**UNIVERSIDADE DE SÃO PAULO
ESCOLA DE COMUNICAÇÕES E ARTES**

ANDRÉA NUNES DE ARAÚJO

**A contribuição da Indexação no processo de seleção de palavras-chave no
Google Adwords**

São Paulo
2018

ANDRÉA NUNES DE ARAÚJO

**A contribuição do processo de Indexação na seleção de palavras-chave para
o Google Adwords**

VERSÃO CORRIGIDA

Dissertação apresentada ao Programa de Mestrado Profissional em Gestão da Informação, da Escola de Comunicações e Artes da Universidade de São Paulo, para obtenção do título de Mestre em Ciência da Informação.

Área de Concentração: Organização, Mediação e Organização da Informação.

Orientação: Profa. Dra. Cibele A. C. Marques dos Santos

São Paulo
2018

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

**Catálogo na Publicação
Serviço de Biblioteca e Documentação
Escola de Comunicações e Artes da Universidade de São Paulo
Dados inseridos pelo(a) autor(a)**

Araújo, Andréa Nunes de
A contribuição do processo de Indexação na seleção de
palavras-chave para o Google Adwords / Andréa Nunes de
Araújo ; orientadora, Cibele A. C. Marques dos Santos. --
São Paulo, 2018.
78 p.: il. + Pen-drive.

Dissertação (Mestrado Profissional) - Escola de
Comunicações e Artes / Universidade de São Paulo.
Bibliografia
Versão corrigida

1. Indexação 2. Recuperação da Informação 3. Google Adwords
4. Marketing Digital I. Santos, Cibele A. C. Marques dos
II. Título.

CDD 21.ed. -
020:658.8

ANDREA NUNES DE ARAÚJO

A contribuição da Indexação no processo de seleção de palavras-chave no Google Adwords

Dissertação apresentada ao Programa de Mestrado Profissional em Gestão da Informação, da Escola de Comunicações e Artes da Universidade de São Paulo, para obtenção do título de Mestre em Ciência da Informação.

Área de Concentração: Organização, Mediação e Organização da Informação.

Orientação: Profa. Dra. Cibele A. C. Marques dos Santos

Aprovado em: __/__/____

BANCA EXAMINADORA

Prof^a Dr^a Cibele A. C. Marques dos Santos (orientadora)
Universidade de São Paulo

Prof^a Dr^a Giovana Deliberali Maimone
Universidade de São Paulo

Prof. Dr. Fabiano Ferreira de Castro
Universidade Federal de São Carlos

Prof. Dr. Fábio Assis Pinho
Universidade Federal de Pernambuco

As energias universais da criação,
que traz luz a novos conhecimentos.

AGRADECIMENTOS

A princípio, o significado de gratidão é o reconhecimento pela ajuda, favor ou auxílio recebido. Portanto, estou extremamente grata a todos aqueles que contribuíram na realização desta pesquisa.

Agradeço aos amigos com suas palavras de apoio, em especial aos colegas de trabalho que me motivaram nesta missão, diante da compreensão de que é impossível fazemos algo sozinhos. Muitas vezes uma simples palavra de incentivo, eleva os nossos sentimentos em busca de superação aos desafios da vida.

Com muito amor agradeço à minha família e amigos, pelo carinho e compreensão da minha ausência nos eventos familiares, em prol da dedicação necessária destinada ao desenvolvimento das tarefas solicitadas.

Em estado de graça, estou feliz pelo encontro com a minha orientadora Cibele A. C. Marques dos Santos, e principalmente pela sua atenção, carinho, compreensão e aos seus conselhos valiosos na concretização deste trabalho.

Agradeço às orientações recebidas dos docentes Fabiano Ferreira de Castro e Giovana D. Maimone participantes da banca examinadora, que a partir dessas contribuições enriqueceram ainda mais essa pesquisa.

Agradecimentos à Universidade de São Paulo, em destaque a Escola de Comunicações e Artes pelo Programa de Pós-Graduação de Ciência da Informação, aos docentes que fazem parte do Mestrado Profissional pela oportunidade e privilégio no compartilhamento de conhecimentos que, este espaço possa se expandir e trazer ainda mais benefícios a nossa sociedade como um todo.

Agradeço e presto à minha admiração pela empresa Google, com toda tecnologia desenvolvida e usufruída por toda humanidade.

E por último, agradeço pelo divino por mais essa oportunidade de expansão da consciência, assim espero estar aberta a novas inspirações e ser um instrumento de utilidade a minha volta.

“...sob o ponto de vista da língua,
o signo é como uma moeda:
este vale por certo bem que permite comprar,
mas vale também com relação a outras moedas,
de valor mais forte ou mais fraco”
(BARTHES, 2012, p. 22).

RESUMO

ARAÚJO, Andréa Nunes. **A contribuição da indexação no processo de seleção de palavras-chave no Google Adwords**. São Paulo, 2018. 78 f. Dissertação (Mestrado Profissional em Gestão da Informação) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2018.

Devido ao estabelecimento de negócios pela Internet, o Google alcançou notoriedade mundial, tornando-se atrativo às empresas como canal do Marketing Digital na utilização da ferramenta Google Adwords. Em vista disso, o presente estudo consistiu no levantamento bibliográfico do Marketing de Busca e da Biblioteconomia e Documentação, no qual o principal objetivo é contribuir através do uso de técnicas de Indexação, no processo de seleção de palavras-chave, e conduzi-las na ferramenta de anúncios do Google. Os métodos desta pesquisa foram estabelecidos pela Representação Documentária, com a finalidade de contribuir na diminuição de palavras-chave que oneram as campanhas no Adwords. A partir da coleta dos assuntos mais acessados no Google Trends, no período de quatro meses (jan-abr) de 2017, e a realização de testes em janeiro de 2018 no planejador de palavras-chave do Google, apresentamos a avaliação dos resultados comparativos entre os dados coletados e os dados da lista, com a finalidade de responder se o uso das técnicas de indexação podem ser ou não favoráveis nas criações de campanhas aos clientes brasileiros neste mecanismo de busca. Concluiu-se que a contribuição da Indexação é favorável as campanhas publicitárias no Google, e também requer a inclusão do bibliotecário, já que se trata do profissional com habilidades e competências necessárias na execução desta tarefa, que ficará atento as influências exercidas pelos algoritmos tanto nos valores de palavras-chave quanto na recuperação desses termos ao público de interesse no mecanismo de busca.

Descritores: 1. Indexação. 2. Recuperação da Informação. 3. Google Adwords. 4. Marketing Digital.

ABSTRACT

ARAÚJO, Andréa Nunes de. **The contribution of indexing in the keyword selection process for Google Adwords.** 2018. 78 f. Dissertation (Master in Information Management) – Escola de Comunicações e Artes, Universidade de São Paulo, 2018.

Due to the establishment of Internet business, Google has achieved world renown, becoming attractive to companies as a channel of Digital Marketing in the use of Google Adwords tool. Therefore, the present study consisted of a bibliographical survey of Search Marketing and Librarianship and Documentation, in which the main objective is to contribute using Indexing techniques, in the process of selection of keywords, and to conduct them in the Google Ad Tool. The methods of this research were established by the Documentary Representation, with the purpose of contributing to the reduction of keywords that affect the campaigns in Adwords. From the most commonly accessed subjects in Google Trends in the four-month period (Jan-Apr), 2017, and testing in January 2018 on Google's keyword planner, we present an assessment of the comparative results between the collected data and the data of the list, in order to answer if the use of the indexing techniques can be or not favorable in the creations of campaigns to the Brazilian clients in this search engine. It was concluded that the contribution of Indexation is favorable to advertising campaigns in Google, and also requires the inclusion of the librarian, since it is the professional with the necessary skills and competences in the execution of this task, which will be attentive to the influences exerted by the algorithms both in keyword values as well as the retrieval of those terms to the public of interest in the search engine.

Keywords: 1. Indexing. 2. Information Retrieval. 3. Google Adwords. 4. Marketing Digital.

LISTA DE GRÁFICOS

Gráfico 1 – Percentual de domicílios com utilização da internet.....	17
Gráfico 2 – Acesso às categorias do Google Trends.....	60
Gráfico 3 – Período mais acessado na categoria Esporte.....	60

LISTA DE QUADROS

Quadro 1 – Cálculo do algoritmo Page Rank	49
Quadro 2 – Organização e testes da lista de palavras-chave	63
Quadro 3 – São Paulo Futebol Clube X São Paulo	65
Quadro 4 – Uso de Grêmio X Grêmio Esportivo Osasco	65

LISTA DE FIGURAS

Figura 1 – As etapas da indexação aplicadas à seleção de palavras-chave em loja virtual.....	31
Figura 2 – Central de Ajuda do Google Adwords para seleção de palavras-chave	37
Figura 3 – Tabela de Correspondência	38
Figura 4 – Passo 1: criação de Conta no Adwords	38
Figura 5 – Passo 2: definição de Idioma e Local	39
Figura 6 – Passo 3: Planejador de palavras-chave	39
Figura 7 – Passo 4: Criação de anúncio	40
Figura 8 – Passo 5: Orçamento de campanha	40
Figura 9 – Relatório do Google Analytics	42
Figura 10 – Tags da lista de palavras-chave sobre futebol.....	62
Figura 11 – Teste com as palavras-chave camisa e camiseta	64
Figura 12 – Etapas de seleção de palavras-chave.....	67
Figura 13 – Mapa Mental do Futebol	67

SUMÁRIO

1	INTRODUÇÃO	14
2	O PROCESSO DA INDEXAÇÃO NA SELEÇÃO DE PALAVRAS-CHAVE.....	23
2.1	Os estágios de Indexação e suas aplicações em ambientes digitais	24
3	<i>SEARCH ENGINE MARKETING (SEM)</i>	36
3.1	As estratégias do Google Adwords	37
4	OS REFLEXOS DOS ALGORITMOS NOS ANÚNCIOS DO GOOGLE	45
4.1	Inteligência Artificial, Algoritmos e Web Semântica no Google.....	46
4.1.2	O contexto cognitivo na recuperação da informação em sistemas de buscas	54
5	MATERIAIS E MÉTODOS	58
5.1	Coleta de dados.....	59
6	RESULTADOS	60
6.1	Análise dos dados	63
7	CONCLUSÕES	69
	REFERÊNCIAS	75
	ANEXO 1 - Análise de Consumo do Futebol.....	00
	APÊNDICE A - Tabela de Aplicação das Palavras-Chave	76
	APÊNDICE B - Relatório <i>Google Trends</i>	00
	APÊNDICE C - Tabelas e Gráficos.....	00
	APÊNDICE D - Lista de Palavras-Chave.....	00
	APÊNDICE E - Aplicação de Lista de Palavras-Chave no <i>Google Adwords</i>	00
	APÊNDICE F - Tabela de Palavras-Chave.....	00

1 INTRODUÇÃO

O objeto desta pesquisa é a inserção da Indexação, com enfoque na Biblioteconomia e Documentação para a seleção de palavras-chave na utilização nos anúncios em links patrocinados no *Google Adwords*¹.

Neste cenário de análise, utilizaremos outra ferramenta do pacote Google, oferecida aos seus anunciantes, conhecida como *Google Trends*². Assim, para a presente pesquisa foram coletados na ferramenta *Trends*, o fluxo dos assuntos, para identificar e selecionar o assunto mais acessado pelos buscadores, com a finalidade de elaborar uma lista de palavras, para testá-las no planejador de palavras-chave, e compará-las com os termos coletados.

Geralmente, as estratégias utilizadas para seleção de palavras-chave pertencem ao Marketing Digital, que não possui o foco da Organização e Representação do Conhecimento, da recuperação da informação, e principalmente dos estudos de indexação da Biblioteconomia e Documentação. Estes estudos são necessários para obtenção de bons resultados de pesquisas em ambientes digitais. Em virtude disto, pretendeu responder se o uso das técnicas de Indexação pode ser ou não favorável na criação de campanhas no Google aos clientes brasileiros, no mecanismo de buscas.

Esse estudo mostra-se relevante, por causa das relações de negócios estabelecidas pela rede através da Internet, onde os mecanismos de buscas são procurados pelos internautas o tempo todo. Por conseguinte, o Google, um buscador que tinha inicialmente a ambição de oferecer a informação contida na rede global aos seus usuários de forma precisa, alcançou a notoriedade mundial, assim atraiu o interesse de empresas como canal do Marketing Digital. Em consequência, houveram mudanças nos resultados de buscas, que sofrem influência dos interesses comerciais do buscador. Portanto, este cenário foi estimulado por algumas mudanças sociais e tecnológicas, apresentados nesta introdução por alguns pesquisadores dos aspectos sociais da Internet.

¹*Google Adwords* é uma ferramenta que permite a uma empresa aparecer em destaque no Google e comprar espaço na página principal, o que funciona como um leilão de palavras-chave, a empresa disposta a oferecer o maior lance por uma determinada palavra, terá maior relevância no resultado de busca. A partir do dia 24 de julho de 2018, o produto passou a ser chamado de *Google Ads*.

² *Google Trends* é uma ferramenta do Google, que apresenta os assuntos mais acessados na Internet, e que impactam na concorrência das palavras-chave, de forma que os assuntos mais interessantes resultarão em palavras-chave de maior custo ao anunciante.

Uma mudança significativa foi a “Terceira Revolução Industrial”, marcada pelas indústrias de serviços, principalmente as do conhecimento. Muitas delas surgiram dentro das instituições de ensino. Como relata Burke (2012, p. 331),

O surgimento da “economia da informação” e de empresas movidas pelo conhecimento como a Apple e a Microsoft, em que a produtividade depende da pesquisa, levou por sua vez ao surgimento da sociedade “pós-industrial” também conhecida como “sociedade da informação” ou sociedade do conhecimento”, em que o cotidiano é permeado de novas formas de conhecimento” (BURKE, 2012, p. 331).

Em efeito cascata, a partir dessa “economia da informação” deflagraram-se outras transformações no mercado. Castells (2003, p. 58) traz uma reflexão sobre a Rede em seu livro “A Galáxia da Internet”, como isto solidificou a Era da Informação, em suas palavras: “redes de informação energizadas pela Internet”, portanto as relações de negócios alcançam maior flexibilidade e adaptabilidade para adequar-se ao ambiente de mudanças contínuas. Então, a sociedade passa a ter uma economia mais aberta, assim tornando global o seu capital, a sua produção e o seu comércio. Nas características sociais nota-se maior liberdade individual e a comunicação mais aberta. Por último, os avanços da computação e telecomunicação propiciaram essa revolução.

Ao findar o século XX, a utilização da Internet aumentou, em consequência as grandes empresas de tecnologias buscaram melhorar as suas interfaces de pesquisas, o que levou ao desenvolvimento de agentes inteligentes para personalização da informação. Pariser (2012, p.13) diz: “a fórmula dos gigantes da internet é simples: quanto mais personalizadas forem suas ofertas de informação, mais anúncios eles conseguirão vender e maior será a chance de que você compre os produtos oferecidos”. Isto culminou em 2009, quando o Google apresentou seu algoritmo de filtros de pesquisa, desenvolvido em função do comportamento do usuário.

A partir disso, começou a disputa pelos dados dos usuários, inclusive surgindo empresas de tecnologias responsáveis pela extração de informações, baseados no perfil de pesquisa. Por isso, estes dados passam a ser um diferencial de mercado, o que estimula as empresas de tecnologias no desenvolvimento de algoritmos de personalização. De fato, essa situação é muito bem aceita pelos consumidores da Internet, pois com o grande volume informacional, existem os filtros para localizar aquilo que é relevante em sua pesquisa.

Há uma década, já previam como os filtros poderiam influenciar as decisões dos indivíduos nas informações encontradas na Internet, e o quanto isto pode tomar uma grande dimensão de mercado. Nesta época, Kerckhove (2009, p.79) declarou: “No futuro as oportunidades permitidas irão para os *gatekeepers*, assistentes inteligentes, caçadores de informação especificamente treinados para saber as últimas sobre qualquer assunto”.

Portanto, Pariser (2012, p. 35) relata que o Google investiu massivamente em tecnologia para conhecer o comportamento de seu público. Assim, desenvolveram aplicativos chamados indicadores de cliques, ou seja, através do Gmail e/ou qualquer outro serviço oferecido pelo mecanismo, a forma como se obtém uma gama de informação, que será analisada para trazer os resultados mais relevantes aos usuários.

E com isso, foi alcançado um nível de detalhamento de perfil, com diversas classificações e categorias que permitem oferecer a informação mais precisa. Todo esse arsenal de dados e os algoritmos, responsáveis por serem os elos do usuário e a informação buscada, tornou o mecanismo muito atraente às empresas, então a informação passou a ter valor e a ser disputada por aqueles que pagam mais para ter maior relevância no resultado de busca do Google.

Anderson (2006, p. 37) explica em seu livro o fenômeno cultural e socioeconômico estimulado pela internet intitulado por ele como “A Cauda Longa”, no qual encontramos uma grande variedade de produtos para atender todo e qualquer nicho de mercado. Filtros são imprescindíveis para evitar algo que acontece nos navegadores da web chamado por Anderson (2006, p. 37) de “Custos de Busca”, isto é, muita informação que pode ocasionar grande perda de tempo, frustração e irritação na busca de um produto, ou até mesmo perda de dinheiro devido ao fato de as opções encontradas serem mais caras.

As estratégias de indexação de palavras-chave podem criar filtros para destacar os produtos em uma pesquisa feita, por exemplo, no Google. Anderson (2006, p.37) destaca da seguinte maneira essa situação: “Quando se conjugam essa grande expansão da variedade e a notável eficácia dos filtros facilitam a seleção, a curva da demanda se torna mais horizontal e mais longa”. O autor referindo-se aos filtros de busca, permite a relação com a Indexação, quanto mais termos selecionados para representar um assunto, maior será a possibilidade de encontrar um desses termos no buscador, o que proporcionará aos usuários a localização de uma grande variedade de produtos.

O artigo de Tim Bourgeois (2016, p. 1), diz que o Google ao lado do Facebook são os dois grandes gigantes do Marketing Digital, e no primeiro semestre de 2016 receberam 85 % do mercado total de publicidade na Internet. Moral e Gonzales (2013, p. 575), falam da procura de empresas de médio e pequeno porte por anunciar no *Google Adwords*, devido à possibilidade de delimitar o orçamento de campanha na ferramenta tornando-se assim mais acessível aos orçamentos mais limitados.

No Brasil, acompanhamos um crescimento vertiginoso no uso da Internet Móvel, por intermédio de planos populares oferecidos pelas operadoras de telecomunicação, e também o

barateamento dos aparelhos celulares, aumentaram o acesso dos indivíduos em geral, principalmente a população de baixa renda. O total de pessoas que acessam a internet pelo microcomputador são 70,1%, pelo telefone móvel celular em 92,1%, pelo tablet em 21,1%, pela televisão em 7,5% e por último 1,0% em outro tipo de equipamento, ou seja, o acesso à Internet pelo telefone celular superou outros meios, como mostra o gráfico do IBGE em 2015 (Gráfico 1).

Gráfico 1 – Percentual de domicílios com utilização da internet

Fonte: IBGE, Pesquisa Nacional por Amostra de Domicílios 2013-2015.

Quando observamos os dados do *Alexa*³ relacionados ao Brasil, o Google representa o site que ocupa o topo dos quinhentos sites mais acessados na Web. O mecanismo de busca é o site mais procurado pelos brasileiros e se o uso da internet aumentar no país, também aumentará a procura por mecanismos de buscas. Por essa razão, quanto mais potencializados forem os tipos de filtros destas ferramentas, mais se obterá como retorno nos mecanismos de buscas.

Dessa maneira, a motivação deste estudo vem da observação do marketing de busca, e da sua utilidade para anunciantes na Internet. Em uma oficina sobre indicadores de métricas, percebe-se as estratégias do Marketing Digital, as quais são utilizadas na criação de campanhas no buscador. No entanto, as métricas e estratégias utilizadas não apresentam as técnicas consagradas da Biblioteconomia e Documentação, como por exemplo, a Indexação, que talvez possa trazer maior precisão e relevância nestes anúncios.

³ O *Alexa* Internet, serviço que monitora ao tráfego dos sites na rede, apresentando as informações divididas em *Global*, *By Country* (países) e *By Category* (categorias). Disponível em: < <https://www.alexa.com/>>. Acesso em: 15 de mar. 2017.

O principal objetivo da pesquisa é estudar o uso da Indexação na seleção de palavras-chave para campanhas no *Google Adwords*. Assim, como estudo de caso, será criada uma lista a partir do assunto mais acessado no *Google Trends*, a qual estará delimitada ao território brasileiro e no idioma português. Posteriormente, pretende-se apresentar os possíveis resultados benéficos a serem aplicados pelos clientes brasileiros em suas campanhas na ferramenta de anúncio do Google. Estão definidos os seguintes objetivos específicos, apresentar as principais estratégias do *Google Adwords*; refletir sobre as influências dos algoritmos nos anúncios do Google; coletar os assuntos mais acessados no *Google Trends*.

As principais questões abordadas nesta pesquisa estão nos capítulos (2, 3 e 4), a teoria de Indexação no processo de seleção de palavras-chave, o funcionamento do *Google Adwords*, os aspectos relacionados aos algoritmos do Google e suas influências nos resultados de busca, e para isto foram selecionados pesquisadores da área de Biblioteconomia e Marketing Digital.

No capítulo 2, partimos aos estudos de Ciência da Informação e da Biblioteconomia, em particular da área de Organização e Representação do Conhecimento, responsável pela organização dos fluxos de informação, em torno de atividades como descrição do documento, indexação e classificação em unidades de informação e ambientes digitais, que incluem os métodos automatizados de recuperação da informação.

Com os autores da Biblioteconomia, Lancaster (2004), Borko (1978), Fungman (1993), Cleveland e Cleveland (2001) faremos uma discussão sobre a teoria da indexação, a sua definição, sua função e aplicabilidade na internet. Iniciaremos o capítulo abordando o tema da Leitura Documentária, em que trazemos algumas pesquisadoras da área para explicar os procedimentos seguidos pelos bibliotecários no processo de leitura e representação de um assunto, e o seu papel de mediação entre a mensagem e o receptor.

Os estudos nos dizem que os índices são produtos da indexação, e surgiram da necessidade humana, como forma de organização da memória com a finalidade de sua posterior recuperação. Portanto, os índices existem desde a antiguidade para facilitar a vida dos indivíduos, pois identificam objetos e informações para algum tipo de uso. Conforme a humanidade evolui, maior é a necessidade de utilização de índices. Assim, com o surgimento da escrita, a comunicação tornou-se mais complexa através de símbolos. A escrita influencia diretamente no arranjo desses índices, pois a ordem alfabética facilita as suas criações de maneira mais elaborada. De acordo com a declaração dos autores, relatam que

A história da indexação e do resumo está intimamente relacionada à história da escrita e da coleção de registros da informação. A primeira evidência sistemática de armazenamento e organização de registros escritos ocorreu na Suméria por volta de 3.000 A.C. (CLEVELAND e CLEVELAND, 2001, p. 06).

Posteriormente, após milênios, o processo de indexação mantém sua importância na recuperação da informação. Já em meados do século XX, pós Segunda Guerra Mundial, duas situações provocaram muitas mudanças, a primeira foi a explosão informacional, e a segunda as tentativas de utilização dos primeiros computadores. No século XXI, início dos anos 2000, adentramos a organização da Internet, com as ferramentas como os mecanismos de buscas, que permitiram a organização e acesso às informações disponibilizadas. Intimamente relacionado a isto, temos os métodos de indexação, que cumprem sua função de preservação da memória e recuperação da informação.

Portanto, podemos entender a indexação, como a seleção de que representam um assunto, para recuperação da informação através de palavras-chave, em sistemas de suportes impressos, analógicos ou digitais. Destacamos a seguinte definição,

A indexação é o processo de identificação da informação em um registro de conhecimento (texto ou não texto) e a organização de pontos para aquela informação dentro de um arquivo pesquisável. A indexação é uma arte, mas uma arte que requer um ordenamento e procedimento metódicos. Os resultados do processo de indexação é um índice que indica tópicos e possíveis usos para um documento e os pontos para localização da informação (CLEVELAND; CLEVELAND, 2001, p. 97).

A indexação é o processo de seleção de termos sobre um tema com a finalidade de sua recuperação em um sistema de busca. Em Lancaster (2004), existem expressões importantes a se pensar na recuperação da informação: “(...) útil, pertinente e relevante para uma necessidade de informação. Ou seja, um documento pertinente (útil) é aquele que contribui para satisfazer a alguma necessidade de informação” (LANCASTER, 2004, p.3). Nesse sentido, quando se pensa em estratégias de busca, procuram-se termos úteis na busca de informação pelos usuários, e por isso considera-se a importância de observar o comportamento do usuário diante dos sistemas de busca.

Dessa maneira, observa-se que os métodos da linguística como as análises semânticas e as análises sintáticas são utilizadas na indexação, que permitem estabelecer as relações hierárquicas, a combinação de termos ou mesmo o uso de sinônimos para obtenção dos melhores resultados na recuperação da informação.

Quando, a Internet, começa a ser utilizada em massa, nos anos 90, passa a existir uma grande necessidade de organização da informação disponível em larga escala, sendo que um dos recursos são as ferramentas de indexação automática, e conforme estas evoluíram, surgiram ferramentas como: os mecanismos de buscas, os diretórios, e os meta-buscadores.

Devido a uma grande quantidade de informação, tornou-se impossível organizá-la por mãos humanas, então foi necessário o uso de robôs para localizar e identificar todo o material, para depois ter condições de recuperá-los através de ferramentas como os buscadores da web.

Portanto, os mecanismos de buscas são programas algoritmos que seguem comandos para indexar algumas palavras-chave, de fato

Mecanismos são programas de computadores, que escaneiam as páginas selecionadas da web para serem indexadas pelo sistema de busca. Eles são frequentemente referenciados como Índices da Web, porque examinam o conteúdo de páginas da web, pois é diferente de apenas recuperar páginas da web como todo, porque encaixam em uma categoria. Aqueles mecanismos consistem de um mecanismo de busca (chamados robôs, aranhas ou rastreadores), periodicamente visitam sites para detectar atualização de informação. (CLEVELAND; CLEVELAND, 2001, p. 217).

O rastreamento é realizado para fazer a cobertura de certas informações nos sites, em que se verifica o documento como todo e seus elementos, palavras específicas no título ou *URLs* – *Uniform Resource Locator* (Localizador Padrão de Recursos), descrição de sites e suas pesquisas, e todos outros componentes do site na Web. Entretanto, Fungman (1993, p. 205) fala que deve existir um trabalho conjunto entre processos mecanizados e processos manuais de indexação, pois a lógica algorítmica nem sempre acompanhará toda amplitude da linguagem natural, principalmente na tradução de termos quando solicitadas nas buscas em sistemas de informação.

Nesta corrente, se anterior à criação de campanha no *Google Adwords*, estiverem estabelecidas nessa regra, e concomitantemente estes termos associados estiverem representados no site rastreados pelas aranhas/rastreadores do Google, a alta revocação e precisão ao final da campanha estarão garantidas, já que os agentes inteligentes são capazes de coletar e organizar informações em grande volume, mas até o momento não consegue interpretar em conformidade ao comportamento do usuário.

No capítulo 3, em Davis (2010), e também em uma das principais pesquisadoras de Marketing Digital no Brasil, Marta Gabriel (2009 e 2010), apresentam as estratégias de *Google Adwords*. Tendo em vista, as agências de publicidade e os setores de marketing de uma empresa, que são os responsáveis pela aplicação de técnicas conhecidas como marketing de busca, em que o foco é mais a venda do que a recuperação da informação. *Search Engine Marketing (SEM)* refere-se às estratégias de mecanismos de busca que consistem na venda de anúncios nos buscadores da web, mais conhecidos como links patrocinados.

O plano do marketing digital envolve o conceito baseado nos 5Ps, ao qual envolve o produto ou serviço oferecido pela empresa anunciante, o preço do produto ou serviço, o ponto de venda como canal de distribuição do produto ou serviço, promoção como conteúdo para atrair público-alvo, posicionamento como a identidade da empresa no mercado, ou melhorar a forma como a marca é reconhecida entre os consumidores.

A aplicação das técnicas dos 5Ps permite atrair mais tráfego e a divulgação de ações específicas da empresa em seu site. Essas técnicas são voltadas ao *off-page* para links patrocinados, otimização de mídias sociais, construção de links no site, participação em programas de afiliados, *press releases*, *links baits* (iscas de links) e publicidade em mecanismos de busca. O *Google Adwords* trata da venda de publicidade para empresas interessadas em anunciar no buscador na procura de maior visibilidade no meio virtual.

O diferencial da propaganda no *Adwords* é a personalização do anúncio. Essa é chamada de campanha, que contém informações de onde e quando os anúncios serão apresentados ao público alvo específico. Se a campanha não estiver funcionando podemos alterá-la para melhor adequar ao público ou até mesmo desativá-la.

A ferramenta funciona quando realizamos uma pesquisa, e recebemos uma lista organizada com uma análise complexa do programa rastreador do Google, cujo resultado chama-se pesquisa natural. No topo da página e no bloco à direita temos os *links* patrocinados, a empresa interessada participa do leilão de palavras e isso determinará a ordem de exibição do seu anúncio na página.

No capítulo 4, utilizaremos o artigo de Larry e Sergey (1999), os fundadores do Google, para explicar como funciona o algoritmo da *Page Rank*. Além, dos autores que discutem sobre este mecanismo de busca, como Davis (2010) empresário de tecnologia, e Vaidhyanathan (2011) estudioso das mídias digitais.

Com o *Page Rank* é possível detectar a distribuição de probabilidades de páginas da Web, e o resultado desta soma será sempre igual a um, porque este algoritmo é um modelo de perfil do usuário, visto pelo conhecido *random surfer*⁴. O cálculo tem a finalidade de identificar a probabilidade de visitas à página de um site e o fator da variável do *random surfer* de solicitação de outra página. Portanto, este algoritmo personaliza o comportamento do usuário, e não permite que se possa enganar o sistema para receber um alto ranqueamento.

O funcionamento deste cálculo vem da citação ou da recomendação de outros sites para determinados links de páginas, quer dizer, quanto mais um site for citado por outros e mais falar sobre um determinado assunto, maior será sua relevância no resultado de busca.

Os algoritmos têm uma grande importância ao Google, assim mantém sua proeminência no mercado digital. Bourgeois (2016, p. 32) diz que,

Embora a principal motivação da empresa em oferecer tecnologia gratuita é promover diversos produtos, o Google está constantemente fazendo mudanças em seu algoritmo de busca. Estes têm efeitos diretos ou indiretos no volume de produtos, que está em

⁴*Random surfer* é uma expressão em inglês, que significa deslizar aleatoriamente. Em ambiente digital, trata-se de visitar páginas diversas na internet.

constante evolução para torná-lo mais fácil aos clientes expender dinheiro em anúncios e em novas formas, como o marketing por e-mail e vídeo. (BOURGEOIS, 2016, p. 32).

No capítulo 4, Green (2013) mostra como os novos algoritmos deste buscador, instalados em 2011 e 2012, alteraram o seu funcionamento, as atualizações dos algoritmos, chamado Panda⁵ e outro algoritmo chamado *Penguin*⁶. Nestas considerações, entendemos que os algoritmos do Google garantem resultados mais precisos e relevantes, e isto também refletirá nas campanhas criadas no *Google Adwords*, pois estas variações nos resultados de buscas são responsáveis pela variabilidade dos valores monetários de cada palavra-chave pesquisada no buscador.

A partir do Capítulo 5, adentramos a parte prática desta pesquisa, onde são relatados os métodos aplicados, as ferramentas e fontes de pesquisadas utilizadas na pesquisa. Sobre o relato da coleta de dados no *Google Trends*, ao qual descobrimos o assunto mais pesquisado no mecanismo de busca.

No capítulo 6, apresentamos as técnicas utilizadas para elaboração da lista de palavras-chave com o tema futebol, e também a sua aplicação no planejador de palavras-chave do *Adwords*, e depois descrevemos e discutimos os procedimentos da indexação de utilidade para seleção de termos, e ainda o desenvolvimento de um mapa mental e uma chuva de *tags* para ilustrar os termos mais representativos para pesquisa do público-alvo.

No capítulo 7, concluímos as questões levantadas durante a realização deste trabalho, com o fechamento de conceitos aplicados e discussões sobre a contribuição oferecida pela área da Biblioteconomia e Documentação. Fazemos também as sugestões de trabalhos futuros prevendo a continuidade desta pesquisa e sua aplicabilidade no mercado digital.

⁵*Panda* é algoritmo do Google, o seu intuito é colocar os sites de alta qualidade acima nos resultados de busca, e utiliza de inteligência artificial para classificação dos sites, que avaliam o fluxo de tráfego, o layout de web design, a velocidade e a confiabilidade da informação.

⁶*Penguin* é algoritmo do Google, sendo responsável pela diminuição de técnicas chamadas “*Black Hat*” (são links ocultos para obtenção de alto ranqueamento), que reduzem sites envolvidos na geração de esquemas de links, duplicações de conteúdo e nos recheios de palavras-chave.

2 O PROCESSO DA INDEXAÇÃO NA SELEÇÃO DE PALAVRAS-CHAVE

Levando em consideração o cerne desta pesquisa, investigamos os conceitos presentes da Indexação como a leitura documentária, um breve histórico, as suas etapas e as suas aplicações no ambiente digital. No decorrer da história, a indexação foi necessária para recuperação e preservação da memória, geralmente apresentada em forma de índices iconográficos, bibliográficos e de assuntos, desenvolvidos perante à necessidade estimulada pelo surgimento da escrita.

Com a invenção da imprensa, aumentou a produção de livros, por isso para facilitar a localização ou de autor ou de assuntos, a informação foi sintetizada no formato de índice. Logo depois, muitos desses livros eram reunidos em bibliotecas, sendo desenvolvidos catálogos de índices dessas coleções para acessá-las facilmente dentro deste espaço físico. Dessa forma, faz parte do processo de Indexação, a seleção das palavras-chave para fazer o controle de vocabulário, com a finalidade de tornar a informação acessível na recuperação de informação específica do usuário em um sistema de busca.

Para isto, existem etapas deste processo compreendidas em três estágios da Indexação, a análise do documento, a identificação de assuntos e a tradução desses em uma linguagem. A norma da ABNT 12676, de agosto de 1992, estabelece os métodos de análise de documentos e define os estágios da Indexação, como “a) exame do documento e estabelecimento do assunto de seu conteúdo; b) identificação dos conceitos presentes no assunto; c) tradução desses conceitos nos termos de uma linguagem de indexação (ABNT, 1992, p. 2)”.

No final do século XX, com a popularização da Internet, verificou-se o crescimento vertiginoso de informação, por isso foram desenvolvidos os sistemas de organização dos dados para facilitar a recuperação da informação, como os mecanismos de buscas organizados por índices de links.

Neste aspecto, a Indexação, pode contribuir ainda mais com a organização de assuntos nos ambientes digitais, por exemplo, a aplicação desta conceituada técnica no processo de seleção de palavras-chave nos anúncios do Google. Tendo em vista, que propomos a intervenção do bibliotecário, um profissional que realiza a leitura documentária, para seleção de palavras-chave na criação de campanhas no *Google Adwords*, com a sua principal atuação no estágio de tradução inserida neste processo.

A ferramenta de anúncios é um leilão de palavras, da qual o menor valor por clique e relevância é executada por um algoritmo. Propomos o bibliotecário, como profissional

capacitado na execução desta tarefa de seleção de palavra orientado pelo contexto sócio cognitivo.

Assim, pode-se dar o uso mais adequado das palavras-chave na ferramenta, e tornando o bibliotecário, mediador entre a necessidade do cliente e usuário do Google, pois evidenciamos nesta investigação, a influência inerente do algoritmo e da complexidade intrínseca da língua. Portanto, na próxima seção deste capítulo, a partir do levantamento bibliográfico sobre os principais aspectos que orientam a Indexação, como as normas, os procedimentos e as abordagens teóricas, será provável acompanhar a sua possível aplicabilidade nas campanhas do *Google Adwords*.

2.1 Os estágios de Indexação e suas aplicações em ambientes digitais

Quando iniciamos o processo de Indexação da seleção de palavras-chave, o primeiro passo é a realização de uma leitura para compreender os assuntos que orientam aquele determinado documento. Fujita (2004, p.1) discorre sobre a leitura como ato social, e este é envolvido por três fatores variáveis de texto-leitor-contexto. Assim define,

A leitura, apesar da individualidade do ato realizado, é um ato social porque existe um processo de comunicação e de interação entre leitor e o autor do texto, ambos com objetivos estabelecidos anteriormente dentro do contexto de cada um. Apesar de, aparente simples e tão natural, o processo de leitura possui uma complexidade que está subjacente porque depende do processamento humano de informações e da cognição de quem lê, de um texto elaborado por um autor e do contexto de ambos, o eu determina os objetivos de leitura (FUJITA, 2004, p. 1).

Portanto, a leitura é um processo de comunicação, em que existe uma mensagem escrita por um emissor, encaminhada para um receptor que interpretará o texto para finalizar o ciclo de comunicação. Para tanto, se deve observar a estrutura de texto, a própria traz as características do contexto a ser transmitido ao leitor, como também aquele que fará a mediação entre o texto e o leitor final. No caso do indexador, que é um leitor profissional por seguir procedimentos da Leitura Documentária, tem habilidade para fazer uma seleção apropriada da representação de assunto para incluí-lo dentro de um sistema de informação. Dessa forma,

Lê-se para condensar textos e representa-los com a finalidade de inserir representações em sistemas que serão interrogados. Isto é, o usuário interroga os sistemas de informação esperando localizar nessas memórias os documentos que possam satisfazer suas necessidades de informação (KOBASHI; SANTOS, 2011, p. 135).

Tendo em vista, que o objetivo é a tradução de assuntos identificados para termos representativos ao universo de pesquisa de um sistema de busca, a análise do assunto depende

de dois contextos presentes, aquele carregado pelo indexador, e um outro presente na comunidade de usuários do sistema de busca. Portanto, o primeiro deve ter repertório anterior para interpretar o texto e traduzi-lo de forma coerente aos usuários do sistema de busca. Já, o segundo também deve ter um conhecimento prévio para localizar a informação pertinente à sua demanda de pesquisa.

Constantemente, na leitura documentária, associamos o seu uso à estrutura de texto, porém no ambiente digital encontraremos outras unidades textuais, que podem estar presentes em uma imagem, um áudio, ou outros objetos digitais. A respeito disso, Fujita (2004, p. 2) diz: “Qualquer processo de compreensão de texto escrito é, portanto, um ato de comunicação que envolve três variáveis: o leitor munido de objetivos para a leitura, o texto contendo as ideias do autor e o contexto composto de elementos influentes na leitura”.

No ambiente de rede, uma página de um site é o texto, no qual existe uma estrutura de ideia representados por links. Neste sentido, um cliente do *Adwords*, pode se beneficiar da atuação do profissional da Indexação, designando-lhe a tarefa de seleção de termos representativos do produto ou serviço oferecido em sua loja virtual, pois o profissional fará a mediação através da seleção de palavras-chave de sua loja, e com o foco de atender às necessidades do público-alvo.

A identificação dos assuntos mais representativos na seleção das palavras-chave durante a Indexação está relacionado ao conceito e/ou significado dessas palavras. De modo que, nos estágios da Indexação, na leitura documentária, se faz necessário compreender o estudo dos significados, pois a tradução exige esse repertório na escolha de uma palavra em detrimento da outra para facilitar a recuperação da informação.

Para tal, a significação é estudada na Linguística e na Terminologia, a partir da linguística estrutural estabelecida por Saussure⁷, na qual se estuda a língua pelas suas oposições, quer dizer, a significação depende de uma relação para defini-la, para assim ocorrerem às diferenças e semelhanças presentes entre dois termos, em que está presente uma estrutura, e o termo é definido por Fiorin (1995, p. 80) como: “... um plano, segundo o qual um objeto é construído, e que esse plano é revelado pela análise interna da totalidade, quando então se mostram os elementos que o compõem, as relações entre eles e a disposição dessas relações”.

Nesta direção, o estudo da língua é baseado numa estrutura que a define como linguagem de uma sociedade. Então, entendidas as relações linguísticas é possível fazer suas

⁷Ferdinand de Saussure foi um linguista e filósofo suíço, nascido no final do século XIX, cujas elaborações teóricas propiciaram o estudo da linguística estrutural, que iniciou o estudo da língua de forma sincrônica e o estabeleceu como Ciência.

distinções, e o significado dos termos. A semântica estrutural foi o estudo inserido dentro da linguística estruturalista que pesquisa as relações entre os termos para defini-los dentro de seus campos semânticos. Assim, se discute a semântica estrutural na maneira pela qual as relações são definidas por estruturas elementares, “a língua não é um sistema de signos, mas uma reunião de estruturas de significação” (GREIMAS, 1973, p. 313).

Em Lopes (1987, p. 24), no domínio da semântica é mais importante a significação do que o significado, pois os valores linguísticos estão nas posições relativas de unidades específicas no interior do sistema. A evidência é que o sentido não é algo isolado, mas surge de uma relação. Em consequência da finalidade da semântica estrutural não sendo conteúdo, mas a forma do conteúdo e as suas oposições determinadas pelo sentido das palavras.

A partir disso, se permite fazer uma indexação mais precisa, porque busca entender o funcionamento do processo cognitivo da linguagem de uma sociedade, na qual Lopes (1987, p. 40) diz que a categoria sêmica é a totalidade, e os semas são as suas partes, com uma visão hierárquica na construção destas relações, o que facilita a extração de assuntos da representação de um tema, ou até mesmo na construção de linguagens para modelos de dados, visto que

A tecnologia mudou e, com ela, a oportunidade de extensão da recuperação de informação no mundo todo, com recursos interconectados. A nova oportunidade nos levaram a reexaminar os princípios descobertos nos anos 70, que não os articulamos de maneira lógica e clara, sendo necessárias nas aplicações atuais dos computadores. Em particular, não esclarecemos a diferença entre os conceitos de uma busca por informação e os termos nos quais expressamos a consulta. Se esta distinção for deturpada, os usuários podem não ser excluídos em absoluto, mas os computadores correm o risco de se debaterem. Para executar dentro da Web Semântica, o software de computador precisa de modelo de dados explícitos que distingue entre os termos e os conceitos (CLARKE; ZENG, 2012, p. 20).

A finalidade de um sistema de busca é oferecer uma boa recuperação, em que se deve representar a estrutura cognitiva de uma comunidade, uma maneira pela qual os indivíduos utilizam um termo para se comunicar, além das possibilidades dessas significações do termo selecionado, assim

O estudo da terminologia é completo porque deve satisfazer três inquietudes distintas: a dos especialistas da matéria, que querem justificar o uso de uma linguagem incompreensível para o público geral e que necessitam de uma ajuda em suas comunicações; a do público geral, que querem saber por quê as representações de dificuldades de compreensão nas trocas com os especialistas; e finalmente, dos mediadores da comunicação, que têm que saber como fazer frente às necessidades de seus clientes. (SERGE, 1993, p. 15).

De modo que, a observação destes acordos sociais para determinar o valor de significação de uma palavra pode trazer resultados mais precisos e relevantes dentro de um sistema de recuperação, pois se antecipará aos termos pesquisados dentro de uma relação de

conceitos. Desta forma, existe acordo social para utilização de uma palavra ao invés da outra, observa-se isto pelo estudo da língua, pois

Trata-se essencialmente de um contrato coletivo ao qual temos de submeter-nos em bloco se quisermos comunicar; além disto, este produto social é autônomo, à maneira de um jogo com as suas regras, pois só se pode manejá-lo depois de uma aprendizagem. Como sistema de valores, a língua é constituída por um pequeno número de elementos de que cada um é, ao mesmo tempo, um vale por e o termo de uma função mais ampla se colocam, diferencialmente, outros valores correlativos; sob o ponto de vista da língua, o signo vale como uma moeda: esta vale por certo bem que permite comprar, mas vale também com relação a outras moedas, de valor mais forte ou mais fraco. (BARTHES, 2012, p. 22).

Portanto, no que concerne o desenvolvimento de anúncios no *Adwords*, a análise do local geográfico e os interesses do público-alvo é uma maneira de conhecer o contexto. Enfim, buscar aspectos culturais para melhor selecionar os termos de campanha, que trará boa revocação em um sistema de recuperação, pois os valores de significação dos termos estará presente no resultado de pesquisa.

A Internet é o encontro de diversos contextos, assim o controle de vocabulário é uma ferramenta necessária para mediar a comunicação entre um site e o seu público, desde a concepção do espaço até sua descrição em mecanismo de busca. Diante do pluralismo cultural da língua presente na Web, dizem as especialistas,

A manipulação semântica não é só a necessidade urgente. A era digital tem encorajado o surgimento de variados tipos de vocabulários, muitas vezes trabalhando lado a lado com o tesouro tradicional. E também, trouxe uma demanda por interoperabilidade para sustentar atividades como serviços da web, a publicação, agregação e troca de dados dos tesouros de múltiplas mídias e formatos; e a exploração dos bastidores de vocabulários controlados na navegação, filtragem e expansão de pesquisas em repositórios na rede (CLARKE; ZENG, 2012, p. 23).

Dessa maneira, esta ferramenta é uma orientação para solucionar a busca no ambiente tão diverso e com públicos distintos, também possibilita focar no público ao qual interessa utilizar um tipo de linguagem. Então, através do controle de vocabulário é possível alcançar dois objetivos, o primeiro é a organização da informação, o segundo na diminuição de ruídos no ciclo de comunicação da informação veiculada no site. Assim, a seguinte colocação diz,

Mesmo que o sistema (o site, por exemplo) possa ser acessado por um público heterogêneo, sua condição de inteligibilidade se relaciona às características particulares que lhe conferem certa homogeneidade. O público de um sistema deve reconhecer ou ter condições de conhecer – o que está sendo veiculado, o que diz respeito eminente a linguagem. (LARA, 2004, p. 234).

Conforme os argumentos apresentados pela especialista, percebemos que a utilização de uma linguagem conhecida aos usuários de um site, o torna eficaz e acessível, seguindo um

critério de representação sócio cognitiva de uma comunidade, isso também refletirá na descrição de conteúdo no processo de rastreamento pelos mecanismos de buscas, o que garantirá a indexação e a recuperação posterior pelos próprios usuários.

Em Ingwersen e Järvelin (2004), dizem sobre a importância da centralização do sistema de recuperação da informação no usuário, em que o domínio de um campo semântico deve estar relacionado a execução de uma tarefa e/ou interesse de uma comunidade a ser atendida pelo sistema de busca. Já dita, esta relevância ao reunir os termos em seus conjuntos significativos, a partir da definição de objetos por conceitos, pode refletir um determinado contexto sócio cognitivo.

Na sociedade da informação, nos deparamos com produtos de informação homogeneizantes da tecnologia, que implica ao acesso dos usuários em relação à informação de interesse, portanto a linguagem é capaz de elaborar estruturas de representação com a finalidade de organizar o fluxo de informação da necessidade e complexidade de uma comunidade. De acordo com as pesquisadoras,

A linguagem como elemento constitutivo da cultura informacional não exerce função meramente instrumental. Para além do seu funcionamento codificador, ela se propõe como representação, seja como insumo do processo social de geração de sentido, seja como resultado textual desse processo (TÁLAMO; LARA, 2006, p. 205).

Por esses meios, o controle de vocabulário é uma proposta para organização da informação na Web, e também uma maneira para diminuir lacunas entre a informação e seus usuários, pois sistemas homogêneos não conseguem abranger a diversidade cultural existente na rede, por este motivo a existência dos índices, que surgiram da necessidade humana, sendo um produto da indexação como forma de organização da memória com a finalidade de sua posterior recuperação. Portanto, os índices existem desde a antiguidade para facilitar a vida dos indivíduos, pois identificam objetos e informações para algum tipo de uso.

Os primeiros registros dessas atividades são as pinturas rupestres. Entre os povos primitivos, o conhecimento era transmitido por esses desenhos, e se pensarmos na palavra índice, as marcas são realizadas para destacar um fato, em que podemos fazer uma comparação ao mundo de hoje como assunto.

Em estudos, os pesquisadores fazem referências aos povos primitivos: Beare (2007, p. 257) “... Arte Paleolítica apresenta a visão moderna de como a vida era a milhões de anos atrás – as imagens são índices delas mesmas”. Em Cleveland e Cleveland (2001, p. 06) “A primeira evidência que nós tivemos do início de tal comunicação são as pinturas nas paredes das

cavernas. O próximo passo foi sistematizar as imagens pictóricas em mensagens mais complexas”.

No ocidente, os gregos, foram os responsáveis pela utilização da ordem alfabética para organizar e facilitar o acesso a uma determinada informação. E também, influenciaram os romanos com este método. Nos séculos seguintes, essa maneira de arranjo foi o protótipo dos índices de livros. No século XV, principalmente, vieram grandes mudanças que refletiram no desenvolvimento da indexação, motivadas pela invenção da imprensa por Gutenberg, pois aumentaram a produção e o consumo literário.

Dois séculos depois, em XVII, surgiram os periódicos que aprimorou a indexação, pois estimulou a criação de índices para facilitar a localização de artigos dentro de cada publicação. Além do mais, esse tipo de publicação passou a ser primordial para comunicação entre as Universidades,

Por quase duzentos anos, entre a invenção da imprensa e o crescimento do jornal, os acadêmicos tinham métodos limitados para comunidade erudita. Eles publicavam os livros e escreviam cartas uns para os outros. Aquilo mudou dramaticamente com o advento do jornal acadêmico (CLEVELAND; CLEVELAND, 2001, p. 07).

Após, a mudança nas pesquisas acadêmicas ocasionadas pelos índices de artigos, que também alterou a forma de índices que trouxe benefícios na organização da informação. Borko (1978, p. 08), relata na história dos índices, que William Frederick Poole, bibliotecário e bibliógrafo americano, em 1882, inventou o índice com a entrada de assunto, precursor da ideia de palavra-chave para representação do texto organizado em lista por ordem alfabética, isto facilitou a busca por informação.

Em Cleveland e Cleveland (2001, p. 09), falam sobre a fundação do “Instituto Internacional de Bibliografia”, pelos acadêmicos Henri La Fontaine e Paul Otlet, no fim do século XIX, o que contribuiu para evolução dos métodos de indexação, pois a finalidade do Instituto era a disponibilização de toda a literatura publicada. Os pesquisadores, também relatam que em meados do século XX, após a Segunda Guerra Mundial, duas situações provocaram muitas mudanças, a primeira foi a explosão informacional, e a segunda as tentativas de utilização dos primeiros computadores.

Nos anos 60, essas condições propiciaram e influenciaram a criação da Internet. A partir disso, começaram a aplicação de novas técnicas de indexação, uma delas é a indexação automática, por meio de tecnologias como programas de computadores que estimularam esse novo tipo de indexação.

No século XXI, iniciamos nossa Odisseia ao espaço da Internet, com uma grande diversidade de dados sem organização, para isto utilizamos os mecanismos de busca, assim temos o acesso às informações disponíveis na rede. Portanto, a indexação é um processo que pode trazer orientações de estruturação da informação, que permite a representação de um contexto social, e assim contribuindo na recuperação da informação. Então, apresentaremos as principais etapas para realização dessa tarefa que resolve às questões relacionadas à linguagem e aos aspectos sociais e cognitivos de uma comunidade.

Na compreensão das etapas de indexação, ter claro a definição de termo auxilia no estágio de identificação de conceitos, para isto utilizamos a norma ANSI/NISO Z39.19 (2015, p. 09), a qual compete a construção de vocabulário controlado monolíngue, que define o termo como “uma ou mais palavras designando um conceito.”

Portanto, a indexação, vem representar os termos descritivos e conceituais de um assunto, para a recuperação da informação em sistemas de suportes impressos ou digitais. Além disso, reforçamos a definição acima referente a indexação ao trazer a NBR 12676 (1992, p.2), que é o “ato de identificar e descrever o conteúdo de um documento com termos representativos dos seus assuntos e que constituem uma linguagem de indexação”.

Quando esta linguagem produz um bom índice, o resultado deve ser uma compreensão do que o autor quer dizer aos seus pesquisadores, ou seja, um índice claro e eficiente. Consequentemente, a indexação não é apenas relacionada à seleção de palavras-chave; mas também, a representação de seus conceitos, do geral ao específico dentro de suas inter-relações.

Contudo, a linguagem é expressiva, e suas variações devem estar refletidas nesta lista de termos selecionados. Em vista disso, quando elaboramos uma lista de termos para aplicação na Internet, esta deve representar o assunto principal do site, ao qual deve refletir o perfil dos consumidores, e também unir os interesses da empresa, através das palavras-chave. Apresentaremos na Figura 1, as principais etapas relacionadas, como exemplo utilizamos a ideia de loja virtual.

Figura 1 – As etapas da indexação aplicadas à seleção de palavras-chave em loja virtual.

Fonte: Da autora (2018).

Um dos problemas observados pelos especialistas, vem da análise do assunto principal, pois compromete todo o desenvolvimento da lista devido à falta de integridade do estabelecimento da relação empresa e consumidor, o que ocasiona um ruído de comunicação, em

Uma indexação inadequada ou uma indexação insuficiente representam 90% das causas essenciais para a aparição de “ruídos” ou de “silêncios” em uma pesquisa. Os 10% restantes serão devidos a causa mecânicas tais como: erro de perfuração, de codificação, de transição etc. (CHAUMIER, 1988, p. 63).

A função da exaustividade é aumentar a probabilidade de recuperação de um assunto, por causa da variação de termos ao cercar um assunto a partir de diversos conceitos que o define.

Esta importância é vista por Lancaster (2004, p. 30) em, “É claro que quanto mais termos forem empregados por documento (isto é, maior for a exaustividade), maior será a probabilidade de ele ser recuperado e maior será o número de características que o distingam de outros documentos”.

Entretanto, esta função da indexação pode acarretar numa poluição informacional, pois traz um alto índice de resultados, conhecido na área como revocação, ou seja, muitos termos recuperados associados a uma palavra-chave utilizada no sistema de busca. Contudo, existe um risco com a especificação, delimita-se muito os termos e diminui os resultados da busca. Assim dito pelo especialista,

As especificidades dos termos asseguram uma qualidade para indexação, pois são palavras mais representativas sobre um assunto, assim aumenta a precisão do resultado de busca. O que pode ocorrer devido a especificidade é a indecisão da classificação de um termo em generalista, e assim perder a relevância em um resultado de busca (CLEVELAND; CLEVELAND, 2001, p. 99).

Subsequente, as seleções de termos específicos e, outros termos abrangentes no processo de exaustividade - é a técnica de ponderação de relevância, torna-se interessante para utilização de toda a lista, pela possibilidade de ter uma revocação de qualidade. A origem dessas técnicas são dos anos 60, época que usavam estudos estatísticos para base da indexação automática,

Análises estatísticas identificam as palavras usadas frequentemente e, pelas hipóteses, as palavras de conteúdo significativa. As análises sintáticas identificam a regra de palavra numa sentença, qual é a sua classe gramatical (e.g., se a palavra funciona como substantivo ou como um verbo) e as relações entre as palavras e a sentença (o cachorro morde o homem versus o homem morde o cachorro). As análises semânticas auxiliam ao estabelecer os paradigmas, ou as relações de classes entre os termos assim como associação com simples conceitos (BORKO, 1978, p. 117).

Observamos nas etapas de Indexação, que a seleção de termos representativos de um assunto deve estar dentro de um estrutura linguística, onde são estabelecidas associações e relações para cumprimento de uma semântica e uma sintática, devido ao fato delas assegurarem o contexto cognitivo de um domínio, portanto aproximando o indivíduo cada vez mais dentro do assunto de interesse. No processo de seleção de palavras-chave no *Adwords* mediada por um especialista em linguagem documentária, teremos palavras mais representativas nas campanhas, já que a seleção pode ser mais eficiente à demanda do produto ou serviço.

Em meados do século XX, muitos estudos desenvolviam “o computador”. Assim, surgem estudos da Ciência da Informação para dar conta da organização da explosão informacional. Estas condições estimulam o uso de novos métodos para indexação automática.

Esta técnica depende de uma linguagem computacional, ou seja, programas que leem o documento, com a finalidade de buscar o número de vezes da palavra mais citada em um arquivo, como também a busca de marcas de identificação do assunto.

Nos anos 90, a Internet, começa a ser utilizada em massa, por este motivo passa a existir uma grande necessidade de organização da informação disponível em larga escala. Um dos recursos são as ferramentas de indexação automática, conforme estas evoluem, surgem as ferramentas de busca, relacionadas pelos especialistas, tais como:

Estas necessidades surgiram conforme o crescimento da World Wide Web, assim cada vez mais com ferramentas de busca, dividem-se em três tipos: mecanismos de buscas, diretórios e meta-buscadores. As buscas gerais como primeira escolha dos diretórios, e depois dos mecanismos de buscas para localizar de maneira mais específica (CLEVELAND; CLEVELAND, 2001, p. 99).

Devido uma grande quantidade de informação, tornou-se impossível organiza-la por mãos humanas, então foi necessário o uso de robôs para localizar e identificar todo o material, e depois ter condições de recuperá-los através de ferramentas como os buscadores da web; assim surgem os mecanismos de buscas, sendo capazes de organizar índices de links em larga-escala de hipertexto, isto torna-se possível com o desenvolvimento de algoritmos para coletar, identificar, classificar, calcular o peso de relevância, e disponibilizar conforme a pesquisa do usuário da Web.

Em outros momentos, neste trabalho se faz uma repetida discussão, sobre a qualidade da indexação, se uma máquina pode superar um ser humano. Em torno disso, pela complexidade da linguagem, suas diversas nuances, dificilmente uma máquina terá tanta primazia nesta ação, até porque ela segue comandos de programas, e não conseguiria tomar uma decisão para uma melhor indexação. Por enquanto, esta função, apenas o ser humano consegue fazer uma avaliação, e assim decidir a pertinência de uma palavra-chave em detrimento de outra, a qual oferecerá um resultado de qualidade de busca, pois estará desenvolvendo uma leitura para melhor descrever um assunto de um determinado contexto. Assim, os autores afirmam:

Frequentemente os mecanismos apresentarão as páginas da web recuperadas pela revisão humana, ou o humano que garantiu a indexação. Claro, existe mais controle de qualidade se feito por humanos do que se feito automaticamente pelos robôs. Tipicamente, quando uma solicitação de busca marcará os termos do usuário contra os substitutos da lista de sites da web, e então utilizando um ranqueamento algoritmo para apresentar a lista de ordem do usuário (CLEVELAND; CLEVELAND, 2001, p. 217).

Para entendermos melhor, Cleveland e Cleveland (2001, p. 220) fazem uma comparação com o índice de livro, ao qual chamam de Índice Web no primeiro é elaborado os pontos dentro do mesmo arquivo; o segundo os pontos de acesso são para milhões de arquivos. Exatamente,

aquilo que é proporcionado pela estrutura do Hipertexto, um link conecta a diversos outros links, cabe ao interessado o desejo de explorar todos esses dados oferecidos pela Web.

Portanto, essa situação é o reflexo do texto digital. O texto impresso, no caso, o livro, oferece uma leitura acabada (início, meio e fim), hierárquica, estável e material. Entretanto, no século XX, com o aparecimento do digital, temos uma nova forma de leitura, o texto digital, que ao contrário do impresso, o ato de ler é inacabado, não hierárquico, instável e imaterial. Isto acontece devido ao hipertexto, um recurso que utiliza áudios, imagens e textos que são conectados por links. Por isso, a leitura se torna aberta a diversas possibilidades de interpretações, o mesmo acontece nas pesquisas por assunto nos sistemas de recuperação da informação.

Em Borko (1978, p.134), nos estudos sobre indexação automática, podemos aplicar a indexação manual que, aprimora e qualifica os índices automáticos, onde essa junção é observada no conceito de mecanismo de busca, como o Google, um processo automático utilizado é a indexação associativa,

[...] esse método é mais complexo que os outros, mas é seguido por um termo mais geral e controlado para expansão de assunto, pelos meios computáveis, a estatística de algoritmo, a coocorrência de dois termos no mesmo documento é calculado e registrado como valor associativo. Isto é, se o Termo A e o Termo B coocorrem num grande número de documentos, aqueles dois termos terão um alto valor associativo. (BORKO, 1978, p. 134).

Portanto, a indexação automática utiliza as associações para determinar relevância de documentos em sistemas de buscas. Partindo da premissa que se em uma pesquisa dois assuntos aparecem com o mesmo peso nos resultados de pesquisa, logo entende-se que existe um valor associativo entre ambos. Em Lancaster (2004, p.55), o especialista utiliza as siglas *KWIC* (*keyword context* – palavras-chave no contexto) e *KWIC* (*keyword out of context* – palavras fora do contexto), pois isto estrutura um índice de palavras que utiliza o termo generalista como cabeçalho, e os termos específicos organizados em ordem alfabética. Essa técnica foi criada por Luhn, em 1959, na *IBM – International Business Machines*, para automatização de índices em sistemas de recuperação da informação.

Diante da versatilidade da linguagem natural, o pesquisador Fungman (1993, p. 205) declara a importância do trabalho conjunto do especialista em Indexação e dos algoritmos no desenvolvimento de índices, o primeiro em virtude da tradução de palavras para inserção nos sistemas, e o segundo na capacidade de coletar e analisar uma grande quantidade de informação.

Portanto, esse conceito pode estar presente na ferramenta *Adwords*, no qual os algoritmos do Google qualificam os links de páginas do site, e ao mesmo tempo os especialistas

da Indexação adequam um conjunto de palavras em conformidade do público-alvo de uma determinada empresa.

3 SEARCH ENGINE MARKETING (SEM)

Abordaremos neste capítulo, as principais estratégias do Marketing Digital relacionadas aos links patrocinados, e como são aplicadas em campanhas de anúncios do Google. Assim, apresentaremos uma visão geral das técnicas que orientam a aplicação da ferramenta *Adwords* e suas particularidades, as quais são baseadas nas estratégias de mercado para atrair consumidores e efetuar vendas dos produtos ou serviços anunciados.

O diferencial é o ambiente digital com suas características, como por exemplo, a possibilidade de ofertar para diversos nichos de mercado em escala mundial, e também um espaço de divulgação aos micro empresários, já que os orçamentos são definidos conforme os valores disponíveis para campanha.

Search Engine Marketing (SEM) são as estratégias do Marketing de Busca, para criação de campanhas em mecanismos de buscas. O objetivo delas é aumentar o tráfego interno e externo para página de lojas virtuais. E sua função é permitir a segmentação de mercado, porque isto facilita o alcance de diversos nichos de interesses presentes na Web.

Hoje, o *Google Adwords*, é a principal ferramenta para divulgação de anúncios nos resultados de busca, principalmente porque é um serviço oferecido pelo Google voltado aos links patrocinados, devido a sua representatividade de mercado na rede mundial. Além da disponibilização da segmentação de mercado, ainda é possível a delimitação de espaço geográfico, do idioma e do período de tempo e orçamento para veiculação de anúncio.

Geralmente, as agências de publicidade e os setores de marketing de uma empresa são responsáveis pela aplicação de técnicas conhecidas como Marketing de Busca, em que o foco é a venda do produto ou serviço, mais do que as técnicas voltadas para recuperação da informação.

No subcapítulo 3.1, serão apresentados os conjuntos de ações que envolvem o *Google Adwords*, as etapas de planejamento, de desenvolvimento e do acompanhamento de campanhas. Para tanto, além do levantamento bibliográfico da área, utilizamos os tutoriais oferecidos aos clientes do *Google Adwords*, que são encontrados em vídeos do *Youtube*, e também na plataforma destinada aos clientes, que contem artigos, vídeos e fórum para troca de informações.

3.1 As estratégias do *Google Adwords*

O *Google Adwords* é uma ferramenta complexa que demanda decisões da forma mais adequada a trabalhar suas estratégias de links patrocinados, por isto antes de montar uma campanha é necessário planejá-la, essa etapa presente nos tutoriais da ferramenta chama-se pré-campanha, identificada na Figura 2,

Figura 2 – Central de Ajuda do *Google Adwords* para seleção de palavras-chave.

Público-Alvo (Intenção)

- Precisamos pensar como o usuário procuraria o produto ou o serviço anunciado ou, se houver a possibilidade, realizar uma pesquisa direto com o público-alvo para saber como pesquisaria determinado produto ou serviço. Essa pesquisa mostrará quem é o público-alvo e sua forma de navegação no Google.

Produto ou Serviço vendido pela empresa (Especificação)

- Devemos especificar pelas palavras-chave o produto ou serviço oferecido pela empresa e também refleti-las na forma de pesquisa do usuário. A seleção de palavras-chave deve ser segmentada em tipos de produtos e/ou serviços que representam a intenção do usuário, ou melhor, a ação do usuário no site (compra, custo, preço, promoção).

Relevância da Seleção de Palavras-Chave no Google (Qualidade da seleção das palavras-chave)

- Dois aplicativos servem para testar a qualidade da seleção das palavras-chave, o Google Trends que indica as tendências de assuntos mais pesquisados no Google, e também mostra-se útil na campanha sazonal, a ferramenta auxilia na seleção de palavras a serem compradas no Google Adwords. E o outro é o Keyword Tools, que testa a relevância da palavra-chave no Google, assim diagnosticamos as palavras-chave de maior retorno para campanha.

Correspondência

- Controle de palavras-chave e anúncio de muito alcance (mais ampla) ou pouco alcance (menos ampla). Com a correspondência, configuramos o retorno da expressão exata na campanha. Em configuração avançada do Adwords, escolhemos a opção ‘detalhes’ e desabilitamos a exibição de seus anúncios para pesquisas com plurais, erros ortográficos e variações aproximadas, porém o seu uso se torna eficaz em palavra-chave exata.

Fonte: *Google Adwords*, 2017.

Apresentamos as principais estratégias para seleção de palavra-chave relatadas pelos especialistas do *Adwords*, que devemos ter em vista o público-alvo (intenção), o produto ou serviço vendido pela empresa (especificação), a relevância dessa seleção de palavras no Google (qualidade da seleção de palavras-chave) e a organização da lista de palavras-chave em grupo

(Correspondência). Neste último caso, a correspondência é a forma como configuramos algumas palavras no anúncio, a partir de algumas regras da computação. Podemos dividi-las em cinco grupos de correspondência, conforme a Figura 3.

Figura 3 – Tabela de Correspondência

Tipo de correspondência	Símbolo especial	Exemplo de palavra-chave	Os anúncios podem ser exibidos em pesquisas que	Exemplos de pesquisas
Correspondência ampla	nenhuma	chapéus femininos	são sinônimos, pesquisas relacionadas e outras variações relevantes	comprar chapéus femininos
Modificador de correspondência ampla	+palavra-chave	+chapéus +femininos	contêm o termo modificado (ou variações aproximadas, mas não sinônimos), em qualquer ordem	chapéus femininos
Correspondência de frase	"palavra-chave"	"chapéus femininos"	são uma frase e variações aproximadas dessa frase	comprar chapéus femininos
Correspondência exata	[palavra-chave]	[chapéus femininos]	são um termo exato e variações aproximadas desse termo exato	chapéus femininos
Correspondência negativa	-palavra-chave	-beisebol	são pesquisas sem o termo	bonés de beisebol

Fonte: *Google Adwords* (2017).

Em vista disso, visualizamos a função da tabela de correspondência, onde se faz o controle das pesquisas, quem acessará os anúncios de campanha, assim evitam-se cliques desnecessários, daqueles sem interesse por determinado produto veiculado.

Apresentaremos nas Figuras de 4 a 8, os passos básicos para desenvolver a campanha na ferramenta de anúncios, tais como a criação de conta no *Adwords*, definição de local e idioma da campanha, elaboração do anúncio, seleção das palavras-chave, delimitação do orçamento para campanha.

Figura 4 – Passo 1: criação de Conta no Adwords

Fonte: *Google Adwords* (2017).

Na criação de conta na ferramenta (Figura 4), podemos perceber que o Google quer conhecer o seu cliente, pois existem duas respostas sobre a pergunta “o que melhor descreve você?”, se já é usuário de outros serviços, ou se é o primeiro contato.

Figura 5 – Passo 2: definição de Idioma e Local

The screenshot shows the 'Definir fuso horário e moeda' (Define time zone and currency) step of Google AdWords account creation. It includes instructions, dropdown menus for time zone (Brazil, GMT-03:00) and currency (Brazilian Real), and a 'Continuar' button.

Defina suas preferências de fuso horário e moeda

Precisamos de mais dois detalhes para configurar sua conta do Google AdWords: a moeda que será usada para pagar os custos de publicidade relacionados ao Google e o fuso horário dos relatórios. Você não poderá alterar esses detalhes depois. Desse modo, escolha suas opções com cuidado.

Selecione um fuso horário permanente para sua conta.
Este será o fuso horário de todos os relatórios e faturamentos da sua conta.

País ou território do fuso horário:

Fuso horário:

Selecione uma moeda permanente para sua conta.
Consulte as [opções de pagamento](#) disponíveis para as moedas locais antes de decidir. Nem todas as moedas estão disponíveis em todas as regiões.

Suas configurações de fuso horário e moeda não poderão ser alteradas depois que você configurar sua conta.
Analisar suas escolhas com cuidado e clique em "Continuar".

Fonte: *Google Adwords* (2017).

Na Figura 5, observamos o segundo passo, tanto a definição de idioma quanto de local, que influenciam em vários aspectos sociais e demográficos, pois fazemos a delimitação geográfica e cultural conforme o público-alvo desejado na campanha. Neste passo, podemos evidenciar o contexto sócio cognitivo a ser alcançado pela seleção de palavras-chave, sendo determinado na fase de pré-campanha.

Figura 6 – Passo 3: Planejador de palavras-chave

The screenshot shows the 'Planejador de palavras-chave' (Keyword Planner) interface. It includes input fields for product/service, destination page, and product category. Below, there are tabs for 'Ideias de grupos de anúncios' and 'Ideias de palavras-chave'. A table displays search results for the keyword 'teor de sódio, pr...'. The table has columns for 'Grupo de anúncios', 'Palavras-chave', 'Média de pesquisas mensais', 'Concorrência', 'CPC méd.', and 'Parcela impr. anúncios'.

Seu produto ou serviço: Sua página de destino: Categoria de seu produto:

Grupo de anúncios (por relevância)	Palavras-chave	Média de pesquisas mensais	Concorrência	CPC méd.	Parcela impr. anúncios
Palavras-chave ...	teor de sódio, pr...	22.220	Baixa	R\$0,13	0%

1 - 1 de 1 grupos de anúncios

Fonte: *Google Adwords* (2017).

Vejamos a Figura 6, selecionamos as palavras-chave ou expressão que melhor descrevem o anúncio, assim toda vez que procurarem o assunto, o anúncio será trazido nos resultados de pesquisa. Nesse passo, fazemos a especificação, pois selecionamos as palavras que trará precisão a campanha.

Figura 7 – Passo 4: Criação de anúncio

✓ Selecionar as configurações da campanha Criar um grupo de anúncios

Tipo: Redes de Pesquisa e Display - Todos os recursos

Nomeie este grupo de anúncios

Um grupo de anúncios contém um ou mais anúncios e um conjunto de palavras-chave relacionadas. Para obter melhores resultados, tente concentrar todos os anúncios e palavras-chave deste grupo de anúncios em um produto ou serviço. [Saiba mais sobre como estruturar sua conta.](#)

Nome do Grupo de anúncios:

Criar um anúncio

☒ Anúncio de texto
 ☐ Anúncio gráfico
 ☐ Criador de anúncios gráficos
 ☐ Anúncio de conteúdo digital/aplicativo
 ☐ Anúncio para celular WAP
 ☐ Anúncio da lista de produtos
 ☐ Anúncio dinâmico da Rede de Pesquisa

Para dar os primeiros passos, crie seu primeiro anúncio abaixo. Lembre-se de que você poderá criar mais anúncios posteriormente. [Saiba como criar um ótimo anúncio de texto](#)

Título

Linha de descrição 1

Linha de descrição 2

URL de visualização

URL de destino

Fonte: Google Adwords (2017).

O quarto passo da campanha (Figura 7), escrevemos o anúncio para explicar o tipo de negócio a ser divulgado. Depois, preenchemos o título, mais duas linhas descritivas (as palavras-chave), o endereço eletrônico (a URL visível e a URL de destino), assim elaboramos o texto de anúncio, e também selecionamos os canais de divulgação da campanha, que refletirá no quinto passo, pois quanto mais forem os canais, maiores serão as probabilidades de cliques e custos, assim dependerá muito da intenção e alcance da campanha.

Figura 8 – Passo 5: Orçamento de campanha

Lance e orçamento

Estratégia de lance [Editar](#)

[Alterar os lances do grupo de anúncios »](#)

Orçamento

☐ Orçamento individual
☒ Aplicar um orçamento da biblioteca compartilhada

Certifique-se de que este orçamento cubra as necessidades de sua campanha.

☒ Atum Gomes da Costa	R\$100,00/dia	1 campanhas
--	---------------	-------------

[Salvar](#) [Cancelar](#)

Fonte: Google Adwords (2017).

No quinto passo (Figura 8), definimos o orçamento da campanha, como o valor máximo pago por um clique em determinada palavra e a forma de pagamento, pois serão cobradas apenas as palavras que recebem cliques. A palavra comprada será exibida dependendo do valor

estipulado pelo cliente para clique, elemento que determinará a relevância nos resultados de pesquisa.

Dessa maneira, após os estágios de pré-campanha e campanha, existem algumas ferramentas do Google para acompanhamento dos anúncios criados no *Google Adwords*, o primeiro é o estimador de tráfego e o segundo é o *Google Analytics*.

O estimador de tráfego é utilizado para avaliar o tráfego de palavras, que possibilita a organização de orçamento para a campanha. Consiste em um relatório que pode ajudar a entender o desempenho da campanha e o rastreamento de conversões, como também a porcentagem dos visitantes enviados ao site que gera conversão. O autor diz o seguinte,

Se você não fizer nenhuma estimativa para o CPC máximo no Traffic Estimator, o Google sugerirá um para você que gere anúncios no topo das posições (1.0) em 85% das vezes. Como você pode imaginar, esta é uma escolha de CPC cara. Por isso, para fazer melhor uso do Traffic Estimator, talvez valha a pena inserir seu próprio CP máximo e mais realista, dependendo do seu orçamento. (DAVIS, 2010, p. 360).

O relatório Estimador de Tráfego, acompanhamos as informações de Cliques (quantidade de cliques por anúncio); de Impressões (quantidade de vezes que uma palavra foi pesquisada); da CTR (a taxa de cliques calculadas pelo cliques e impressões da página, gera um valor em porcentagem); de CPC médio (custo médio da palavra pesquisada no Google); de Custo (valor total gasto pelo cliente com determinada campanha).

Já, o *Google Analytics*, é uma plataforma de monitoramento de tráfego que podemos instalar em sites, lojas virtuais e blogs. Davis (2010, p. 387) afirma sobre a importância de trabalhar com contas vinculadas de *Adwords* e *Analytics*, já que “ambos os programas têm como objetivo ajudar os proprietários de sites a entender a dinâmica dos sites para que possam tomar melhores decisões de negócio e aumentar o desempenho”. Nesse sentido, de acordo com o autor a utilização do *Google Analytics* permite a identificação das áreas mais problemáticas do site, possibilitando mudanças necessárias para correções e aumento do tráfego.

Dessa forma, conhecemos o perfil e comportamento do usuário que acessa a página monitorada, como idade, gênero, horário de navegação, cidade e o tempo de permanência no espaço acompanhado pela plataforma. Um dado oferecido pelo relatório é a taxa de rejeição, com isto é possível acompanhar o tempo de permanência do visitante na página, caso a taxa venha ser alta, o que pode indicar uma interface nada amigável. Vejamos na Figura 9, um exemplo de relatório gerado pelo *Google Analytics*,

Figura 9 – Relatório do *Google Analytics*

Fonte: Google Imagens (2018).

Nos dois tipos de acompanhamento da campanha é importante observar o índice de qualidade, quer dizer, o posicionamento do anúncio no resultado de busca, o que depende do leilão de palavras (valor máximo pago por clique), da relevância da palavra na pesquisa, do texto do anúncio, se a seleção de palavras-chave é específica e se o site do anunciante tem boa navegabilidade.

A importância de uma página bem formulada é enfatizada por especialistas, uma empresa pode dobrar a “taxa de conversão” do seu Website (realizando uma venda ou obtendo uma solicitação de informações sobre um produto por parte de um novo usuário) por meio de um bom projeto de usabilidade.

Nielsen e Loranger (2007, p. 43) dizem, “Descobrimos em um estudo de 42 projetos de reformulação que as medições dos sites melhoraram 135% depois de um redesenho que levou em conta as diretrizes de usabilidade”. Portanto, não fará sentido ter um anúncio muito acessado, se caso os visitantes saírem da página frustrados, pois a loja virtual não conseguirá a conversão, quer dizer, não conseguirá vender o produto e/ou serviço disponível na página.

A eficácia da campanha depende do acompanhamento de conversões, que pode ser realizada por meio dos aplicativos do *Google Adwords* para descobrir quais anúncios, palavras-chave e campanhas estão gerando negócios, o que facilita uma análise para delinear as ações e a busca do retorno do investimento.

Existem dois tipos de conversão, então precisamos saber a mais pertinente ao nosso interesse, para depois configurar a métrica mais adequada ao caso, ou a conversão uma por clique, que a ferramenta rastreia o clique e o Google cria um cookie para um período de 30 dias, por exemplo, se nesse período alguém fizer uma ou mais visitas ao site contará somente uma conversão, o que representa no máximo 100% de retorno; ou as conversões muitas por clique, que no período de 30 dias são contabilizadas todas as conversões originadas por um clique, nesse caso, a taxa de conversão ultrapassará os 100%.

Davis (2010, p. 357), afirma que, “(...) a taxa geral de conversões no *Adwords* e na Web, dependendo de várias questões de definições oscila aproximadamente entre 1% e 2%”, o que depende da quantidade de acessos ao site da empresa. Se for aplicada essa média em 1% ou 2% de 100 teremos um resultado, de 1000 teremos outro tipo de resultado, pois muda a perspectiva em relação à campanha.

O Estudo ‘Performance no E-Commerce – e-mail, *search*, social e portal⁸, realizado no Brasil pela *Experian Marketing Services*, apresenta dados de conversão média do varejo de 1,44%. Outro dado sobre a taxa de conversão, o chamado *search*, quer dizer, técnicas de marketing para atrair público através de mecanismos de busca, fica em torno de 2%.

Há variáveis que indicam o sucesso de uma campanha no *Adwords*, então só é possível mensurar esses resultados se entendermos primariamente o objetivo central de cada campanha, como, por exemplo, atrair tráfego de visitantes ao site da empresa ou promover uma ação específica de anúncio sazonal.

Dessa maneira, teremos os indicativos sobre a campanha, permitindo a criação de uma medida-padrão para avaliação dos resultados obtidos e definição dos valores para investimento nas palavras-chave nessa ferramenta do Google.

A maioria das campanhas espera alcançar alta taxa de conversão a um baixo custo, até mesmo aquelas que desenvolveram todas as estratégias para aproximar a marca de seu público-alvo, pois pretendem, de forma indireta, aumentar a receita de sua empresa.

⁸Conteúdo acessado por Slideshare: <http://www.slideshare.net/HitwiseBR/estudo-indito-performance-no-ecommerce-14085751>, material produzido por Experian Marketing Services. Disponível em: <<http://www.experian.com/intl/brazil.html>>. Acesso em: 08 mar. 2016.

De fato, o que determinará o sucesso da campanha são as estratégias pensadas antes da efetivação da compra de palavras-chave. Por isso, o público-alvo é fundamental, no caso, se pensamos em uma forma mais abrangente, pensamos nos seus gostos, nos seus hábitos e na localização geográfica, enfim, aquilo que envolve o perfil de público-alvo para fazê-los acessar o anúncio divulgado.

Precisamos analisar o tipo de tráfego que o site atrairá, caso tenha muitas visitas é do tipo mais amplo, ou se for mais concentrado é do tipo mais específico, e se for mais estrito pode ser mais útil para uma publicidade mais segmentada, garantindo maiores resultados, pois encontrará de fato mais interessados no produto. Nesse sentido, de acordo com Davis, “um bom conteúdo posicionado em um nicho muito estrito pode também ser lucrativo, mas é preciso ter anunciantes dispostos a pagar por palavras-alvo relacionadas a esse nicho” (DAVIS, 2010, p. 23).

Nas décadas que antecederam o fenômeno da internet havia a impossibilidade de oferecer em escala mundial um tipo de produto para um tipo específico de mercado. Hoje, todos estão mais conectados, assim é possível oferecer uma grande variedade de produtos para agradar qualquer tipo de segmento do mercado. Nesse contexto, muitos recorrem ao Google para localizar o seu produto e/ou serviço, por isso, ferramentas como o *Google Adwords* possibilitam o elo entre marca e consumidor.

4 OS REFLEXOS DOS ALGORITMOS NOS ANÚNCIOS DO GOOGLE

Em toda história da humanidade, a tecnologia foi desenvolvida pelo ser humano com o intuito de aprimorar suas funcionalidades nas execuções de suas tarefas. Sempre com a visão de fazer um processo aumentar as potencialidades de força, de inteligência e de comunicação.

Um dos maiores estudiosos do século XX, afirma no seu livro *“Os meios de comunicação como extensões do homem”* que, “isto apenas significa que as consequências sociais e pessoais de qualquer meio – ou seja, de qualquer uma das extensões de nós mesmos – constituem o resultado do novo estalão introduzido em nossas vidas por uma nova tecnologia ou extensão de nós mesmos” (MCLUHAN, 2007, p. 21).

Outro pensador, Logan (2012, p. 194), ainda vai adiante ao dizer que não só a tecnologia são as nossas extensões, mas também fazem parte de nós mesmos, como algo orgânico para nosso funcionamento em sociedade. Uma condição relacionada a nossa evolução e sobrevivência no planeta.

Nesse aspecto, os argumentos nos fazem refletir que a tecnologia é desenvolvida pelo ser humano, como meio de estender as potencialidades humanas, ao ponto que se torna orgânico na interação com o mundo. Assim visto, um mecanismo de busca atua na forma de um cérebro ao lidar com o dado. Conforme as ferramentas e padrões de conceitos possibilitarem a sua interpretação, culminará em conhecimento a ser aplicado na otimização das tarefas diárias.

Por este motivo, fez-se necessário buscar compreender alguns conceitos relacionados à Inteligência Artificial, principalmente sobre algoritmos, com a finalidade de discernir o funcionamento do Google. Até mesmo, perceber que as atualizações realizadas pelo mecanismo de busca, tem a pretensão de uma hegemonia de mercado buscando atender às necessidades de buscas de informação na Internet. Já, que existe uma grande produção e fluxo de informação nesse ambiente, ao qual exige a organização desses dados para diversos usos e interesses pelos indivíduos.

Neste capítulo teremos dois subcapítulos (4.1 e 4.1.2), a primeira é mais descritiva e traz o conceito de Inteligência Artificial, Algoritmo, Web Semântica, pois estão presentes na composição do Google, e o quanto essa formação influenciam nos resultados de busca. A segunda é uma discussão teórica sobre a representação do contexto sócio cognitivo em sistemas de busca, o que possibilita uma revocação qualitativa na recuperação da informação, uma aplicação possível na seleção de palavras-chave aos anunciantes do Google, sem depender apenas dos algoritmos do mecanismo.

4.1 Inteligência Artificial, Algoritmos e Web Semântica no Google

Neste estudo, perceberemos as influências dos algoritmos nos resultados de busca do Google, e quanto reverberam nas campanhas de seus anunciantes, sendo um diferencial na elaboração dos anúncios na ferramenta. Por conseguinte, para alcançar esta compreensão, buscamos alguns estudiosos que definem e explicam a Inteligência Artificial, sendo a área responsável pelo desenvolvimento de algoritmos.

Norvig e Stuart (2010, p. 22), definem a Inteligência Artificial como estudo da Ciência Empírica, que envolvem diversos campos científicos. Assim, os estudos foram iniciados posterior à segunda guerra mundial, e buscam compreender como funcionam o pensamento e a ação humana, como também a automatização e/ou a mecanização desses processos.

Historicamente, a necessidade surge com a finalidade de aumentar a produtividade, devido à abertura do mercado econômico estimulada pelas inovações tecnológicas, pois alteraram a percepção de tempo e espaço com a conexão dos indivíduos em rede mundial, ou seja, aumenta a demanda de mercado e para supri-la automatizam os processos industriais e os intelectuais. De modo que, a Inteligência Artificial estuda procedimentos lógicos para Representação do Conhecimento, e por estes meios buscam desenvolver processos cognitivos do pensamento e da ação humana, porém com a finalidade de serem realizados pelas máquinas.

Os pesquisadores (NORVIG; SUTART, 2010, p. 22), dizem que existem quatro abordagens principais da área, o pensamento humano trata-se da automação de atividades associadas aos humanos, quer dizer, máquina com mentes, no sentido completo e literal. Já o pensamento racional busca transformar os mapas mentais em modelos computacionais. Por outro lado, a ação humana trata da ação pela máquina aprimorada pela mão humana, e a ação racional é a Internet das Coisas, ou seja, tornar os artefatos tecnológicos mais inteligentes aos quais fornecem informações que agilizam o dia-a-dia.

De acordo com os pesquisadores, entendemos que a Inteligência Artificial veio para resolver problemas do cotidiano. Assim, a máquina passa a ser desenvolvida para executar as tarefas mecânicas e do pensamento humano. Norvig e Stuart (2010) dizem que, “Um agente racional é aquele que age assim como para alcançar o melhor resultado ou, quando existir a incerteza, o melhor resultado esperado”. Na abordagem do Agente Racional oferecem duas vantagens em detrimento às outras. A primeira, ela é mais generalista do que as leis do pensamento, porque a inferência correta é uma das várias possibilidades de mecanismos para alcançar a racionalidade. A segunda, é mais acessível ao desenvolvimento científico do que são as abordagens baseadas no comportamento e pensamento humano.

O algoritmo, faz parte da Ação Racional, sendo um agente racional, quer dizer, agente é aquele que faz alguma coisa, por exemplo, o computador deve ser um agente da ação que faz algo além do esperado, como operar automaticamente, perceber o seu desenvolvimento, persistir por um longo período de tempo, também se adaptar a mudança, criar e perseguir objetivos. Diante disto, Thomas Cormen (2014, p. 2) define algoritmo como: “... um conjunto de etapas para executar uma tarefa descrita com precisão suficiente para que computador possa executá-la”, ou seja, o principal objetivo de um algoritmo é resolver um problema com recursos computacionais para aprimorar ou facilitar uma tarefa humana.

Dessa maneira, trazemos o artigo de Sergei e Page (1999), “*A anatomia de um mecanismo de busca da Web em larga-escala hipertextual*”, documento de lançamento do protótipo do Google, onde na época relatam a importância do algoritmo *Page Rank*, assim podemos mensurar o quanto isso faz parte do sucesso atual do mecanismo ao trazer os resultados mais relevantes de busca na Internet.

Em vista disso, o Google, quando foi lançado pelos seus fundadores tinha o objetivo de resolver o problema das técnicas tradicionais de busca na web. Devido ao fato, que nos anos 90, existiam alguns mecanismos de buscas, porém a indexação destes eram realizadas manualmente, o que representava um alto custo para manutenção desses índices, ao qual a cada dia aumentava-se o fluxo informacional, sendo um esforço hercúleo na execução dessa tarefa.

Por isso, Larry e Page (1999), desenvolveram um produto que utilizasse a estrutura de hipertexto, ou seja, a partir de links presentes na Web, elaboraram um sistema de organização desses através de um índice. Portanto, há um volume muito grande e descontrolado da informação em rede global, e por isso, os fundadores do Google, buscaram a solução na Inteligência Artificial ao desenvolverem o algoritmo “*Page Rank*”, sendo um agente racional para resolução do problema na organização de dados na estrutura do índice de links, com a finalidade de trazer a precisão no resultado de busca, minimizando o tempo de resposta, propondo um mecanismo mais eficiente aos usuários da Internet.

A metáfora mais adequada ao que acontece no Google, seria ao ambiente de biblioteca, onde existem os mesmos processos técnicos da organização da informação, que realiza a coleta, identifica, classifica e indexa o documento, por fim ele é ordenado em forma de índice. A missão de uma biblioteca e do Google é a mesma, oferecer uma informação precisa e de qualidade diante da necessidade de seu usuário.

Dessa forma, os resultados de busca são marcados pelo *Page Rank*, que trará um assunto pesquisado de acordo com sua relevância, ou seja, quanto mais for buscado e recomendado, mais bem colocado estará na página do mecanismo, e é exatamente esta condição

que influencia os valores monetários agregados ao termo, em que estes valores oscilarão dependendo do comportamento do usuário. Diz o autor que,

O Page Rank é essencialmente um algoritmo recursivo, o que significa um processo que invoca a si mesmo. O Page Rank de uma determinada página é a soma dos Page Rank das páginas que se ligam a ela (ponderado pelo número total de links obviamente). Nesse esquema, um link proveniente de uma página com Page Rank alto claramente conta mais do que um link proveniente de uma página de baixa classificação. (DAVIS, 2010, p. 122).

Para entendermos a relevância no Google, precisamos saber como funciona o processo de precisão da busca pelo “*Page Rank*”, que é a forma de calcular a qualidade de cada página dentro da estrutura de link da web, na qual é utilizada a quantidade de citação sobre determinado assunto que os indivíduos costumam buscar no mecanismo.

Dessa forma é possível priorizar os resultados de buscas de palavras-chave na web, estes resultados são rastreados pelos robôs do Google nas marcações de texto que são extraídos dos títulos de páginas, e também de textos completos. De acordo com o especialista,

O algoritmo do PageRank é usado, em parte, pelo Google para ordenar os resultados retornados por consultas de busca específicas. Como tal, é fundamental entender o PageRank em termos de empreendimento de SEO para melhorar os resultados de buscas naturais. (DAVIS, 2010, p.120).

O funcionamento deste cálculo vem da citação ou da recomendação de outros sites para determinados links de páginas, quer dizer, quanto mais um site for citado por outros e mais falar sobre um determinado assunto, maior será sua relevância no resultado de busca. Além disso, pode-se normalizar a quantidade dos números de links de páginas, assim evita-se as páginas repetidas. Assim,

Desse modo, cada website copiado no servidor Google traz consigo um conjunto de pontos relativos, calculados instantaneamente, de modo a colocá-lo num lugar específico de uma página de resultados, e supõe-se que essa classificação reflita sua importância para a pesquisa. (VAIDHYANATHAN, 2011, p. 601).

Apresentamos no Quadro 1, a lógica presente no algoritmo *Page Rank*, a programação realizada pelos fundadores do Google, o cálculo que apresenta as páginas mais acessadas e confiáveis da web em formato de índice links.

Quadro 1 –Cálculo do algoritmo *Page Rank*

$$\text{"PR (A) = (1-d) + d (PR (T1)/C (T1) + ... + PR ((Tn) / C (Tn))"}$$

Lê-se:

PR = PageRank;

A = Página A (página principal de um site);

d = fator de redução da oscilação de uma palavra-chave (a configuração é entre 0 e 1 – para este cálculo é utilizado 0,85);

T1 = início da quantidade total de páginas do site;

Tn = final da quantidade total de páginas do site;

C (A) = número de recomendações externas desta determinada página;

Fonte: BRIN, S.; PAGE, L. *The anatomy of a Large-Scale Hypertextual Web Search Engine*, 1999.

A partir disso, o algoritmo determina as páginas mais relevantes e mais citadas em muitos lugares na web, por exemplo, se a página for citada no Yahoo!, isto garantirá a indexação, e terá uma boa posição no índice. Todavia, se caso a página não tiver alta qualidade e seu link estiver quebrado, o Google não apresentará em seus resultados de busca. O cálculo é realizado esporadicamente, o que facilita a revisão de detalhes em páginas, tornando possível melhorar o seu índice de ranqueamento. A autora relata,

O pagerank não é calculado com a mesma frequência com que os robots dos buscadores visitam as páginas. Na realidade, ele é atualizado apenas algumas vezes ao ano e é usado atualmente para determinar a frequência de visita dos robots, a profundidade que eles devem atingir no site e a seleção da base de dados. (GABRIEL, 2009, p. 50).

De modo que obedecerá uma cadeia de acontecimentos, por exemplo, se diversos indivíduos passam a buscar um determinado termo, este terá uma alta relevância. E se estiver relacionado ao produto ou serviço oferecido por algumas empresas, o termo terá uma alta concorrência para sua utilização em anúncios. Porque quanto mais um termo for pesquisado, mais este termo custará no *Google Adwords*.

Entretanto, no processo de recuperação da informação, nos sistemas de sua organização, na seleção de termos é imprescindível escolher aqueles que mais representam um assunto pesquisado, e se possível relacioná-lo dentro de uma cadeia hierárquica de conceito.

Dessa maneira, observamos que os clientes interessados em criar campanhas no *Google Adwords*, precisam fazer estas relações para alcançar resultados salutareos dentro do mecanismo de busca, pois a precisão depende da escolha de um termo ou de uma expressão que melhor

represente os produtos e/ou serviços de uma empresa na recuperação de seu site. Nesta condição, o critério de citação é um método antigo, utilizado por bibliotecários para determinar a relevância de um documento, assim comparada pelo autor,

A ideia por trás do PageRank é antiga e tem sido usada por bibliotecários do mundo pré-web para oferecer um método objetivo de pontuar a importância relativa de documentos eruditos. Quanto mais citações outros documentos fizerem a um determinado documento, mais “importante” será esse documento, maior será a sua classificação do sistema e mais probabilidade ele terá de ser acessado primeiro. (DAVIS, 2010, p. 121).

O algoritmo executa a avaliação de suas ações, pois o agente racional é programado para validar seus próprios procedimentos, por isto a *Page Rank* faz avaliação pela análise das solicitações de busca. O termo recebe um ID de palavras-chaves, que segue para uma lista de documentos, que são classificados, encaixados em blocos (categorias), depois realiza uma leitura para localizar todos os termos buscados, o mecanismo computa o *rank* de cada solicitação. Se estiver no início dos blocos categorizados, retorna para o procedimento de leitura, e os documentos classificados são marcados pelo *rank*, aos quais retornam para o topo da lista.

Com isso, conseguem a solução do problema da busca de informação na Web, que há necessidade de dados estruturados para construção de uma árvore de busca, que é construída por nós, ou seja, as ligações de um determinado dado para outro. A construção é realizada pelo elo - o “nó pai” (origem), o estado (espaço no mundo), a ação (etapa), custo de caminho, e finalmente o elo “nó filho” (final). Para os especialistas é importante fazer a distinção de nó e estado:

Um nó é uma escrituração⁹ de dados estruturados usados para representar a árvore de busca. Um estado corresponde à configuração de mundo. Assim, os nós são caminhos particulares, como definidos pelos ponteiros Pais, enquanto que não são os estados. Além do mais, dois nós diferentes podem conter o mesmo estado de mundo, se aquele estado é gerado via dois diferentes caminhos de busca (NORVIG; STUART, 2010, p. 79).

No Google, referente aos dados coletados, o “estado” da *Page Rank* é realizado pelos blocos categorizados, onde as informações são classificadas, e o nó são as informações organizadas em listas pela meta-informação rastreada na Web, que concomitantemente rodam o indexador e o rastreador, porém o indexador é mais rápido que o rastreador, isto quer dizer, o mecanismo de busca organiza de maneira mais rápida do que coleta as informações na Web.

⁹A forma como os dados são registrados metódica e sistematicamente das ações ocorridas no sistema, com a finalidade de fixar e fornecer dados quando necessários a sua verificação. O termo foi emprestado da contabilidade financeira, em inglês *bookkeeping* (escrituração contábil).

Uma característica do mecanismo de busca são as atualizações esporádicas de suas ações. Em 2011, uma delas foi o algoritmo “*Panda*”, que tinha a finalidade de colocar os sites de alta qualidade mais acima nos resultados de busca. Este executa procedimentos de classificação dos sites, que avaliam o fluxo de tráfego, o layout de web design, a velocidade e a confiabilidade. O *Panda* é responsável pela qualidade de arquitetura dos sites, teve a sua criação estimulada pelo aumento do acesso à Internet Móvel, que exige adaptação do conteúdo da Web para outros dispositivos de informação.

No ano seguinte, o Google lançou também, o *Penguin*, responsável pela diminuição de técnicas chamadas ‘*Black Hat*’ (são links ocultos para obtenção de alto ranqueamento no mecanismo de busca), no qual objetivamente reduzem sites envolvidos na geração de esquemas de links, como duplicações de conteúdo, os recheios de palavras-chave e ocultação de links no texto âncora.

Os dois algoritmos foram desenvolvidos para garantir conteúdo de qualidade para usuários do Google. Então, o buscador classifica os sites e procura oferecer uma experiência gratificante, pois retorna as informações mais relevantes dentro de um interesse de pesquisa. Assim relatado pelo autor,

Como parte da mudança de algoritmo, o Google percebeu que as pessoas podem manipular o sistema para ganhar mais backlinks, porém o que é muito difícil manipular é um bom conteúdo, isto é, sem simplesmente copiá-lo de qualquer outro lugar. (GREEN, 2013, p. 403).

A finalidade desses algoritmos é aumentar a qualidade da informação disponível nos sites indexados, por isso os seus agentes são desenvolvidos para diminuir ou eliminar dos índices, aquele conteúdo considerado irrelevante, com uma interface inadequada nos resultados de busca, o Google espera manter um padrão inatingível diante de outros buscadores.

No “*Adwords*”, existe o algoritmo “*Ad Rank*”, que determina a ordem dos anúncios nos resultados de busca. Os atores envolvidos nos procedimentos realizados pelo mecanismo são o buscador, o anunciante e o usuário. O buscador é o mediador na seleção e qualificação do anúncio mais relevante, com a finalidade de satisfazer a necessidade do usuário. Então, o procedimento desse algoritmo é garantir um leilão das palavras-chave. O leilão é a disputa dos espaços oferecidos nos resultados de busca, assim os lances são decididos em grande parte pela taxa de cliques (valor a ser pago por clique), relevância (popularidade dos anúncios), e por último o resultado da qualidade de página (medida pela experiência do usuário).

Até mesmo, o cliente escolhendo uma alta taxa de cliques por seu anúncio, só garantirá espaço relevante, caso atenda aos critérios embutidos no *Ad Rank*. Por isso, não basta estipular

um alto orçamento pela palavra-chave, ainda deve conformar o anúncio com o público-alvo, e garantir ao usuário uma experiência satisfatória no site.

Desde 2011, o *Panda* atinge diretamente o *Ad Rank*, porque ele classifica as melhores páginas nos resultados de busca, ou seja, as páginas que tem design adaptado para cada tipo de dispositivo. Isto também, influencia diretamente os resultados dos anúncios, pois o valor é cobrado para divulgação em cada dispositivo. Claro, que pode-se delimitar os canais de alcance da campanha. Entretanto, o cliente do *Adwords* deve estar preparado não apenas para pagar o seu anúncio, mas também nas estratégias de usabilidade de seu site. O suporte da ferramenta oferecido aos seus clientes, garante quanto maior for a qualidade do anúncio, tanto será o sucesso da campanha. Inclusive, informando-os que o fato é responsável pela queda de valor pago por clique.

No período de 34 meses, Jansen, Liu e Simon (2013, p. 2), analisaram o desempenho do *Ad Rank*, e discutiram as influências do algoritmo nos resultados de pesquisa do Google. Tendo em vista, a teoria do efeito de posição serial, baseada no comportamento do usuário, enviesada pela sob posição dos links patrocinados, a qual reflete a interação humana no mecanismo de busca, que está sempre condicionada ao clique dos primeiros links de resultado da página.

Diante das singularidades de cada algoritmo, existem outras características que compõe o Google. Larry e Sergey (1999), utilizam a meta informação para organização dos dados, ou seja, as informações inferidas sobre um documento, tais elencadas como a reputação de busca (os termos pesquisados considerados confiáveis pelo buscador); a frequência atualizada (conteúdo atualizado da página); qualidade da informação indexada; a popularidade (as páginas mais populares ou mais acessadas); a citação (a recomendação de páginas).

A princípio, o mecanismo de busca era o *Page Rank*, contava com um banco de 14 milhões de palavras, e o uso de meta informação para organização dos dados armazenados. Nesse ínterim, os fundadores do Google, também buscaram a estrutura da Web Semântica. O conceito criado pelos especialistas propõe que,

A Web Semântica não é separada da Web, mas uma extensão da atual, em que a informação é dado de significado bem definido, possibilitando melhor cooperação no trabalho entre pessoas e computadores. Os primeiros passos para entrelaçar a Web Semântica dentro da estrutura existente da Web já estão a caminho. No futuro próximo, esses desenvolvimentos trarão novas funcionalidades significativas à medida que as máquinas se tornarem mais capazes de processar e "entender" os dados que eles exibem no momento (BERNES-LEE; HANDLER; LASSILA, 2001, p.1).

Consideramos, que esse conceito é uma das ambições dos fundadores do Google, não só trazer o conteúdo mais preciso e relevante como resultado de busca, mas também a estruturação do mecanismo dentro da Web Semântica. Assim, até mesmo antecipando a

necessidade de informação de seu usuário. Em Bernes-Lee, Hendler e Lassila (2001, p.2), consideram “O desafio da Web Semântica, portanto, é oferecer uma linguagem que expresse ambos dados e regras para raciocinar sobre os dados e, que admite padrão de qualquer sistema de representação do conhecimento a ser exportado para Web”, assim desenvolveu-se padrões de informações tácitas localizadas tanto fora quanto dentro da rede.

Nesse aspecto, existem muitas pesquisas para superar o desafio da linguagem dentro da web, podemos encontrá-las no W3C¹⁰, sendo um consórcio com o objetivo de ampliar ainda mais o potencial da Web. A WWW idealizada por Tim Bernes-Lee visou estruturar e organizar a informação da Internet, desenvolvendo padrões como de endereço para localização de sites.

Atualmente, o consórcio desenvolve a Web Semântica do conteúdo, isto auxilia os estudos da Internet das Coisas, tornando a máquina capaz de interpretar dados e interagir melhor com as pessoas. A criação de um código para cada linguagem viabilizou a troca de informação entre base de dados, por exemplo,

A linguagem humana prospera quando se usa o mesmo termo para significar coisas um pouco diferentes, mas não a automação. Imagine que eu contrate um serviço de palhaços mensageiros para entregarem os balões aos meus clientes em seus aniversários. Infelizmente, o serviço transfere os endereços do meu banco de dados para o seu banco de dados, sem saber que os "endereços" no meu são os locais onde as contas são enviadas, e que muitas delas são caixas postais. Meus palhaços contratados acabam entretendo vários funcionários dos correios - não necessariamente uma coisa ruim, mas certamente não é o efeito pretendido. Usar um URI diferente para cada conceito específico resolve esse problema. Um endereço que é um endereço de correspondência pode ser distinguido de um que é um endereço de rua e ambos podem ser distinguidos de um endereço que é um discurso (BERNES-LEE; HENDLER; LASSILA, 2001, p. 3).

Isto posto, a Web Semântica é estruturada em código de linguagem de programação, por isso exige a estruturação de léxico para seu uso em sistemas de rede. Assim sendo uma ferramenta de grande valia para identificação e padrão de termos aos webmasters. Nesta esfera, o website colaborativo chamado *schema.org*¹¹ faz parte dessas pesquisas, que é patrocinado pelo Google, Microsoft, Yahoo e Yandex. Os vocabulários são desenvolvidos por um processo de comunidade aberta, utilizando o site: <http://lists.w3.org/Archives/Public/public-schemaorg> para envio e teste de listas através do *GitHub*.

¹⁰ O Consórcio *World Wide Web* é um consórcio internacional no qual organizações filiadas, uma equipe em tempo integral e o público trabalham juntos para desenvolver padrões para web. Liderado pelo inventor da web Tim Bernes-Lee e o CEO Jeffrey Jaffe, o W3C tem como missão conduzir a *World Wide Web* para que atinja todo seu potencial, desenvolvendo protocolos e diretrizes que garantam seu crescimento de longo prazo. Essa informação foi retirada do site W3C, disponível em: <http://www.w3c.br/Sobre>.

¹¹ SCHEMA. Disponível em: <http://schema.org/>. Acesso em: 17 jul. 2017.

Agora, sabemos que, o Google utiliza os dados estruturados da Web Semântica, e existem dois tipos de tecnologias desenvolvidas pelos pesquisadores da Web. XML¹² e RDF¹³, a primeira padroniza a linguagem de classificação dos documentos por etiquetas, a segunda permite através de separação sintática, as sentenças das informações contidas no conteúdo. Ainda, no RDF os dados recebem uma URI (identificador de pesquisa universal), para que assim as máquinas possam interpretar aos dados, trazendo uma melhor execução em uma tarefa de busca.

Os padrões de RDF são encontrados no *schema*, sendo utilizado no mecanismo de busca para rastrear, coletar e organizar os dados encontrados na Web. As propriedades de RDF são *vocab*(vocabulário), *property* (propriedade), *typeof* (tipo), *content* (conteúdo). A partir delas, os rastreadores são capazes de ler determinadas informações contidas nos dados, e garantir a interoperabilidade deles, tanto no intercâmbio quanto na representação em diversas páginas da Internet. Por meio disso e de outras ferramentas, se tornou possível trazer a informação com mais qualidade ao usuário.

Em virtude do exposto, a próxima seção deste capítulo, apresentará uma discussão voltada para importância da representação do contexto de usuários nos sistemas de recuperação da informação. O Google é um mecanismo de busca, a na execução da tarefa é cumprida pela competência e aplicação algorítmica, porém quando trata da ferramenta *Adwords*, existe a lacuna da mediação entre o cliente e o usuário.

4.1.2 O contexto cognitivo na recuperação da informação em sistemas de buscas

Em Ingwersen e Willet (1995), a partir do que foi apresentado sobre o Google, faremos uma discussão, pois falam sobre as abordagens algorítmicas e cognitivas para recuperação da informação. Assim sendo, a junção dessas abordagens podem garantir um sistema de recuperação da informação qualitativa.

A abordagem algorítmica, cerca a ordenação do fluxo de informação com métodos estatísticos para ranqueamento, processamento de vetor, cálculo de medida de aproximação dos

¹² XML – *Extensible Markup Language*: são etiquetas adicionadas sobre os dados de páginas da Web ou seções de texto da página.

¹³ RDF – *Resource Description Framework*: os significados são expressados por essa linguagem, organizadas em triplas, quer dizer, por sujeito, verbo e objeto elementar de uma sentença.

termos de buscas. A abordagem cognitiva, leva em consideração o comportamento do usuário, em que se desenvolve os métodos de análise das necessidades de informação.

De acordo com os autores, a abordagem algorítmica não consegue acompanhar o contexto sócio organizacional dos usuários, portanto entendem que se deve utilizar ambas para sistema de recuperação da informação, “acreditamos que o algoritmo e as abordagens centrada no usuário são complementares em natureza, em que técnicas algorítmicas são necessárias, se deseja ter plena consideração dos contextos em que operam as pesquisas e as informações recuperadas devam ser usadas” (INGWERSEN; WILLET, 1995, p. 175).

Concordamos com isto, visto que esta pesquisa, mostra que são desenvolvidos procedimentos matemáticos para trazer uma determinada informação, porém nem sempre o sistema está programado para estabelecer relações cognitivas ao trazer os resultados de busca ao usuário, o que dependerá da experiência de pesquisa daquele indivíduo ao decidir sobre um determinado resultado.

Adiante, em Ingwersen e Jarvelin (2004), os pesquisadores relatam a importância de pesquisas sobre recuperação e busca de informação, na qual afirmam que essas dependem da Organização Social e o Contexto Cultural que envolve o cognitivo do ator da busca, e a interface disponibilizada entre o Sistema e os Objetos de Informação. Esse esquema evidencia que, a excelência de um sistema de busca está na presença do contexto cultural dos usuários, devido ao retorno mais preciso e relevante no resultado de busca. Assim, argumentam:

Apreciamos os esforços até agora buscando informações de exploração nos diversos contextos de atores e tarefas, mas também pensamos que a diversidade de contexto está distante da exaustão. Por exemplo, a Web é uma marca popular para muitas pesquisas de buscas de informação. Contudo, não é a única unidade coerente, porém aparece bastante diferenciada pelos diversos atores, tarefas e domínios. Assim sendo, não é suficiente analisar apenas o intermédio dos aspectos das atividades pesquisadas. Estas necessidades precisam ser relatadas na dimensão de tarefas (INGWERSEN; JARVELIN, 2004, p. 11).

Em vista desta discussão, os especialistas em recuperação da informação, refletem sobre a importância de um sistema de busca, que atenda ao contexto cultural e social de um indivíduo, não apenas estabelece as relevâncias perante os cálculos estatísticos, mas também a função do algoritmo quanto à cognição determinará a revocação e a precisão do mecanismo, em resposta a uma necessidade informacional.

Conforme observamos nesta explanação, existem algumas particularidades descritas no mecanismo de busca do Google, e a forma que acontece a recuperação da informação executada pelo buscador, na qual são perceptíveis as influências de seus algoritmos para o funcionamento do sistema, e os reflexos nos anúncios publicitários dos serviços oferecidos aos clientes através da ferramenta *Adwords*.

Aliás, o Google, desenvolve diversos aplicativos pelos quais obtém muitas informações a respeito de seus usuários. Nesse aspecto, os estudiosos Bernes-Lee, Hendler e Lassila (2001, p. 3) já antecipavam como forma de aprimorar o agente inteligente, com a colaboração dos próprios usuários:

O poder real da Web Semântica será percebido quando as pessoas criarem muitos programas que colem conteúdo da Web de diversas fontes, processam as informações e trocam os resultados com outros programas. A eficácia desses agentes de software aumentará exponencialmente à medida que mais conteúdo da Web legível por máquina e serviços automatizados (incluindo outros agentes) se tornarem disponíveis (BERNES-LEE; HENDLER; LASSILA, 2001, p. 3).

Tendo em vista, que os anunciantes devem estar atentos não só as estratégias do Marketing Digital para criação de campanhas, como também as implicações referentes aos agentes presentes no mecanismo, diante dos processos de coleta, identificação, classificação, enfim da organização de informações existentes na Web, assim coloca-se em perspectiva os procedimentos apresentados no desenvolvimento dos sites comerciais.

Por outro lado, buscar conhecer o seu público alvo atenderá aquilo que os especialistas dizem sobre a excelência da recuperação da informação, quer dizer, o contexto ao qual seu público está imerso, refletirá no momento da seleção de palavras-chave na campanha, onde são decididos o retorno e valor pretendido no alcance de anúncio. Até mesmo, porque a ferramenta oferecida apenas apresenta a competitividade em termo, ou a expressão de utilização em campanha.

Entretanto, o Google, por enquanto não apresentou uma tecnologia que emule por seus agentes as relações cognitivas. Nesse caso, a precisão do serviço oferecido aos anunciantes, apenas contam com os dados estatísticos da concorrência de palavras-chave compradas entre eles, e ainda estão envolvidos pela disputa da relevância no resultado de busca, quando devem oferecer o maior valor associado à palavra-chave que representa uma importância no anúncio.

Desse modo, os algoritmos refletem nessas disputas os termos ou expressões chaves mais buscados. Afirma-se, desse fato, que há evidência de reflexo dos algoritmos nos valores de palavras-chave na criação de campanhas, já que a concorrência é determinada pela procura de uma palavra em comparação aos outros anunciantes, e isto é a influência dos resultados de busca mais relevantes no mecanismo.

Se os avanços do mecanismo de busca forem além do esperado, ou seja, quando conseguirem incorporar a Web Semântica de conteúdo, atenderão e anteciparão as necessidades informacionais do usuário, inclusive resolverão o problema proposto por pesquisadores da recuperação de informação, em que os aspectos cognitivos e os algoritmos desempenharão juntos a tarefa de recuperação da informação, o que aprimorará a interação homem-máquina.

Entretanto, isto significa ter uma influência sem medida sobre os seus usuários. Ainda mais, diante de seus anunciantes, que terão intenso interesse na ferramenta, e estarão dispostos a aumentar a receita do Google. Enquanto, não se concretiza esse agente superinteligente, os especialistas em seleção de palavras-chave podem fazer a diferença nos resultados de campanha, estabelecendo a ponte entre o algoritmo e o contexto do usuário.

5 MATERIAIS E MÉTODOS

Tálamo (2004, p. 9) apresenta em seu artigo, duas fundamentações sobre métodos. A primeira, o desenvolvimento para se obter uma determinada finalidade, na qual existem distintas ideias sobre o mesmo assunto, assim sendo orientadas para uma abordagem adequada ao assunto discutido. A segunda, um planejamento que prevê etapas a serem realizadas no decorrer da pesquisa, para obter um resultado mais eficiente e econômico. Portanto, esta pesquisa se encaixa neste último fundamento, pois foram estabelecidos passos com o objetivo de uso da Indexação para tornar o processo de seleção de palavras-chave mais eficiente e econômico na criação de campanhas no *Adwords*, nos quais estão relacionados neste capítulo.

A pesquisa tem caráter exploratório, em que o método vai do dedutivo ao indutivo. As fontes utilizadas para a realização deste estudo foram empíricas e documentais. No primeiro, trazendo a experiência da arte de indexar palavras-chave; no segundo, a bibliografia da área de Biblioteconomia e Documentação, e também do Marketing Digital.

O levantamento bibliográfico foi realizado em base de dados internacionais: Web Science, Scopus, GALE, LISA, LISTA e ISTA, através dos portais SIBI e CAPES que cercam o cenário proposto pelo estudo da indexação envolvendo os clientes do Google. A seguir são relatadas as três etapas, que foram necessárias para realização desta investigação e os instrumentos para sua realização:

1ª etapa — o levantamento bibliográfico, buscou-se o amadurecimento dos conceitos, sobre a contribuição da Indexação conforme os pesquisadores da área de Biblioteconomia e Documentação, as estratégias do marketing digital aplicados ao *Google Adwords*, e a Inteligência Artificial, os algoritmos e a Web Semântica e a maneira como influenciam o funcionamento do Google, e por último apontamentos de especialistas da recuperação de informação em sistemas de busca.

2ª etapa — a coleta de dados, para amostra dos assuntos mais acessados no *Google Trends*, e depois a seleção do assunto mais pesquisado na ferramenta, no caso, o futebol, para elaboração de uma lista com palavras-chave, utilizando as orientações da Indexação (leitura do texto, exaustividade, especificidade e relações dos termos).

3ª etapa — a análise dos resultados da lista de termos testadas no planejador de palavras-chave do Google, a partir do comparativo entre os dados coletados e os dados da lista. Em vista disso, buscou-se responder se o uso das técnicas de indexação podem ser ou não favoráveis nas criações de campanhas aos clientes deste mecanismo de busca.

5.1 Coleta de dados

No *Google Trends*, a coleta foi realizada em dois momentos, no primeiro período de quatro meses (jan.-abr.), em 2017, conforme o cronograma estabelecido para pesquisa, sendo que a ferramenta foi visitada trinta e oito vezes, das quais foram selecionadas as cinco notícias mais acessadas, em variados dias da semana. Devido uma grande oscilação da busca, as pesquisas foram realizadas em horários diferentes, pois foi necessário acompanhar o comportamento dos usuários.

As notícias no Google estavam divididas em cinco categorias – Ciência/Tecnologia, Entretenimento, Esportes, Negócios, Notícias Principais e Saúde. Dessa maneira, um relatório de coleta foi elaborado, organizado por dia, mês e horário com as cinco primeiras notícias da lista, apenas foram considerados os resultados no território brasileiro, como delimitado no objeto de pesquisa, apresentados nos Apêndices B e C.

A segunda etapa da coleta seguiu com a elaboração de lista de palavras-chave sobre o futebol, a qual incluiu os nomes de times brasileiros, os nomes dos principais jogadores do momento, os nomes das roupas e acessórios utilizados na prática do esporte. No segundo período, em dezembro de 2017, foi elaborada uma lista de palavras-chave com o tema futebol. Já, em janeiro de 2018, foram realizados vinte testes no planejador de palavras-chave, também em variados períodos do dia, apresentado nos Apêndices D, E e F.

6 RESULTADOS

Após a realização das coletas, apresentamos os resultados de pesquisa, sendo que as expressões de pesquisas selecionadas foram testadas no planejador de palavras-chave do *Google Adwords*. O assunto mais representativo foi a categoria Esportes, como podemos ver no Gráfico 2,

Fonte: Da autora (2017).

A categoria ‘Esportes’ obteve a maior relevância em relação às outras categorias. Em segundo, as ‘Notícias Principais’, depois ‘Entretenimento’, ainda o índice sem grande expressão ‘Ciência/Tecnologia’, já as categorias ‘Negócios’ e ‘Saúde’ apresentaram resultados nulos. E também, foram analisados os índices por período, com a finalidade de descobrir o período mais acessado para cada tipo de categoria, tanto ‘entretenimento’ quanto ‘esporte’ foram os mais acessados pela manhã, e ‘notícias principais’ no período da tarde.

Fonte: Da autora (2017).

No Gráfico 2, a maioria dos acessos foram sobre Esportes, em todos os períodos do dia. O grande destaque foi o assunto “futebol”, não só relacionado aos campeonatos brasileiros, como também aos campeonatos europeus.

Os resultados da coleta de dados estão presentes na lista de palavras-chave, que foi aplicada no planejador do *Adwords*, porém buscou reforçar a seleção com dados de comportamento de consumo observados no público interessado em artigos esportivos. Na análise dos dados foi a aplicação na ferramenta do Google, aparada nos principais procedimentos da Indexação, também foram utilizados recursos figurativos da criação de mapas mentais e nuvem de *tags* para representar as principais palavras e suas relações no perfil de pesquisa do mecanismo de busca.

A elaboração da lista refletiu o público alvo de lojas virtuais de artigos esportivos, com ênfase no futebol. Por este motivo foram pesquisados dados em relação aos consumidores desse esporte. De acordo com a pesquisa do “Mercado de Consumo do Futebol Brasileiro em 2016”¹⁴ realizado pelas instituições (CNDL – Confederação Nacional de Dirigentes Lojistas) e (SPC Brasil – Serviço de Proteção ao Crédito do Brasil), observamos informações interessantes sobre o perfil analisado por eles.

Nesta pesquisa, o total de público entrevistado, 80% dos homens e 49,7% das mulheres têm interesse por futebol. As pessoas entrevistadas estão em três faixas etárias; de 18 a 34 anos, de 35 a 54 anos, de 55 anos acima. As classes sociais as quais pertencem é 77,7% da Classe A/B, e 64,5% da Classe C/D/E. Os produtos de suas preferências para consumo estão com 33,8% bebidas, 31,9% camisetas, 29,5% ingressos e 28,0% em aperitivos, em média são gastos R\$ 260,00 por mês com produtos ligados ao futebol, de acordo com a citada pesquisa.

Ainda segundo esta pesquisa, os itens mais comprados dos times são camisetas (39,9%), canecas ou copos (23,3%), chaveiros (19,9%), roupas para crianças (11,4%). Geralmente, essas compras são realizadas em lojas físicas especializadas em esporte (21,4%), loja física própria do time (21,2%), lojas virtuais/internet (16,8%), essas compras são estimuladas pelas lojas especializadas (64,5%), por propaganda na internet (15,5%), e os comerciais de TV (9,7%).

Na análise de consumo do futebol brasileiro de 2016 apresentou um dado importante, que foi comprovado pela presente pesquisa. As mídias pelas quais os torcedores acompanham as notícias sobre o time preferido, em primeiro lugar pela TV aberta (63,6%), em segundo são pelos sites (61,9%), em terceiro pelas redes sociais (42,7%), se relacionamos os sites e as redes

¹⁴ SERVIÇO DE PROTEÇÃO AO CRÉDITO DO BRASIL. Mercado de Consumo do Futebol Brasileiro em 2016. Disponível em: <www.spcbrasil.org.br/wpimprensa/wp-content/uploads/2016/09/Analise_Consumo_Futebol-1.pdf>. Acesso em: 11 de jun. 2018.

sociais superam à televisão. Então, podemos concluir que grande parte do público apaixonado pelo futebol está presente na internet, provável motivo de realizarem tantas pesquisas no Google sobre o assunto.

Em vista no perfil, foi desenvolvida a lista de palavras-chave em dezembro de 2017, baseada nos sites Super Placar (72 palavras) para as divisões de campeonatos brasileiros, loja virtual TKA (74 palavras) para artigos esportivos de futebol, ranking da FIFA – Federação Internacional de Futebol Associado (48 palavras) para pesquisar os dez melhores jogadores de futebol no mundo. Apresentamos as palavras-chave ou expressões chave mais representativas identificadas por esta pesquisa, por nuvem de *tags* na Figura 10.

Figura 10 - Tags da lista de palavras-chave sobre futebol

Fonte: Da autora (2018).

A partir de janeiro de 2018, foram iniciadas as aplicações de algumas dessas palavras-chave, nas quais resultaram em 20 testes (**Apêndice A**), sendo finalizados em 02/02/2018. Para evitar o viés de pesquisa, os testes ocorreram no período de férias das competições dos campeonatos brasileiros, desviou-se dos altos custos das palavras no planejador do *Adwords*. Estes foram aplicados nos três períodos do dia (manhã, tarde e noite), dado que há alternância no resultado do mecanismo de busca durante o dia.

Os termos foram classificados em categoria e subcategoria, os times de futebol e os melhores jogadores no ranking da FIFA foram classificados em categoria, e a subcategoria dos artigos esportivos. A especificação da indexação é fundamental para obter um resultado mais preciso no buscador, assim a classificação hierárquica de palavras-chave permite esta ordenação nos resultados mais enfáticos das campanhas.

6.1 Análise dos dados

Os dados foram coletados (**Apêndices B, C, D, E e F**), no planejador de palavras-chave do *Google Adwords*, organizados em planilhas no Excel, foram apresentados por dados comparativos no relatório do *Google Trends*, as informações colocadas foram o número de teste, as palavras-chave foram ordenadas por categoria ou subcategoria colocadas lado a lado, data, horário, preço e concorrência, veja no Quadro 2.

Quadro 2 – Organização e testes da lista de palavras-chave

Número de Teste	Palavras-Chave	Data	Horário	Preço	Concorrência
1	Camisa Cristiano Ronaldo	03/01/2018	10h30	R\$ 0,83	Alta
1	Camiseta Cristiano Ronaldo	03/01/2018	10h30	R\$ 0,74	Alta
1	Uniforme Cristiano Ronaldo	03/01/2018	10h30	R\$ 1,60	Média
2	Camisa Cristiano Ronaldo Portugal	03/01/2018	11h00	R\$ 0,76	Alta
2	Camiseta Cristiano Ronaldo Portugal	03/01/2018	11h00	R\$ 1,16	Média
2	Uniforme Cristiano Ronaldo Portugal	03/01/2018	11h00	R\$ 0,00	X
3	Chuteira Lionel Messi	04/01/2018	10h40	R\$ 0,44	Alta
3	Camisa Lionel Messi	04/01/2018	10h40	R\$ 0,73	Alta
3	Camiseta Lionel Messi	04/01/2018	10h40	R\$ 1,15	Alta
4	Corinthians Bola	06/01/2018	20h30	R\$ 0,10	Média
4	Sport Club Corinthians Paulista	06/01/2018	20h30	R\$ 0,92	Baixa
4	Corinthians	06/01/2018	20h30	R\$ 0,61	Baixa
5	Palmeiras Bola	10/01/2018	10h20	R\$ 0,34	Média
5	Palmeiras Luvas de goleiro	10/01/2018	10h20	R\$ 0,00	X
5	Sociedade Esportiva Palmeiras	10/01/2018	10h20	R\$ 0,88	Baixa
5	Palmeiras	10/01/2018	10h20	R\$ 0,41	Baixa

Fonte: Da autora (2018).

Dessa maneira, a tabela é a miscelânea do relatório da aplicação da lista de palavras-chave e os dados retirados do *Google Adwords*. Então, o número de teste é a quantidade de acesso ao planejador. As palavras-chave são sempre retiradas da primeira expressão do *Google Trends*, e as seguintes da lista de palavras-chave, sendo que a primeira é a categoria e a segunda é a subcategoria, em algumas situações existem a inversão das posições hierárquicas. A data é do acesso ao planejador e, o horário é o momento deste acesso. Também tem o preço de lance mínimo para utilização da palavra-chave na elaboração de campanha no *Google Adwords*.

Os dados foram coletados e relacionados em conformação com o público-alvo imaginado nessa situação. Conforme a lista apresentada (**Apêndice A**), observamos algumas situações dos testes, principalmente relacionada à Categoria de Artigos Esportivos, precisamente a variação de indumentária, por exemplo, as utilizações ou de camisa ou de camiseta, as duas tratam-se de vestimenta dos times de futebol, ao qual o torcedor faz uso para mostrar sua predileção de time. A palavra camisa manifesta o resultado sempre de alta concorrência no planejador de palavras-chave, ou seja, o termo é de alta popularidade na busca do Google, e também competitiva aos clientes, veja no Figura 11.

Figura 11 – Teste com as palavras-chave camisa e camiseta

Número de Teste	Palavras-Chave	Data	Horário	Preço	Concorrência
1	Camisa Cristiano Ronaldo	03/01/2018	10h30	R\$ 0,83	Alta
1	Camiseta Cristiano Ronaldo	03/01/2018	10h30	R\$ 0,74	Alta
1	Uniforme Cristiano Ronaldo	03/01/2018	10h30	R\$ 1,60	Média

Fonte: Da autora (2018).

Portanto, quando uma loja de artigos esportivos for criar sua campanha no *Google Adwords*, deverá utilizar camisa ao invés de camiseta, pois garantirá a conversão de público ao seu *e-commerce*, possivelmente efetuando a venda de produto. A concorrência é o resultado da procura pela palavra-chave e a competitividade dela entre os clientes Google conforme sua utilização, o que possibilita a definição do orçamento para aplicação de campanha. A concorrência alta é a necessidade de aumentar o lance mínimo que, aparece no planejador de palavra-chave; a média pode aumentar o lance mínimo ou manter o valor da ferramenta; a baixa também pode aumentar o lance mínimo para garantir o sucesso da campanha, porém é indício que a palavra-chave não é tão pesquisada no Google.

Em alguns testes, houve a inexpressividade de algumas palavras-chave, nesta condição é melhor não utilizar a expressão, pois não trará nenhum resultado em sua aplicabilidade no anúncio, isto pode representar desperdício de tempo no acompanhamento da campanha.

Do mesmo modo, também apresentou no planejador de palavra-chave resultados interessantes, alguns nomes de times de futebol que têm significados variados, tais como: Palmeiras, Vitória, São Paulo, São Caetano e América. Nestas situações, deve-se negatizar aqueles significados que o cliente não quer trazer nos resultados de sua campanha. Como Palmeiras significado de planta; ou Vitória como condição de vencedor; ou São Paulo como cidade; ou São Caetano como cidade; ou América como cidade e continente, a lista completa no Apêndice A.

Em todos, deve-se deixar claro que tratamos de times de futebol na criação de campanha. Assim, selecionamos **São Paulo Futebol Clube** ao destaca-los na condição de time de futebol, uma forma de especificar os resultados de campanha da loja virtual de artigos esportivos, exemplificado no Quadro 3.

Quadro 3 – São Paulo Futebol Clube X São Paulo

8	São Paulo Futebol Clube	19/01/2018	16h00	R\$ 0,51	Baixa
8	São Paulo	19/01/2018	16h00	R\$ 0,89	Baixa

Fonte: Da autora (2018).

E também, existem os times de futebol que possuem a mesma palavra em seu nome principal, a palavra Grêmio faz parte do nome de um time (Grêmio Futebol Porto Alegre) do Rio Grande do Sul, e de outro time (Grêmio Esportivo Osasco) de São Paulo. No segundo, usar o termo Grêmio para campanha de impressão, pode ser uma boa solução para atrair público alvo. Então, elabora-se duas campanhas, uma utiliza-se a palavra Grêmio com concorrência alta, porém com a finalidade de sensibilização. E a outra campanha, a expressão Grêmio Esportivo Osasco com objetivo da conversão, como mostra o Quadro 4.

Quadro 4 – Uso de Grêmio X Grêmio Esportivo Osasco

6	Gremio	13/01/2018	11h00	R\$ 0,38	Baixa
6	Gremio Kit	13/01/2018	11h00	R\$ 1,05	Baixa
7	Gremio	18/01/2018	17h45	R\$ 0,36	Alta
7	Gremio Esportivo Osasco	18/01/2018	17h45	R\$ 0,05	Baixa

Fonte: Da autora (2018).

Além da diminuição de custo de campanha, devido ao fato de evitar os resultados de pesquisa que não trará a conversão, quer dizer, não resultará na venda do produto, mas no máximo o aumento de impressão, que para o Marketing Digital é a quantidade de vezes da palavra-chave pesquisada no Google. A impressão não é negativa em uma campanha, pois nas estratégias de reconhecimento da marca, isto pode atrair mais público para uma determinada loja virtual.

Entretanto, na condição descrita, a impressão aumentará o custo de campanha, e nem sempre será finalizada com a venda. No Marketing Digital, os especialistas recomendam a definição do alcance de campanha, se o objetivo for a sensibilização, pode-se elaborar campanhas para obter maiores resultados com a impressão. Agora, se o objetivo for a conversão, quanto mais especificar as palavras-chave, melhor serão os resultados de venda de produto e/ou serviço do cliente, pois alcançará o público-alvo definido.

Os nomes completos de times de futebol foram retirados dos resultados do *Google Trends*, mas os resultados mais favoráveis para campanhas são as palavras nucleares dos times de futebol, aquelas palavras que os times de futebol são reconhecidos popularmente.

Lancaster (2004, p.186) discorre sobre dispositivos da linguagem de indexação, que seriam a Indexação Ponderada, Elos entre termos, os Indicadores de Função e os Sub cabeçalhos, todas elas têm o objetivo de aumentar a precisão da recuperação em sistemas de buscas. A Indexação Ponderada é a adoção de um peso numérico para cada termo, aquele ao receber um peso maior estará mais distante no resultado de busca, embora seja da indexação automática, no contexto do Google pode ser aplicado. Os Elos entre Termos são as relações que um termo tem ao outro. Os Indicadores de Função são fundamentais para relacionar as suas funções sintáticas. E por último, os Sub cabeçalhos farão a coordenação dos Elos entre Termos e os Indicadores de Funções.

Dessa maneira, possibilita a eliminação de ambiguidades, enfatizando o público alvo a ser alcançado na campanha. Como na lista de palavras-chave, a aplicabilidade da Indexação Ponderada dar um peso maior a palavra camiseta, com a finalidade de eliminá-la para aproximar um público mais interessado no produto da loja de Artigos Esportivos. Os Elos entre Termos e os Indicadores de Funções são associações dos Artigos Esportivos aos Times de Futebol, desenvolvendo o elo entre a camisa e o time da preferência de determinado torcedor. Os Sub cabeçalhos são a coordenação entre os Indicadores de Funções elaborados para as relações sintáticas entre os termos.

Ainda nesse momento, ressalta-se, que os algoritmos de mecanismos de buscas como o buscador em questão, não alcançaram o nível da busca exímia de inteligência artificial. Atualmente, estão sendo aplicados aos sistemas algorítmicos os esquemas de redes neurais, que executam de forma artificial esses dispositivos descritos acima, para trazer maior precisão aos resultados de busca. No entanto, o controle de vocabulário serve para facilitar a escolha de uma palavra ao invés da outra.

Muitos especialistas da área de Indexação, consideram improvável a substituição do trabalho humano pelas máquinas por completo nesta atividade de indexação na internet. Inclusive Lancaster e Warner (2001, p. 366) dizem: “... provavelmente ainda decorrerá muito tempo até que as máquinas sejam suficientemente inteligentes para substituir por completo os seres humanos nessas importantes atividades, se é que de fato um dia farão”.

Apresentamos as etapas da Indexação, utilizadas na seleção de palavras-chave para elaboração da lista sobre futebol, vejamos na Figura 9.

Figura 9 - Etapas de seleção de palavras-chave

Fonte: Da autora (2018).

As etapas relacionadas são evidenciadas na Figura 9, nesta escolha de palavras-chave são testadas a qualificação delas para utilização em campanhas no *Google Adwords*. A seleção foi a escolha de palavras mais adequadas do tema futebol. Já, a exaustividade era exploração máxima das palavras sobre o assunto. Assim, o conjunto delas foram conformadas diante do processo de tradução. A especificação determinou filtros de perfil dos consumidores. Outro processo foi aplicação de ponderação, no qual adotou-se um peso para cada palavra representativa do assunto, com maior probabilidade de busca pelo consumidor. O último processo foi determinar as relações semânticas neste conjunto de palavras-chave para aumentar a efetividade na busca.

Na seleção foi identificada e escolhida as palavras-chave, relacionadas à estrutura hierárquica sobre o assunto esporte. A organização foi determinada pelo mapa mental baseado em loja virtual de artigos esportivos, como na Figura 10.

Figura 10 - Mapa Mental do Futebol

Fonte: Da autora (2018).

Neste esquema, a categoria é Esporte, depois como subcategoria Futebol e Artigos Esportivos. E também, a relação da Performance de Jogadores que influenciam o relacionamento entre as subcategorias.

A partir disso, uma pesquisa foi realizada em lojas virtuais e portais de notícias esportivas, com a finalidade de buscar o máximo de referência para localizar os termos mais conhecidos. Assim, cumpriu-se a etapa de exaustividade.

No caso, a tradução foi verificada nos testes, pois como já apresentado nos termos “Camisa” ou “Camiseta” ficou definido que o primeiro é a melhor decisão para obter um resultado favorável na campanha, devido ao fato que camisa é o termo preferido dos pesquisadores no buscador.

A especificação foi determinada pela junção das subcategorias futebol e artigos esportivos, desenvolvendo filtros para atrair os consumidores dos artigos esportivos. Então, passamos a próxima etapa de ponderação, ou seja, os termos de maior relevância para obter o melhor êxito, como por exemplo, camisa, uniformes, chuteiras, os melhores jogadores do mundo dentre outros.

A última etapa da indexação foi encaixar as palavras-chave dentro das relações construídas, que estão evidentes no mapa mental (**Figura 10**). As subcategorias Futebol e Artigos Esportivos estão ligadas à categoria Esporte, a relação entre elas é Performance de Jogadores, portanto os filtros dependem dessas relações para definição de elos pensados no comportamento dos consumidores de artigos esportivos de futebol.

7 CONCLUSÕES

No final do século XX, as diversas transformações da sociedade provocadas pelas tecnologias da comunicação, e também estimuladas por grandes fluxos de informação, ocasionaram a “Era da Informação”, que se tornou o núcleo de mudanças sociais e econômicas. Nessa última, a informação é a propulsora da produtividade e da abertura de novos negócios. A Internet é o meio pelo qual perpassa toda a informação. Então, nos anos 90, quando inicia a expansão da rede mundial, há uma exponencial crescente na produção de informação.

Portanto, em muitos lugares no mundo empresas de tecnologias buscam inovações para organizar toda a informação disponível, torná-las acessível aos usuários da Internet. Assim, estas desenvolvem agentes inteligentes capazes de coletar, analisar, testar e disseminar a informação, portanto são desenvolvidos os mecanismos de buscas para filtrar aquilo que era interessante para cada usuário da Web.

Neste aspecto, algumas empresas de produtos ou serviços enxergaram oportunidade para promoção da marca e um grande alcance em venda, desenvolvendo espaços de e-commerce no ambiente virtual, e suas participações tornaram cada vez mais ativas na rede. O impacto disto foram as grandes empresas de tecnologias passaram a ter um papel influenciador na decisão de compra, tais desenvolveram algoritmos capazes de criar filtros aos usuários. Dessa maneira, oferecendo as informações consideradas mais úteis e precisas aos interesses dos indivíduos. Os envolvidos passam a ter grande satisfação, pois localizam com mais eficiência e eficácia o dado pesquisado.

No capítulo 2, acompanhamos pelos especialistas os procedimentos relacionados à Indexação, muito do que foi elaborado em conformidade com uma tecnologia disponível no respectivo momento histórico. Portanto, os índices estão intimamente ligados com os aparatos tecnológicos, porque partem do princípio de uma extensão da memória, assim auxiliam a leitura de um texto, que podem estar presentes em diversas formas, que em tempos primitivos eram criadas pinturas para informar sobre o espaço geográfico, clima, rituais e etc., assim cumprindo a funcionalidade do acesso à informação precisa.

O bibliotecário nesse processo é o especialista na mediação entre a informação e o usuário, sendo também um leitor proficiente e com objetivo profissional que facilita na escolha das palavras-chave mais representativas nas necessidades informacionais. O diferencial é a utilização de estudos estratégicos e técnicas de Indexação, ao desenvolver meios para essas representações. Assim, pesquisamos os principais critérios da indexação na representação de

assunto como, a exaustividade, a especificidade, a tradução, a ponderação pela relevância e a relação (associação) de assuntos.

No caso, os critérios de ponderação pela relevância e a relação são princípios utilizados nos índices automáticos, inclusos nos agentes inteligentes em sistemas de informação. Devido ao grande volume de informação, existe a necessidade da extração de muitos assuntos, e também de organizá-los em menor tempo. Sobretudo, quando nos referimos à rede mundial, diante da infinidade de dados desestruturados e senão houvesse estas ferramentas, as informações seriam inacessíveis.

Assim, consolidamos que a leitura do texto deve ser realizada por um profissional da área, já que este está apto a considerar os aspectos sócio cognitivos de um indivíduo ou de uma comunidade no ambiente informacional. Os procedimentos intimamente ligados à leitura, fazem parte da escolha e da representação de assuntos, ou seja, as execuções da exaustividade, especificidade, tradução, e por último, mesmo no caso do Google já sendo executados por máquinas, o estabelecimento de relações entre um termo e outro.

Portanto, a contribuição da Indexação no processo de seleção de palavras-chave está em seus estágios, e nesta atuação, o bibliotecário é o especialista que oferece o aperfeiçoamento necessário na escolhas de palavras ou expressões mais representativas, o que favorece às campanhas no *Google Adwords*. Assim, assegura-se ao cliente da ferramenta, uma seleção baseada nos interesses de seu público-alvo, perante a especialidade de estratégias na recuperação da informação em sistemas de buscas.

No capítulo 3, apresentamos as estratégias do Marketing Digital para elaboração dessas campanhas no *Adwords*, muitas delas são baseadas no posicionamento da marca, consequentemente a venda do produto ou serviço. Entretanto, essas não asseguram a relevância do anúncio, pois é tratado de maneira aparente diante da complexidade da seleção de palavras-chave, e um dos fatores delas representarem a eficiência e o baixo custo de campanha.

À vista disso, quando tratamos de criação das campanhas no *Adwords* precisamos seguir as estratégias do Marketing Digital, porém buscar outras técnicas complementares, que serão propícias na perspicácia dos resultados desejados na ferramenta de anúncios, como a Usabilidade da Web, e também da Organização do Conhecimento, principalmente da Representação de Assuntos. Dado que, o exposto nesta pesquisa, mostrou a importância de ter qualidade na experiência do usuário no site do cliente, e ainda o estudo da seleção de palavras-chave, que podem reduzir custos e trazer conversão a página.

Lidamos com influenciadores humanos e mecanismos delineados por lógica de mercado. As grandes empresas de tecnologias como Google, Amazon e Facebook são sustentadas por seus algoritmos, sempre programados em oferecer um produto personalizado. Tal lógica tende a produzir um comportamento corporativo competitivo, em busca de informações de possíveis clientes, transpondo as barreiras do privado desses indivíduos, que

Algoritmos são capazes de quantificar o comportamento humano em uma escala sobre-humana. Aí está a fonte de sua utilidade e seu poder. Porém, esse distanciamento da dimensão humana é também um de seus maiores defeitos – e um dos principais motivos pelo qual, desde a criação do Facebook, em 2004, e do Twitter, em 2006, os dois serviços somados conseguiram reunir mais de um bilhão de usuários (CHATFIELD, 2012, p. 93).

Os algoritmos fazem recomendações aos usuários conforme as suas buscas realizadas nos serviços da Internet. Dessa forma, as pessoas são colocadas em bolhas, por isso, não acessam outros tipos de serviços que não estejam contemplados no perfil ao qual o mecanismo captura no rastreamento.

O Google, em época de seu lançamento, conseguiu diminuir com seu principal algoritmo, o *Page Rank*, o custo de manutenção do índice, pois os outros mecanismos trabalhavam com índices realizados manualmente. Após uma década, o buscador alcançou uma hegemonia de mercado, devido aos seus agentes inteligentes serem capazes de criarem filtros do perfil de usuários, o que trouxe predomínio neste mercado.

Ademais, há a popularização de outros meios de comunicação, como por exemplo, o smartphone, esse aparato capaz de reunir outros equipamentos, suas funções vão além de um telefone, permite acesso à Internet em diversos lugares do mundo. Os dois fatores facilitaram essa situação, o barateamento de celulares e dos pacotes de telecomunicações, com isso as empresas observam o fenômeno com grande interesse, na possibilidade de expansão dos negócios.

Diante do exposto, as transformações tecnológicas promoveram mudanças socioeconômicas, por isso, a internet está a cada dia mais presente na vida das pessoas, o que representa aumento no uso de dispositivos móveis, lojas on-line e representações virtuais de serviços utilizados no mundo para além dela. Assim, o Google é um ícone daquilo que as pessoas mais acessam quando se fala sobre o assunto, o que atrai atenção de empresas como meio de anunciar os seus serviços e produtos.

Esses argumentos sustentam a importância desta pesquisa, e por este motivo, buscamos compreender as influências dos algoritmos do Google, tanto na relevância quanto nos anúncios apresentados nos resultados de busca. Para isto, pesquisamos sobre Inteligência Artificial, já que é a área responsável pelos desenvolvimentos de agentes inteligentes, assim é cumprido o

objetivo no Capítulo 4, referente à execução dos algoritmos e seus impactos no valor das campanhas criadas no *Adwords*.

Conclui-se que os algoritmos estudados influenciam nos valores de campanhas do *Google Adwords*, pois são responsáveis pela computação e verificação de relevância de assuntos pesquisados, como também na garantia de qualidade e organização das páginas conforme os interesses dos indivíduos. Até mesmo, reduzem os valores de cliques, caso o cliente atenda aos seus padrões estabelecidos no desenvolvimento de páginas do site.

Portanto, os anunciantes estarem atentos à parceria do Google com os desenvolvedores da Web Semântica, assegurará a relevância dos seus anúncios criados dentro do espaço do *Adwords*, pelo menos estará no padrão de qualidade exigido pelo mecanismo. Agora, o que não é oferecido pelo Google, seria a oferta de palavras-chave de qualidade no planejador da ferramenta de anúncios, que permitiriam maior conversão, além de menor custo em campanha.

Aliás, as discussões são reforçadas por pesquisadores de recuperação da informação em sistemas de buscas, que afirmam a necessidade de alinhar o contexto sócio cognitivo aos algoritmos, porque o primeiro traz a subjetividade da língua de um indivíduo. E o segundo tem a capacidade de organizar e disponibilizar os dados em larga-escala, tarefa executada pelo Google com grande eficiência. Agora, em 2018, a empresa divulgou inovações que prometem resolver as interações humano-máquina, ao desenvolver um agente inteligente com capacidade de aprendizagem, a partir dos dados estruturados da Web Semântica, que foi idealizado pelos seus desenvolvedores em 2000.

Assim, elaboramos dois capítulos teóricos de discussão (Capítulo 2 e 4), o primeiro, trazemos as contribuições da Biblioteconomia e Documentação, com ênfase das discussões na Indexação. O segundo, expõe os conceitos sobre Inteligência Artificial, que inclui a função e aplicabilidade dos algoritmos, linguagens descritivas de páginas da internet desenvolvidas pelos especialistas da Web Semântica, e também apresentação dos principais algoritmos do Google, fechando com os especialistas de recuperação da informação em sistemas de buscas automatizados.

Em vista disso, o levantamento bibliográfico mostrou as influências dos algoritmos do Google na concorrência de palavras-chave no *Adwords*, e que os agentes inteligentes executam tarefas programadas com interesses de seus desenvolvedores. Nesta condição, a empresa Google quer alcançar o nível de excelência nos resultados de busca, até mesmo antecipando as necessidades de seus usuários. Então, colocam-se como mediadores entre os anunciantes e os seus usuários, porém as ferramentas do *Adwords* não auxiliam diretamente na escolha das palavras-chave pelos anunciantes, ou seja, pressupõe um ambiente competitivo entre os

anunciantes, pois disputam em leilões os termos mais representativos de seus produtos ou serviços.

Contudo, esta pesquisa oferece uma segunda opção como mediação entre anunciante e público-alvo, em torno da seleção de palavras-chave, o bibliotecário, porque é o especialista para representação de assuntos, como mostrado nos procedimentos da Indexação, em que resolve um outro problema, a seleção realizada conforme os aspectos cognitivos do público-alvo numa campanha de *Adwords*.

Os métodos desta pesquisa foram estabelecidos pela Representação Documentária, com a finalidade de contribuir na diminuição de palavras-chave que oneram as campanhas no *Adwords*, baseadas na observação das estratégias do Marketing de Busca em links patrocinados, das quais estão voltadas para venda do produto ou serviço de e-commerce, sem muitas vezes perceber as singularidades em torno da recuperação da informação em mecanismos de buscas.

Nos capítulos (5 e 6) temos a parte aplicada desta pesquisa, que conta com a coleta de dados realizadas no *Google Trends* durante um certo período, no qual resultou no assunto mais pesquisado do Google. Tendo em vista disso, o resultado foi futebol classificado na categoria Esportes na ferramenta *Trends*, a partir desse assunto foi elaborada uma lista aplicada no *Google Adwords*.

Em análise da lista aplicada, as palavras-chave camisa e camiseta mostraram os resultados mais interessantes. Até mesmo, porque são aquelas que expressaram grande representatividade do público-alvo, quando relacionadas aos produtos oferecidos por lojas virtuais de artigos esportivos, pois é o produto com o segundo maior índice de venda nas lojas especializadas no assunto. Ademais, foi identificado na literatura que, os consumidores buscam sites de notícias e as redes sociais para obterem informações sobre os campeonatos de futebol, e também de comprarem as camisas e outros produtos em lojas virtuais especializadas.

Os procedimentos da Indexação foram utilizados para seleção dessas palavras-chave, o que facilitou para verificação desses fenômenos da língua. Além, da elaboração de um mapa mental para desenhar o fluxo de informação relacionado a Categoria Esporte, que traz como sub assuntos Futebol, Performance dos Jogadores e Artigos Esportivos. Em especial, a performance dos jogadores influenciam nos valores de palavras-chave associadas à intenção de venda de camisas.

Esperamos com esta pesquisa mostrar a contribuição necessária da Indexação, para diminuir os custos de campanhas no *Google Adwords*, com a intenção da inclusão do bibliotecário neste processo de seleção das palavras-chave, já que se trata do profissional com habilidades e competências necessárias na execução desta tarefa, que também estará atento as

influências exercidas pelos algoritmos tanto nos valores de palavras-chave quanto na recuperação desses termos ao público de interesse no mecanismo de busca. Inclusive, com orientações de arquitetura da informação para garantir uma experiência excelente ao público desejado pelo e-commerce, que foi mostrado na pesquisa ser capaz de diminuir a concorrência e o valor de uma palavra-chave.

Diante da complexidade desse assunto, deixamos algumas sugestões de trabalhos futuros com enfoque nas implicações decorrentes das aplicabilidades do desenvolvimento da Inteligência Artificial na recuperação da informação, e a possibilidade de alinhamento das práticas consolidadas da Biblioteconomia e Documentação no fazer do bibliotecário, pois ambas podem andar juntas para aperfeiçoamento da interação homem-máquina.

Portanto, são sugeridos estudos como, o acompanhamento de aplicativos com inovações relacionadas aos assistentes de buscas automatizadas e a relação da máquina com o homem; a verificação dos conceitos de semiótica a respeito da teoria de signo e aplica-la no mundo digital; o desenvolvimento de ontologias para representação do conhecimento para construção de site e testar sua efetividade em links patrocinados; a proposição da representação de assuntos na descrição de páginas da Internet, principalmente considerando os padrões de linguagem aplicados na W3C; a realização do estudo de caso, em aplicação das contribuições apresentadas nesta pesquisa, assim observar dados quantitativos dos resultados destas aplicações.

REFERÊNCIAS

- AMAZON. Alexa: keyword research, competitive analysis, & website ranking. Disponível em: <<https://www.alexa.com/>>. Acesso em: 15 de mar. 2017.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: Métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.
- ANDERSON, C. **A cauda longa**. Rio de Janeiro: Elsevier, 2006.
- BARTHES, R. **Elementos de semiologia**. 19 ed. São Paulo: Cultrix. 2012.
- BEARE, G. Past, present and future. **The Indexer**, Norwich, v. 25, n. 4, p. 257-264, out. 2007. Disponível em: <<http://www.ingentaconnect.com/content/index/tiji/2007>>. Acesso em: 27 abr. 2017.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Web Semantic. **Scientific American**, mai. 2001. Disponível em: <<https://www.scientificamerican.com/author/tim-berners-lee-james-hendler-and-ora-lassila/>>. Acesso em: 19 out. 2017.
- BERNIER, C.; BORKO, H. **Indexing Concepts and Methods**. New York: Academic Press, 1978.
- BOURGEOIS, T. Understanding the Google Marketing Stack. **Econtent**, Wilton, v. 39, n. 6, 2016. Disponível em: <<https://search.proquest.com/docview>>. Acesso em: 05 jun. 2017.
- BRIN, S.; PAGE, L. **The anatomy of a Large-Scale Hypertextual Web Search Engine**. In: Seventh International World-Wide Web Conference (WWW 1998), Stanford Infolab Publication Server, Brisbane, apr. 1998. Disponível em: <<http://ilpubs.stanford.edu:8090/361/>>. Acesso em: 28 mar. 2016.
- BURKE, P. **Uma história social do conhecimento II**: da enciclopédia à Wikipédia. Rio de Janeiro: Zahar, 2012.
- CASTELLS, M. **A galáxia da internet**: reflexões sobre a internet, os negócios e a sociedade. Rio de Janeiro: Zahar, 2003.
- CHATTIFIELD, T. **Como viver na era digital**. Rio de Janeiro: Objetiva, 2012.
- CHAUMIER, J. Indexação: conceito, etapas e instrumentos. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 21, n. 1, p. 63-79, 1988.
- CLARKES, S.; ZENG, M. Standard Spotlight, from ISO 2788 to ISO 25964: the evolution of thesaurus standards towards interoperability and data modeling. **Information Standards Quarterly**, v. 24, n. 1, p. 20-26, 2012.
- CLEVELAND, D.; CLEVELAND, A. **Introduction to indexing and abstracting**. 3 ed. Colorado: A division of Greenwood Publishing Group, 2001.
- CORMEN, T. H. **Desmistificando algoritmos**. Rio de Janeiro: Elsevier, 2014.
- EXPERIAN MARKETING SERVICE. Disponível em: <<http://www.experian.com/intl/brazil.html>>. Acesso em: 08 mar. 2016.
- DAVIS, H. **Ferramentas de publicidade do Google**: como ganhar dinheiro com a AdSense, o Adwords e as Apis do Google. São Paulo: Novatec Editora, 2010.

FUNGMAN, R. **Subject analysis and indexing: theoretical foundation in practice advice**. Frankfurt: Indeks Verlag, 1993.

FIFA. Disponível em: < <https://www.fifa.com/fifa-world-ranking/>>. Acesso em: 08 de dez. 2017.

FUJITA, M. S. L. A Leitura Documentária na perspectiva de suas variáveis: leitor-texto-contexto. **Data Grama Zero**, v.5, n.4, 2004. v. 5, n. 4, p. A01, 2004. Disponível em: <<http://www.brapci.inf.br/v/a/7547>>. Acesso em: 27 de Jul 2018.

GABRIEL, M. **SEM e SEO dominando o marketing de busca**. São Paulo: Novatec Editora, 2009.

GABRIEL, M. **Marketing na era digital: conceitos, plataformas e estratégias**. São Paulo: Novatec Editora, 2010.

GONZALES, P.; MORAL, P. Methodologies for monitoring website performance. **Online Information Review**, v. 38, n. 4, p. 575-588, 2014. Disponível em: <<https://doi.org/10.1108/OIR-12-2013-0267>>. Acesso em: 23 de jun. 2017.

GOOGLE. Disponível em: < www.google.com.br>. Acesso em: 26 de jul. 2018.

GOOGLE ADS. Disponível em: < ads.google.com/GoogleAds>. Acesso em: 03 de jan. 2017.

GREEN, J. **How to rank in Google: SEO strategies post panda and penguin**. New York: Kindle Edition, 2013.

INGWERSEN, P.; WILLET, P. An introduction to algorithmic and cognitive approaches for information retrieval. **Libri**, Copenhagen, v. 45, p. 160-177, 1995.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Pesquisa Nacional por amostra de domicílios: sínteses de indicadores 2013**. Rio de Janeiro: IBGE, 2013. 288 p. Disponível: < <https://biblioteca.ibge.gov.br/visualizacao/livros/liv94414.pdf>>. Acesso em: 22 de mar. 2018.

JANSEN, B. J.; LIU, Z.; SIMON, Z. The effect of Ad Rank on the performance of keyword advertising campaigns. **Journal of the american society for information science and technology**, Pennsylvania, 2013.

JÄRVELIN, K.; INGWERSEN, P. Information seeking research needs extension towards tasks and technology. **Information Research**, v. 10, n.1, 2004.

KERCKHOVE, D. **A pele da cultura: investigando a nova realidade eletrônica**. São Paulo: Anna blume, 2009.

KOBASHI, N.; SANTOS, C. A. M. Leitura documentária: aspectos pragmáticos da recepção e condensação de textos para indexar e resumir. **Revista EDICIC**, v.1, n.4, p.130-140, Oct./Dez. 2011. Disponíbleen: <<http://www.edicic.org/revista/>>. Acesso em: 11 de Jun. 2018.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2ed. Brasília: Briquet de Lemos, 2004.

LARA, M. G. L.; TÁLAMO, M.F.G.M. O campo da Linguística Documentária. **Transinformação**, Campinas, v.3, n.18, p.203-211, 2006.

LARA, M. G. L. Linguagem documentária e terminologia. **Transinformação**, Campinas, v.3, n.16, p.231-240, 2004.

LOGAN, R. **Que é informação:** a propagação da informação na biosfera, na simbologosfera, na tecnosfera e na econosfera. Rio de Janeiro: Contraponto, 2012.

LOPES, E. A semântica estrutural de Greimas. In: _____. **Fundamentos da lingüística contemporânea.** São Paulo: Cultrix, 1987.

MCLUHAN, M. **Os meios de comunicação como extensões do homem.** São Paulo: Cultrix, 2007.

NATIONAL INFORMATION STANDARDS ORGANIZATION. **ANSI/NISO Z39.19:** guidelines for the construction, format, and management of monolingual controlled vocabularies. Maryland, 2005.

NORVIG, P.; RUSSEL, S. **Artificial Intelligence: a modern approach.** 3 rd ed. New Jersey: Pearson, 2010.

PARISER, E. **O filtro invisível:** o que a internet está escondendo de você. Rio de Janeiro: Zahar, 2012.

SAGER, J. Prólogo: la terminologia, ponte entre varios mundos. In: CABRÉ, M. T. **La terminologia:** teoría, metodologia, aplicaciones. Barcelona: Ed. Antártida; Empúries, 1993. p.11-17.

SCHEMA. Disponível em: <<http://schema.org/>>. Acesso em: 17 de jul. 2017.

SUPER PLACAR. Disponível em: <<http://www.superplacar.com.br/>>. Acesso em: 08 de dez. 2017.

SERVIÇO DE PROTEÇÃO AO CRÉDITO DO BRASIL. Mercado de Consumo do Futebol Brasileiro em 2016. Disponível em: < https://www.spcbrasil.org.br/wpimprensa/wp-content/uploads/2016/09/Analise_Consumo_Futebol-1.pdf>. Acesso em: 11 de jun. 2018.

TÁLAMO, M. F. T. G. A. M. A pesquisa: recepção da informação e produção do conhecimento. **Data Grama Zero**, v. 5, n. 2, p. A01, 2004. Disponível em: <<http://www.brapi.inf.br/v/a/7748>>. Acesso em: 09 Jun. 2018.

TKA. Disponível em: <<https://www.tkaesportes.com.br/>>. Acesso em: 08 de dez. 2017.

VAIDHYANATHAN, S. **A Googlelização de tudo.** São Paulo: Cultrix: Kindle Edition, 2011.

W3C BRASIL. Disponível em: <<http://www.w3c.br/Sobre/>>. Acesso em: 17 de jul. 2017.

APÊNDICE A - Tabela de Aplicação das Palavras-Chave

Número de Teste	Palavras-Chave	Data	Horário	Preço	Concorrência
1	Camisa Cristiano Ronaldo	03/01/2018	10h30	R\$ 0,83	Alta
1	Camiseta Cristiano Ronaldo	03/01/2018	10h30	R\$ 0,74	Alta
1	Uniforme Cristiano Ronaldo	03/01/2018	10h30	R\$ 1,60	Média
2	Camisa Cristiano Ronaldo Portugal	03/01/2018	11h00	R\$ 0,76	Alta
2	Camiseta Cristiano Ronaldo Portugal	03/01/2018	11h00	R\$ 1,16	Média
2	Uniforme Cristiano Ronaldo Portugal	03/01/2018	11h00	R\$ 0,00	X
3	Chuteira Lionel Messi	04/01/2018	10h40	R\$ 0,44	Alta
3	Camisa Lionel Messi	04/01/2018	10h40	R\$ 0,73	Alta
3	Camiseta Lionel Messi	04/01/2018	10h40	R\$ 1,15	Alta
4	Corinthians Bola	06/01/2018	20h30	R\$ 0,10	Média
4	Sport Club Corinthians Paulista	06/01/2018	20h30	R\$ 0,92	Baixa
4	Corinthians	06/01/2018	20h30	R\$ 0,61	Baixa
5	Palmeiras Bola	10/01/2018	10h20	R\$ 0,34	Média
5	Palmeiras Luvas de goleiro	10/01/2018	10h20	R\$ 0,00	X
5	Sociedade Esportiva Palmeiras	10/01/2018	10h20	R\$ 0,88	Baixa
5	Palmeiras	10/01/2018	10h20	R\$ 0,41	Baixa
6	Gremio	13/01/2018	11h00	R\$ 0,38	Baixa
6	Gremio Kit	13/01/2018	11h00	R\$ 1,05	Baixa
7	Gremio	18/01/2018	17h45	R\$ 0,36	Alta
7	Gremio Esportivo Osasco	18/01/2018	17h45	R\$ 0,05	Baixa
8	São Paulo Futebol Clube	19/01/2018	16h00	R\$ 0,51	Baixa
8	São Paulo	19/01/2018	16h00	R\$ 0,89	Baixa
9	Esporte Clube Vitória	20/01/2018	15h30	R\$ 0,34	Baixa
9	Vitória	20/01/2018	15h30	R\$ 2,12	Baixa
9	Camiseta (camisa) Vitória	20/01/2018	15h30	R\$ 0,27	Alta
10	Clube de Regatas Flamengo	22/01/2018	11h20	R\$ 0,45	Baixa
10	Flamengo	22/01/2018	11h20	R\$ 0,31	Baixa
10	Flamengo Camisa	22/01/2018	11h20	R\$ 0,50	Alta
10	Flamengo camiseta	22/01/2018	11h20	R\$ 0,55	Alta
11	Botafogo de Futebol e Regatas	23/01/2018	13h15	R\$ 0,67	Baixa
11	Botafogo	23/01/2018	13h15	R\$ 0,10	Baixa
12	Clube Atlético Mineiro	24/01/2018	14h00	R\$ 0,46	Baixa
12	Atlético Mineiro	24/01/2018	14h00	R\$ 0,23	Baixa
13	Santos Futebol Clube	26/01/2018	13h20	R\$ 0,31	Baixa
13	Santos	26/01/2018	13h20	R\$ 0,77	Baixa
14	Fluminense Futebol Clube	27/01/2018	12h00	R\$ 0,51	Baixa
14	Fluminense	27/01/2018	12h00	R\$ 0,27	Baixa
14	Fluminense Uniforme	27/01/2018	12h00	R\$ 0,79	Baixa
14	Uniforme Fluminense	27/01/2018	12h00	R\$ 0,45	Média
15	Vila Nova Futebol Clube	28/01/2018	10h00	R\$ 0,71	Baixa
15	Vila Nova	28/01/2018	10h00	R\$ 1,49	Baixa
15	Vila Nova Camisa	28/01/2018	10h00	R\$ 0,35	Alta
16	Sport Futebol Clube	29/01/2018	14h40	R\$ 0,67	Baixa
16	Sport	29/01/2018	14h40	R\$ 2,82	Baixa
16	Sport Camiseta	29/01/2018	14h40	R\$ 0,00	X
16	Sport Camisa	29/01/2018	14h40	R\$ 0,75	Alta
17	Fortaleza Futebol	30/01/2018	21h15	R\$ 0,00	X
17	Fortaleza Camisa	30/01/2018	21h15	R\$ 0,27	Alta
17	Fortaleza Camiseta	30/01/2018	21h15	R\$ 0,00	X
18	Associação Desportiva de São Caetano	31/01/2018	18h20	R\$ 0,00	X
18	São Caetano	31/01/2018	18h20	R\$ 0,10	Baixa
18	São Caetano Camisa	31/01/2018	18h20	R\$ 0,26	Alta
19	Cruzeiro Esporte Clube	01/02/2018	22h00	R\$ 0,67	Baixa
19	Cruzeiro	01/02/2018	22h00	R\$ 0,63	Baixa
19	Cruzeiro Camisa	01/02/2018	22h00	R\$ 0,56	Alta
19	Cruzeiro Camiseta	01/02/2018	22h00	R\$ 0,68	Média
20	América de Futebol Clube	02/02/2018	20h45	R\$ 0,43	Baixa
20	América	02/02/2018	20h45	R\$ 0,63	Baixa
20	América Futebol Camisa	02/02/2018	20h45	R\$ 0,00	X
20	América Camisa	02/02/2018	20h45	R\$ 0,59	Alta