçoar a linguagem que os usuários utilizam no mecanismo de busca chamado Processamento da Linguagem Natural.

Mas o desafio inicial de mostrar ao usuário o que realmente é relevante nos resultados de sua pesquisa ainda continua. A dificuldade consiste em dizer ao buscador (máquina), o que é relevante, uma vez que até mesmo os usuários (humanos) não sabem dizer o que é realmente relevante para eles no processo de busca, talvez por se tratar de questões pessoais e de cunho subjetivo. O ser humano usa, intuitivamente, a noção de relevância no processo de comunicação para que seja produtivo e eficaz (SARACEVIC, 2016), da mesma forma, nos SRI, a relevância tem papel fundamental na comunicação com o usuário, sendo preciso responder às suas necessidades informacionais de forma produtiva e eficaz.

Embora os mecanismos de busca considerem aspectos culturais, sociais e contextuais na apresentação de seus resultados, ainda existe a fragilidade que a noção de relevância é humana e não técnica. Esse fato dificulta a formalização por ter muitas variáveis. Nesse sentido, (SILVA; SANTOS; FERNEDA (2013) consideram que a relevância consiste em:

mostrar os resultados possivelmente mais relevantes em forma de ranque (ranking), do mais relevante ao menos relevante. Entretanto, o conceito de relevância é subjetivo e inexato, não podendo ser definido por fórmulas matemáticas e implementadas em sistemas computacionais(SILVA; SANTOS; FERNEDA, 2013, p. 37).

Os pioneiros reconheceram que nem tudo que será recuperado será relevante (MOOERS, 1950; PERRY, 1951; TAUBE, 1955). A preocupação deles era com a não relevância, com a recuperação indesejada, e não com a relevância. Diagnosticaram que a falta de relevância se deve a ineficácia de qualquer representação documental utilizada e/ou a inadequação na forma como estes foram aplicados. Segundo Saracevic (1975, p. 327, tradução nossa):

a relevância foi concebida em termos de indexação, codificação, classificação, manipulações linguísticas, organização de arquivos e, eventualmente, estratégias de análise e busca de questões. Este pensamento levou ao desenvolvimento de uma miríade de esquemas, e à atenção ao processamento e manipulação de entradas, quase com exclusão de outros aspectos. Em trabalhos teóricos, a linguística tem sido objeto de muita atenção, pois acredita-se que ela levará a melhores esquemas de representação (SARACEVIC, 1975, p. 327, tradução nossa).

Dessa maneira, a construção de SRIs efetivos somente será possível quando traduzir a noção de relevância intuitiva em formalismo escrito (LAVRENKO; CROFT, 2009). De qualquer forma, os humanos usam o contexto para solucionar problemas, e os mecanismos utilizam algoritmos e processos tecnológicos associados à probabilidade de apresentação de seus resultados (SARACEVIC, 2016).

Dos primeiros SRIs automáticos ao surgimento dos mecanismos de busca, ocorreram mudanças como as apresentadas anteriormente, e observa-se hoje, que os usuários buscam recuperar informações, “[...] mas a recuperação de informações relevantes para seus propósitos de pesquisa”(ARAUJO, 2012, p. 142).

2.1 A relevância no mecanismo de busca Google

O mecanismo de busca Google permitiu o uso dos seus serviços pela primeira vez em 1998, com buscas fundamentadas em palavras-chave e usando o algoritmo *PageRank*, elaborado por seus fundadores. No entanto, apenas recentemente a ferramenta passou a solidificar sinais mais claros de elementos semânticos, especialmente por meio de seu algoritmo *Hummingbird*. Este algoritmo é uma modificação que afasta a ferramenta do uso de apenas termos singulares para a percepção geral de toda uma frase de consulta (HENSHAW, 2013).

O *Hummingbird* surgiu com o propósito de determinar a "intenção" do usuário e retornar as páginas que melhor correspondem a essa intenção, em oposição ao que era usado anteriormente que considerava a melhor correspondência das palavras-chave na frase da pesquisa (WILLIAMS, 2019). Está principalmente focado em semântica, empenhando-se em entender melhor o propósito do usuário por trás de sua busca. O algoritmo emprega o processamento de linguagem natural e consultas complexas para esta finalidade (HENSHAW, 2013).

Dado que o *Hummingbird* é um dos principais elementos na ferramenta de busca, e a percepção total de seu desempenho não é divulgada publicamente, assim como ocorre com a maior parte do que se reproduz na empresa Google (YSASI, 2016), temos apenas dicas sobre como as buscas no Google acontecem.

As pesquisas por palavra-chave são úteis para recuperar documentos que já foram visualizados, mas geralmente não são adequadas para localizar documentos relacionados em tópicos que ainda não apareceram nos resultados de pesquisa (NAGYPÁL, 2005). Esse tipo de pesquisa, por exemplo, terá dificuldade em distinguir entre as consultas "futebol na escola" e "escola de futebol". O primeiro caso envolve o futebol nas escolas, enquanto o segundo caso envolve as instituições que constituem o esporte. Isso se dá em decorrência dos mecanismos de busca tradicionais não poderem resolver a ambiguidade da consulta porque não conhecem o contexto da pesquisa. Tais mecanismos se concentram na palavra de consulta (sintaxe) sem prestar atenção ao seu significado (semântica), o que afeta negativamente a precisão do documento retornado (QU et al., 2011). Nesse caso, é cada vez mais necessário que os mecanismos de busca incorporem elementos semânticos em seus mecanismos para aprimorar o processo de compreensão das intenções dos usuários por trás de suas pesquisas.

As pesquisas semânticas e os mecanismos de busca que fornecem essas pesquisas são chamados de 'mecanismos de busca semântica'. Os mecanismos de busca semântica visam melhorar resultados da pesquisa por meio da compreensão da terminologia e da intenção relacionado à pesquisa (KALAIVANI; DURAIWAMY, 2012).

Já "[...] Web Semântica não é uma Web separada, mas uma extensão da atual. Nela a informação é dada com um significado bem definido, permitindo melhor interação entre os computadores e as pessoas". (BERNERS-LEE; HENDLER; LASSILA, 2001). Foi assim que os autores definiram o plano de sua equipe de trabalho no *World Wide Web Consortium (W3C)* em 2001 para operar e modificar a rede para o que nós conhecemos hoje. 'Web Semântica' (BERNERS-LEE; HENDLER; LASSILA, 2001; DECKER et al., 2000) é o nome do projeto liderado pelo W3C que pretende incorporar Inteligência e contexto em código *Extensible Markup Language (XML)*. Dessa forma, criando páginas otimizadas nas quais os programas podem interagir.

Embora 'semântica' signifique 'estudar o significado das palavras', Guiraud (1975) reconhece três sequências principais de problemas semânticos:

- a ordem dos problemas psicológicos, que relaciona os estados fisiológicos e psíquicos dos interlocutores nos processos de comunicação de signos;
- a ordem dos problemas lógicos, que estabelece as relações dos signos com a realidade no processo de significação;
- a ordem dos problemas linguísticos, que estabelece a natureza e as funções dos vários sistemas de signos.

Nessa direção, em meados de maio de 2012 o Google lançou o '*Knowledge Graph*', onde foi possível perceber traços semânticos nos resultados de busca nos mecanismos. Também foi descrito como um imenso mapa de elementos do mundo real e as ligações entre eles, uma gigantesca enciclopédia virtual (SULLIVAN, 2020).

Os componentes de pergunta e resposta desenvolvidos para fornecer semântica na pesquisa do Google são combinados com o '*Knowledge Graph*'. O componente de respostas a perguntas é um método que usa consultas baseadas em modelo para mapear a intenção do usuário (STARR, 2013). Esses modelos assumem a forma de perguntas como 'Quem é W?', 'O que é X?', 'Quando Y acontece?' e 'Onde Z aconteceu?'. Nesse caso, W geralmente corresponde a uma pessoa; X corresponde a coisas ou lugares; Y e Z correspondem a eventos. Depois que a consulta do usuário corresponder a um desses modelos, você poderá entender melhor o tipo de informação e a entidade que está procurando (ROZSA; GODOY VIERA; DUTRA, 2019).

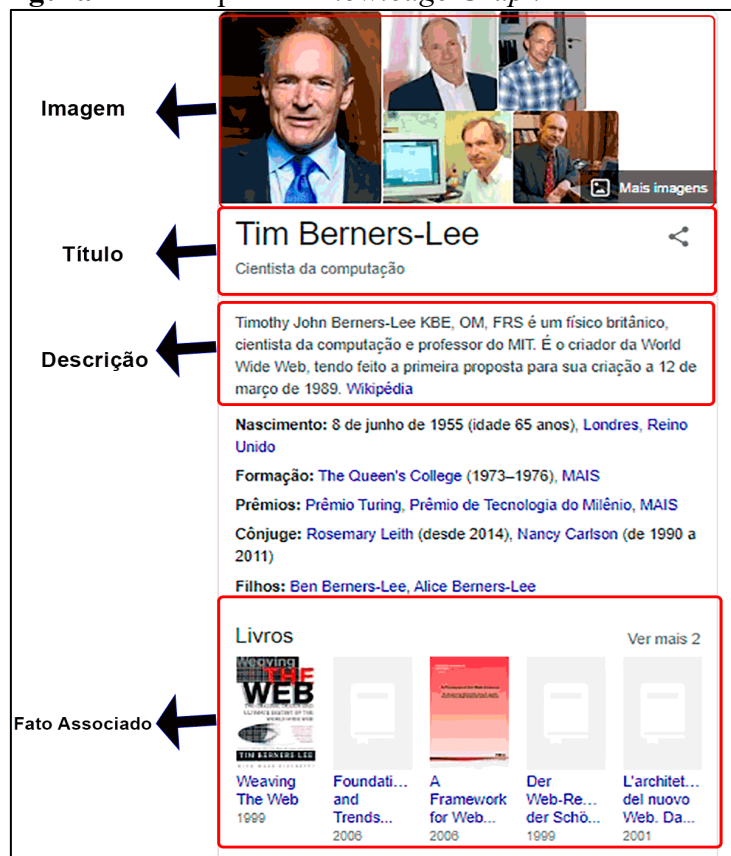
Para (2020), o *Knowledge Graph*, como citado, é uma importante ferramenta que utiliza padrões semânticos e tem como objetivo levar o leitor a confiar nas informações presentes na *Search Engine Result Pages* (SERP) do Google, sem precisar acessar links ou consultar muitos dos resultados apresentados.

Podemos comparar o *Knowledge Graph* com a ontologia da Web Semântica (WS) (incluindo entidades, atributos e seus relacionamentos). Dessa forma, o Google pode visualizar não apenas palavras-chave, mas também identidades em consultas (STARR, 2014). Em geral, as respostas às perguntas adequadas aos modelos mencionados anteriormente foram selecionadas, verificadas e armazenadas no *Knowledge Graph*, fornecendo assim uma base de informações confiável (STARR, 2013).

Esta base de informações compreende em fatos e informações sobre entidades a partir de materiais compartilhados através da Web, bem como de bancos de dados de código aberto e licenciados. A *Wikipédia* é uma fonte comumente citada, mas não é a única. Informações são extraídas de centenas de fontes de toda a Web, por exemplo, de provedores médicos para criar conteúdo cuidadosamente examinado para painéis de informações sobre questões de saúde (BLOG GOOGLE, 2020, tradução nossa). Nesse caso, o Google pode fornecer uma resposta mais específica, extrair conteúdo diretamente dos resultados e apresentar o conteúdo ao usuário na forma de um resumo.

A Figura 1 mostra uma sinopse gerada pelo Google para a pergunta "Quem é Tim Berners-Lee?".

Figura 1 - Exemplo do *Knowledge Graph* em um resultado de busca



Fonte: Imagem gerada pelos autores a partir de um recorte de uma pesquisa feita no Google.

O conteúdo da sinopse é exibido no lado direito da lista de resultados, que contém referências diferentes ao conteúdo relacionado. Além desta sinopse, também pode-se obter outros tipos de resultados para consultas específicas, por exemplo, um resumo das condições climáticas de uma região ao consultar a temperatura ou um diagrama explicativo ao consultar as estatísticas do país / região.

Durante a redação ou edição de um conteúdo, compreende-se que esse trabalho ocorre com foco no cidadão; mas diante de uma fonte algorítmica, outros fatores passam a ter relevância para que a página construída se torne visível em meio a outras inúmeras opções disponíveis (PEDROSA, p.24, 2020).

O uso das técnicas de SEO como uma das ferramentas pode trazer benefícios e influenciar o *Knowledge Graph*, como o crescimento: do alcance, da reputação, da confiança e autoridade do ambiente informacional digital (MARSHALL, 2021).

Para destacar uma conexão potencialmente benéfica entre *Knowledge Graph* e SEO para aumentar visibilidade e alcance, Pecánek (2020), defende o uso de SEO no Google Meu Negócio para aumentar os resultados individuais dos *Knowledge Graph*. A relevância está associada aos mecanismos de busca em cada consulta realizada pelo usuário, os mecanismos avaliam quais os ambientes informacionais digitais em seu índice são capazes de atender os requisitos dos usuários e retomam uma lista de resultados na SERP.

Conteúdos relevantes são de extrema importância para os mecanismos de busca como o Google, já que os resultados de pesquisas devem corresponder a diversas buscas de forma a

satisfazer os usuários. Deste modo algumas características individuais dos ambientes informacionais digitais são consideradas, como corpo do texto, título da página, imagens, entre outros elementos.

Entre esses elementos e os termos principais como as palavras-chave, *metatags*, *title tag*, *meta keywords tag*, entre outros precisam complementar o conteúdo de uma página, a fim de certificar-se de que o título e todo o conteúdo do ambiente informacional digital correspondem exatamente ao que o usuário busca, esses elementos são essenciais para o *Search Engine Optimization* (SEO).

3 SEARCH ENGINE OPTMIZATION (SEO)

Ledford (2015) definiu o *Search Engine Optimization* (SEO) como um conjunto de técnicas que, uma vez aplicadas adequadamente, tornarão um website mais visível e melhor posicionado nos motores de busca. O principal objetivo do SEO é otimizar o ambiente informacional digital para que os mecanismos de busca possam encontrá-los, lê-los, avaliá-los e indicá-los.

Nesse contexto, Ledford (2015) acredita que o SEO é um processo para melhorar a visibilidade de um ambiente informacional digital ou página web nos mecanismos de busca. A otimização desses ambientes para mecanismos de busca está geralmente relacionada a pequenas alterações no conteúdo. Individualmente, essas alterações podem parecer irrelevantes, mas quando combinadas com outras técnicas de otimização, podem dar origem a melhorias consideráveis na qualidade e relevância da informação, na eficácia das pesquisas e na forma como um determinado ambiente informacional digital é encontrado pelos usuários. Existem dois métodos para a aplicação das técnicas de otimização. O método “*White Hat SEO*”, que foi utilizado neste estudo, segue as boas práticas incentivadas pelos mecanismos de busca; e o método “*Black Hat SEO*”, que utiliza técnicas de otimização pouco éticas, podendo levar os mecanismos de busca a indexar uma página pouco relevante como sendo uma página de grande relevância.

Nesse sentido, Corcoglioniti et al., (2016) apontaram que os sistemas de recuperação de informações têm como finalidade determinar o conteúdo da consulta, os documentos relacionados no conjunto de textos, de acordo com sua relevância.

A relevância dos mecanismos de busca (em especial o Google) significa menos carga de trabalho de processamento de informações e mais efeitos cognitivos para satisfação do usuário. De acordo com as teorias da relevância de Yus (2017), Wilson e Sperber, (2005), (CALDEIRA, 2015, p. 92), a função cognitiva envolve:

Implicação textual: ocorre quando suposições resultam da síntese de uma informação antiga com uma informação nova;

Fortalecimento ou enfraquecimento de suposições: ocorre quando a informação nova fornece evidências que fortalecem ou enfraquecem uma informação antiga;

Eliminação de suposições contraditórias: ocorrem quando a informação nova fornece evidências que levam ao abandono de suposições antigas.

A técnica de SEO inclui dois processos principais: otimização *off-page* (fora da página) e otimização *on-page*. A otimização fora da página envolve a construção de *backlinks* em outros ambientes informacionais digitais de boa reputação e, assim, aumenta a autoridade no nível do domínio e no nível da página. Otimização *on-page* SEO requer a otimização de páginas da Web usando palavras-chave alvo no título, trechos e na URL. A inserção de termos adicionais, semanticamente relacionados à palavra-chave alvo, é considerada uma técnica avançada de SEO e vem ganhando popularidade (GENNARO, 2015).

3.1 Os métodos de SEO – *on page*

Esta técnica refere-se tanto ao conteúdo quanto ao código fonte *HyperText Markup Language* (HTML) da página que precisa ser otimizado. O principal fator que influencia esta técnica é como a página web é considerada relevante para a consulta enviada pelo usuário ao mecanismo de busca.

Desenvolver estratégias de SEO podem aumentar as visitas de um determinado ambiente informacional digital e aumentar a sua participação nos resultados do mecanismo de busca. A seguir estão algumas tecnologias classificadas em SEO *on-page*.

- a) **Metatags:** As *metatags* representam um elemento muito importante do SEO porque permitem reconhecer o título (*title tag*), as palavra-chave (*meta keywords tag*), e realizar a descrição (*meta description tag*) do ambiente informacional digital para os motores de busca. Existe uma ampla quantidade de *metatags*. Para Ledford (2015), as mais relevantes são as seguintes: (1) “Título da página”: o título das páginas aparece no topo do navegador e é habitualmente o título que aparece nos resultados de uma pesquisa; ele deve descrever brevemente os conteúdos da página, possuir entre 10 e 70 caracteres com espaços incluídos, conter a palavra-chave e o nome da página deve constar em todos os títulos; (2) “Palavras Chave”: as palavras-chave (*keywords*) servem para definir o conteúdo do ambiente informacional digital e; (3) “Descrição”: a descrição, embora não ofereça grande vantagem na otimização nos mecanismo de busca, deve ser objetiva e clara, ter mais de 50 e menos de 149 caracteres com espaços incluídos para que possa ser mostrada corretamente nos resultados *Search Engine Results Page* (SERP) dos mecanismos de busca. O conteúdo deve estar relacionado com esta página e é importante que a descrição tenha *keywords* incluídas no seu texto;
- b) **Imagens e Atributos Alt:** as imagens são um elemento importante em qualquer comunicação, principalmente para os visitantes do ambiente informacional digital. Elas precisam ter boa resolução e excelente qualidade, mas não devem prejudicar a velocidade do carregamento da página, sendo obrigatória a sua otimização. As imagens devem conter o título referente ao assunto que mostram, precisando também incluir a descrição, para comunicar o público do que se trata (preencher o atributo Alt) e, por questões de acessibilidade, auxiliando na navegação de deficientes visuais, lendo o conteúdo da página (LEDFOURD, 2015);
- c) **Domínio:** a nomeação de um domínio deve ser um processo primordial e muito bem examinado quando se pretende ser descoberto rapidamente na Internet (MEYER, 2006);
- d) **URL do Domínio:** a URL deve ser curta e de fácil memorização (MEYER, 2006);
- e) **Conteúdo:** o conteúdo deve ser escrito com clareza, com parágrafos curtos e com imagens otimizadas e de boa qualidade. O uso de cabeçalhos auxilia no entendimento do conteúdo e contribui também no posicionamento nos mecanismos de busca. É importante utilizar cabeçalhos e as hierarquias de páginas organizadas como, por exemplo, h1, h2, h3, para que os mecanismos de busca percebam que se trata de partes importantes do conteúdo. Para dar destaque a partes importantes é recomendado também formatar certos trechos do texto em itálico e negrito (MACDONALD; WORKBOOK, 2013) e;
- f) **Links Internos:** como parte do conteúdo, podem ser utilizados links para o próprio website (*interlink*) ou para páginas externas (*link building*); para tanto, utilizam-se textos âncora visíveis e clicáveis que, uma vez acionados, direcionam o usuário para as páginas especificadas (LEDFOURD, 2015).

4.2 Os Métodos de SEO –*off page*

Essa técnica refere-se em aumentar a autoridade em seu domínio através do ato de obter links de outros ambientes informacionais digitais. Os maiores fatores SEO *off-page* de página são o número e a qualidade dos *backlinks* para o Website do proprietário.

Quanto aos Métodos de SEO *Off page*, consideram-se os seguintes:

- a) **Idade do Domínio:** A longevidade dos ambientes informacionais digitais é um indicador muito importante para os mecanismos de busca. Geralmente os domínios mais antigos são classificados com uma melhor nota perante os ambientes informacionais digitais mais recentes. No entanto, isso não é suficiente para serem bem classificados. Na verdade, se os seus conteúdos não forem atualizados com frequência, esses ambientes informacionais digitais podem ser prejudicados por darem a entender ao mecanismo de busca que são estáticos e que o seu conteúdo não é interessante (LEDFOURD, 2007);
- b) **Número de Links Externos (*link building*):** Trata-se de uma técnica que tem como objetivo angariar links para um ambiente informacional digital, no sentido de que este cresça nos rankings dos mecanismos de busca (MACDONALD; WORKBOOK, 2013). O número de links externos é muito importante para a criação de *link building*, dado que quanto mais links estiverem a apontar para a uma página, maior importância ela terá para os mecanismos de busca. Essa importância é calculada por um conjunto de algoritmos (PAGE et al., 1999) que inferem que, se um ambiente informacional digital é muito referenciado por outro com boa reputação então é importante, e como tal, merece maior destaque;
- c) **Qualidade de Links Externos:** A qualidade e reputação dos ambientes informacionais digitais que referenciam o ambiente informacional digital a otimizar é muito importante. Esses ambientes informacionais digitais deverão ter uma boa classificação do *PageRank*, para que o *PageRank* do ambiente informacional digital também possa aumentar. (PAGE et al., 1999).

A encontrabilidade (ou *findability*) será o elemento central de todas as estratégias de *Search Engine Optimization* (SEO), pois “é um elemento que se situa entre as funcionalidades de um ambiente informacional tradicional, digital ou híbrido e as características dos sujeitos” (VECHIATO; VIDOTTI, 2014, p. 169). Os autores afirmam que a encontrabilidade está relacionada a processos que compõem o fluxo infocomunicacional e que possibilitam o encontro da informação adequada às necessidades dos usuários em diferentes situações de busca.

Sendo assim, o objetivo do SEO, é proporcionar que a experiência do usuário, se torne gratificante ao realizar uma busca, trazendo aos usuários resultados de busca mais relevantes e precisos.

4 CONSIDERAÇÕES FINAIS

Este trabalho se apresenta como uma fonte de conhecimentos e procedimentos relacionados com a otimização orgânica de ambientes informacionais digitais de mecanismos de busca.

Conclui-se que os mecanismos de busca estão fazendo progressos em prestar assistência ao usuário na formulação da consulta e na interpretação dos resultados. Com a utilização das técnicas de SEO e as práticas advindas da Web Semântica, como a estruturação de dados na

Web, os mecanismos de busca estão mostrando resultados cada vez mais eficientes e precisos, contribuindo assim para uma melhor experiência do usuário no processo de busca.

Podem ser citados como fatores importantes para definir a relevância de um ambiente informacional digital as técnicas de SEO apresentadas e a criação de conteúdos adequados para atrair o público. Com isso, as organizações estarão aptas a criar um relacionamento duradouro e satisfatório com seu público.

Destaca-se que o uso das técnicas de SEO, não asseguram a indexação das páginas pelos mecanismos de busca, já que os mecanismos não revelam as alterações e as melhorias que são realizadas em seus algoritmos, isso se faz necessário para combater as aplicações de *Black Hat* SEO. Portanto, para saber qual técnica obterá os melhores resultados, é necessário nomear e medir algumas delas, pois algumas são viáveis e outras inúteis, é preciso tempo, dedicação, pesquisa e análise para construir e desconstruir o SEO.

O conceito de relevância aplicado junto com as técnicas de SEO, podem incentivar o processo de encontrabilidade e a recuperação da informação em ambientes digitais, pois não basta ser somente encontrado, se faz necessário proporcionar conteúdo relevante para os usuários, através da aplicação das técnicas de SEO.

Logo, entende-se que o uso das técnicas de SEO contribui para a CI, uma vez que corrobora para a encontrabilidade da informação e esta, por sua vez, não está apenas associada à organização, mas pode ser aplicada também a repositórios digitais, bibliotecas e outros ambientes informacionais digitais, promovendo, portanto, informação adequada aos usuários, considerando suas características e limitações no processo de busca de informação.

Para trabalhos futuros, recomenda-se a aplicação dessas técnicas de SEO juntamente com as tecnologias da Web semântica um ambiente informacional digital, a fim de verificar se essas novas formas de representar e organizar as informações nos ambientes digitais contribui no processo de Recuperação de informações.

REFERÊNCIAS

AMERICAN MARKETING ASSOCIATION (AMA). **Definition of marketing**. 2017.

Disponível em: <https://www.ama.org/the-definition-of-marketing-what-is-marketing/>. Acesso em: 10 dez. 2021

ARAÚJO, Vera Maria Araujo Pigozzi De. Sistemas de recuperação da informação: uma discussão a partir de parâmetros enunciativos. **Transinformação**, [S. l.], v. 24, p. 137–143, 2012.

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Recuperação de Informação:- Conceitos e Tecnologia das Máquinas de Busca**. [s.l.] : Bookman Editora, 2013.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic web. **Scientific american**, [S. l.], v. 284, n. 5, p. 34–43, 2001.

BONIFACIO, Everton Lopes. Ciência da informação e marketing: uma interdisciplinaridade possível. **Ciência da Informação**, [S. l.], v. 44, n. 3, 2015.

BORKO, Harold. Information science: what is it? **American documentation**, [S. l.], v. 19, n. 1, p. 3–5, 1968.

CALDEIRA, Fátima Hassan. O mecanismo de busca do Google e a relevância na relação sistema-usuário. **Letrônica**, [S. l.], v. 8, n. 1, p. 91–106, 2015.

CENDÓN, Beatriz Valadares. Ferramentas de busca na Web. **Ciência da Informação**, [S. l.], v. 30, p. 39–49, 2001.

CONEGLIAN, Caio Saraiva; DA COSTA, Ana Maria Jensen Ferreira; MONTEIRO, Silvana Drumond; VIDOTTI, Silvana Aparecida Borsetti Gregório; SEGUNDO, José Eduardo Santarem. A Experiência do Usuário nos mecanismos de busca Knowledge Graph e o Knowledge Vault. **Informação@ Profissões**, [S. l.], v. 6, n. 2, p. 35–59, 2017.

CORCOGLIONITI, Francesco; DRAGONI, Mauro; ROSPOCHER, Marco; APROSIO, Alessio Palmero. Knowledge extraction for information retrieval. *Em: EUROPEAN SEMANTIC WEB CONFERENCE 2016*, **Anais [...]**. : Springer, 2016. p. 317–333.

DAVIS, Harold. **Search engine optimization**. [s.l.] : “ O’Reilly Media, Inc.”, 2006.

DECKER, Stefan; MELNIK, Sergey; VAN HARMELEN, Frank; FENSEL, Dieter; KLEIN, Michel; BROEKSTRA, Jeen; ERDMANN, Michael; HORROCKS, Ian. The semantic web: The roles of XML and RDF. **IEEE Internet computing**, [S. l.], v. 4, n. 5, p. 63–73, 2000.

DOMENE, Fernando Maciá. **Técnicas avanzadas de posicionamiento en buscadores**. [s.l.] : Anaya multimedia, 2014.

ENGE, Eric; SPENCER, Stephan; FISHKIN, Rand; STRICCHIOLA, Jessie. **The art of SEO**. [s.l.] : “ O’Reilly Media, Inc.”, 2012.

FACHIN, Gleisy Regina Bories. Recuperação inteligente da informação e ontologias: um levantamento na área da Ciência da Informação. **BIBLOS**, [S. l.], v. 23, n. 1, p. 259–283, 2009.

FIGUEIREDO, Laura Maia De. O conceito de relevância e suas implicações. [S. l.], 1977.

GABRIEL, Martha; KISO, Rafael. **Marketing na Era Digital - Conceitos, Plataformas e Estratégias**. [s.l.] : Novatec Editora, 2020.

GENNARO, Susan. Brevity and clarity: titles, key words, and search engine optimization. **Journal of Nursing Scholarship**, [S. l.], v. 47, n. 3, p. 195, 2015.

GUIRAUD, P. **A semântica**. 2. ed. Rio de Janeiro: Difel, 1975.

HENSHAW, J. **What Google’s Hummingbird Update Means for Content Marketers**. 2013. Disponível em: <https://smallbusiness.yahoo.com/advisor/google-hummingbird-means-content-marketers-143119302.html>. Acesso em: 22 de jun. 2020.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. [s.l.] : Atlas São Paulo, 2002. v. 4

GIOMELAKIS, Dimitrios; VEGLIS, Andreas. Employing search engine optimization techniques in online news. **Studies in media and communication**, [S. l.], v. 3, n. 1, p. 22–33, 2015.

HJØRLAND, Birger. The foundation of the concept of relevance. **Journal of the american society for information science and technology**, [S. l.], v. 61, n. 2, p. 217–237, 2010.

IGLESIAS-GARCÍA, Mar; CODINA, Lluís. Los cibermedios y la importancia estratégica del posicionamiento en buscadores (SEO). **Opción**, [S. l.], v. 32, n. 9, p. 929–944, 2016.

JERKOVIC, John I. **SEO warrior: essential techniques for increasing web visibility**. [s.l.] : “O’Reilly Media, Inc.”, 2009.

KALAIVANI, S.; DURAI SWAMY, K. Comparison of question answering systems based on ontology and semantic web in different environment. *Em: JOURNAL OF COMPUTER SCIENCE 2012*, **Anais [...]**. : Citeseer, 2012.

LAVRENKO, Victor; CROFT, W. Bruce. **A generative theory of relevance**. [s.l.] : Springer, 2009. v. 26

LEDFORD, Jerri L. **Search engine optimization bible**. [s.l.] : John Wiley & Sons, 2015. v. 584

LIMA, Gercina Ângela De. Organização e representação do conhecimento e da informação na web: teorias e técnicas. **Perspectivas em ciência da informação**, [S. l.], 2020.

MACDONALD, J.; WORKBOOK, S. E. O. Fitness. The Seven Steps to Search Engine Optimization Success on Google. **EUA: JMInternet Group**, [S. l.], 2013.

MARSHALL, Oliver. **Search Engine Optimization and the connection with Knowledge Graphs**. , 2021.

MEYER, L. **Pagerank and beyond-The Science of Search Engine Rankings**. EUA: Princeton University Press, , 2006.

MINAYO, Maria Cecília de Souza. Ciência, técnica e arte: o desafio da pesquisa social. **Pesquisa social: teoria, método e criatividade**, [S. l.], v. 21, p. 9–29, 1994.

MONTEIRO, Silvana Drumond; RODAS, Cecílio Merlotti; VIDOTTI, Silvana Aparecida Borsetti Gregório. A busca e o eye tracking: um olhar semiótico sobre o knowledge graph. **Em Questão**, [S. l.], v. 26, n. 1, p. 304–326, 2020.

MOOERS, Calvin N. Zatoncoding applied to mechanical organization of knowledge. **American documentation**, [S. l.], v. 2, n. 1, p. 20–32, 1951.

MOOERS, Calvin S. EDITOR’S CORNER: " Coding, Information Retrieval, and the Rapid Selector". **Journal of the American Society for Information Science**, [S. l.], v. 1, n. 4, p. 225, 1950.

NAGYPÁL, Gábor. Improving information retrieval effectiveness by using domain knowledge stored in ontologies. *Em: OTM CONFEDERATED INTERNATIONAL CONFERENCES" ON THE MOVE TO MEANINGFUL INTERNET SYSTEMS" 2005, Anais [...].* : Springer, 2005. p. 780–789.

NEVES, Barbara Coelho; SANTANA, Ramon Davi; DE ASSUNÇÃO GOMES, Dulcinéia Vieira; DE JESUS REIS, Makson. Se estou no Google, logo existo: técnicas de alavancagem e visibilidade de um periódico científico em motores de busca por meio de técnicas de SEO. **Informação & Informação**, [S. l.], v. 25, n. 4, p. 402–430, 2020.

ORDUNA-MALEA, Enrique; ALONSO-ARROYO, Adolfo. **Cybermetric techniques to evaluate organizations using web-based data**. [s.l.] : Chandos Publishing, 2017.

PAGE, Lawrence; BRIN, Sergey; MOTWANI, Rajeev; WINOGRAD, Terry. **The PageRank citation ranking: Bringing order to the web**. [s.l.] : Stanford InfoLab, 1999.

PECÁNEK, M. How to Use Google Analytics to Improve SEO Performance. ahrefs. com. **Lainattu**, [S. l.], v. 15, p. 2020, 2020.

PEDROSA, L. L. C. **SEO on-page no jornalismo: fatores algorítmicos como lide aos buscadores**. Tese (doutorado) – Faculdade de Arquitetura, Artes e Comunicação, Universidade Estadual Paulista Julio de Mesquita Filho. Bauru, 2020. Disponível em: <https://repositorio.unesp.br/handle/11449/193479>. Acesso em: 27 jan. 2022.

PERRY, James W. Superimposed punching of numerical codes on hand-sorted punch cards. **American documentation**, [S. l.], v. 2, n. 4, p. 205–212, 1951.

QU, Junhua; WEI, Chao; WANG, Wenjuan; LIU, Fei. Research on a retrieval system based on semantic web. *Em: 2011 INTERNATIONAL CONFERENCE ON INTERNET COMPUTING AND INFORMATION SERVICES 2011, Anais [...].* : IEEE, 2011. p. 543–545.

RAYWARD, W. Boyd. The history and historiography of information science: some reflections. **Information processing & management**, [S. l.], v. 32, n. 1, p. 3–17, 1996.

ROTHER, Edna Terezinha. Systematic literature review X narrative review. **Acta Paulista de Enfermagem**, [S. l.], v. 20, p. v–vi, 2007.

ROZSA, Vitor; GODOY VIERA, Angel Freddy; DUTRA, Moisés. Aplicação de tecnologias da Web Semântica em motores de busca na Internet. **Investigación bibliotecológica**, [S. l.], v. 33, n. 78, p. 165–191, 2019.

SARACEVIC, Tefko. Introduction to information science. [S. l.], 1970.

SARACEVIC, Tefko. Relevance: A review of and a 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, Tefko. Ciência da informação: origem, evolução e relações. **Perspectivas em ciência da informação**, [S. l.], v. 1, n. 1, 1996.

SARACEVIC, Tefko. The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really? **Synthesis lectures on information concepts, retrieval, and services**, [S. l.], v. 8, n. 3, p. i–109, 2016.

SILVA, Renata Eleuterio Da; SANTOS, Plácida Leopoldina Ventura Amorim da Costa; FERNEDA, Edberto. Modelos de recuperação de informação e web semântica: a questão da relevância. **Informação & Informação**, [S. l.], p. 27–44, 2013.

STARR, B. **Demystifying the Google Knowledge Graph**. 2014. Disponível em: <http://searchengineland.com/demystifying-knowledge-graph-201976>. Acesso em: 15 jun. 2018.

STARR, B. **Google Hummingbird**: when evolutionary becomes revolutionary. 2013. Disponível em: <http://searchengineland.com/google-hummingbird-when-evolutionary-becomes-revolutionary-173740>. Acesso em: 15 jun. 2018.

SULLIVAN, Danny. **Google's Knowledge Graph and Knowledge Panels**. 2020. Disponível em: <https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/>. Acesso em: 5 jan. 2023.

TAUBE, Mortimer. Storage and retrieval of information by means of the association of ideas. **American Documentation (pre-1986)**, [S. l.], v. 6, n. 1, p. 1, 1955.

VECHIATO, Fernando Luiz; VIDOTTI, Silvana Aparecida Borsetti Gregorio. Encontrabilidade da informação: atributos e recomendações para ambientes informacionais digitais. [S. l.], 2014.

WILLIAMS, Andy. **SEO 2019: Actionable, Hands-on SEO, Including a Full Site Audit:[the Ultimate Step-by-step Visual Guide to Top 10 Rankings]**. [s.l.] : ezSEONews.com, 2019.

WILSON, Deirdre; SPERBER, Dan. Teoria da relevância. **Linguagem em (Dis) curso**, [S. l.], v. 5, p. 221–268, 2005.

YSASI, E. **What is Google's Semantic Search?** 2016

YUS, Francisco. Putting relevance at centre stage in research on human activity on the Internet. **Applications of Relevance Theory: From Discourse to Morphemes**, ed. by Agnieszka Piskorska, and Ewa Walaszewska, [S. l.], p. 86–102, 2017.