



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
E GESTÃO DO CONHECIMENTO

Antônio Carlos Picalho

Das bibliotecas aos buscadores: testando técnicas avançadas para a recuperação da
informação em pesquisas por documentos na web

Florianópolis
2023

Antônio Carlos Picalho

**Das bibliotecas aos buscadores: testando técnicas avançadas para a recuperação da
informação em pesquisas por documentos na web**

Dissertação submetida ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Engenharia e Gestão do Conhecimento.

Orientadora: Prof^ª. Luciane Maria Fadel, Dra.

Coorientador Prof. Alexandre Leopoldo Gonçalves, Dr.

Florianópolis

2023

P585b Picalho, Antonio Carlos.

Das bibliotecas aos buscadores: testando técnicas avançadas para a recuperação da informação em pesquisas por documentos na web / Antonio Carlos Picalho. – Florianópolis, SC, 2023.

124 p. : il. color.

Orientadora: Luciane Maria Fadel.

Dissertação (mestrado) – Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, 2023.

Inclui referências

Inclui índice

1. Buscadores. 2. História dos buscadores. 3. Lógica aplicada. 4. Ferramentas de busca na web. 5. Expressões de busca. I. Fadel, Luciane Maria, orient. II. Golçalves, Alexandre Leopoldo. III. Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento. IV. Título.

CDU 025.4.036

Antônio Carlos Picalho

Das bibliotecas aos buscadores: testando técnicas avançadas para a recuperação da
informação em pesquisas por documentos na web

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca
examinadora composta pelos seguintes membros:

Profa Ketry Gorete Farias dos Passos, Dra.
Universidade do Estado de Santa Catarina

Prof. Richard Perassi Luiz De Sousa, Dr.
Universidade Federal de Santa Catarina

Prof. Francisco Antonio Pereira Fialho, Dr.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi
julgado adequado para obtenção do título de mestre em Engenharia e Gestão do Conhecimento.

Prof. Dr. Roberto Carlos dos Santos Pacheco
Coordenação do Programa de Pós-Graduação

Prof^a Luciane Maria Fadel, Dra.
Orientadora

Florianópolis, 2023.

Este trabalho é dedicado a todos os bibliotecários e bibliotecárias que acreditam e compreendem a nossa responsabilidade social enquanto profissionais da informação.

AGRADECIMENTOS

Agradeço primeiramente a minha mãe Magali e ao meu pai Antônio que desde sempre se preocupam com meu bem-estar e me apoiam muito nos estudos. Mesmo a mais de 700km de distância vocês nunca deixaram de estar presentes. Graças ao esforço de vocês, comprometimento e responsabilidade adquirida por meio dos ensinamentos de vocês e com auxílio de políticas de ações afirmativas, tive a oportunidade de ser o primeiro do núcleo familiar a concluir um ensino superior e também ingressar na pós-graduação. Não serei o último. Obrigado por tudo, eu amo vocês.

A minha orientadora Luciane Fadel pela paciência, atenção, disponibilidade em ensinar —muitas vezes repetir o que já havia sido dito—, pelas correções e principalmente pela confiança no meu trabalho. Não poderia ter estado sob melhor orientação, foi um presente tê-la conhecido. Conte comigo para o que precisar.

Um muito obrigado para Lani Lucas e Jane Lecardelli. Mais que colegas que admiro imensamente, vocês são amigas especiais e caminharam junto comigo nesses dois anos.

Obrigado a Fernanda Ribeiro, bibliotecária e colega de turma. Fazer cada artigo de disciplina contigo foi um privilégio, uma das melhores pessoas com quem já trabalhei.

Obrigado a Maicon Puppi e Árlon Lima pela amizade durante esse período. Compartilhar o dia a dia da pós-graduação com vocês tornou tudo mais leve.

Obrigado as minhas amigas Nathália Tonetto, Fernanda Pimentel e Thays Pagani. Vibramos juntos pelas conquistas uns dos outros e isso para mim não tem preço. Amo nossa amizade.

Agradeço a Universidade Federal de Santa Catarina (UFSC) na figura dos seus docentes e técnicos que trabalham para oferecer o melhor do ensino público, gratuito e de qualidade. Agradeço também ao Laboratório de Ambientes Hiperídia para Aprendizagem (Hiperlab) e a Biblioteca Universitária (BU) por terem sido minha segunda casa durante um ano de mestrado.

Por fim, o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“Os buscadores se tornaram tão presentes na vida das pessoas que existe o perigo de elas acreditarem que, se uma coisa não aparece no Google, ela não existe” (JANES, 2004).

RESUMO

Pesquisar on-line é uma atividade presente na vida de muitas pessoas, principalmente no que diz respeito às pesquisas de rotina para resolver problemas do cotidiano. Ao contrário de uma pesquisa em uma biblioteca física, a pesquisa de rotina em um buscador pode levar menos de um minuto, a depender do grau de dificuldade em obter o resultado esperado. Além da rapidez, a busca on-line também oferece outras vantagens sobre a pesquisa física, como o número de resultados, os vários idiomas, diversas mídias resultantes e diversas fontes. Simultaneamente, o volume de dados disponíveis na web cresce dia após dia, e filtrar os resultados pode determinar a eficácia desta busca. Na mediação entre os valores da pesquisa e seus resultados estão os buscadores. Dentre uma vasta gama de buscadores, o Google é o mais utilizado no Brasil e dispõe de operadores avançados que servem para refinar até as pesquisas mais simples. Porém, o uso de operadores avançados depende de um conhecimento técnico o qual, muitas vezes, não pertence ao domínio de um usuário leigo. Os buscadores também retornam outros elementos além dos resultados, compondo uma interface complexa. Esta complexidade pode dificultar o uso e desviar a atenção do usuário para estes elementos. Diante disso, o presente trabalho buscou examinar quais características nos resultados de pesquisa se destacam a partir o uso de operadores avançados na busca por documentos na web. Partindo do pressuposto de que a utilização desses operadores pode sugerir maior agilidade na busca e precisão aos resultados, a pesquisa identificou os elementos e outras características apresentadas junto aos resultados de uma busca. Por meio de uma pesquisa experimental, foram coletados e analisados os resultados presentes em 380 capturas de tela de usuários que reproduziram uma busca no Google em cinco cenários diferentes: busca livre; busca normalizada; busca com o operador booleano AND e acréscimo de aspas; busca com o operador filetype: e busca com o operador site:. Os resultados ressaltaram características como anúncios de pesquisa, sugestões de pesquisas já realizadas por outros usuários e quantidade aproximada de resultados. Entende-se que o uso de operadores avançados de pesquisa pode interferir não apenas nos resultados apresentados, mas também na sua forma. Considera-se ainda que a interface do Google o qual preza pela simplicidade, não privilegia os operadores. Portanto, que o design da interface do Google poderia usufruir da remediação para que a oscilação entre a transparência e opacidade dos operadores contribua ativamente para instruir os usuários na sua aplicação. Este design estaria agindo para o movimento contínuo e alternado de influência das camadas culturais e computacionais.

Palavras-chave: História dos buscadores. Mecanismos de busca. Ferramentas de busca na Web. Expressões de busca. Recuperação da informação. Mídia do conhecimento.

ABSTRACT

Searching online is an activity present in many people's lives, especially when it comes to routine searches to solve everyday problems. Unlike a search in a physical library, a routine search in a search engine can take less than a minute, depending on the degree of difficulty in obtaining the expected result. In addition to speed, online searching also offers other advantages over physical searching, such as the number of results, multiple languages, multiple output media, and multiple sources. At the same time, the volume of data available on the web is growing day by day, and filtering the results can determine the effectiveness of this search. Mediating between the values of the search and its results are the search engines. Among a wide range of search engines, Google is the most used in Brazil and has advanced operators that can refine even the simplest searches. However, the use of advanced operators depends on technical knowledge which often does not belong to the domain of a lay user. Search engines also return elements other than results, making up a complex interface. This complexity can hinder the use and divert the user's attention to these elements. In light of this, the present work sought to examine which features in search results stand out from the use of advanced operators in the search for documents on the Web. Based on the assumption that the use of these operators may suggest greater search agility and precision to the results, the research identified the elements and other characteristics presented in the results of a search. Through an experimental research, the results present in 380 screenshots of users who reproduced a Google search in five different scenarios were collected and analyzed: free search; normalized search; search with the Boolean operator AND and the addition of quotation marks; search with the operator filetype: and search with the operator site:. The results highlighted features such as search ads, suggestions of searches already performed by other users, and approximate amount of results. It is understood that the use of advanced search operators may interfere not only in the results presented, but also in their form. It is also considered that Google's interface, which prizes simplicity, does not privilege operators. Therefore, that Google's interface design could take advantage of remediation so that the oscillation between the transparency and opacity of operators actively contributes to instruct users in their application. This design would be acting towards the continuous and alternating movement of influence of cultural and computational layers.

Keywords: Search engine history. Search engines. Web search tools. Search expressions. Information retrieval. Knowledge media.

LISTA DE FIGURAS

Figura 1 - Meme sobre o acesso aos resultados do Google durante uma pesquisa	20
Figura 2 - O método dialético de Sócrates	32
Figura 3 - Modelo mental de pesquisa antes dos buscadores	39
Figura 4 - Modelo mental de pesquisa em 2022 com o uso de buscadores.....	40
Figura 5 - Esboço do Memex proposto por Vannevar Bush	42
Figura 6 - Archie	46
Figura 7 - Janela do GOPHER	47
Figura 8 - Navegação nos itens do GOPHER através do Veronica	48
Figura 9 - W3 Catalog	49
Figura 10 - Aliweb.....	50
Figura 11 - JumpStation	51
Figura 12 - WebCrawler.....	52
Figura 13 - Lycos.....	53
Figura 14 - Lycos em 2022.....	53
Figura 15 - Infoseek, Excite e SAPO	54
Figura 16 - Yahoo.....	55
Figura 17 - AltaVista	56
Figura 18 - Ask Jeeves	57
Figura 19 - Cadê	59
Figura 20 - Aonde.....	60
Figura 21 - Achei, Prokura e Radix.....	60
Figura 22 - Fusões e aquisições de motores de busca de 1994 a 2005.....	61
Figura 23 - Google no início.....	62
Figura 24 - Doodle demonstrativo de booleanos no 200º aniversário de George Boole.....	63
Figura 25 - Vestido responsável pela criação do Google Imagens.....	64
Figura 26 - Interface do Google em 2022	65
Figura 27 - Exemplo de fatores presentes nas categorias de um perfil de busca	69
Figura 28 - Posição dos anúncios nos primeiros resultados do Google	70
Figura 29 - Pesquisa realizada na caixa de pesquisa do site da UFSC.....	74
Figura 30 - Cenário um do questionário de coleta de dados	83
Figura 31 - Campo de resposta do cenário um	83

Figura 32 - Cenário dois do questionário de coleta de dados.....	84
Figura 33 - Cenário três do questionário de coleta de dados.....	85
Figura 34 - Cenário quatro do questionário de coleta de dados	86
Figura 35 - Cenário cinco do questionário de coleta de dados.....	87
Figura 36 - Execução da busca livre.....	95
Figura 37 - Nuvem de palavras com os termos da busca livre	96
Figura 38 – Execução da busca normalizada.....	97
Figura 39 - Erro de pesquisa com o operador booleano AND	98
Figura 40 - Pesquisa bem-sucedida com o operador booleano AND.....	99
Figura 41 - Guia de pesquisa avançada do Google	100
Figura 42 - Erro no uso do operador filetype:	101
Figura 43 - Etiqueta com extensão de arquivo	101
Figura 44 - Único anúncio com o operador site:	102
Figura 45 - Anúncios de apostilas usando o operador filetype:	105
Figura 46 - Exemplo de sugestões de pesquisas já realizadas.....	106

LISTA DE QUADROS

Quadro 1 - Teses e Dissertações do EGC relacionadas aos temas desta pesquisa	26
Quadro 2 - Descrição e função dos operadores avançados selecionados	71
Quadro 3 - Classificação da pesquisa	79
Quadro 4 - Matriz de consistência do projeto de pesquisa do Mestrado	81

LISTA DE GRÁFICOS

Gráfico 1 - Gênero dos respondentes	91
Gráfico 2 - Faixa etária dos participantes	92
Gráfico 3 - Dados de geolocalização dos participantes.....	92
Gráfico 4 - Tipo de dispositivo utilizado pelos participantes.....	93
Gráfico 5 - Navegador de internet utilizado pelos participantes.....	93
Gráfico 6 – Presença de bloqueador de anúncios no navegador.....	94

LISTA DE TABELAS

Tabela 1 - Conjunto de dados pré e pós seleção.....	89
Tabela 2 - Porcentagem de anúncios nos cenários normalizados.....	104
Tabela 3 - Porcentagem das sugestões de pesquisa nos cenários normalizados	105
Tabela 4 - Mediana dos resultados nos cenários normalizados.....	106

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
ALA	American Library Association
ALIWEB	Archie-Like Indexing in the Web
CEP	Comitê de Ética e Pesquisa com Seres Humanos
CESAR	Centro de Estudos e Sistemas Avançados do Recife
CERN	European Organization for Nuclear Research
CIn UFPE	Centro de Informática da Universidade Federal de Pernambuco
CMC	Comunicação Mediada por Computador
DNS	Domain Name System
DOC	Document
DOI	Digital Object Identifier
FTP	File Transfer Protocol
HTML	HyperText Markup Language
IBGE	Instituto Brasileiro de Geografia e Estatística
IHC	Interação Humano-Computador
JUGHED	Jonzy's Universal Gopher Hierarchy Excavation and Display
LGPD	Lei Geral de Proteção de Dados
PDF	Portable Document Format
PNG	Portable Network Graphics
PPGEGC	Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento
MIT	Massachusetts Institute of Technology
MP3	MPEG 1 Layer-3
SAPO	Servidor de Apontadores Portugueses Online
SMART	System for the Mechanical Analysis and Retrieval of Text
SRI	Sistemas de Recuperação da Informação
SRT	SubRip Subtitle
TCD	Transmission Control Protocol
UFSC	Universidade Federal de Santa Catarina
URL	Uniform Resource Locator

VERONICA Very Easy Rodent-Oriented Net-wide Index to Computerized Archives
VSM Vector Space Model
W3C World Wide Web Consortium
WWW World Wide Web

LISTA DE SÍMBOLOS

“ ”	Aspas
*	Asterisco
()	Parêntese
+	Sinal de adição
-	Sinal de subtração
X	Sinal de multiplicação

SUMÁRIO

1	INTRODUÇÃO	19
1.1	OBJETIVOS	23
1.1.1	Objetivo Geral	23
1.1.2	Objetivos Específicos	24
1.2	JUSTIFICATIVA	24
1.3	ADERÊNCIA CONCEITUAL AO EGC	25
1.4	ESCOPO	27
1.5	ORGANIZAÇÃO DA DISSERTAÇÃO.....	28
2	DO RACIOCÍNIO LÓGICO ÀS BIBLIOTECAS E OS BUSCADORES ...	29
2.1	DE SÓCRATES A GEORGE BOOLE	29
2.2	A BUSCA EM BIBLIOTECAS	36
2.2	PRECEDENTES NO DESENVOLVIMENTO DE BUSCADORES	42
2.2.1	Os buscadores e suas evoluções	45
2.2.2	Google	62
2.3	OPERADORES DE PESQUISA.....	71
2.3.1	AND	71
2.3.2	Aspas “ ”	72
2.3.3	filetype:	73
2.3.4	site:	73
2.4	BUSCADORES COMO MÍDIA DO CONHECIMENTO	75
2.5	SÍNTESE	77
3	OPÇÕES METODOLÓGICAS	79
3.1	CARACTERIZAÇÃO DA PESQUISA	79
3.2	PESQUISA EXPERIMENTAL	82
3.2.1	Participantes	82

3.2.2	Procedimento	82
3.2.3	Coleta de dados	88
4	ANÁLISE DOS RESULTADOS	90
4.1	CONTRIBUIÇÕES PARA A CIÊNCIA ABERTA	90
4.2	DADOS DEMOGRÁFICOS	91
4.3	BUSCA LIVRE	94
4.4	BUSCA NORMALIZADA	97
4.5	BUSCA COM OPERADOR BOOLEANO AND.....	98
4.6	BUSCA COM OPERADOR AVANÇADO filetype:	100
4.7	BUSCA COM OPERADOR AVANÇADO site:.....	102
4.8	OBSERVAÇÕES GERAIS	103
5	CONSIDERAÇÕES FINAIS	107
	REFERÊNCIAS	111
	ÍNDICE ONOMÁSTICO	122

1 INTRODUÇÃO

A lógica na qual se apoia a formulação de uma pesquisa evoluiu com base nas diferentes teorias filosóficas e matemáticas construídas ao longo da história. Esta evolução se forma desde Sócrates e Aristóteles, perpassando George Boole e Claude Shannon e chegando até os buscadores¹ on-line existentes atualmente. A forma prática como uma pesquisa de rotina é realizada, modificou-se muito através dos séculos, principalmente pelos avanços tecnológicos que resultaram no surgimento de novas mídias (MANOVICH, 2001).

Em pleno ano de 2023, realizar uma pesquisa de rotina é, para muitos, algo trivial, com práticas tão inseridas no cotidiano que por mais que se repitam dia após dia, passam despercebidas, levando em consideração o quão automático elas se tornaram. Wu, Zhu e Wu (2022) citam a busca como uma das atividades mais comuns inerentes ao ser humano, que constantemente está tentando satisfazer suas necessidades em diversos assuntos.

Para elencá-las é necessário diferenciar uma pesquisa de rotina de uma pesquisa científica. Prodanov e Freitas (2013, p. 43) dizem que “nós pesquisamos a todo momento, em nosso cotidiano, mas, certamente, não o fazemos sempre de modo científico”. Enquanto na pesquisa científica exige-se um rigor e há maior grau de complexidade na forma como é efetuada, numa pesquisa de rotina não há regras que a validem enquanto processo. A pesquisa de rotina irá ser aplicada quando um usuário tem uma demanda informacional relativamente simples e precisa saná-la com rapidez.

Para aqueles que possuem acesso à internet, cerca de 82,7% dos domicílios no Brasil (IBGE, 2019), realizar uma pesquisa de rotina em um buscador pode levar menos de um minuto, a depender do grau de dificuldade em obter o resultado esperado. Seja em um smartphone, notebook, tablet ou outro dispositivo, em pouco tempo, com alguns cliques ou até mesmo por comando de voz, é possível expressar uma demanda e obter resultados relacionados na tela.

Líder mundial, o Google é o mecanismo de busca de maior relevância no mercado nas mais diferentes plataformas com acesso à internet. Dados da empresa irlandesa de análise de tráfego web StatCounter (2023) mostram que globalmente, de janeiro de 2009 a dezembro de 2022 ele permaneceu no topo com uma média de 92.58% de uso, sem ameaças ao pódio já que os demais colocados dividem médias abaixo dos 4%, sendo eles Bing, Yahoo!, Yandex, Baidu e DuckDuckGo, respectivamente. No Brasil, no mesmo período, o percentual sobe para 96.32%.

¹ Nesta dissertação optou-se por utilizar o termo ‘buscador’ ao invés de ‘mecanismo de busca’ ou ‘motor de busca’ principalmente pelo foco na forma de ferramenta contemplando algumas discussões acerca da interface.

A própria gigante de buscas define a primeira página de resultados de pesquisa como uma posição valiosa (GOOGLE, 2021c). Não é à toa que para isso existem inúmeras técnicas orgânicas de otimização como o Search Engine Optimization (SEO), além do tráfego pago. Aparecer bem-posicionado nos resultados do buscador pode garantir mais acessos e engajamento com o conteúdo disponibilizado.

A Figura 1, traz um meme² que representa de forma despretensiosa como os resultados de pesquisa no Google são encarados pelos usuários.

Figura 1 - Meme sobre o acesso aos resultados do Google durante uma pesquisa



Fonte - MEME sobre o acesso aos resultados do Google durante uma pesquisa (2012)

Tomando por exemplo a caixa de pesquisa disponibilizada pelo Google, o usuário digita alguns termos (podendo receber inclusive sugestões do próprio buscador sobre como completá-los), clica em pesquisar e pronto. O que retorna na tela são inúmeros resultados, que, segundo o buscador, estão relacionados com aquilo que foi pesquisado. Podem ser muitos, podem ser poucos, podem ser exatamente aquilo que o usuário procura ou também podem nada ter a ver com o que ele precisava. São vários os cenários possíveis.

Em paralelo a esse cenário, o volume de dados disponíveis na web cresce dia após dia. Se hoje uma pesquisa qualquer de um documento em formato PDF buscou um dentre 1.000 conteúdos, no próximo ano poderá buscar esse mesmo único documento em um montante de 10.000 conteúdos. Mesmo os buscadores que captam preferências de buscas dos usuários e as utilizam para otimizar os resultados e a forma como cada um vê os conteúdos pesquisados, exigirá de seus robôs uma varredura mais abrangente e exaustiva.

² Referência visual com imagem e textos cujo conteúdo é irônico ou humorístico.

A informação que é recuperada por um usuário mediante o uso de um buscador, para além de suprir as suas necessidades informacionais, contribui para a produção de novos conhecimentos (CHOO, 2006). Assim sendo, Santos e Godoy Viera (2015) refletem que se a informação almejada for de fato encontrada, oportunizando a geração de conhecimentos, há chances desse novo conhecimento ser disponibilizado em rede. Isto corrobora o aumento no volume de dados disponíveis na web discutido anteriormente.

Considerando a facilidade em pesquisar qualquer tipo de conteúdo na internet, sugere-se que a precisão — ou falta dela — nos resultados, levanta questões sobre quão criterioso é a construção e a execução das pesquisas de rotina. Afinal, basta digitar algumas palavras soltas, para que muitas possibilidades surjam segundos depois.

Para exemplificar melhor as variáveis do processo de busca, é possível fazer uma analogia com uma enciclopédia física. O usuário pegava o livro em mãos e por meio de uma breve análise visual e textual, sabia qual era seu começo e seu final, além de identificar qual tipo de conteúdo estava sendo disponibilizado naquele volume, e caso não fosse de seu interesse, poderia simplesmente deixá-lo de lado e partir para outro livro, repetindo o processo. Na internet, não se sabe qual é o começo e final de grande parte das fontes que se pesquisa. Apesar de uma estimativa da quantidade de resultados encontrados para cada pesquisa, ainda assim, a ordem na qual os resultados são apresentados (mesmo que estes sejam descritos pelos próprios buscadores em páginas dedicadas a explicar como funciona a pesquisa) ela ainda segue critérios desconhecidos por grande parte da população leiga que utiliza tais recursos.

Diferentemente do ambiente de uma biblioteca digital, onde apesar do grande volume de informações disponíveis, elas são limitadas e há, em grande parte, um processo de curadoria de conteúdo que envolve seleção, organização, classificação, catalogação, indexação e resumo. O ambiente web no todo, com acesso regido por buscadores, é obviamente muito maior em termos de busca e acesso. Ao refletir sobre a biblioteca versus o fenômeno da Googlização, Sayão (2009, p. 11) não acredita na “concepção de que o Google representa a apoteose da informação digital e que os problemas existentes nesse domínio já foram resolvidos ou serão resolvidos pelo serviço ou por outra ferramenta semelhante”. Tal afirmação reforça a ideia incerta de que as pessoas leigas (sem formação específica em computação ou pesquisa) dominem a pesquisa de rotina. Num cenário ideal, bastaria inserir alguns termos e o buscador resolveria a busca trazendo a informação necessária.

Entretanto, mesmo a pesquisa por meio de buscadores tenha se tornado algo relativamente simples de ser realizado, muitas vezes o usuário não encontra aquilo que precisa, depara-se com informações incorretas ou até mesmo perde muito tempo para chegar em algo que seria relativamente simples de encontrar. Aparentemente, pesquisar na internet, por mais mecanicamente simples que seja, ainda continua sendo um “exercício de sobrevivência do usuário mais apto” que, ao possuir maiores conhecimentos sobre a busca, se destaca por conseguir extrair o melhor dos buscadores (VIDMAR, 1999, p. 1, tradução nossa).

Os primeiros sistemas de recuperação da informação dispunham basicamente de um conjunto de fichas, em que, para recuperar alguma informação, os dados fornecidos pelo usuário se limitavam a título do documento ou nome do autor. No entanto, avanços tecnológicos permitiram que outros campos de busca fossem acrescentados de acordo com a necessidade de cada sistema e público-alvo, surgiam neste momento as buscas por assunto.

Posteriormente, o melhoramento das funcionalidades dos programas e o avanço em questões de interface gráfica, possibilitaram que alguns sistemas contassem com a possibilidade de combinar palavras-chave com operadores de pesquisa específicos a uma determinada função. Operadores são símbolos ou palavras que são adicionados à pesquisa para tornar os resultados mais precisos (GOOGLE, 2013). Portanto, o uso de operadores permite a formulação de buscas complexas se necessário.

A engenharia do buscador emprega um rastreador (também conhecido como robô, bot ou indexador) que vasculha a web em busca de páginas novas e atualização de páginas já conhecidas, ao encontrá-las incorpora-as ao seu índice (GOOGLE, 2021b). O usuário realiza uma pesquisa e o buscador apresenta resultados correspondentes a consulta que estejam em seu índice. Ao utilizar operadores avançados, o usuário se aproxima de comandos por vezes muito específicos e instrucionais, como uma linguagem de programação. Assim, entende-se os operadores como funções da linguagem computacional, os quais não pertencem ao domínio de grande parte dos usuários

Esta evolução no design e engenharia do buscador (camada do computador) continua sendo influenciada pela camada cultural, ou seja, como o usuário a utiliza e entende o processo de busca. A este processo Manovich (2001) chamou de transcodificação. Ou seja, enquanto os buscadores através de sua interface, traduzem os dados computadorizados numa apresentação compreensível pelo usuário, o tratamento computacional dos dados requer que o buscador seja

uma mídia computacional, onde os dados seguem uma estrutura de banco de dados e são tratados por algoritmos.

Porém, parece existir um descompasso nesta mútua influência, pois grande parte dos usuários não domina as propriedades do buscador para alcançar resultados precisos. Wu, Zhu e Wu (2022, p. 22) diz que “a tendência geral da evolução da busca é desenvolver habilidades eficientes para encontrar um alvo específico a partir de um grande conjunto de itens” e que isso requer planejamento e aprimoramentos

Com o intuito de possibilitar uma pesquisa com delimitações e resultados mais precisos, os buscadores permitem o uso de símbolos ou palavras como auxiliares no processo de busca. No entanto, tais operadores de pesquisa, se utilizados de forma arbitrária, podem, ou não funcionar diante do sistema ou atrapalhar os resultados. Ao enxergá-los enquanto recursos de pesquisa avançados também se faz necessário entender cada um de seus significados, funções e quando podem ser potencialmente relevantes à pesquisa, o que implica num grau específico de letramento digital. Enquanto isso, no intuito de torná-la cada vez mais simples, a interface omite as opções para uma busca mais produtiva, o que pode dificultar sua aplicação.

Neste contexto, tem-se um problema de pesquisa a ser explorado: quais características nos resultados de pesquisa se destacam a partir do uso de operadores avançados na busca por documentos na web?

O pressuposto é que a utilização desses operadores pode apresentar características mais uniformes o que pode sugerir maior agilidade na busca e precisão aos resultados e contribuir para uma emancipação do usuário durante uma pesquisa on-line, possibilitando a ele maior controle e direcionamento dos resultados apresentados.

1.1 OBJETIVOS

Para responder ao problema de pesquisa elucidado, as seções abaixo descrevem o objetivo geral e os objetivos específicos desta dissertação.

1.1.1 Objetivo Geral

Examinar as características resultantes do uso de operadores avançados de busca na realização de pesquisas em um buscador.

1.1.2 Objetivos Específicos

Para alcançar o objetivo geral foram necessários alcançar os seguintes objetivos específicos:

- a) Descrever o processo lógico de uma pesquisa;
- b) Selecionar os operadores de pesquisa presentes num buscador que auxiliem na recuperação de documentos em pesquisas on-line;
- c) Identificar quais são os princípios fundamentais no funcionamento de um buscador;
- d) Testar a partir da estruturação de dados o uso dos operadores de pesquisa por diferentes usuários.
- e) Categorizar as resultantes do uso de operadores avançados de busca online por diferentes usuários, de acordo com suas características.

1.2 JUSTIFICATIVA

Buscadores estão em constante aprimoramento para melhorar a experiência de pesquisa dos usuários. Muitas vezes a forma simples funciona, pois, o usuário insere um ou mais termos, clica em buscar e navega pelos resultados apresentados em tela. Todavia, é importante conhecer recursos para que, caso venha a precisar, possa recorrer a fórmulas específicas visando encontrar um corpo menor ou mais padronizado de resultados para avaliar, fazendo com que tal tarefa possua mais assertividade e precisão.

Considerando que um dos papéis da ciência é contribuir para o processo de geração do conhecimento de uma sociedade, este pode ser aprimorado, à medida que as pessoas entendam melhor o funcionamento de ferramentas tecnológicas que permitem o acesso a informação. Assim, este trabalho se justifica na medida que futuros pesquisadores podem empregar expressões de busca com maior diligência, desde o momento em que forem formuladas até analisarem os resultados apresentados a partir dela.

É fato que grande parte do referencial teórico científico das pesquisas da academia advém de bases de dados científicas, e nesse contexto, buscadores atuam como auxiliares e fios condutores do processo. No entanto, dois fatores são importantes aqui: (1) a lógica de construção de uma expressão de busca, em sua concepção, se assemelha tanto à pesquisa em bases de dados

científicas quanto à pesquisa em buscadores. Sendo ambos sistemas de recuperação da informação; e (2) esses sistemas podem dispor de operadores avançados de pesquisa e dependem de uma movimentação a partir de escolhas do usuário para construir uma expressão assertiva e logicamente correta.

Além disso, o corpus de pesquisa de muitos trabalhos científicos está alocado em endereços eletrônicos recuperados por buscadores tradicionais. Ou seja, incluem sites institucionais, redes sociais, notícias, documentos nas mais diversas extensões de arquivos como imagens, vídeos, estatísticas, entre inúmeros outros recursos informacionais que possam integrar uma pesquisa científica. Diante disso, é relevante tirar o proveito de ferramentas como buscadores.

Para a sociedade, as ferramentas tecnológicas disponíveis a partir do surgimento das novas mídias, permitem que todos os que possuem mínimas condições de literacia digital e possibilidade de utilizar um dispositivo com acesso à internet, possam navegar em conteúdos web a partir das suas necessidades informacionais.

Sendo os buscadores os intermediários principais entre uma pesquisa on-line e um resultado, conhecer seu funcionamento, bem como possibilidades avançadas de realização de uma pesquisa, permite que o usuário adquira uma emancipação durante as pesquisas on-line, no sentido de tornar-se mais independente em suas buscas. A partir disso, o domínio da busca pode direcionar os resultados para os interesses de pesquisa e promover o desenvolvimento de competência em informação. Atributo este que se faz presente desde a percepção da necessidade informacional, busca, avaliação e uso da informação, indo ao encontro do que a American Library Association (ALA) (1989) entende como uma pessoa competente em informação. Portanto, este trabalho contribui como o acesso ao conhecimento, objetivo de todo e qualquer percurso investigativo.

1.3 ADERÊNCIA CONCEITUAL AO EGC

A pesquisa está diretamente relacionada ao Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento (PPGEGC) ao levar em consideração o **conhecimento**, principalmente quando este é visto enquanto processo. O método presente na realização de uma pesquisa on-line permite uma abordagem precisa dentro da área de concentração de Mídia do Conhecimento.

O objeto principal do estudo é, além de muitas coisas, uma mídia do conhecimento. Buscadores podem ser entendidos como ferramentas essenciais para alcançar a informação e o conhecimento, suas funcionalidades vão ao encontro disso.

Inserida na área de Mídia do Conhecimento há a linha de Mídia e Disseminação do Conhecimento que engloba a captação da informação por meios tecnológicos, que trata justamente de uma das principais finalidades de um buscador (UFSC, 2017; BRANSKI, 2004).

No que diz respeito aos trabalhos já realizados por alunos do PPGEGC, ao consultar o Banco de Teses e Dissertações do Programa, conforme demonstrado no Quadro 1, sobressaem duas dissertações de mestrado que proporcionam subsídios para o avanço desta pesquisa.

Dos termos pesquisados estão: buscadores; mecanismos de busca; Google; busca integrada e linguagem natural. Os dois últimos foram os responsáveis por recuperar respectivamente os trabalhos descritos no Quadro 1. Outros trabalhos foram recuperados na pesquisa individual dos termos, mas após análise, foram considerados como não relacionados a proposta desta dissertação.

Quadro 1 - Teses e Dissertações do EGC relacionadas aos temas desta pesquisa

ESTUDOS ANTERIORES		
Autor(a)	Título	Tipo
Fagundes (2018)	Proposta de uso de engenharia do conhecimento para revisão sistemática	Dissertação
Silva (2011)	Uma Arquitetura de <i>business intelligence</i> para processamento analítico baseado em tecnologias semânticas e em linguagem natural	Dissertação

Fonte: Banco de Teses e Dissertações do EGC (2022).

Silva (2011) descreve quais foram as limitações do seu trabalho expondo que “o estudo da língua expõe diversas problemáticas e variações que é evidente a dificuldade em interpretar e formalizar a escrita e a comunicação humana em sistemas de conhecimento” (SILVA, 2011, p. 146). Esta dissertação explora esta lacuna, ao levar em consideração o foco na mudança das linguagens na comunicação entre usuários e buscadores na criação de expressões de busca, bem como a construção sintática e semântica por trás delas.

Ainda em relação às linguagens, posteriormente na dissertação de mestrado de Fagundes (2018) o autor propõe um modelo conceitual para a revisão sistemática. No referido modelo, a primeira etapa descrita por ele, intitulada como preparação, tem relação direta com a formulação de expressões de busca em bases de dados científicas.

Esta dissertação reflete as razões que fundamentam a estruturação de tais expressões e inclui indivíduos externos ao contexto acadêmico, ao trabalhar com mídias mais comuns do cotidiano do pesquisador leigo, como os buscadores.

Ademais, os buscadores enquanto mídia do conhecimento e sua utilização em pesquisas de rotina por usuários leigos é um assunto que foi pouco explorado no rol dos trabalhos desenvolvidos pelo PPGEGC/UFSC. Considerando este cenário, o presente trabalho vem contribuir também com a temática no contexto da interdisciplinaridade do Programa.

1.4 ESCOPO

Como delimitação conceitual, apesar do processo de pesquisa passar por toda a espiral do conhecimento proposta por Nonaka e Takeuchi (1997), a parte prática da proposta de pesquisa na dissertação está diretamente relacionada com o conhecimento explícito. Mais precisamente nos processos de combinação (explícito → explícito) ao unir a linguagem textual de pesquisa aos operadores de pesquisa identificados e internalização (explícito → tácito) a partir dessa apropriação individual do conhecimento e a identificação de em quais pesquisas eles podem se encaixar.

Há inúmeros buscadores, alguns com funções e propósitos específicos, mas para esta pesquisa o buscador selecionado foi o Google, que foi o buscador mais utilizado dos últimos dez anos (STATCOUNTER, 2022). Mesmo sendo este um buscador que assume uma posição de utilização em escala global, o nível de análise do presente estudo foi realizado em âmbito nacional, a partir de pesquisas formuladas em português, mesmo que algum participante esteja numa localidade fora do território brasileiro ou utilize o buscador com as configurações em outro idioma que não o português.

As pesquisas também foram delimitadas em temas rotineiros, não científicos, que pudessem aproximar os testes a situações de busca recorrentes de um público leigo em relação a pesquisa avançada em buscadores. Os operadores avançados foram selecionados de modo a atender a demanda criada pela temática da expressão de busca, que por ter como foco a busca de provas por concurso público de um cargo específico, compreendia dimensões de intersecções dos termos, entre eles termos simples e compostos; aspectos de formato de arquivo e por último, dimensões acerca da fonte de informação pesquisável.

1.5 ORGANIZAÇÃO DA DISSERTAÇÃO

A seção dois do trabalho trata do raciocínio lógico como base para a forma como os buscadores funcionam e como os usuários formulam uma pesquisa. Desde Sócrates a George Boole, passando também pelas buscas em bibliotecas. Posteriormente são apresentados os primeiros buscadores e suas evoluções durante a década de 90 até chegar ao Google. Na mesma seção ainda são apresentados os operadores de pesquisa utilizados na presente pesquisa.

A seção três descreve todas as opções metodológicas definidas para o delineamento do trabalho desde a caracterização da pesquisa até a descrição de todas as etapas do experimento realizado.

A seção quatro compreende a análise dos resultados e a disponibilização dos dados abertos da pesquisa.

As considerações finais apresentam a retomada do problema de pesquisa e dos objetivos. Por fim estão as referências e o índice de autores para ajudar o leitor na localização de conceitos no texto.

2 DO RACIOCÍNIO LÓGICO ÀS BIBLIOTECAS E OS BUSCADORES

Até chegarmos de fato aos buscadores e assim como em todo e qualquer outro campo do conhecimento, inúmeros outros fatores contribuíram para alcançar o desenvolvimento tecnológico atual. Em relação aos buscadores, enquanto uma ferramenta de inserção de questionamentos e resolução de dúvidas, passaremos a seguir desde o agir de Sócrates, a lógica clássica de Aristóteles até a álgebra booleana de George Boole. Na sequência, a busca em bibliotecas pavimentou o caminho para os buscadores tal qual conhecemos hoje, do Archie ao Google.

2.1 DE SÓCRATES A GEORGE BOOLE

Grande parte das pesquisas, seja ela realizada em uma mídia física (enciclopédias, livros, documentos históricos, entre outros) ou em uma mídia eletrônica (buscador, base de dados científica, ficheiro de computador, entre outros), parte de uma demanda informacional do usuário. O objetivo da pesquisa é encontrar algo que supra essa demanda. Esse ‘algo’ podem ser dados, informações ou até uma orientação de como proceder a própria pesquisa.

A demanda informacional pode ser entendida como uma formulação de questionamento. Neste sentido a filosofia foi uma das responsáveis em permitir que o ser humano as expressasse.

Chauí (2000) apresenta o que é chamado por atitude filosófica como o momento em que nos distanciamos das afirmações e ‘verdades’ pré-estabelecidas e passamos a nos questionar, a fim de compreender melhor o nosso entorno, antes de assumir um parecer em relação ao que está em questão naquele momento. Assim sendo, numa possível definição do que vem a ser filosofia, a autora a descreve como:

A decisão de não aceitar como óbvias e evidentes as coisas, as idéias, os fatos, as situações, os valores, os comportamentos de nossa existência cotidiana; jamais aceitá-los sem antes havê-los investigado e compreendido. (CHAUÍ, 2000, p. 9)

Sócrates (470 a.C.-399 a.C.) é considerado um dos filósofos precursores desse processo de questionar. A insistência de Sócrates em suas perguntas, desenvolveu o que é conhecido hoje como método dialético, no qual uma investigação tem como abordagem central o uso da argumentação e da provocação (GODOY, 2003). Sócrates não deixou nada por escrito, e foi Platão (428 a.C.-347 a.C.) o responsável pelo registro das suas ideias e ações (O LIVRO DA FILOSOFIA, 2016).

Ao contrário de uma pergunta lançada com o intuito de obter uma resposta definitiva, os questionamentos de Sócrates, apesar de refletirem também um interesse pessoal em obter melhor compreensão do tema, eram questionamentos voltados principalmente a uma pesquisa interna, a partir de reflexões pessoais acerca do que cada indivíduo considerava como sendo algo ‘bom’, ‘ruim’, ‘justo’, entre outros conceitos básicos da humanidade (O LIVRO DA FILOSOFIA, 2016). Esse perfil questionador de Sócrates foi um dos fatores responsáveis por suscitar desavenças e acusações, que culminaram em um julgamento sob acusações de não reconhecer os deuses da cidade, introduzir novas divindades e de corromper a juventude” (GODOY, 2003, p. 13).

Essa não aceitação inicial das coisas, ideias, fatos, situações, valores e comportamentos, envolve duas faces, uma negativa e outra positiva. A face negativa está relacionada com uma atitude crítica, a qual consiste em “dizer não ao senso comum, aos pré-conceitos, aos pré-juízos, aos fatos e às ideias da experiência cotidiana, ao que ‘todo mundo diz e pensa’, ao estabelecido” (CHAUÍ, 2000, p. 9). A outra face, positiva, relaciona-se com a busca do entendimento do “porquê disso tudo e de nós [...] sobre como tudo isso é assim e não de outra maneira” (CHAUÍ, 2000, p. 9). Ambas constituem a ‘atitude filosófica’ e para que sejam entendidas como tal, são indissociáveis. Não há atitude filosófica se houver apenas uma negativa de forma arbitrária, sem uma interrogação posterior.

Nesse movimento, Sócrates pôde observar a forma como a atribuição de significado para diferentes termos funcionava em diferentes indivíduos. Atribuição essa que por vezes era derivada de uma real falta de questionamento interno acerca de alguns conceitos, como se o entendimento herdado pelo que a sociedade sobre algo fosse verdade e ponto final. Sendo este incômodo um dos principais motivadores para ele (NUNES; FREITAS, 2018).

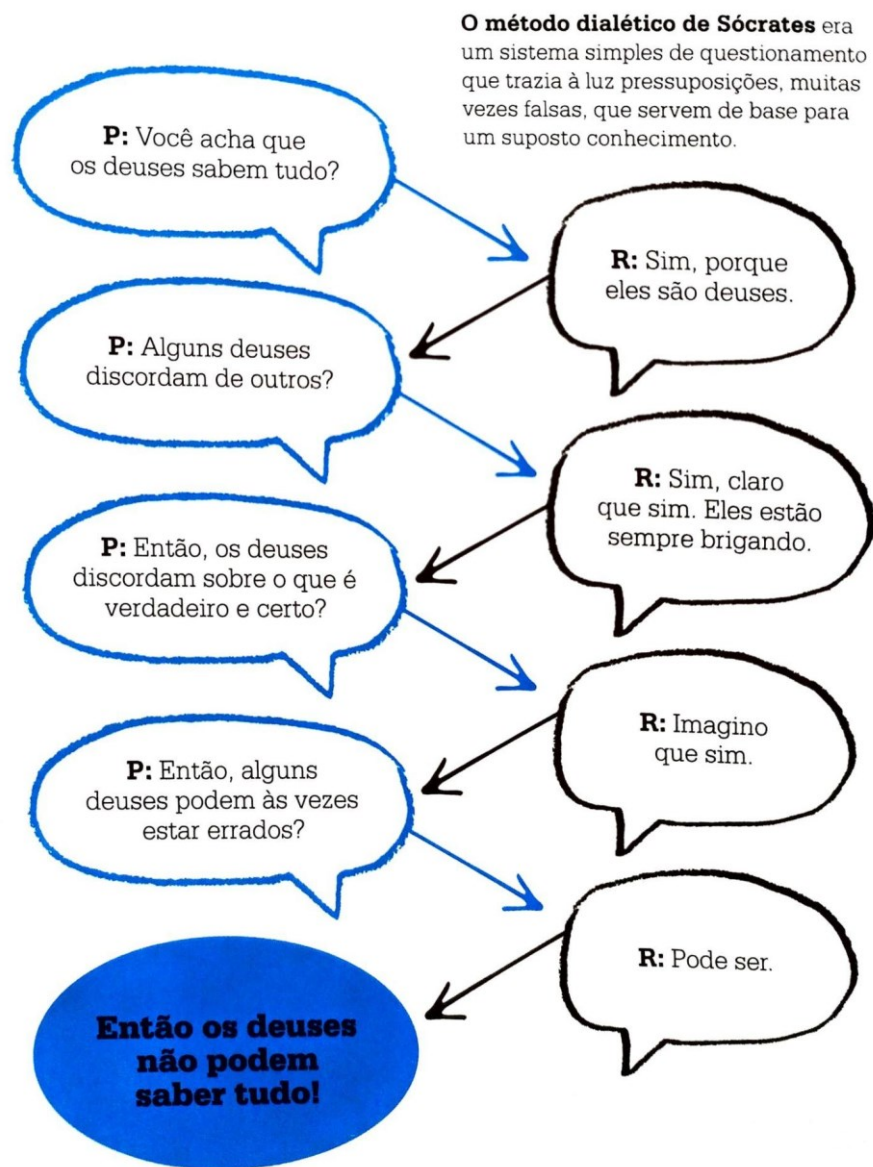
Séculos depois, Descartes (1596-1650) experimenta um processo parecido ao assumir uma posição em que o propósito seria duvidar de tudo. Ao menor sinal de dúvida haveria uma rejeição e aquilo seria considerado falso. Na teoria do conhecimento, essa prática ficou conhecida como ‘dúvida hiperbólica’ (DESCARTES, 2009). Aqui, diferentemente de Sócrates, o objetivo era encontrar alguma coisa da qual não pudesse se duvidar, algo que precisasse ser aceito como verdade e ponto final. Por fim, concluiu-se que “mesmo duvidando de tudo, não podia duvidar de que ele mesmo era alguma coisa” (MATOS, 2016, p. 46). Mais do que a conclusão pessoal na afirmação anterior, o maior ganho a partir dessa experiência foi o entendimento de que “a dúvida torna-se um instrumento de busca do conhecimento. Seu valor é reconhecido na medida em que,

pelo estabelecimento da dúvida, uma atitude de investigação e busca imediatamente se instala na mente do sujeito que pensa” (MATOS, 2016, p. 46). Posteriormente, a noção da dúvida como instrumento veio a ser conhecido por ‘dúvida metódica’.

No caso dos experimentos de Sócrates, os questionamentos provocados por ele, despertavam nos indivíduos questionados, uma reflexão sobre a resposta a partir da referência e repertório de cada um. Uma pesquisa como resposta a um estímulo externo, causando um processo interno desempenhado por cada indivíduo questionado, o que segundo Nunes e Freiras (2018, p. 8) trouxe uma nova perspectiva ao campo da filosofia, no entendimento que “a busca pelo conhecimento passa por primeiro no conhecimento de si”.

A partir das respostas retornadas, Sócrates buscava expor contradições ao assumir uma posição de quem nada ou pouco sabia sobre o tema questionado. No exemplo de diálogo apresentado na Figura 2, é possível observar o método dialético de Sócrates.

Figura 2 - O método dialético de Sócrates



Fonte: O LIVRO DA FILOSOFIA (2016, p. 48)

Esse processo ilustrado na Figura 2 demonstra uma falha nas premissas que precedem um argumento. Visto que, se os deuses sabem tudo, eles não podem discordar entre eles sobre o que é certo e verdadeiro. Caso discordem, um dos deuses obrigatoriamente precisará estar errado e com isso, não saberá tudo. Esse exemplo em que Sócrates propunha uma negação aos fatos ditos pelo indivíduo questionado expôs o primeiro uso conhecido do que viria a ser intitulado posteriormente como método indutivo.

Algumas décadas depois, Aristóteles (384 a.C.-322 a.C.) organizou de forma concreta o que conhecemos hoje como método indutivo e método dedutivo. Estabelecendo parâmetros aos diferentes tipos de argumentos gerados por cada um deles (SMITH, 2009).

O método indutivo, conforme demonstrado na Figura 2, parte de um conjunto de premissas individuais, consideradas verdadeiras, para concluir algo como verdadeiro, de forma universal. O método dedutivo segue o caminho inverso, partindo de um conjunto de premissas gerais, consideradas verdadeiras, para concluir algo particular como verdadeiro.

Para exemplificar, de forma recorrente utiliza-se o silogismo mais famoso da história.

Todo homem é mortal
Sócrates é homem
Logo, Sócrates é mortal

Ao utilizar-se de duas premissas, sendo uma construção do pensamento que vai do geral ao particular, temos a constatação de que Sócrates é mortal, baseada em um argumento advindo do método dedutivo.

Seguindo o exemplo de diálogo representado pela Figura 2, novamente um exemplo, de um raciocínio, dessa vez apoiado no método dedutivo.

Sócrates tem duas pernas e dois braços
Platão tem duas pernas e dois braços
Sócrates e Platão são humanos
Logo, todos os humanos têm duas pernas e dois braços

Esse padrão de raciocínio, baseado em duas premissas e a partir delas uma conclusão, é conhecido por silogismo. Essência do primeiro sistema formal de lógica do qual se tem registro, transformando-a em ciência (ABAR, 2004; BARNES, 2013).

Entretanto, de acordo com Mortari (2001) essa forma de diferenciação de argumentos — indutivos e dedutivos — como afirmações que seriam ampliadas ou não ampliadas respectivamente, é uma explicação costumeira, mas ligeiramente insatisfatória. O autor afirma que se formos pensar de forma rigorosa “um argumento é dedutivo se e somente se ele for válido” (MORTARI, 2001, p. 24). De forma menos restritiva, ao ampliar as possibilidades, um

argumento mesmo que logicamente inválido, poderia ser ainda considerado dedutivo “quando há intenção, por parte de quem constrói ou apresenta o argumento, de que sua conclusão seja consequência lógica das premissas, ou seja, a pretensão de que a verdade das suas premissas garanta a verdade da conclusão” (MORTARI, 2001, p. 24).

Todo esse processo é realizado de modo sistemático, no qual estão fundamentadas todas as indagações filosóficas (CHAUI, 2000). Ao pensar em uma pesquisa na qual os enunciados que a precedem sejam de fato válidos, o resultado, por si só, deve permitir a sustentação de uma afirmação também válida. Isso está longe de um “eu acho que” ou um “na minha opinião”:

A Filosofia trabalha com enunciados precisos e rigorosos, busca encadeamentos lógicos entre os enunciados, opera com conceitos ou idéias obtidos por procedimentos de demonstração e prova, exige a fundamentação racional do que é enunciado e pensado. (CHAUI, 2000, p. 13)

Cerca de 2000 anos depois, o matemático George Boole (1815-1864) desenvolveu um sistema que possibilitaria que argumentos lógicos pudessem ser resolvidos pela matemática.

Para tanto, Boole se apoiou nos princípios de Aristóteles e nos estudos de Gottfried Leibniz (1646-1716), que mais de 100 anos antes havia tentado fazer uso da álgebra numa tentativa de formalização moderna da lógica. Boole concebeu a lógica, a construção formal que ela precisava, para que pudesse ser trabalhada de modo matemático (MOREIRA, 2007; O LIVRO DA MATEMÁTICA, 2020).

Com início das discussões na obra *Mathematical analysis of logic* (Análise lógica da matemática) e mais tarde em seu escrito mais importante, intitulado *The laws of thought* (As leis do pensamento), Boole introduziu letras como formas de representar um subconjunto de coisas. Assim, desenvolveu uma linguagem artificial que podia “expressar matematicamente as proposições lógicas e fosse capaz de formalizar as regras da lógica tradicional (aristotélica) a partir de um cálculo algébrico satisfatório” (SOUZA, 2005, p. 108).

Blanché e Dubucs (1996, p. 274) afirmam que:

Para permitir o tratamento algébrico do pensamento tal qual ele se exprime na nossa linguagem, Boole procura primeiro, partindo do raciocínio algébrico que opera sobre signos, classificar esses signos segundo a sua função, e em seguida encontrar o análogo dessas funções nas formas da linguagem comum, de maneira a poder traduzir estas em signos análogos aos signos algébricos, e prestando-se como eles a um cálculo.

A álgebra booleana está pautada em dois tipos de resultados, SIM ou NÃO, VERDADEIRO ou FALSO. Tais resultados refletem de forma binária no sistema numérico,

sendo representados por 0 para NÃO/FALSO e 1 para SIM/VERDADEIRO (MOREIRA, 2007; O LIVRO DA MATEMÁTICA, 2020).

As ideias de Boole demoraram várias décadas para serem empregadas de modo prático. Até que Claude Shannon (1916-2001) percebeu que haveria uma maneira de implementar a lógica em circuitos eletrônicos, demonstrando a possibilidade de automatizações em atividades antes realizadas apenas de forma manual. Na ocasião, as chamadas telefônicas, eram operadas principalmente por mulheres, as quais recebiam demandas de conexão entre pessoas e as realizavam manualmente alternando circuitos. A ideia era que essa ação fosse realizada com o apoio de operações lógicas, usando relés, conhecidos em sua versão moderna como transistores (SOBRAL, 2015).

Shannon ficou conhecido como o pai da Teoria da Informação, tendo como principal publicação o estudo denominado Teoria Matemática da Comunicação em 1948. Seus estudos foram responsáveis por estender o poder comunicacional da época e as tecnologias criadas a partir dessa ruptura tecnológica foram fundamentais para sustentar o patamar tecnológico no qual o mundo se encontra atualmente (TOTH, 2017).

Muitos outros estudiosos como Charles Babbage (Máquina Diferencial e Máquina Analítica), Ada Lovelace (primeiro algoritmo), Alan Turing (Máquina de Turing), entre inúmeras posteriores contribuições (SOBRAL, 2015; TOTH, 2017) contribuíram com a evolução tecnológica para que ano após ano, processos fossem aprimorados ou substituídos e avanços significativos ocorressem, afetando direta ou indiretamente as formas como as pesquisas na internet são realizadas.

Boole segue constituindo um dos pilares das operações computacionais e servindo de apoio a construção de expressões de busca tanto em bases de dados científicas como em buscadores tradicionais. Nesse caso as operações booleanas são aplicadas de forma visível aos usuários com o uso dos operadores geralmente conhecidos por AND, OR e NOT.

Esses operadores funcionam da seguinte forma: AND é utilizado para obrigatoriamente unir dois ou mais termos. OR para unir dois ou mais termos e garantir que sejam recuperados caso também apareçam de forma isolada. NOT para excluir um ou mais termos de uma busca. A depender da forma como são combinados, trazem maior precisão ou abrangência de resultados, agindo como conectivos lógicos para com os termos utilizados dentro da expressão de busca (PICALHO; LUCAS; AMORIM, 2022).

Ocasionalmente podem aparecer também num formato diferente como por exemplo: AND ou + (sinal de mais); OR ou * (asterisco) ou x (sinal de multiplicação); NOT ou AND NOT ou – (sinal de menos)

O uso dos operadores booleanos está atrelado a lógica do pensamento aristotélico à medida que, quando o usuário estabelece uma estratégia de pesquisa e a transcreve na forma de uma expressão de busca que contempla o uso desses operadores, está fazendo isso com base nos princípios instituídos na lógica clássica, elementos do cálculo proposicional clássico: conjunção, disjunção e negação, representados por AND, OR e NOT respectivamente (D’OTTAVIANO; FEITOSA, 2009).

Com ou sem o uso de operadores booleanos, a lógica clássica ainda se mantém por trás da formulação de pensamento ao desenhar uma estratégia de pesquisa que faça sentido na mente do usuário. “Cada caso de formulação de expressões de busca com o auxílio de operadores booleanos é um caso diferente. É preciso analisar a lógica de formulação para então montar uma expressão busca eficiente” (PICALHO; LUCAS; AMORIM, 2022, p. 10). Como estratégias de pesquisa, Goulart e Hetem Junior (2007, p. 56) definem esse processo como “um plano contemplando uma série de ações visando encontrar uma informação”, sendo necessário ainda que os usuários entendam que nem sempre esse plano contará com apenas uma estratégia simples e suficientemente acertada para encontrar aquilo que precisam, tendo que combinar esforços, unindo diferentes estratégias, recursos como os operadores e demais possibilidades.

2.2 A BUSCA EM BIBLIOTECAS

No fim dos anos 90 muitos se referiam a internet como uma vasta biblioteca contendo todo tipo de informação conhecida pelos seres humanos, de navegação relativamente simples em que bastava o usuário “inserir uma ou mais palavras-chave em um mecanismo de busca comercial em um esforço para localizar sites sobre um assunto de interesse” (RENO, 1997, item 9, tradução nossa). Isso acontecia em paralelo a uma discussão que envolvia compreender se a web poderia ser considerada uma espécie de biblioteca digital (KOEHLER, 1999).

É fato que mesmo que bibliotecas e internet tenham seguido por caminhos conceituais diferentes, o propósito de ambas sempre foi o mesmo. Enquanto em Kuny e Cleveland (1998) afirmavam que o negócio das bibliotecas era conectar pessoas a informações e que a tecnologia não iria mudar isso, em 2021 o Google aponta para a mesma direção quando descreve que sua

missão é “organizar as informações disponíveis no mundo todo e torná-las acessíveis e úteis para todos” (GOOGLE, 2021d).

Bibliotecas e buscadores precisam mapear as fontes de informação, organizar o conteúdo, classificá-lo e disponibilizá-lo para os usuários. Ambos evoluíram e melhoraram seus processos para acompanhar as mudanças tecnológicas e da sociedade, e, mesmo se tratando de ‘ambientes’ completamente diferentes, de certa forma se complementam, sem riscos da substituição total do ato de pesquisa na web por uma pesquisa em bibliotecas (GOOGLE, 2021d; SHERMAN, 2007). Prova disso é a atuação dos/as bibliotecários/as fora do ambiente físico da biblioteca, o insumo de trabalho destes profissionais nunca foi o livro, mas sim, o conteúdo dele, a informação, independentemente do suporte.

Os bibliotecários atuam diretamente no funcionamento de uma biblioteca física ou digital. Porém, diferentemente do espaço físico eles não estão à disposição para ajudar com uma pesquisa no Google, sobretudo as pesquisas de rotina realizadas com maior frequência. Neste contexto as questões semânticas dos buscadores merecem serem consideradas.

Como semântica, entende-se a “linguagem natural e formal (signos, termos, palavras) e sua função de representação, o que significa ou a que se refere”. A web semântica é aquela que visa fornecer estruturas com significado ao conteúdo presente nas páginas web, “criando um ambiente onde agente de software e usuários possam trabalhar de forma cooperativa” (TOUTAIN, 2006, p. 21, 23; BERNERS-LEE; HENDLER; LASSILA, 2001).

A web semântica foi idealizada por Tim Berners-Lee em 2001 como uma extensão do conceito anterior de World Wide Web (WWW) cunhado por ele. Na época a Comissão Europeia (2003) considerou a internet como algo verdadeiramente transformador para o mundo, podendo ser considerada tão importante quanto a revolução industrial dos séculos XVIII e XIX. O modo como as empresas funcionavam, os estudantes estudavam, os cientistas realizam suas pesquisas, as administrações públicas ofereciam seus serviços aos cidadãos e conseqüentemente o público em geral realizava buscas por informações diversas foi modificado.

A web semântica implica também numa mudança na forma da comunicação mediada por computador (CMC) e particularmente na forma da busca de informações e realização de pesquisas cotidianas. Em meados dos anos 2000, durante a ascensão do Google enquanto

buscador, alguns autores traziam ponderações entre a forma de recuperação dos catálogos em linha³ das bibliotecas:

Por uma década e meia, começando no início dos anos 1980, o catálogo em linha da biblioteca era a jóia da coroa, quando pessoas ávidas aguardavam nas filas dos terminais para procurarem informações escritas por especialistas mundiais. Fico abismada em relação ao modo como as pessoas ávidas agora abraçam o Google, devido a fonte de informação que o Google recupera. [esse período] foi a idade de ouro do catálogo em linha, porque **os usuários da biblioteca dependiam exclusivamente dele para encontrar o tópico que lhes interessava** (MARKEY, 2007, p. 2, tradução nossa, grifo nosso)

Uma década e meia depois, os usuários dependem menos ainda. Com pouca literacia digital, um indivíduo, por meio de um dispositivo com acesso à internet consegue buscar por um tópico de seu interesse. E aqui entende-se por literacia ou letramento digital “um conjunto mínimo de capacidades que habilitem o usuário a operar com eficiência os softwares, ou a realizar tarefas básicas de recuperação de informações” (BUCKINGHAM, 2010, p. 47-48).

Importante salientar que para além de um uso instrumental dos buscadores para a realização de uma pesquisa, mais do que literacia digital, é interessante que o usuário possua competência em informação, que compreende também o uso adequado das informações encontradas após a busca. O buscador é uma ferramenta de acesso, o letramento digital seria a capacidade de usá-la e a competência em informação a habilidade e consciência de realizar uma leitura crítica antes de partir para uma possível geração e disponibilização de novos conhecimentos (SPUDEIT, 2021).

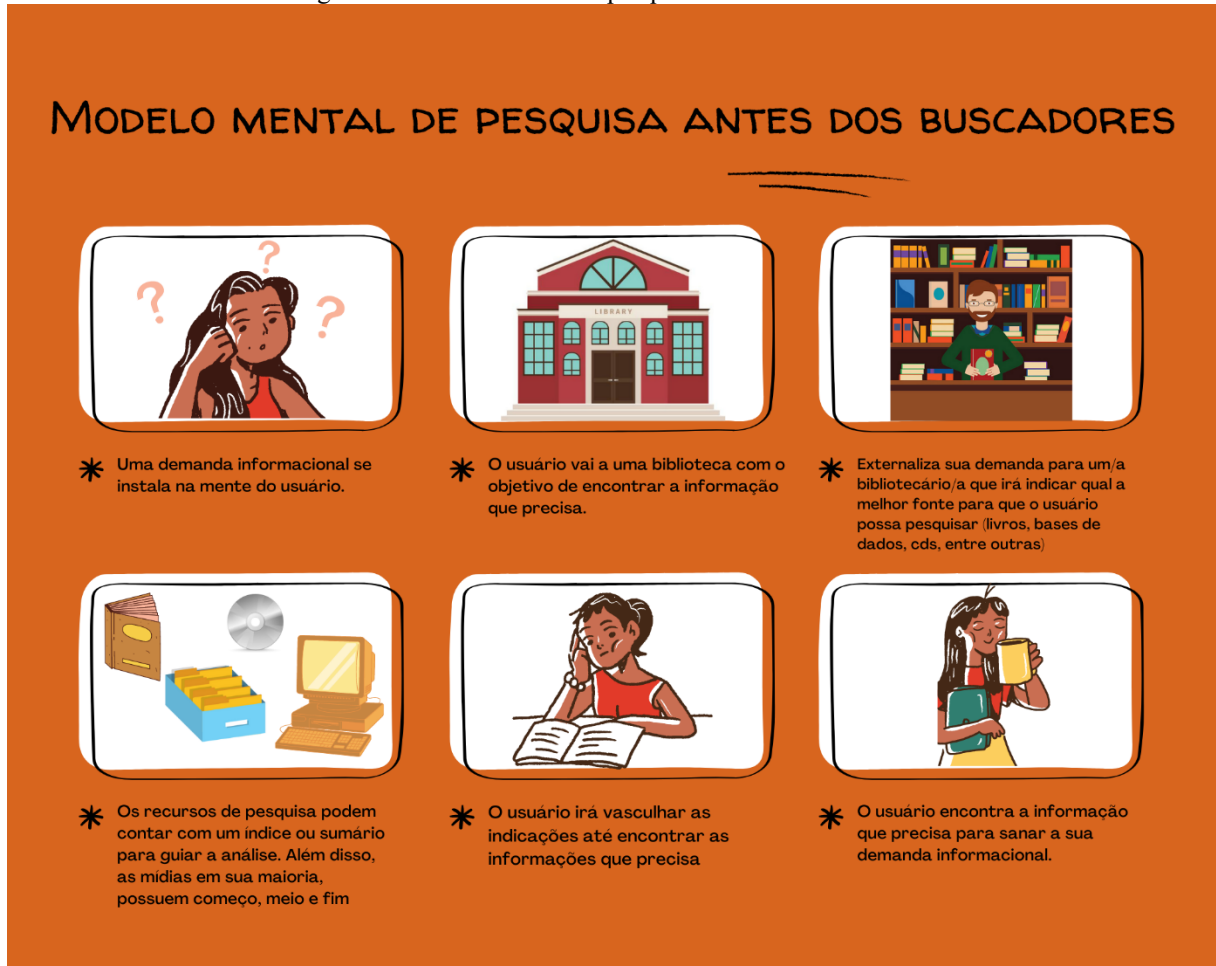
Essa possibilidade de autonomia na pesquisa on-line tida por um usuário letrado digitalmente, em momento algum indica que os catálogos em linha das bibliotecas perderam seu valor. Apenas são utilizados para finalidades diferentes em pesquisas realizadas por usuários com necessidades também diferentes. Um usuário não vai à biblioteca para saber se vai chover no dia seguinte, mas certamente pesquisa isso na internet. Ao mesmo tempo que um usuário que precisa de ajuda para encontrar uma obra específica recorrerá à biblioteca em busca de uma consulta local ou a realização de empréstimo do material.

A fim de elucidar as diferenças de um processo de pesquisa nesses dois ambientes, a Figura 3 e a Figura 4 retratam um modelo mental representativo do momento antes dos buscadores nas

³ Catálogo automatizado no qual o usuário faz o acesso direto, sem necessidade de intermediário, utilizando interfaces amigáveis. A maioria dos catálogos de bibliotecas disponíveis na internet é desse tipo. (CUNHA; CAVALCANTI, 2008, p. 73).

bibliotecas e num momento pós buscadores, em que o acesso a dispositivos eletrônicos com internet é uma realidade comum.

Figura 3 - Modelo mental de pesquisa antes dos buscadores



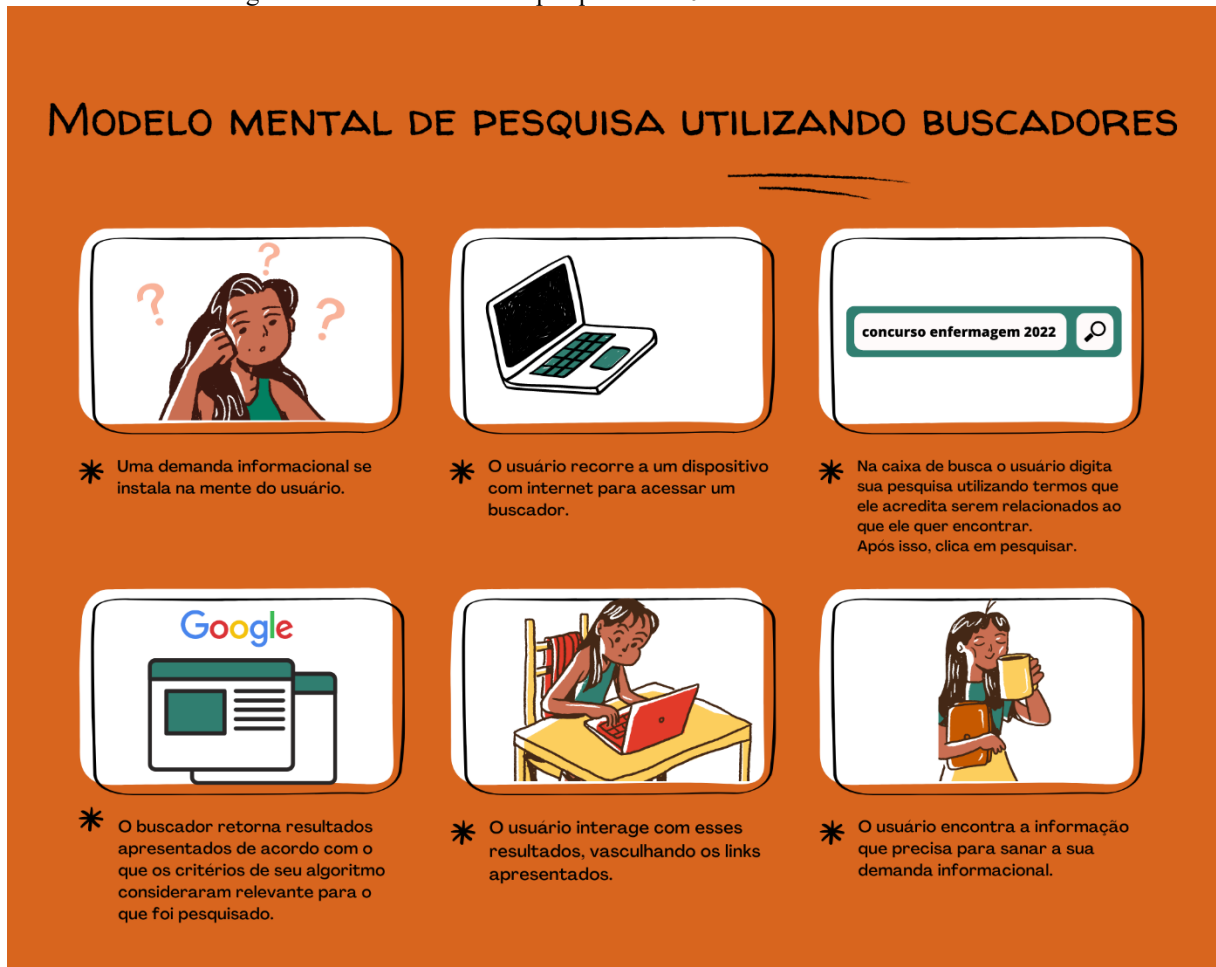
Fonte: Dados da pesquisa (2022)

O modelo mental representado na Figura 3 é um resumo de uma pesquisa realizada antes dos buscadores. Nele a personagem Verônica de 22 anos, acaba de se formar como técnica de enfermagem e encontra-se com uma necessidade de informação em que precisa descobrir editais de concursos públicos para o cargo e materiais de estudo da área.

Ao perceber que possui essa demanda informacional, Verônica decide ir até a biblioteca pública da sua cidade em busca de encontrar informações que resolvam seu problema. Ao chegar no local e narrar o seu problema, um bibliotecário lhe fornece acesso a jornais locais, apostilas de concurso e livros da área da saúde. Ao vasculhar os materiais fornecidos, Verônica encontra a publicação de dois editais em cidades vizinhas a qual ela reside, uma apostila atualizada e dois livros relacionados a temáticas que ela gostaria de revisar durante o período de estudos para a realização do concurso. Com isso, sua demanda informacional foi sanada.

A Figura 4 apresenta um modelo mental da realização de pesquisas em 2022 após a ampla utilização de buscadores e para uma usuária que possui em casa dispositivos eletrônicos com acesso à internet.

Figura 4 - Modelo mental de pesquisa em 2022 com o uso de buscadores



Fonte: Dados da pesquisa (2022)

A personagem Verônica, 22 anos, acaba de se formar como técnica de enfermagem e encontra-se com uma necessidade de informação em que precisa descobrir editais de concursos públicos para o cargo e materiais de estudo da área.

Ao perceber que precisa sanar essa demanda, ela pega o seu notebook que já está conectado à internet da residência, inicia o navegador e acessa o endereço do buscador Google para digitar sua pesquisa.

Na caixa de pesquisa Verônica insere os termos 'concurso', 'enfermagem' e '2022' na sequência e clica no ícone da lupa que aciona o comando pesquisar. O buscador entende que a usuária gostaria de recuperar resultados relacionados a concursos da área de enfermagem em

2022 e lhe apresenta milhares (ou até milhões), dos quais o algoritmo do sistema julgou relevante a pesquisa executada.

Com os resultados em tela, Verônica vasculha entre as primeiras opções e encontra a publicação de nove editais nas regiões sul e sudeste, anúncios de cursos preparatórios para concursos no cargo de técnico de enfermagem com materiais próprios de estudo e provas anteriores comentadas. Com isso, sua demanda informacional foi sanada.

Naturalmente, esses modelos mentais são apenas exemplos a caráter de diferenciação e representam somente a superfície dessas duas situações e muitos outros aspectos podem modificar o processo.

Uma pesquisa realizada em bibliotecas traz como prós alguns pontos que devem ser levados em consideração, tais como: atendimento especializado; interação social, e material pré-selecionado.

Ir a uma biblioteca permite que o usuário seja atendido por um bibliotecário que conhece profundamente o acervo e sua organização, ele poderá compreender de maneira específica a demanda informacional do usuário e recomendar os materiais mais apropriados para resolvê-la.

A interação social também é um fator importante para se levar em conta pois, ao expor suas perguntas o usuário tem a possibilidade de conversar sobre o problema em questões e pensar em soluções talvez antes não cogitadas. Além disso, ouvir a si mesmo relatando um problema faz com que essa fala reverberada possa ser refletida e talvez, se necessário, reformulada.

Em uma biblioteca todo material ali disponível para consulta foi anteriormente pré-selecionado e são fontes de informação seguras. Assim sendo, o usuário não precisa se preocupar se aquele conteúdo é seguro ou não para realizar sua pesquisa.

Enquanto os prós de uma pesquisa em buscadores, foco desse trabalho, consistem em: amplitude de conteúdo e recursos; e quantidade de resultados.

Nesse sentido, com o uso de buscadores, é possível visualizar “a dinâmica do conhecimento como processo que, anteriormente, era apenas individualmente observada nos processos de pensamento da própria pessoa” conforme apresentado nos processos de investigação praticados por Sócrates e Descartes por exemplo (SOUSA, 2019, p.16).

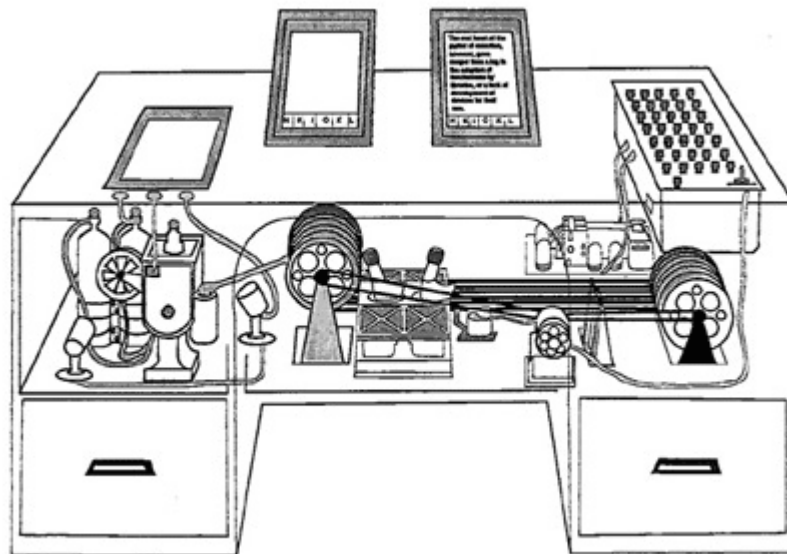
Com isso, a partir da progressão de todos os delineamentos descritos, em paralelo ao avanço tecnológico, se inicia no próximo tópico um apanhado acerca da trajetória dos buscadores que contribuíram para que os utilizados atualmente, sejam como são.

2.2 PRECEDENTES NO DESENVOLVIMENTO DE BUSCADORES

A história dos buscadores se inicia em 1945 com o ensaio publicado pelo estadunidense Vannevar Bush intitulado *As We May Think* (Como nós podemos pensar). Entre discussões acerca da guerra e atuação científicas, Bush introduziu pela primeira vez o conceito de hipertexto ao sugerir esforços para coletar e reunir o conhecimento humano a fim de torná-lo mais acessível a todos.

Nesse âmbito, a sua maior contribuição foi a criação de um dispositivo conhecido por Memex. Baseado nos processos de pensamento executados pela mente humana, que, diferente de um índice convencional, relacionam e distinguem informações por meio da associação. O Memex foi um dispositivo que armazenava conteúdos em microfilmes, nos quais esses conteúdos eram indexados com base nas associações entre si (BUSH, 1945; MURRAY, 2003).

Figura 5 - Esboço do Memex proposto por Vannevar Bush



Fonte: Time Inc (1945)

Como é possível observar na Figura 5, o Memex dispunha de uma mesa com duas gavetas que serviam de armazenamento dos microfilmes. Sobre a mesa, havia um conjunto de teclas e botões nos quais o usuário poderia informar códigos que solicitavam a apresentação de um conteúdo nas duas telas disponíveis. Esses conteúdos poderiam receber marcações que por sua vez, se relacionariam e estabeleceriam associações de caráter permanente, sendo, portanto, registradas junto ao conteúdo dos microfilmes. Esse fluxo contínuo de uma ou mais atribuições de associações levava o nome de trilha, e eram as trilhas que permitiam que um usuário em potencial pudesse recuperar as informações posteriormente.

Dito isso, e para diferenciar o índice das associações, no Memex os conteúdos eram, portanto, passíveis de recuperação de duas formas. A primeira delas ao fornecer um código de identificação que seria comparado as informações dispostas no índice convencional e se houvesse um *match*, o conteúdo era retornado às telas. Já na segunda forma, a navegação dos conteúdos era feita pelas trilhas desenvolvidas pelas associações estabelecidas, tornando possível que o usuário refizesse trajetórias que anteriormente foram percorridas e marcadas.

Pode-se dizer que de certa forma o Memex foi a materialização do Mondotheque idealizado pelo belga Paul Oltet, como um dispositivo capaz de armazenar, organizar e compartilhar documentos, referenciado por ele como uma espécie de ‘cérebro global’, algo que compartilhassem informações de todo o mundo. Em menor escala, o Memex foi o início da concretização desse pensamento, ao ligar um documento ao outro (TOTH, 2017).

Alguns anos depois, o alemão Gerard Salton contribuiu para os buscadores atuais no que diz respeito ao conceito de relevância. Junto a sua equipe, desenvolveu um sistema de recuperação conhecido por SMART (System for the Mechanical Analysis and Retrieval of Text), no qual o conceito de relevância foi aplicado. A partir de um modelo computacional conhecido por VSM (Vector Space Model), Salton utilizou de vetores matemáticos para desenvolver um sistema que extraísse palavras-chave do documento, analisasse essas mesmas palavras com base na frequência em que eram citadas, de forma que fosse possível atribuir um peso para cada um dos termos e por último classificá-los como mais ou menos relevante de acordo com as palavras-chave buscadas pelo usuário durante a consulta (SALTON, 1975; TOTH, 2017).

Um conteúdo é considerado relevante se for apropriado ao usuário que o solicitou. Essa relevância pode variar para mais ou menos, e a medição derivava da forma como o usuário formulou sua consulta. O SMART funcionava para que relevância versus correspondência contextual resultassem em êxito para o usuário em relação ao conteúdo retornado.

Salton também escreveu um livro chamado A Teoria da Indexação, que contava com muitas informações acerca dos testes contínuos das pesquisas realizadas por ele. Mais à frente, Frederick Wilfrid Lancaster com a obra Indexação e Resumos: teoria e prática, também faria uma enorme contribuição no campo da recuperação da informação, detalhando as formas de indexação de assuntos, sobretudo em relação a documentos.

Outra contribuição importante para a história do desenvolvimento dos buscadores foi a do filósofo e sociológico estadunidense Ted Nelson. Em seu projeto intitulado Xanadu, Nelson consolidou os termos hipertexto e hipermídia em meados dos anos 60.

Diferentemente do projeto encabeçado por Bush com o Memex, o Xanadu apresentava uma noção mais reduzida e individualista no desenvolvimento de um sistema de arquivos na rede de computadores.

O projeto se estendeu ao longo de décadas e nos anos 90 teve seu código publicado como aberto (*open source*). Mesmo de fato nunca decolando, é possível visualizar alguns dos seus preceitos no WWW (World Wide Web) de Tim Bernes-Lee.

Bernes-Lee uniu toda a parte teórica advinda do conceito de hipertexto com a parte técnica dos protocolos TCD (Transmission Control Protocol) e DNS (Domain Name System) para então chegar ao famoso WWW.

Em sua parceria de trabalho com o CERN (European Organization for Nuclear Research) durante a década de 80, Bernes-Lee encontrava-se em meio as melhores condições possíveis para contribuir com a história dos buscadores. O CERN precisava de novas condições de transferência de conhecimento, que não as já realizadas por papel, para que pudesse melhorar a logística das suas pesquisas. Na busca por uma tecnologia de transferência que resolvesse esse problema, a solução encontrada foi padronizar a codificação de todos os documentos utilizados pelos envolvidos, para que a partir de um único formato e a atribuição de códigos de controle baseados em regras específicas, intitulados posteriormente como HTML (HyperText Markup Language).

O HTML ou Linguagem de Marcação de Texto, permitiu que documentos pudessem ser interpretados por navegadores, informando-os qual é o estilo e a estrutura do documento em questão.

Posteriormente o HTML se tornou o padrão oficial da internet e recebe atualizações e orientações específicas do W3C (World Wide Web Consortium), consórcio internacional, também fundado por ele, que trabalha constantemente no desenvolvimento de padrões para a web.

O físico inglês ficou conhecido como pai da web e como já dito, trabalha atualmente para a expansão do conceito de web semântica.

Antes de adentrar na apresentação dos buscadores em si, é importante frisar que antes dos buscadores, o acesso a sites disponíveis na internet se dava por um link já conhecido, na esperança de que, durante a navegação fosse encontrado um outro link que levasse o usuário para outro site, depois outro, depois mais outro e assim sucessivamente. Uma experiência parecida com o que acontece hoje ao acessar alguns conteúdos da Dark Web, já que sem o link de acesso

direto é quase impossível chegar a um determinado site. Mesmo que haja ferramentas de busca voltadas a esse ambiente, ainda assim existe uma grande parte de conteúdo que provavelmente não é indexada por nenhum buscador.

2.2.1 Os buscadores e suas evoluções

Nas palavras de Cendón (2001) os motores de busca começaram a surgir quando o número de recursos na web estava crescendo e adquiriu proporções que impediam a coleta manual e busca somente por meio de uma navegação simples. “A maioria deles derivou do trabalho de estudantes de pós-graduação, professores, funcionários do departamento de sistemas de empresas ou outras pessoas interessadas na web” (CENDÓN, 2001, p. 41).

Um mecanismo de busca, ou buscador, como adotado preferencialmente nesse texto, é segundo Buchinger, Cavalcanti e Hounsell (2014, p. 109) “a composição de site, motor de busca, recursos de interface e conjunto de bases de dados que cobrem uma determinada área de conhecimento, visando a facilitar a identificação de materiais específicos e relevantes”. Tal definição vai de encontro com a estruturada por Gabriel (2012, p. 36) que considera um buscador como um sistema de recuperação de informações que tem como finalidade “auxiliar na busca de informações armazenadas em ambientes computacionais”, ou seja, é um intermédio entre o usuário e os dados e informações que ele precisa.

Além de mídia do conhecimento, os buscadores são Sistemas de Recuperação da Informação (SRIs). Estes, segundo Rowley (2002) passaram por três gerações, sendo a primeira geração nomeada de metadados (com interfaces para especialistas e mediadores da informação, ou seja, não acessível a todos e com um número muito limitado de sistemas on-line nas instituições. A segunda geração era caracterizada por dados com texto integral. As interfaces eram baseadas em menus e comandos, e o texto era recuperado de forma integral e sua utilização mais abrangente, estando disponível também para usuários. Por último, a terceira geração multimídia (com interface gráfica, on-line, onde o mediador da informação tem papel de instrutor de uso já que esta é voltada para o usuário final.

O primeiro buscador surgiu em 1990, intitulado Archie, o programa foi desenvolvido por Alan Emtage, Bill Heelan e Peter Deutsch, ambos da Universidade McGill em Montreal no Canadá. Consistia em uma ferramenta de busca em sites públicos de FTP (File Transfer Protocol) e retornava ao usuário uma lista de diretórios reunindo os arquivos localizados nesses sites. Como

é possível observar na Figura 6, a caixa de pesquisa superior indicava a possibilidade de buscar arquivos por nome (MOREIRA; ROMÃO, 2010; SEYMOUR; FRANTSVOG; KUMAR, 2011).

A interface do Archie trazia a possibilidade de escolher a base de dados, o tipo de busca, diferenciar maiúsculas e minúsculas, se o resultado seria o site ou apenas a expressão e ainda como a busca se daria. Todas estas possibilidades pertencem a uma linguagem lógica e se afastam do que o usuário leigo entende sobre uma busca, ou seja, de seu modelo mental sobre como realizar uma busca, como representado na Figura 3.

Argumenta-se, portanto, que a interface transparecia parte das decisões, e as apresentava como responsabilidade do usuário. O *layout*⁴, porém, favorecia o aceite das opções *default*, uma vez que a ordem de apresentação e área de foco privilegia a entrada de texto e restringe a leitura e seleção das opções.

Figura 6 - Archie

Archie Query Form

Search for: Search

Database: Worldwide Anonymous FTP Polish Web Index
 Search Type: Sub String Exact Regular Expression
 Case: Insensitive Sensitive

Do you want to look up strings only (no sites returned):
 NO YES

Output Format For Web Index Search: Keywords Only
 Excerpts Only
 Links Only

[Wersja polska](#) | [Help](#) | [Advanced Search](#)

Last changed on 11/20/2017 18:40:18 by [Wojtek Sylwestrzak](#)

Fonte: Web Design Museum, 1996

Em seguida surgiu o GOPHER (ver Figura 7), o qual nasceu como um protocolo de comunicação. Um buscador desenvolvido por Mark McCahill na Universidade de Minnesota que organizava arquivos em um formato de texto simples, priorizando a apresentação por este tipo

⁴ A presença ou ausência da estilização em itálico para termos estrangeiros nesta dissertação seguiu as recomendações dispostas no Manual de Comunicação da SECOM do Senado Federal. Disponível em: <https://www12.senado.leg.br/manualdecomunicacao/estilos/estrangeirismos-grafados-sem-italico-ou-aspas>.

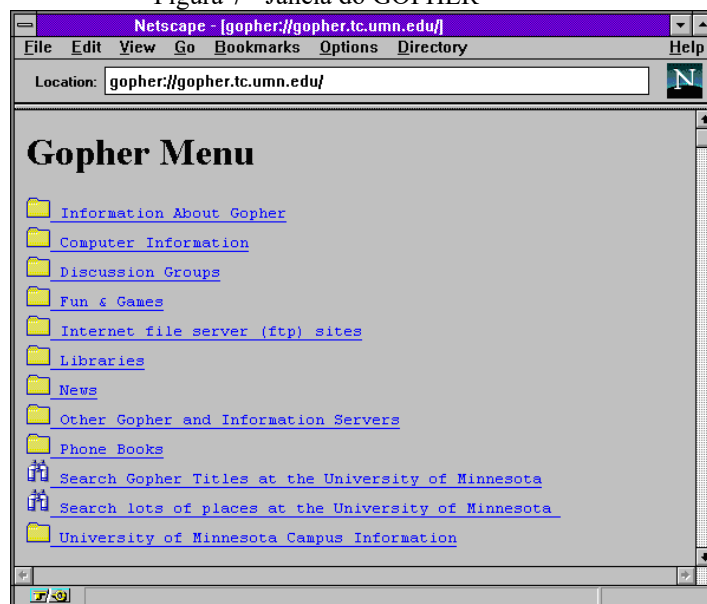
de documento, como um menu, o que o tornava ágil na execução de suas tarefas. Ele permitia que o usuário visualizasse uma breve descrição do conteúdo disponibilizado e pudesse fazer o download se assim o desejasse (MOREIRA; ROMÃO, 2010; SEYMOUR; FRANTSVOG; KUMAR, 2011).

É possível observar na Figura 7 que o campo de busca não era centralizado e dividia espaço com a logo, filtros e informações do rodapé. Aproximadamente 40% de espaço para filtros, 20% para busca, 15% para o logo no cabeçalho e 25% para o rodapé.

Comparado ao Archie, a área do campo de entrada de texto para busca é reduzida, entendendo-se que o foco de atenção pretendido é a navegação a partir do Gopher Menu. As categorias listadas neste menu, organizam serviços como Discussão em grupos, jogos, etc., e junto com estes o serviço FTP.

A interface do Gopher procura destacar as categorias, e como no Archie apresenta um design centrado na tecnologia uma vez que atribui ao usuário a responsabilidade de entender como a informação é armazenada e catalogada.

Figura 7 - Janela do GOPHER

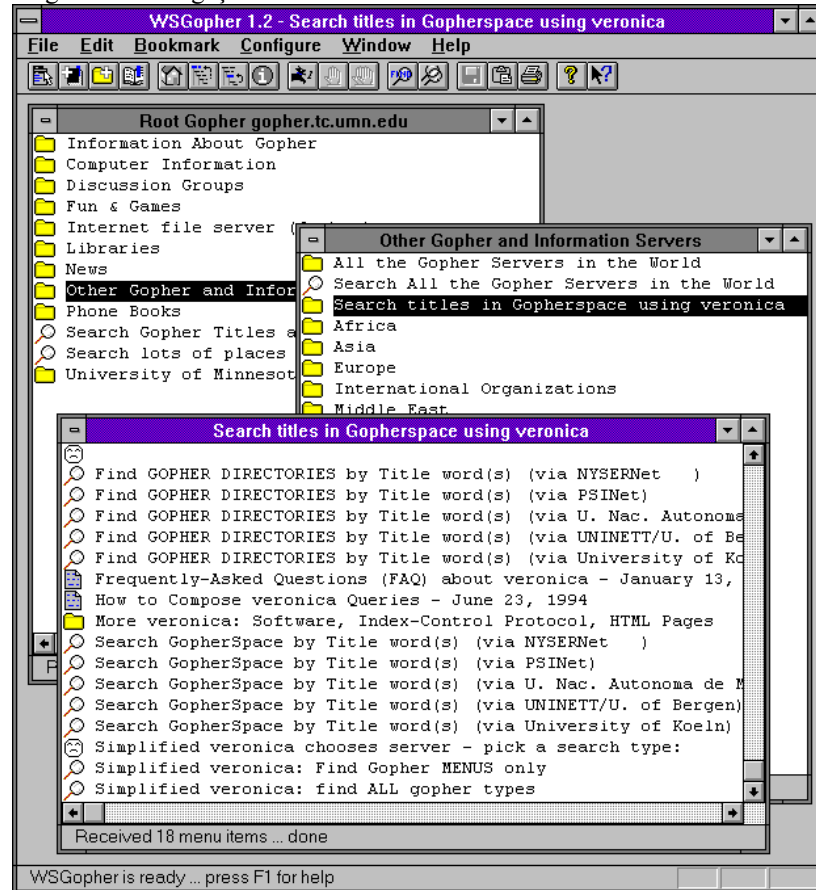


Fonte: Web Design Museum, 1991

Nesse meio tempo, outras duas ferramentas também aproveitaram do sistema de índice do GOPHER e buscavam arquivos armazenados nele. ‘Subcriações’ como o Veronica (Very Easy Rodent-Oriented Net-wide Index to Computerized Archives) (ver Figura 8) e Jughead (Jonzy's Universal Gopher Hierarchy Excavation and Display) tiveram seu valor ao proporcionar

experiências significativas a partir do modelo de índice do GOPHER (MOREIRA; ROMÃO, 2010; SEYMOUR; FRANTSVOG; KUMAR, 2011).

Figura 8 - Navegação nos itens do GOPHER através do Veronica

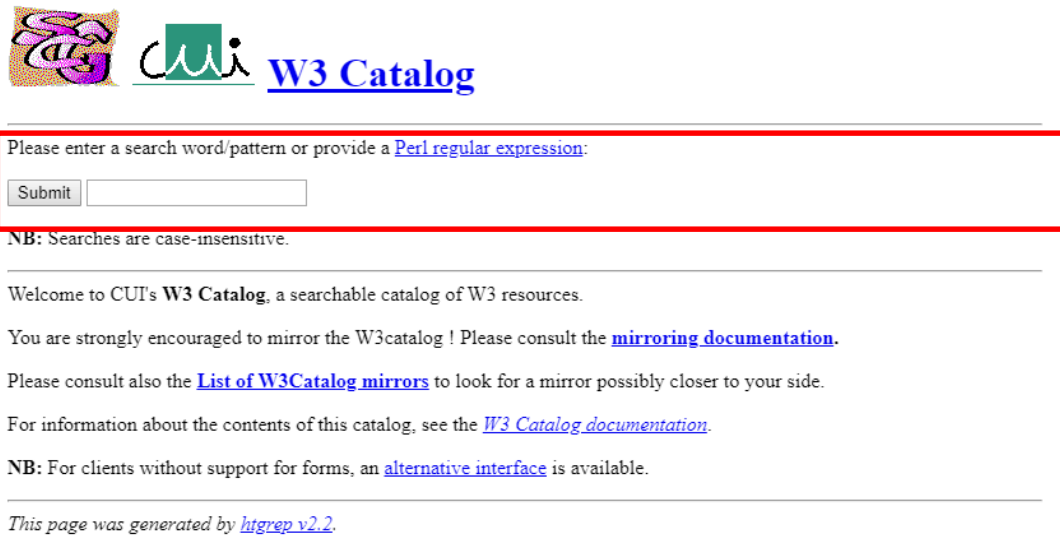


Fonte: Web Design Museum, 1992

Importante frisar que tanto o GOPHER quanto o Veronica não possuíam caixas de busca como as que existem hoje ou a do Archie por exemplo. A navegação ocorria por menus.

O W3 Catalog (ver Figura 9), surgiu em 1993 e foi desenvolvido na Universidade de Genebra na Suíça, por Oscar Nierstrasz, mesmo autor do Jughead, como uma evolução desse. Sua contribuição foi na replicação de páginas web com listas de recursos já existentes, para que pudesse torná-las pesquisáveis de uma forma dinâmica a partir da atribuição de entradas individuais (SEYMOUR; FRANTSVOG; KUMAR, 2011).

Figura 9 - W3 Catalog



Please enter a search word/pattern or provide a [Perl regular expression](#):

Submit

NB: Searches are case-insensitive.

Welcome to CUI's **W3 Catalog**, a searchable catalog of W3 resources.

You are strongly encouraged to mirror the W3catalog ! Please consult the [mirroring documentation](#).

Please consult also the [List of W3Catalog mirrors](#) to look for a mirror possibly closer to your side.

For information about the contents of this catalog, see the [W3 Catalog documentation](#).

NB: For clients without support for forms, an [alternative interface](#) is available.

This page was generated by [htgrep v2.2](#).

Fonte: Web Design Museum, 1995

A caixa de pesquisa agora está mais ao topo da página, contudo ainda não possui destaque e divide espaço com várias outras informações de consulta, uso e apresentação da ferramenta.

Tanto o Archie, como o GOPHER, Veronica, Judhead e W3 Catalog, são de um período pré WWW, portanto, não são considerados buscadores web, mas contribuíram essencialmente para os que vieram depois da tecnologia.

Assim sendo, dois anos após o nascimento do World Wide Web, em 1993, surge o Wanderer (também conhecido por World Wide Wanderer ou Wandex, nome do seu índice). Projetado por Mathew Gray e desenvolvido dentro do Instituto de Tecnologia de Massachussets (MIT), é considerado o primeiro buscador web (INCE, 2009). Capturava o endereço das páginas e teve como principal objetivo efetuar a medição do quantitativo de conteúdo presente na internet até o momento. Seu avanço foi indexar os endereços das páginas da internet (URLs), e sua tecnologia se tornou uma das bases de todos os buscadores atuais (SEYMOUR; FRANTSVOG; KUMAR, 2011)

O Aliweb (Archie-Like Indexing in the Web) (ver Figura 10) surgiu ainda em 1993. Inovou na época ao permitir que os próprios usuários enviassem as páginas que seriam indexadas. Donos de site ajudavam a construir o acervo enviando descrições e palavras-chaves. Se o responsável por um site quisesse integrar ao Aliweb, teria que escrever um arquivo no formato solicitado e registrá-lo no servidor do buscador (SEYMOUR; FRANTSVOG; KUMAR, 2011).

Figura 10 - Aliweb

ALIWEB Search Form

A [Public Service](#) provided by [NEXOR](#)

This form queries the [ALIWEB](#) database. You can provide multiple search terms separated by spaces, and the results will be displayed in a best-match order.

Search term(s):

There are several types of search:
 Case Sensitive

Which type records would you like to search:

Which fields would you like to search?
 Titles Descriptions Keywords URL's

What fields would you like displayed in addition to the title?
 Descriptions Keywords URL's Other fields

You can restrict the results to a domain (e.g. "uk"):

Stop after the first **matches.**

[Return to NEXOR home page](#)

For questions or comments regarding our web service, contact webmaster@nexor.co.uk
 Copyright NEXOR Ltd 1995, All rights reserved, PO BOX 132, Nottingham, NG7 2UU, UK.

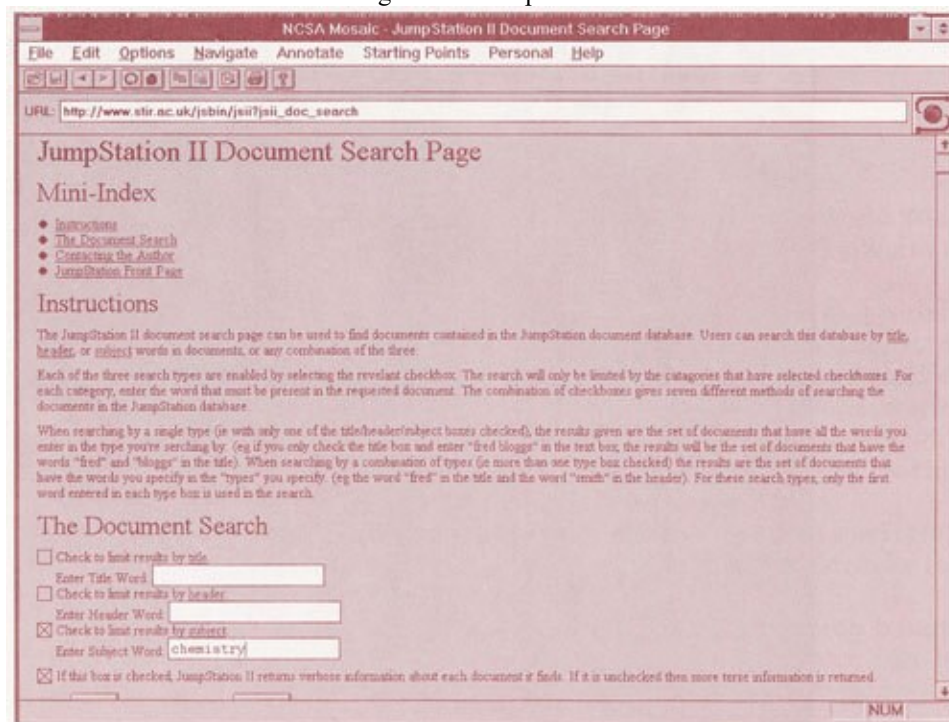
Fonte: Web Design Museum, 1995

O campo de busca no Aliweb apresentava na interface do buscador um grau de importância equivalente aos filtros de direcionamento das buscas. Esses já incluíam possibilidades similares aos operadores avançados de pesquisa disponíveis hoje, como por exemplo, a viabilidade em direcionar o domínio dos sites nos quais os usuários desejassem obter resultados. O design da interface continua centrado na tecnologia.

Ainda em 1993, surgiram outros buscadores como o JumpStation (ver Figura 11) que utilizava um robô que vasculhava a web atrás de novos conteúdos, por meio de palavras-chave. O que para a época era algo bastante inovador, ao levar em consideração que os anteriores não realizavam indexação automática de sites. Surgia nesse momento o primeiro buscador que combinava o uso, do que viria a ser o tripé do funcionamento dos buscadores: rastreamento, indexação e pesquisa. As limitações até então estavam em pesquisas limitadas a somente títulos e cabeçalhos das páginas web indexadas pelo JumpStation e sem uma classificação conveniente desses resultados, o que poderia suceder uma tarefa árdua em encontrar aquilo que realmente o

usuário estava buscando (DOERNHOEFER, 2010; SEYMOUR; FRANTSVOG; KUMAR, 2011).

Figura 11 - JumpStation



Fonte: Smith, 2019

Jonathan Fletcher, criador do Jumpstation, disse em uma entrevista ao jornal escocês *The Scotsman* como outros webmasters reagiram ao seu robô de rastreamento e indexação de sites. Segundo ele, na época recebeu uma mensagem de um deles contendo o seguinte pedido: “Não sabemos quem você é ou o que está fazendo, mas por favor, pare com isso” (THE SCOTSMAN, 2009). É até ligeiramente cômico ler essa mensagem em 2022, numa realidade em que o processo de indexação de sites é automático e há inúmeras técnicas orgânicas e pagas para ‘melhorar’ o posicionamento de determinado site dentre os resultados apresentados.

Um ano depois, em 1994 o WebCrawler (ver Figura 12) inovou ao possibilitar que o usuário buscasse em seu sistema a partir de qualquer palavra em qualquer página da internet. Como o próprio nome sugere, utilizava robôs para localizar documentos na web, indexar páginas inteiras, ou seja, textos completos e não apenas índices (SEYMOUR; FRANTSVOG; KUMAR, 2011). Seu nome virou o termo oficial para designar a tecnologia que realizava o serviço de indexação de sites automaticamente.

Figura 12 - WebCrawler



Fonte: Web Design Museum, 1996

Desenvolvido por Brian Pinkerton na Universidade de Washington, inicialmente foi projetado para ser um aplicativo de *desktop*, e posteriormente se tornou a versão web (Figura 12). Dos buscadores citados aqui, o WebCrawler é um dos poucos que ainda está ativo, e o mais antigo (DOERNHOEFER, 2010). Já não possui seu próprio banco de dados, mas agrega resultados vindos do Google e do Yahoo. O que o torna hoje um metabuscador¹.

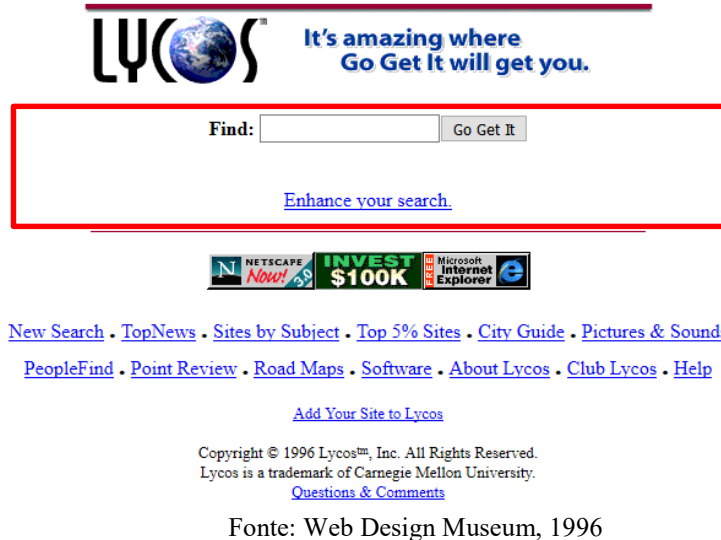
A interface do WebCrawler começa a ter um design centrado no usuário, pois oferece opções de como apresentar o resultado, e principalmente pelo exemplo de como escrever o texto para busca. Também apresenta uma composição por hierarquia, privilegiando a busca, pela área ocupada e pela posição superior.

O ano de 1994 seguiu com grandes inovações. O Lycos (ver Figura 13), criado na Carnegie Mellon University e tendo como responsável Michale Mauldin, trouxe como novidade a correspondências de prefixos e apresentação de palavras-bônus de acordo com um critério de proximidade da busca realizada pelo usuário (MAULDIN, 1997).

Dois anos depois ele apresentava em seu catálogo um total de pelo menos 60 milhões de documentos indexados. O que para a época, representava algo maior do que qualquer outro buscador havia atingido (MAULDIN, 1997).

Figura 13 - Lycos

[Serious Sports Fans Only \\$1,000,000 in Cash and Prizes!](#)
[For serious sports fans only! Play Fantasy Football!](#)



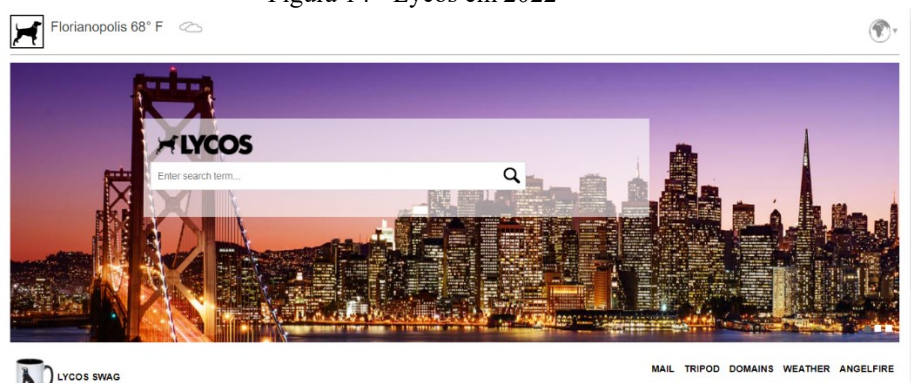
Fonte: Web Design Museum, 1996

O Lycos também foi importante para construir a imagem que o mercado começaria a ter dos buscadores na época ao iniciar uma oferta pública de ações. Investindo em anúncios para o *mainstream* e se tornando um dos primeiros negócios lucrativos do ramo (A HISTÓRIA..., 2022 DAVIS, 2007).

A caixa de pesquisa começa a dividir espaço com informações referentes à publicidade. Sai de cena um pouco a questão de informar ao usuário como o buscador funciona ou pode funcionar por meio de buscas avançadas e entra a questão de como eles podem ser atrativos para o mercado, apresentando sua tela de busca inicial como uma espécie de vitrine para anúncios.

Assim como o WebCrawler, o Lycos (ver Figura 14) ainda segue em funcionamento.

Figura 14 - Lycos em 2022



Fonte: Lycos, 2022

Na sequência chegam outros buscadores como o Infoseek sob responsabilidade de Steve Kirsch na Califórnia, o Excite, direto da Universidade de Stanford elaborado por seis estudantes, e o Sapo (Servidor de Apontadores Portugueses Online) que foi desenvolvido em Portugal na Universidade de Aveiro (ver Figura 15) (KRAUS, 2007; COUVERING, 2008; SEYMOUR; FRANTSVOG; KUMAR, 2011).

Figura 15 - Infoseek, Excite e SAPO

Infoseek
Once you know, you know.™

Type a specific question, phrase or Name.

 the Web [Tips](#)

To explore the Web's largest directory, click a topic below.
[Arts](#) - books, photography, fashion, museums ...
[Business](#) - business tools, find a job, small business...
[Computers](#) - desktop computers, hardware, software...
[Education](#) - colleges, teaching, adult education...

excite
 search reviews city.net live! NEW tours
 people finder maps yellow pages news

Excite Search: twice the power of the competition.
 What:
 Where: World Wide Web
 INTEGRATED BROWSING, EMAIL, NEWSGROUPS AND PAGE CREATION.

Excite Reviews: site reviews by the web's best editorial team.
 • [Arts](#) • [Entertainment](#) • [Money](#) • [Regional](#)
 • [Business](#) • [Health](#) • [News & Reference](#) • [Science](#)
 • [Computing](#) • [Hobbies](#) • [Personal Pages](#) • [Shopping](#)
 • [Education](#) • [Life & Style](#) • [Politics & Law](#) • [Sports](#)

SAPO
 Informações
 TOP SAPO

 Ainda não registou o seu endereço de E-Mail? [De que é que está à espera?](#)

- **Novidades**
[Novos Links](#), [Congressos](#), ...
- **Ensino e Investigação**
[Universidades](#), [Institutos](#), [Escolas](#), ...
- **Comunicação Social**
[Jornais](#), [Rádios](#), [Televisão](#), ...
- **Comércio, Indústria e Serviços**
[Serviços](#), [Informática](#), [Saúde](#), [Lojas](#), ...
- **Páginas Pessoais**
[Páginas pessoais](#), [Lista de E-Mails](#)
- **Sociedade e Cultura**
[Museus](#), [Hospitais](#), [Religião](#), [Governo](#), ...

Fonte: Web Design Museum, 1996, 1997; SAPO, 2018

Logo depois veio o Yahoo! (ver Figura 16) que ficou famoso por categorizar e descrever as páginas indexadas.



Fonte: Web Design Museum, 1997

Fundado por David Filo e Jerry Yang, ambos egressos da Universidade de Stanford, o Yahoo! surgiu como um diretório em 1994 e teve até o final da mesma década uma das maiores evoluções no quesito interface de pesquisa, inclusive firmado o nome do seu buscador como Yahoo! Search (COUVERING, 2008; SEYMOUR; FRANTSVOG; KUMAR, 2011).

O Yahoo! foi muito popular enquanto buscador e também por oferecer outros serviços integrados como e-mail, grupos e Messenger (DOERNHOEFER, 2010). Com a ideia original de servir como um guia, a facilidade de utilização do buscador permitia que usuários leigos navegassem com clareza e simplicidade. A empresa trabalhou muito bem o marketing a fim de que fosse reconhecida como uma marca divertida e de interface amigável.

Hoje existe um ditado popular —principalmente para negócios— de que ‘se você não está no Google você não existe’. Na época o Yahoo causou um frisson equivalente. Era essencial estar indexado pelo Yahoo e aparecer em seus resultados. A partir do Lycos ao surgimento do Yahoo, pode-se dizer que os buscadores entraram numa era de ouro (ZHAO, 2018).

O Alta Vista (ver Figura 17) surgiu em 1995 e se tornou o preferido de muitos usuários por possuir uma estrutura de funcionamento que o tornava mais rápido que seus concorrentes na época. Foi desenvolvido por Louis Monier, responsável pela tecnologia de rastreamento do buscador, e Michael Burrows, no comando do seu indexador. Foi um dos pioneiros em permitir a utilização de linguagem natural na busca, permitindo pesquisas por frases na forma como são faladas entre pessoas no dia a dia, como por exemplo ‘onde fica Londres?’, ao invés de galgar essa mesma informação apenas com o uso de palavras-chave na caixa de pesquisa como ‘Londres’ e ‘Localização’ (SEYMOUR; FRANTSVOG; KUMAR, 2011).

Figura 17 - AltaVista

The image shows the AltaVista search engine homepage. At the top left is the 'alta vista: SEARCH' logo with the tagline 'smart is beautiful'. To the right is a banner for 'Earn \$50 gift certificates!'. Below the logo is a navigation bar with links for 'Search', 'Live', 'Shopping', 'Local', 'Free Access', and 'Email'. A search bar is located below the navigation bar, with a 'Search' button and a dropdown menu for 'any language'. To the right of the search bar are 'SEARCH TOOLS' including 'Yellow Pages', 'People Search', and 'Directions'. Below the search bar is a 'Find Results on:' section with radio buttons for 'The Web', 'News', 'Discussion Groups', and 'Products'. The main content area is divided into several columns. The left column features 'AltaVista Live!' with links to 'Money, News, Sports, Translation', 'Travel, Careers, Health', 'Entertainment, Local, Portfolio', and 'Email, Chat, Alert'. It also lists 'Today's Features' and 'Top Stories'. The middle column lists various categories like 'Arts & Entertainment', 'Autos', 'Business & Finance', 'Computers', 'Games', 'Health & Fitness', 'Home & Family', 'Internet', and 'News & Media'. The right column features 'Recreation & Travel', 'Reference', 'Regional', 'Science', 'Shopping', 'Society & Culture', 'Sports', and 'World'. At the bottom right, there is a 'Bienvenido al AltaVista!' section with a 'Translate for me' link. The footer contains copyright information for AltaVista Company and a list of links for 'About AltaVista', 'Help', 'Contact Us', 'Advertise With Us', 'Business Solutions', 'Job Openings', 'Press Room', 'Privacy', 'Terms of Use', 'A CMGI Company', 'Shopping', 'Money', 'News', 'Sports', 'Travel', 'Careers', 'Health', and 'Entertainment'.

Fonte: Web Design Museum, 1999

Tanto o Infoseek, Excite, SAPO, Yahoo e AltaVista possuem interfaces bem próximas a um modelo visual que remete às enciclopédias e a busca em bibliotecas. A classificação por assuntos principais dividia espaço com a busca livre, indicando ao usuário equivalência entre uma busca com possibilidades do uso de filtros para com uma navegação por categorias por meio do índice de assuntos desenvolvido pelo buscador.

Essa questão do uso da linguagem natural nas pesquisas foi fortalecida pelo AskJeeves (ver Figura 18). O buscador lançado em 1997, priorizava perguntas feitas em linguagem natural, dispunha de uma nova pergunta principal todo dia, tida como a ‘pergunta do dia’ e retornava aos usuários perguntas relacionadas de acordo com a cada categoria (DOERNHOEFER, 2010).

Figura 18 - Ask Jeeves



Fonte: Web Design Museum, 1999

O AskJeeves centralizou a questão da caixa de pesquisa e da inserção de uma expressão de busca como função principal do seu buscador, incentivando e destacando-a na interface. O design também busca centrar no usuário pelo uso de um personagem e de uma linguagem na forma de diálogo.

A ilustração humana na Figura 18 fazia o papel de representar hipoteticamente uma pessoa real responsável por responder os questionamentos trazidos pelo usuário. Na época, um comercial de TV fortalecia tal ideia.

Vídeo 1 - Comercial de TV do Ask Jeeves



Fonte: Jim Margle, 1999

O comercial reforça a humanização do buscador e fez com que o processo de Interação Humano-Computador (IHC) se assemelhasse mais ainda a uma consulta de forma tal qual é realizada por meio do atendimento ao público em bibliotecas.

Partindo para a esfera nacional de buscadores na década de 90, o Brasil também obteve êxitos em serviços como o Cadê? e o Aonde? (ver Figura 19 e Figura 20).

O Cadê? nasceu em 1995 e foi o primeiro buscador brasileiro. Funcionava inicialmente como um diretório, nos mesmos moldes do Yahoo (tanto que foi absorvido pela empresa anos depois). A primeira versão contava com 300 páginas catalogadas manualmente de acordo com os critérios de seleção dos próprios fundadores, Gustavo Viberti e Fabio Oliveira (VIEIRA, 2003).

A interface do Cadê segue o formato da busca livre e navegação pelas categorias, sugerindo ao usuário uma estrutura de portal.

Atualmente, ao acessar o antigo endereço eletrônico cade.com.br, o usuário é direcionado para o Yahoo! Search.

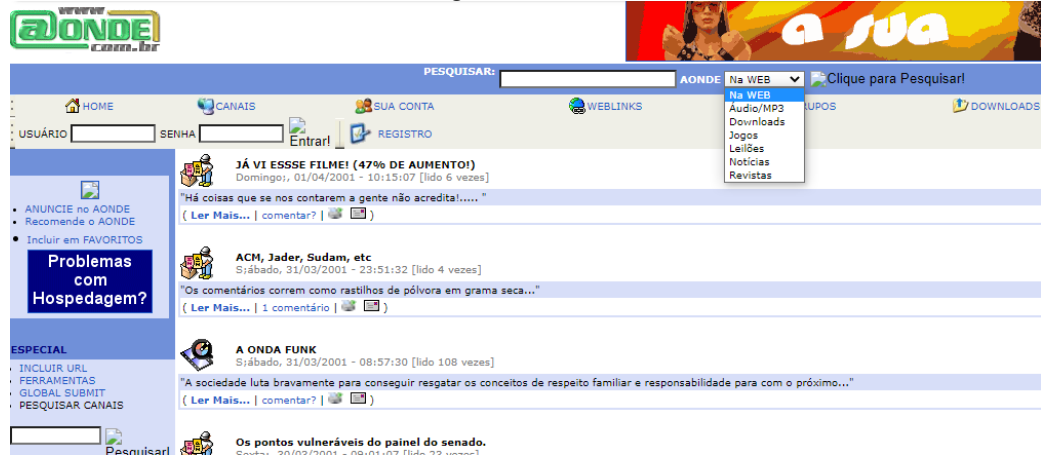
Figura 19 - Cadê



Fonte: Internet Archive, 1996

Já o Aonde, veio anos depois. Agregou outros serviços ao modelo de negócios como o desenvolvimento e hospedagem de sites, mas, segue em pleno funcionamento apesar de não ser tão popular no país em relação a outros buscadores mais utilizados.

Figura 20 - Aonde



Fonte: Internet Archive, 2001

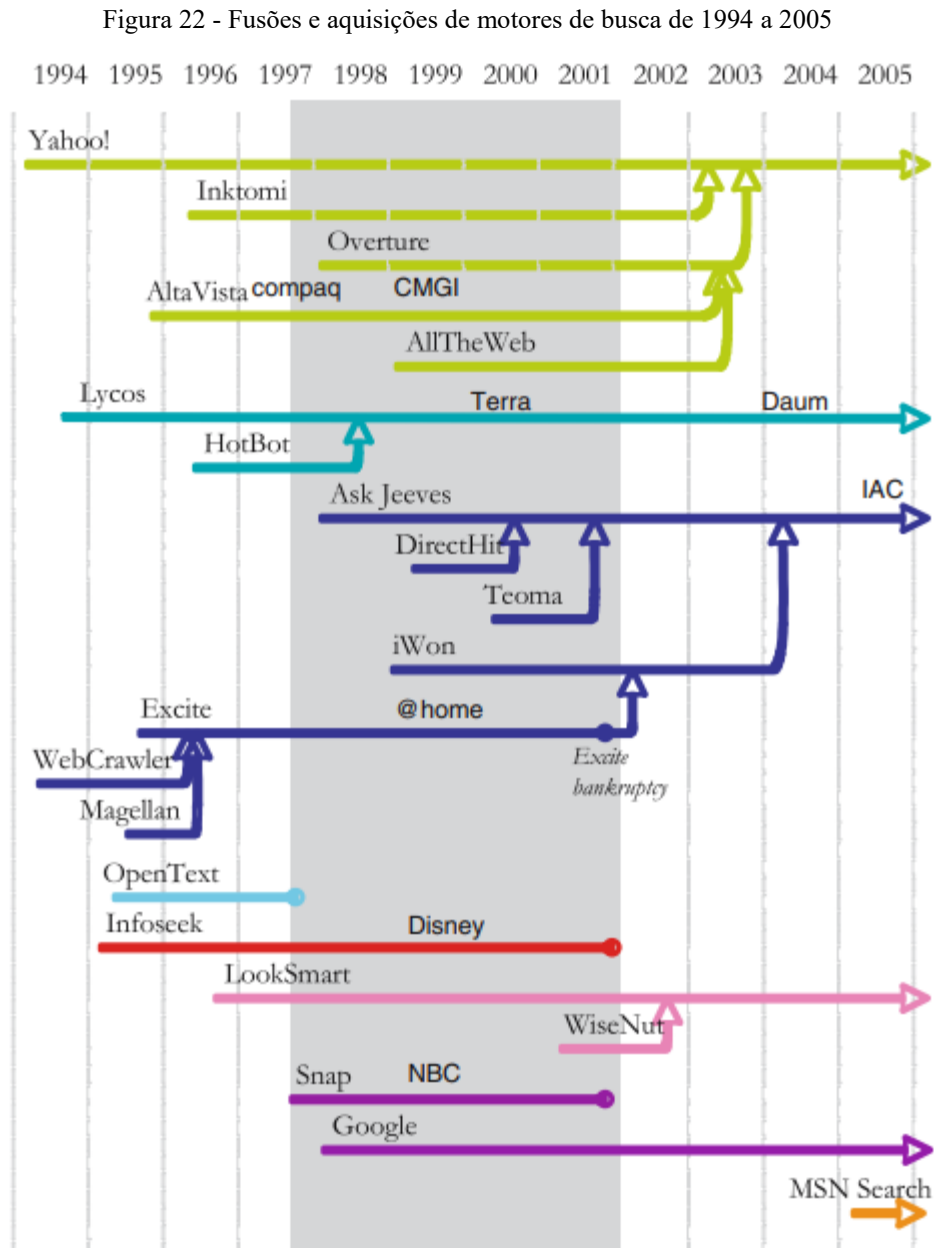
Ainda é possível listar outras iniciativas como o Achei, o Prokura e o Radix (ver Figura 21). Esse último desenvolvido numa parceria entre o CIn UFPE (Centro de Informática da Universidade Federal de Pernambuco) e o CESAR (Centro de Estudos e Sistemas Avançados do Recife) (RADIX, 1999).

Figura 21 - Achei, Prokura e Radix



Fonte: Web Design Museum, 1996, 1997; SAPO, 2018

Entre os buscadores ainda ativos, inativos e absorvidos, a Figura 22 elaborada por Couvering (2008) pode exemplificar melhor a linha cronológica a partir de 1994 até 2005.



Até a consolidação do Google como buscador mais popular em grande parte do mundo a partir de 2002, o Yahoo, Lycos, Excite, Infoseek e Altavista disputaram de forma acirrada a liderança de mercado, alternando posições no ranking de maiores pesquisas realizadas e total de usuários ativos. Desses cinco, o Yahoo foi o que mais obteve êxito comercial (GANDAL, 2001).

De 1994 a 1997 diversas outras empresas aperfeiçoaram os modelos anteriores a fim de encontrar seu espaço no mercado ainda inicial de buscadores. Mais precisamente em 1998 surge o Google, objeto de aplicação da pesquisa dessa dissertação.

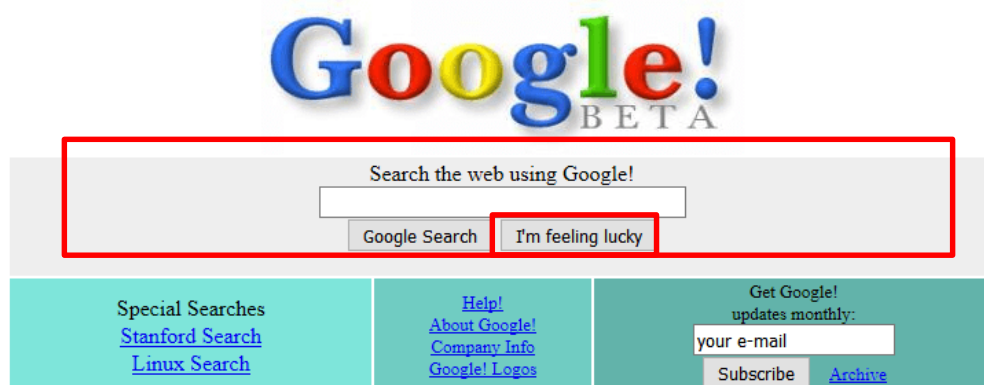
2.2.2 Google

Em 1998, os estudantes da Universidade de Stanford, Larry Page e Sergei Brin fundaram a Google. Empresa que tinha como principal carro chefe um buscador de mesmo nome.

Do final da década de 90 até início da década 20, o buscador só cresceu em popularidade, tornando-se em grande parte do mundo praticamente um sinônimo de pesquisa na internet. Frases derivadas da famosa expressão ‘*Google it*’, como ‘dá um Google’, ‘joga no Google’ e ‘procura no Google’ não são difíceis de ouvir quando há uma dúvida relativamente simples de ser resolvida (LOWE, 2009).

Na primeira interface do Google (ver Figura 23), a caixa de pesquisa já apareceu centralizada e inclusive dando ao usuário a possibilidade de ao inserir uma expressão de busca, executar a função ‘estou com sorte’ e deixar por conta do buscador a escolha da fonte de informação mais adequada para a pesquisa indicada, geralmente, o primeiro resultado elencado na pesquisa clássica.

Figura 23 - Google no início



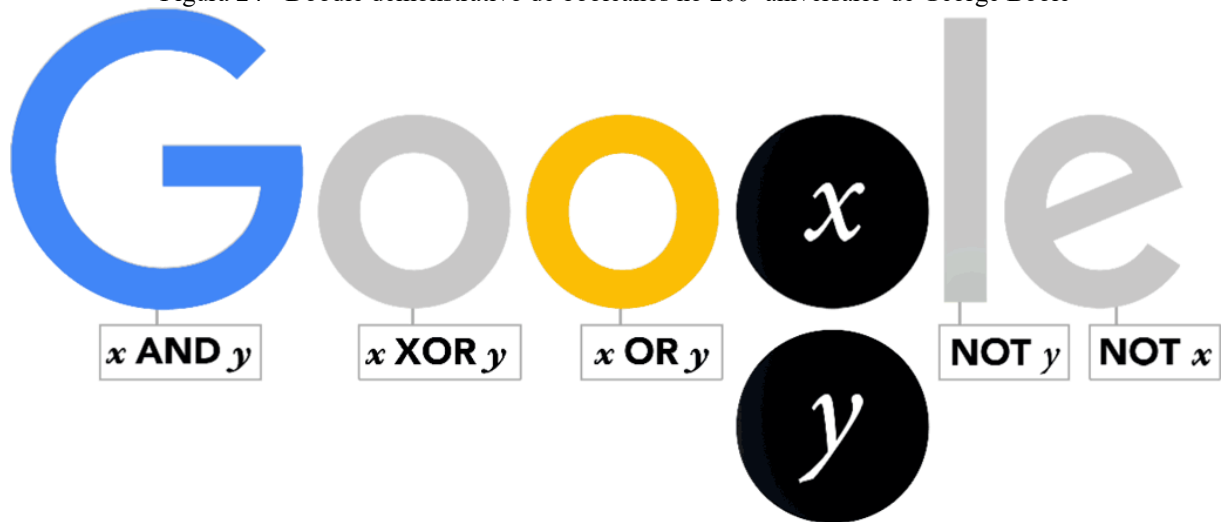
Copyright ©1998 Google Inc.

Fonte: Web Design Museum, 1998

A logo icônica do buscador, buscou tanto nas letras e cores selecionadas, consolidar a ideia de que uma busca na internet poderia ser simples e divertida. A imagem se consolidou e ao

falar Google as pessoas imediatamente veem a imagem colorida da marca. Ao longo dos anos, projetos como o Google Doodle, tornaram o ato de ‘brincar’ com a logomarca em ocasiões especiais uma tradição, como no caso da Figura 24 que celebra o aniversário do já citado George Boole (LOWE, 2009).

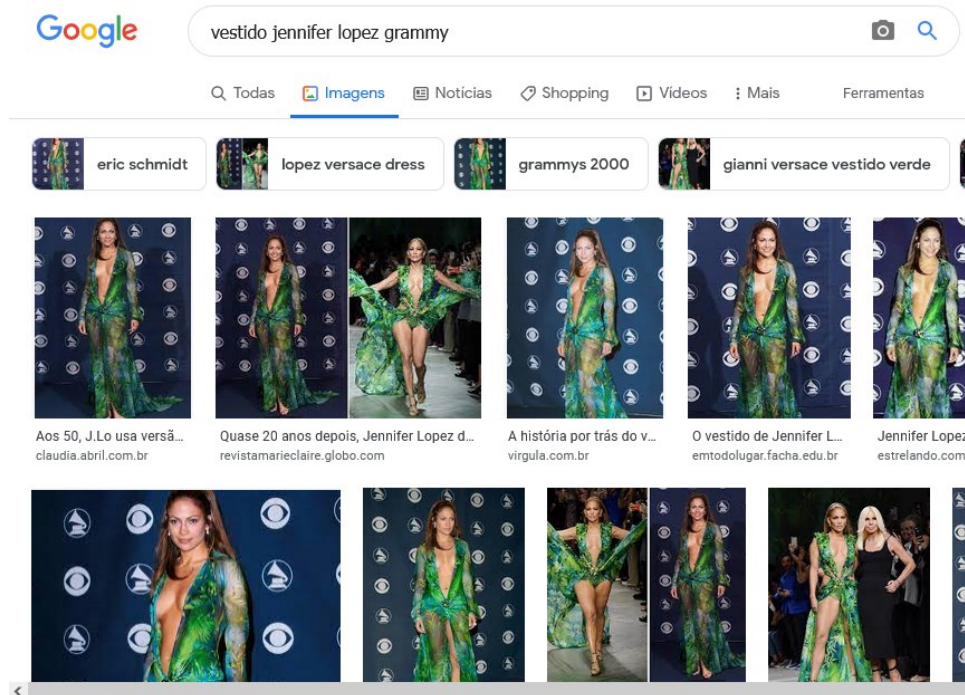
Figura 24 - Doodle demonstrativo de booleanos no 200º aniversário de George Boole



Fonte: Google, 2015

O Google Imagens, por exemplo, teve a cantora Jennifer Lopez como principal motivação para seu surgimento. Em 2000 na cerimônia da 42ª edição do Grammy Awards, Lopez compareceu ao evento com um vestido verde da grife de luxo Versace e causou muito alvoroço. No dia seguinte, todos queriam ver uma imagem desse vestido (ver Figura 25), e para isso, recorreram ao Google. No entanto, ao pesquisarem de diferentes formas, os usuários estavam sendo direcionados a resultados que continham apenas texto. Dessa demanda, surgiu uma aba dedicada à busca por imagens. Posteriormente a busca por imagens também evoluiu, com a disponibilização de filtros, operadores de pesquisa específicos e até a possibilidade de realizar uma pesquisa reversa, na qual o insumo na caixa de pesquisa não são expressões textuais e sim um próprio arquivo de imagem (GRAMMY AWARDS, 2000).

Figura 25 - Vestido responsável pela criação do Google Imagens



Fonte: Google, 2023b

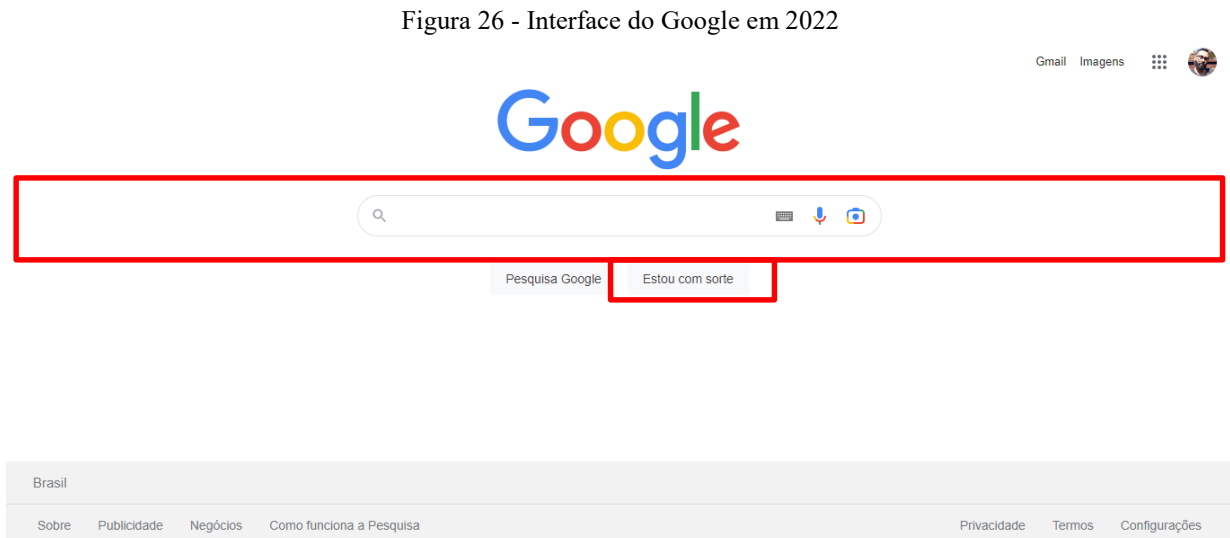
Antes do Google Imagens, outros buscadores ofereciam aos usuários recursos para pesquisas por imagens, dentre eles os já citados Alta Vista, Lycos e Image Surfer do Yahoo, além de vários outros como WebSEEK, HotBot, Amazing Picture Machine e Web Places Clip Art Searcher. No entanto, esse processo baseava-se especialmente em texto, buscado por termos contidos na legenda e nos metadados das imagens dispostas na web. Recursos mais avançados reconheciam cores e extensões de arquivos (BERINSTEIN, 1998). Nessa circunstância, destaca-se o AltaVista que oferecia possibilidades de busca simples e avançadas mais eficazes, inclusive com uso de operadores booleanos (PACK, 2002).

O Google revolucionou ao desenvolver uma tecnologia de análise visual. Hoje se um usuário pesquisar por vestido verde, mesmo que ambos os termos não sejam citados uma única vez sequer no arquivo da imagem, a ferramenta é capaz de retornar imagens relacionadas e atender a demanda de pesquisa solicitada.

Outras formas de pesquisar conteúdo foram surgindo com o passar dos anos a fim de tornar a página inicial do buscador cada vez mais cômoda ao usuário. É possível utilizar a própria página inicial do buscador como calculadora, conversor de moedas, temperatura, entre outras medidas, e até fugir de propagandas ao arriscar uma total confiabilidade em seus resultados,

utilizando o botão ‘Estou com sorte’, que é responsável por direcionar o usuário ao primeiro site resultante da pesquisa realizada (LOWE, 2009).

Em 2022 a interface do Google se apresenta da seguinte maneira (ver Figura 26).



Fonte: Google, 2023a

A caixa de pesquisa assume destaque e divide espaço com outras duas possibilidades recentes de pesquisa: por voz e por imagem de forma reversa. A interface é minimalista, o foco está totalmente na inserção de uma expressão de busca por parte do usuário.

Comandos como o ‘estou com sorte’ permanecem na interface mais como um saudosismo por parte do Google do que como uma ação útil ao usuário. Afinal, levar um usuário direto para uma página fora da empresa sem que ele passe pela página dos resultados e visualizasse anúncios de pesquisa, gera prejuízo financeiro.

Para além disso, o rodapé apresenta informações de publicidade, do histórico da empresa, privacidade digital, uso de dados e configurações adicionais. Em todo caso, as informações adicionais em tela não remetem mais a filtros de busca ou possibilidades de realizar uma pesquisa avançada com o uso de operadores.

2.2.2.1 Como a pesquisa no Google funciona?

Em sua página de explicação sobre o funcionamento da busca, o Google se intitula como “a maior biblioteca do mundo” e afirma que seu índice de pesquisa “é como uma biblioteca que contém mais informações do que todas as bibliotecas do mundo somadas” (GOOGLE, 2021a).

Anteriormente os buscadores ranqueavam os resultados da busca de acordo com a quantidade de vezes que a palavra-chave pesquisada aparecia naquele documento. Sendo esse o único critério de ranqueamento de resultados, é possível afirmar que essa ordem pode ser facilmente manipulável e conseqüentemente não relevante para o usuário, afinal basta que um documento tenha a palavra-chave em questão citada uma única vez a mais que os adversários e o primeiro lugar nos resultados era quase garantido.

Diferente da situação anterior, Page e Brin perceberam que o critério de ranqueamento no sistema de apresentação de resultados poderia funcionar tal como ocorria no meio acadêmico em relação aos artigos científicos. Quanto mais citações e referências um artigo tem, e quanto mais importantes essas citações e referências forem, mais ele deve ser relevante (BIANCHINI; GORI; SCARSELLI, 2005). Esta é uma noção aproximada de funcionamento, já que se parte do pressuposto que diferente de uma página web que qualquer um pode criar a qualquer momento, em um artigo científico, há uma série de protocolos para que este seja publicado e considerado de fato um bom artigo.

Além disso, detalhes específicos do funcionamento do algoritmo do Google não são de domínio público ou estão disponíveis aos usuários do serviço.

Sabe-se que grande diferencial foi o algoritmo que recebeu o nome de *PageRank*. O algoritmo realiza um ranqueamento das páginas web de acordo com diferentes pesos atribuídos, verificados por meio de uma equação que envolve os seguintes critérios (SIQUEIRA, 2013):

- a) Probabilidade de um link/página ser clicado;
- b) *Damping Factor* (Fator de redução) que é a probabilidade de o visitante abandonar a navegação em decorrência da frustração com o conteúdo retornado pela busca;
- c) Quantidade de páginas conectadas na página de origem;
- d) Qualidade atribuída pelo próprio buscador as páginas conectadas na página de origem

A popularidade do Google, em grande parte se deve ao fato da utilização de um algoritmo que estuda os costumes e aproveitamento dos usuários em relação a suas buscas e interação com resultados apresentados, produzindo resultados personalizados de acordo com tais

medições (YUS, 2011). Ao julgar o que pode ou não ser potencialmente relevante para cada usuário.

O trabalho do buscador começa antes mesmo do usuário realizar a sua pesquisa. Os *crawlers* ou robôs de busca, vasculham constantemente a web, a fim de atualizar conteúdos já indexados e encontrar novas páginas. Quando o usuário digita ou realiza sua pesquisa por meio de comando de voz, automaticamente o buscador irá rastrear todo o conteúdo indexado e relacioná-lo com a pesquisa executada. Isso envolve análise dos termos utilizados, uso ou não de operadores avançados, geolocalização, entre outros.

Na sequência os algoritmos farão as combinações e os resultados retornarão ao usuário. Sendo ranqueados também de acordo com mais combinações e inferências realizadas pelo sistema, no que, como já dito anteriormente, é conhecido como relevância.

De acordo com as informações descritas pelo Google (2021b) em sua Central de Pesquisa, o processo de pesquisa no buscador funciona em três etapas:

1. **Rastreamento:** o Google faz pesquisas na Web com programas automatizados chamados rastreadores, buscando páginas novas ou atualizadas. O Google armazena os endereços dessas páginas (ou os URLs) em uma grande lista para analisar posteriormente. Encontramos páginas de muitos métodos diferentes, mas o principal é seguir links de páginas que já conhecemos.
2. **Indexação:** o Google acessa as páginas que rastreou e tenta analisar o conteúdo de cada uma. Analisamos o conteúdo, as imagens e os arquivos de vídeo da página, tentando identificar sobre o que ela trata. Essa informação fica registrada no índice do Google, um grande banco de dados armazenado em diversos computadores.
3. **Exibição dos resultados da pesquisa:** quando um usuário faz uma pesquisa, o Google tenta identificar os resultados mais qualificados. Os "melhores" resultados têm muitos fatores, incluindo localização, idioma, dispositivo (computador ou smartphone) e consultas anteriores do usuário. Por exemplo, em uma pesquisa por "oficinas de conserto de bicicletas", os resultados seriam diferentes para um usuário de Paris e outro de Hong Kong. O Google não aceita pagamento para atribuir aos sites uma classificação mais alta. A classificação é feita por algoritmos.

A etapa 1 de rastreamento é um processo cíclico e contínuo. Todos os dias novas fontes surgem na web e conteúdos já existentes são modificados, esse mapeamento constante é o que atualiza o índice do buscador e aprimora a conexão entre as pesquisas do usuário para com os resultados apresentados pelo buscador.

Na etapa 2 ocorre o processo de indexação. De acordo com Lancaster a indexação de assuntos implica numa das etapas necessárias para a “representação do conteúdo temático dos documentos” (LANCASTER, 2003, p. 6). O resultado desse processo é justamente um índice

que irá indicar os tópicos e os possíveis usos para um documento, além de registros de acesso para localização da informação. Exatamente como realizado pelo Google, que irá extrair das páginas rastreadas, informações acerca do seu conteúdo, a fim de representá-las em seu índice, possibilitando pontos de acesso para posterior recuperação.

Em muitos documentos, a indexação será dar por termos previamente escolhidos pelos autores e empregados em diversos campos como título, palavras-chave, resumos, entre outros. Apesar dos conceitos de indexação expressos por Lancaster (2003) e Cleveland e Cleveland (2013) estarem ligados a área de recuperação da informação dentro da Biblioteconomia e quando se trata de marketing de conteúdo o processo pelo qual a seleção de palavras-chave estarem ligados à área de Marketing Digital, algumas técnicas convergem e são relevantes, sendo este processo automático ou não e com foco na organização e representação do conhecimento ou otimização e chamariz para o aumento de visitas as páginas web e acesso a determinado conteúdo (ARAÚJO, 2018).

A etapa 3 é quando o buscador irá apresentar os resultados da pesquisa realizada pelo usuário. Os resultados serão dispostos e classificados de acordo com uma série de critérios algoritmos que ele julgar relevante para o usuário. Nesse sentido, sabe-se que “os motores de busca procuram compensar o excesso de itens recuperados com seus mecanismos internos de ordenação por relevância, mostrando em primeiro lugar os que, de acordo com seus critérios, deveriam ser os mais importantes” (CENDÓN, 2001, p. 46).

É necessário ter em mente que cada usuário poderá visualizar os resultados de uma forma diferente, mesmo que a pesquisa executada seja idêntica. Fatores como a geolocalização do usuário detectada pelo tipo de dispositivo utilizado no momento da pesquisa influenciam na forma como o algoritmo entende que cada resultado deve ser classificado. Se houver um histórico de busca do usuário, este também pode ser levado em consideração. São todas características que servem como *feedback* para a forma como veremos o resultado apresentado pelo buscador.

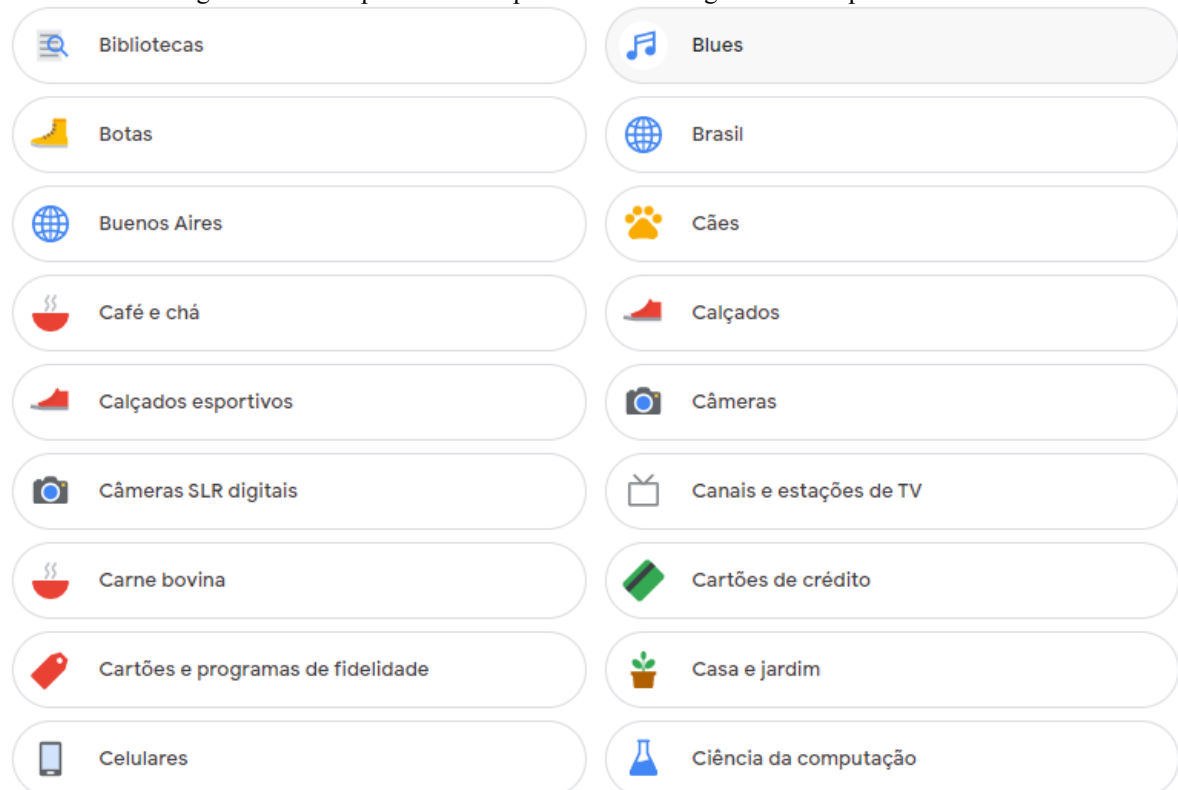
2.2.2.2 Anúncios nos resultados da pesquisa

Os resultados de pesquisa do Google, assim como de inúmeros outros buscadores, podem apresentar anúncios. Inicialmente, os anúncios pouco ou nada tinham a ver com a pesquisa executada pelo usuário, com aprimoramento e ao passar dos anos isso mudou e os anúncios passaram a ter relação ou com a pesquisa executada pelo usuário naquele momento ou

com o seu perfil e interesses já demonstrados anteriormente em plataformas integradas. Sendo essa uma das principais receitas das empresas gerenciadoras de buscadores.

No Google, ao acessar as configurações de conta do Gmail, na aba de dados e privacidade, o usuário consegue obter detalhes acerca da personalização dos anúncios, como por exemplo as categorias nas quais foi enquadrado.

Figura 27 - Exemplo de fatores presentes nas categorias de um perfil de busca



Fonte: Google (2022a)

Essas e outras categorias da Figura 27 contribuem para a personalização dos resultados de busca de um determinado usuário e são personalizadas de acordo com três fatores distintos: informações pessoais que o usuário adicionou a conta do Google; estimativas do buscador com base nas atividades em serviços da empresa como um todo (como a pesquisa no Youtube) e dados de anunciantes que possuem parceria com o buscador dos quais são baseadas nos rastros digitais deixados pelos usuários, resultantes dos famosos *cookies*¹ (ROSA, 2021; GOOGLE, 2022a).

Intituladas por Campanhas de pesquisa, palavras-chave específicas são marcadas automaticamente pelo sistema ou manualmente pelo responsável pela campanha (GOOGLE, 2018). Os anúncios aparecem sempre com a *tag* ‘Anúncio’ diferenciando-os dos resultados orgânicos da pesquisa.

Figura 28 - Posição dos anúncios nos primeiros resultados do Google

The figure consists of two screenshots of Google search results. The top screenshot shows a search for "provas anteriores concurso público auxiliar administrativo" with approximately 4,690,000 results. It features five sponsored ads for exam materials: "Apostila Prefeitura de PALHOÇA" (R\$ 25,20), "FASE-RS (Assistente ...)" (R\$ 144,00), "Apostila Prefeitura de PALHOÇA/SC" (R\$ 27,00), "Apostila CRBM 3 2022" (R\$ 29,40), and "Apostila Prefeitura de BALNEÁRIO PIÇARRAS/SC" (R\$ 40,00). The bottom screenshot shows a search for "provas de concurso para cargo de auxiliar administrativo" with approximately 7,000,000 results. It displays organic search results for "Provas de Auxiliar Administrativo - PCI Concursos" and two additional sponsored ads: "Apostila Prefeitura d..." (R\$ 36,00) and "Livro Série Provas &..." (R\$ 83,00).

Fonte: Dados da pesquisa – coletado dos participantes

A Figura 28 demonstra a diferenciação dos resultados. Estes podem se apresentar nos primeiros resultados em duas posições distintas: em cima ou ao lado dos resultados orgânicos.

O estudo de Rosa (2021) reforça a ideia de que maioria dos usuários entendem que os anúncios nos resultados de pesquisa são personalizados e apresentados com base em seus interesses, entretanto, possuem conhecimentos mínimos acerca da personalização desses interesses.

Desse modo, apesar de certa ciência na forma como esses anúncios aparecem e sendo os resultados orgânicos e pagos diferenciados pela plataforma entre si, para o usuário, que anseia por uma informação, resultado é resultado e essa diferenciação pode passar despercebida. Questão essa que envolve vários aspectos acerca de como a informação é organizada e apresentada ao usuário na web.

2.3 OPERADORES DE PESQUISA

Operadores de pesquisa são caracteres especiais que ao serem adicionados a uma expressão de busca, permitem criar uma maior complexidade a fim de obter resultados mais precisos. Em tese, um sistema de recuperação da informação pode transformar qualquer palavra ou símbolo em um operador de pesquisa de acordo com uma finalidade específica de busca dentro do seu banco de dados. Os operadores apresentados aqui são o AND, Aspas, filetype: e site: conforme o Quadro 2.

Quadro 2 - Descrição e função dos operadores avançados selecionados

OPERADOR	FUNÇÃO RESUMIDA
AND	Unir dois ou mais termos a fim de recuperar resultados que representem uma intersecção entre eles
Aspas “ ”	Blindar termos compostos ou frases completas a fim de recuperar resultados com todos os termos na ordem exata em que são indicados
filetype:	Indicar ao buscador o formato exato de arquivo que deseja
site:	Indicar ao buscador o site exato no qual você quer ele busque e apresente resultados

Fonte: Dados da pesquisa (2022)

Nas subseções a seguir há um aprofundamento acerca de como cada um destes operadores funciona.

2.3.1 AND

O AND, conforme já apresentado, é um operador booleano que corresponde à **intersecção** de dois ou mais termos.

Numa busca, serve para garantir que o sistema apresentará resultados que contenham obrigatoriamente todos os termos ou expressões ligadas por ele. O resultado também exige que

a ordem de todos os termos seja respeitada conforme foi descrita na expressão de busca. Ao empregar o operador AND em uma expressão de busca no Google, a URL do resultado discrimina seu uso

Exemplo: google **AND** bing **AND** duckduckgo.

No caso da expressão acima, os resultados retornarão documentos que citam os três termos descritos.

2.3.2 Aspas “ ”

As aspas “ ” podem ser utilizadas sempre que existir um termo composto, frase completa ou até mesmo um parágrafo todo de texto, desde que seja de interesse do usuário recuperar documentos que constem todos os termos exatamente na ordem nas quais foram informados.

Termos compostos como por exemplo: auxiliar administrativo; ciência da informação; ou técnico em assuntos educacionais. Podem ser acrescidos de aspas “ ” antes do primeiro e depois do último caractere da expressão para que um buscador entenda que a pesquisa requer que ele retorne resultados exatamente como as expressões pedem. “Auxiliar administrativo”; “ciência da informação”; ou “técnico em assuntos educacionais”.

O uso de aspas “ ” neste caso visa evitar que o sistema retorne resultados que não contenham algum dos termos contidos na expressão, resultados que invertam a ordem dos termos ou aqueles em que o sistema interpretou como sinônimos. O resultado compreenderá apenas aqueles que apresentarem a expressão completa e exatamente na ordem na qual foi apresentada.

Esse é o operador de pesquisa que objetiva alcançar maior precisão nos resultados e é indicado aos usuários que sabem exatamente os termos ou trechos que desejam encontrar nos documentos recuperados.

Enquanto o operador booleano AND encontra resultados pertencentes a intersecção de diferentes termos, as aspas concentram uma única expressão completa e ordenada. Além do que, podem ser utilizados numa única expressão se necessário.

Exemplo: **“auxiliar administrativo” AND “assistente comercial”**

A expressão acima retornaria resultados que sejam citados ambos os cargos por causa do uso do AND ao mesmo tempo que mantém a ordem dos termos compostos, não permitindo que sejam retornados resultados como ‘auxiliar comercial’, ‘assistente administrativo’ ou somente ‘auxiliar’ por exemplo.

2.3.3 filetype:

O filetype: é um operador utilizado para pesquisar documentos que estejam em um formato de arquivo específico. Por exemplo: PDF, DOC, SRT, XLS, MP3, PNG, entre inúmeros outros.

É útil para evitar que documentos com uma extensão de arquivo diferente da requerida pelo usuário apareça entre os resultados. Se ele estiver buscando por provas de concurso em busca de arquivos editáveis, pode utilizar a extensão DOC e posteriormente editar o arquivo no Word que é o software aplicativo responsável pela extensão indicada.

Pode ser utilizado antes ou depois da expressão de busca, entretanto, pode haver variação na quantidade e apresentação de resultados.

Exemplo 1: **modelos de currículo filetype:doc**

Exemplo 2: **filetype:doc modelo de currículo**

Tematicamente, os resultados em ambas as expressões são os mesmos. O buscador retornará modelos de currículo editáveis em Word já que o formato indicado para a extensão de arquivo requerida foi DOC.

2.3.4 site:

O operador site: é utilizado para direcionar a pesquisa realizada dentro da caixa de pesquisa para apenas um site em específico. Todos os resultados apresentados na página do buscador virão obrigatoriamente de dentro do site indicado.

É útil quando a navegação dentro do próprio site em questão apresenta muitas opções de caminhos a seguir ou a arquitetura da informação dele não é tão boa quanto deveria para o usuário encontrar a informação que precisa.

Ao empregar o uso desse operador na página inicial do Google, o resultado é similar a ação de pesquisar dentro da área de busca de sites que não possuem um mecanismo de pesquisa próprio e incorporam os efeitos desse buscador na pesquisa interna. Aos menos atentos, tem se uma impressão de que o usuário está buscando num sistema próprio quando na verdade se trata de uma busca melhorada pelo Google.

A Figura 29 representa uma pesquisa por meio de uma funcionalidade personalizável nos sites que leva o nome de Mecanismo de pesquisa programável (GOOGLE, 2022b). Gratuita se permitir a exibição de anúncios.

Figura 29 - Pesquisa realizada na caixa de pesquisa do site da UFSC

concurso público para técnicos

Aproximadamente 22,300 resultados (0,18 segundos)

Edital nº 001/2022/DDP – Concurso Público para carreira Técnico ...
 001ddp2022.concursos.ufsc.br > edital
 Contatos. CAC Coordenadoria de Admissões, **Concursos Públicos** e Contratação Temporária E-mail: concurso.ddp@contato.ufsc.br.

Edital nº 001/2022/DDP – Concurso Público para carreira Técnico ...
 001ddp2022.concursos.ufsc.br
 A Universidade Federal de Santa Catarina (UFSC) torna público o **concurso público ... e seis** vagas para a carreira **técnico** administrativo em educação.

Portal da Coordenadoria de Admissões, Concursos Públicos e ...
 concursos.ufsc.br
 Sorteio público de reserva de vagas para o próximo **concurso público** da ... da carreira **Técnico-Administrativo em Educação (TAE)** para os Campi da UFSC, ...

Edital nº 136/2018/DDP – Concurso Público da UFSC (TAE)

UFSC

UFSC abre **concurso público para Técnico-Administrativo em ...**

Notícias da UFSC > 2022/03 > ..._trashed-3
 03/03/2022 ... A Universidade Federal de Santa Catarina (UFSC) informa a abertura de **concurso público** para selecionar candidatos para provimento em ...

Coperve abre cadastramento para compor equipes de ...
 Notícias da UFSC > 2022/03 > coperve-abre-cadastramento-para-compor-equ...
 11/03/2022 ... Tags: **concurso público** copervedocentesestudantesSeleção de colaboradores servidores **técnicos**-administrativosUFSCUniversidade Federal de Santa ...

Abertura de concurso público para a carreira Técnico-Administrativo ...
 concursos.ufsc.br > 2022/03/03 > abertura-de-concurso-publico-para-a-car...
 03/03/2022 ... O Departamento de Desenvolvimento de Pessoas (DDP) informa a abertura de **concurso público** para selecionar candidatos para provimento em ...

Prazo de Inscrição no concurso de técnico administrativo da UFSC ...
 Notícias da UFSC > 2022/03 > prazo-de-inscricao-no-concurso-de-tecnico-ad...
 há 8 dias ... Termina no dia 6 de abril, às 23h59min, o prazo para inscrição no **concurso público** para selecionar candidatos em diversos cargos da carreira ...

Portal da Coordenadoria de Admissões, Concursos Públicos e ...
 concursos.ufsc.br > editais-tecnico-administrativo-em-educacao
 Editais de **Técnico-Administrativo em Educação (TAE)**. Editais destinados a seleção de candidatos para o provimento dos cargos da carreira de TAE para o quadro ...

Portal da Coordenadoria de Admissões, Concursos Públicos e ...
 concursos.ufsc.br > editais-tecnico-administrativo-em-educacao
 Editais de **Técnico-Administrativo em Educação (TAE)**. Editais destinados a selecionar candidatos para provimento dos cargos da carreira de TAE para o quadro ...

Portal da Coordenadoria de Admissões, Concursos Públicos e ...
 concursos.ufsc.br > 2022/02/22 > sortelo-publico-de-reserva-de-vagas-para...
 Sorteio público de reserva de vagas para o próximo **concurso público** da carreira **Técnico-Administrativa em Educação (TAE)**. 22/02/2022 08:17.

Edital nº 136/2018/DDP – Concurso Público da UFSC (TAE)
 1362018ddp.concursos.ufsc.br
 A Universidade Federal de Santa Catarina (UFSC) torna público o **concurso público ... e seis** vagas para a carreira **técnico**-administrativo em educação.

UFSC

Pesquisar **concurso público para tè...** no Google

MELHORADO PELO Google

© 2022 Google

Fonte: UFSC (2022)

Por ser personalizável, além de questões relacionadas a design e publicidade on-line, o responsável pelo site decide se a pesquisa realizada dentro do site retornará resultados apenas do próprio site ou dele e de outros.

2.4 BUSCADORES COMO MÍDIA DO CONHECIMENTO

Mídias do conhecimento, segundo Sousa (2019, p. 10) “reúnem ou alteram informações e produzem conhecimento, na percepção de usuários humanos”. Enquanto um sistema mediador do conhecimento, os buscadores, apesar do alto poder dos algoritmos de impressionar na execução de inúmeras funções, não são autoconscientes, e, assim como qualquer outro ambiente digital, entram numa propriedade dita citada por Murray (2003) como procedimentais, ou seja, possuem uma grande capacidade de executar uma série de regras pré-estipuladas.

Os buscadores constituem o mesmo princípio de armazenar, organizar e disponibilizar informações, no entanto, se diferenciam de mídias tradicionais analógicas por meio de cinco princípios específicos das mídias digitais, elencados por Manovich (2001) como: representação numérica; modularidade; automatização; variabilidade e transcodificação.

O princípio de representação numérica diz respeito a mídias digitais compostas de bits e passíveis de manipulação por algoritmos. Um buscador consiste em um bom exemplo de mídia digital a medida que por ser descrito matematicamente e possui algoritmos que comandam seu funcionamento, tornando-o uma mídia programável.

O segundo princípio diz respeito a “estrutura fractal das novas mídias” (MANOVICH, 2001) intitulado por modularidade. Um buscador é constituído por inúmeros objetos de mídia, como código fonte, rastreadores e indexadores, além de armazenarem em seus índices documentos, fotografias, páginas web, vídeos, gifs, códigos, entre outros inúmeros elementos. As mídias digitais são constituídas de partes integradas de elementos menores que possuem singularidade quando separados, e podem ser combinados e recombinaados para criar algo maior se necessário.

A automação envolve tarefas que podem ser executadas automaticamente por essas mídias, e no caso dos buscadores, compreende exemplos tanto de baixo como de alto nível. A correção de termos escritos com erros ortográficos em caixas de pesquisa é um exemplo de baixo nível de automação da mídia. Enquanto o trabalho semântico da tecnologia de um buscador para tentar compreender em qual contexto está a pesquisa que um usuário deseja, configura alto nível de automação.

Variabilidade é o item que enquadra toda e qualquer atualização de uma mídia. Nada está definitivamente encerrado e o surgimento de novas versões é tido como algo potencialmente infinito. Os buscadores sofrem constante atualização, sendo que na década de 90 esses saltos tecnológicos eram maiores e a variabilidade desse tipo de mídia ficava evidente à medida que esse tipo de mídia se consolidava como essencial no mundo web. Atualmente os aprimoramentos, mesmo que extremamente significativos, parecem ser mais sutis. Manovich (2001, p. 57, tradução nossa) diz que “várias interfaces diferentes podem ser criadas para os mesmos dados”, o que é visível em buscadores como o Ecosia que utiliza o motor de buscas do Bing ou metabuscadores como o metaGer que utiliza o índice de motores de busca do Yahoo, Bing e Scopia e Infotiger.

Por último, e mais importante para a presente pesquisa, há o princípio de transcodificação, que de maneira sucinta significa traduzir algo para outro formato. Dentre os cinco, este pode ser definido como sendo “a consequência mais substancial da informatização da mídia” (CORREA; FARIA; FADEL; VAZ, 2022, p. 92).

Em uma primeira instância se discute o quanto a camada computacional afeta a camada cultural. Para buscadores e especificamente se tratando de operadores de pesquisa, a forma como estes se apresentam na busca e funcionam dentro de uma expressão de busca tendem a afetar a forma como os usuários entendem, constroem e executam uma pesquisa on-line.

Conceitua-se aqui camada cultural e camada computacional como características distintas, mas interdependentes. Como camada cultural entende-se aquela com características reconhecíveis e próprias da cultura humana. Já a chamada camada computacional é representada pelos computadores e seus dados operáveis, em suma, os elementos que são próprios do universo computacional (MANOVICH, 2001; CORREA; FARIA; FADEL; VAZ, 2022).

Manovich (2001) também deixou claro em sua obra, que o que ele chama por camada do computador não é algo fixo e sim mutável. Mais de 20 anos depois isso se confirma quando computadores fazem parte cada vez mais atividades no cotidiano, entre elas, a pesquisa, que se aperfeiçoou por meio de buscadores sofisticados devido ao avanço tecnológicos tanto em hardware e principalmente em software, além é claro, da internet.

Assim como a camada do computador influencia a camada cultural, o inverso também ocorre. As diferentes necessidades dos usuários em recuperar documentos na web exigiram que a camada computacional se adaptasse e fornecesse a possibilidade de que o usuário realizasse uma pesquisa avançada sem grandes dificuldades tecnológicas, por meio de uma linguagem

próxima da utilizada em termos e demais palavras-chaves presentes em expressões de busca. Daí nascem os operadores avançados de pesquisa.

Tais influências, do ponto de vista do usuário para o computador, somente são possíveis devido a mais uma propriedade de ambientes digitais citada por Murray (2003), o fato deles serem participativos. Um buscador reage a informações que inseridas, seja por texto ou voz, cada bit de informação introduzido em sua caixa de pesquisa gera um comportamento diferente por parte do sistema.

Ademais, ainda existem outras duas propriedades elencadas por Murray (2003) em ambientes digitais. Eles são espaciais, ou seja, em um buscador há a possibilidade de navegar pelo espaço de diferentes maneiras e de forma nem sempre linear. O usuário pode avançar em um resultado, voltar, partir dele para outro e mais inúmeras outras possibilidades. Mesmo que por vezes nem tudo esteja ao alcance do usuário em uma busca, e enfrente problemas como dificuldade em encontrar o que precisa, estratégias de pesquisa ineficazes, informações falsas e grandes quantidades de documentos para vasculhar, ainda assim um buscador com uma caixa de pesquisa em branco esperando para ser preenchida faz com que todos os recursos do mundo pareçam “acessíveis recuperáveis e imediatos” (MURRAY, 2003).

Por fim, ambientes digitais também são enciclopédicos e segundo Murray (2003), induzem a uma expectativa enciclopédia. Com tanta informação reunida para ser acessada em um só local por meio de um buscador, a ideia deste como uma grande biblioteca global se apresenta como algo extremamente razoável

2.5 SÍNTESE

A habilidade de questionar define a humanidade e permitiu gerar conhecimento e desenvolvimentos tecnológicos, sociais, morais, políticos, sendo estes mesmos questionáveis. O conhecimento gerado é passado entre gerações através de um sistema de armazenamento e catalogação que foi transformado pela mídia digital em várias de suas dimensões como em grau, quantidade, distribuição, produção e acesso. A distribuição, por exemplo, seguiu uma arquitetura de rede, no formato da World Wide Web.

O acesso a este conhecimento na Web, é mediado por buscadores que simplificam a participação do questionador, para uma participação reduzida. Não é preciso mais uma pergunta, mas apenas uma palavra. Esta redução implica numa sugestão de subordinação.

Percebe-se que os buscadores evoluíram muito mais a nível de algoritmo quanto a nível de interface, e se estabeleceram como a ‘primeira página’ da Web. Assim, se apresenta para o usuário como a “cara” da Internet, com a qual pode ter acesso ao seu conteúdo. Por isso, Gandal (2001), ainda no início do século XXI, entendia que os buscadores eram peça central para organizar a riqueza de informações presente na web. Duas décadas depois, são praticamente primordiais nessa tarefa, e de fato uma mídia indispensável no campo da pesquisa on-line.

3 OPÇÕES METODOLÓGICAS

Visando atingir os objetivos propostos e obter uma resposta ao problema de pesquisa apresentado, nesta seção são descritas as escolhas metodológicas adotadas ao longo da pesquisa. Primeiramente em relação a caracterização do estudo e num segundo momento a descrição detalhada dos materiais e métodos da pesquisa de campo empregada.

3.1 CARACTERIZAÇÃO DA PESQUISA

O estudo proposto procura avançar o conhecimento em relação à pesquisa on-line em buscadores, e qualifica-se como uma pesquisa com abordagem qualitativa com objetivo descritivo.

Procurando um melhor detalhamento da classificação dessa pesquisa, o Quadro 3 apresenta o posicionamento metodológico do estudo segundo as definições elencadas nas obras de Freire (2013) e Prodanov e Freitas (2013).

Quadro 3 - Classificação da pesquisa

Abordagem	Finalidade	Ângulo do Problema	Objetivos	Procedimentos
Qualitativa	Aplicada	Técnico	Descritiva	Pesquisa experimental
		Cultural		Pesquisa bibliográfica

Fonte: Adaptado de Freire (2013); Prodanov; Freitas (2013)

Com uma abordagem qualitativa, trata-se de uma pesquisa com finalidade aplicada à medida que o conhecimento gerado irá possibilitar uma aplicação prática dos processos. Na ocasião, o problema levantado foi com base em aspectos técnicos e sociais.

Reforçado por Gil (2008), indica-se como de natureza qualitativa e possui caráter descritivo, ao preocupar-se com uma atuação prática das características apresentadas após análise do objeto.

Os procedimentos se encaixam como uma pesquisa bibliográfica para a construção do referencial teórico e como uma pesquisa experimental em sua aplicação pois, foram provocadas alterações no ambiente digital do buscador, de forma sistemática, para que fosse possível observar se cada intervenção realizada com o uso dos operadores de pesquisa funcionasse ou produzisse algo nos resultados de pesquisa. Wazlawick (2021) diz que em alguns casos é possível

tornar a pesquisa experimental um pouco mais flexível, tratando-se do que o autor denomina como quase experimento.

Para tanto, não houve um grupo de controle e sim uma situação de controle representada pela busca normalizada (cenário dois) e outros três cenários (cenário três, quatro e cinco) com diferentes intervenções a fim de observar alterações a partir delas.

As variáveis independentes neste estudo foram os operadores de pesquisa AND, aspas, site: e filetype: que pretendiam influenciar a variável dependente representada pelos resultados visualizados a partir das capturas de tela. Marconi e Lakatos (2017) conceituam as variáveis dependentes como fatores que aparecem, desaparecem ou variam à medida que as variáveis independentes são introduzidas ou modificadas, portanto, anúncios, sugestões de pesquisas já realizadas e quantidade aproximada de resultados englobam as características dos resultados enquanto variável dependente.

Este trabalho se insere dentro do paradigma funcionalista de Morgan (1980) com base numa visão de mundo que considera que as características do objeto a serem levantadas possuem significados específicos e funcionam de forma concreta perante a sociedade.

Em relação aos objetivos específicos, o objetivo a) que se refere a descrever o processo lógico de uma pesquisa com as relações dos primórdios filosóficos que influenciaram o ato de pesquisar até a sua dinâmica atual on-line em buscadores, compôs a seção dois do trabalho.

O objetivo específico b) selecionou os operadores de pesquisa presentes num buscador que auxiliam na recuperação de documentos em pesquisas on-line. Ao identificá-los, eles formam o Quadro 2 com a descrição das funções de cada um deles.

Na sequência, o resultado do objetivo específico c) integrou a revisão de literatura que buscou identificar os princípios fundamentais que compõem a construção de um buscador

Por fim, o objetivo d) foi resolvido com os procedimentos de uma pesquisa experimental. Esta pesquisa foi criada para coletar dados junto aos usuários sobre o uso dos operadores definidos no objetivo b). Também confronta os aspectos de funcionamento do objetivo c), a fim de levantar dados da ação de busca sem e com a presença de diferentes operadores por uma amostra contemplada com diferentes usuários.

Com intuito de apresentar uma visão panorâmica do projeto, o Quadro 4 traz a matriz de consistência desenhada desde o início da dissertação e consolidada ao longo do trabalho.

Quadro 4 - Matriz de consistência do projeto de pesquisa do Mestrado

Operadores de pesquisa em buscadores: testando técnicas avançadas para a recuperação da informação durante pesquisas por documentos na web				
PROBLEMA DE PESQUISA	OBJETIVOS	OPÇÕES METODOLÓGICAS	VARIÁVEIS	RESULTADOS ENTREGÁVEIS
Quais características nos resultados de pesquisa se destacam a partir do uso de operadores avançados na busca por documentos na web?	Objetivo geral	Pesquisa bibliográfica: Revisão narrativa	Variáveis independentes - booleano AND - aspas - filetype: - site:	A dissertação completa + Dados de pesquisa disponibilizados no Zenodo
	Examinar as características resultantes do uso de operadores avançados de busca na realização de pesquisas em um buscador.			
PRESSUPOSTOS	Objetivos específicos	Pesquisa bibliográfica: Revisão narrativa	Variáveis dependentes Resultados do experimento e análise dos anúncios, sugestões de pesquisas já realizadas e quantidade aproximada de resultados	
<ul style="list-style-type: none"> • A utilização desses operadores pode apresentar características mais uniformes aos resultados • A utilização desses operadores pode contribuir para uma emancipação do usuário durante uma pesquisa on-line, dando a ele mais controle e direcionamento dos resultados apresentados. 	a) Descrever o processo lógico de uma pesquisa	Análise dos operadores disponibilizados pelo buscador Google.		
	b) Selecionar os operadores de pesquisa presentes num buscador que auxiliam na recuperação de documentos em pesquisas on-line			
	c) Identificar quais são os princípios fundamentais no funcionamento de um buscador	Pesquisa bibliográfica: Revisão narrativa		
	d) Testar a partir da estruturação de dados, o uso dos operadores de pesquisa por diferentes usuários	Pesquisa experimental: Realização do experimento a partir dos dados coletados		
e) Categorizar as resultantes do uso de operadores avançados de busca online por diferentes usuários, de acordo com suas características.				

Fonte: Dados da pesquisa (2022)

A matriz de consistência é um instrumento que reúne os elementos mais significativos do processo de pesquisa em questão, a fim de torná-lo mais compreensível em relação as conexões lógicas entre problema de pesquisa, objetivos, hipótese (no caso aqui, pressuposto), variáveis e metodologias empregadas (VERA-PÉREZ; LUGO-ORTIZ, 2016).

3.2 PESQUISA EXPERIMENTAL

Os tópicos a seguir contemplam o passo a passo da realização da pesquisa experimental a partir da coleta de dados por meio de questionário.

3.2.1 Participantes

Os participantes da pesquisa de campo foram convidados por meio de divulgação do link de acesso ao questionário em grupos de pesquisa e mídias sociais (Instagram e Facebook) do pesquisador.

Não foram impostos critérios relacionados ao perfil pessoal, como idade ou gênero, para qualificar os usuários como aptos a participarem da pesquisa. Os únicos pré-requisitos foram que o respondente tivesse acesso a um dispositivo com acesso à internet (smartphone, tablet, notebook ou computador de mesa) e possuísse uma conta no Gmail que estivesse ativa no momento das buscas e fosse informada para que o questionário fosse respondido apenas uma vez.

3.2.2 Procedimento

Os operadores escolhidos para integrar os cenários presentes no questionário durante a coleta de dados foram o operador booleano **AND**, as aspas “ ”, o operador **filetype:** e o operador **site:**. A seleção destes operadores obedeceu a adequação para complementar a expressão de busca criada a partir da narrativa de pesquisa e com base na experiência do autor enquanto bibliotecário. Como a busca visava encontrar documentos, o foco dos operadores foi auxiliar na união das palavras-chave (**AND** e aspas “ ”), na definição do formato de arquivo (**filetype:**) e na fonte de informação (**site:**).

Foram desenvolvidos cinco cenários distintos para testar as buscas, sendo que todos eles tratavam do mesmo tema: a busca por provas de concurso público para o cargo de auxiliar administrativo. A escolha ocorreu por conta da necessidade de criar uma busca fictícia que tivesse por objetivo a recuperação de algum tipo de documento na web, neste caso, modelos de provas para um concurso público. O cargo de auxiliar administrativo foi definido, por ser

geralmente um cargo de nível médio, ampliando as possibilidades de resultados se comparado a outros cargos mais específicos de nível superior.

A seguir são apresentados cada um dos cenários desenvolvidos e seus objetivos.

O cenário um (ver Figura 30 e Figura 31) consistia em apresentar a temática ao respondente e instruí-lo a se imaginar na situação de uma pessoa interessada em encontrar provas de concurso públicos já realizados, para o cargo de auxiliar administrativo.

Figura 30 - Cenário um do questionário de coleta de dados

Leia abaixo o cenário 1 e siga as instruções

Imagine a seguinte situação: Você quer se preparar para um concurso público com foco em conseguir uma vaga para o cargo de auxiliar administrativo.

Passo 1 - Abra uma nova guia do navegador e procure no Google alguns exemplos de provas.

Passo 2 - Tire um print da página inicial da pesquisa de forma que a caixa de busca apareça e volte a responder esse questionário.

*não role a página com o objetivo de excluir propagandas do print caso apareçam.

Fonte: Dados da pesquisa (2022)

Nesse momento, o objetivo era coletar quais palavras-chave o usuário considerava relevantes e essenciais para realizar a busca diante do cenário apresentado. Após a busca era solicitado que o usuário inserisse a expressão de busca executada em sua página de buscas do Google e realizasse o upload da tela com os resultados apresentados ao executá-la.

Figura 31 - Campo de resposta do cenário um

Insira o texto que você utilizou na caixa de pesquisa do Google *

exemplos de prova de auxiliar adminitrativo

Insira o print de tela relacionado aos passos 1 e 2 *

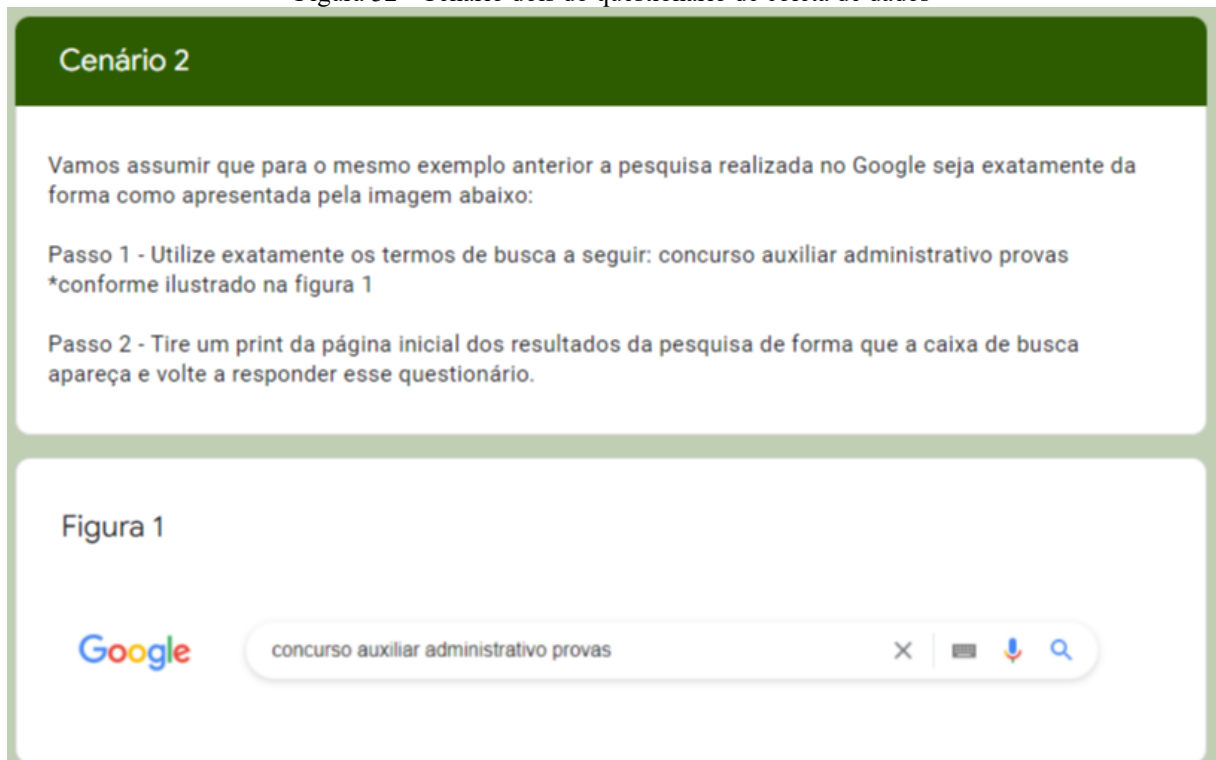
Captura de Tela ...

Fonte: Dados da pesquisa (2022)

Em todos os cenários foi orientado que o usuário tirasse uma captura de tela sem considerar o aparecimento de anúncios, não sendo, portanto, necessário que ele rolasse a página de forma que só aparecessem resultados orgânicos em relação a busca realizada.

No cenário dois da Figura 32, ao invés de deixar a busca livre para que o usuário definisse as palavras-chave e a formulação das expressões de busca de acordo com seus próprios critérios, foi fornecida uma expressão única consolidada, para que ele copiasse, colasse em seu buscador e executasse a pesquisa. A prática uniformizou a busca para todos os participantes e teve por objetivos comparar os primeiros resultados de cada um deles.

Figura 32 - Cenário dois do questionário de coleta de dados



Fonte: Dados da pesquisa (2022)

A partir do cenário dois apenas a captura de tela era requerida ao participante e a forma como a pesquisa deveria ser executada vinha acompanhada de um exemplo visual de como a caixa de pesquisa do Google deveria estar preenchida quando ele a realizasse.

O cenário três (ver Figura 33) introduzia o uso de operadores ao determinar que o usuário utilizasse o operador booleano AND na expressão de busca apresentada.

Figura 33 - Cenário três do questionário de coleta de dados

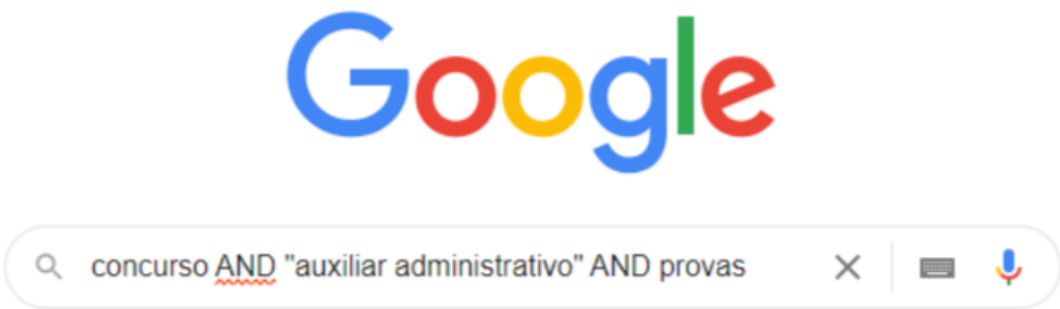
Cenário 3

Os próximos cenários seguirão a mesma lógica da questão anterior:

Passo 1 - Utilize exatamente os termos de busca a seguir: concurso AND "auxiliar administrativo" AND provas
*conforme ilustrado na figura 2 abaixo

Passo 2 - Tire um print da página inicial da pesquisa de forma que a caixa de busca apareça e volte a responder esse questionário.

Figura 2



The image shows the Google search interface. The search bar contains the text 'concurso AND "auxiliar administrativo" AND provas'. The word 'AND' is underlined in red. To the right of the search bar are icons for clearing the search, voice search, and image search.

Fonte: Dados da pesquisa (2022)

Mesmo que toda busca no Google, de certa forma aplique o operador booleano AND em sua execução, a ideia aqui era testá-lo de forma mandatória. Visando garantir somente resultados que obrigatoriamente apresentem as três expressões (concurso; auxiliar administrativo; provas) juntas.

Além do uso do operador booleano AND, nesse cenário, outro operador também foi empregado de forma combinada. As aspas “ ”.

Neste caso, as aspas “ ” funcionam como compilador de um termo composto, forçando ao sistema a recuperação de somente resultados com o termo “auxiliar administrativo” de forma completa e exatamente na ordem apresentada. Descartando resultados que apresentem somente os termos ‘auxiliar’, ‘administrativo’ ou juntos, porém, na ordem oposta, como por exemplo: ‘administrativo auxiliar’.

No cenário quatro (ver Figura 34) era requerido que os participantes utilizassem o operador filetype:pdf para que, ao aplicar a mesma pesquisa proposta pelos cenários dois e três,

fossem restringidos os resultados para somente aqueles que apresentassem documentos de extensão no formato PDF (Portable Document Format).

Figura 34 - Cenário quatro do questionário de coleta de dados

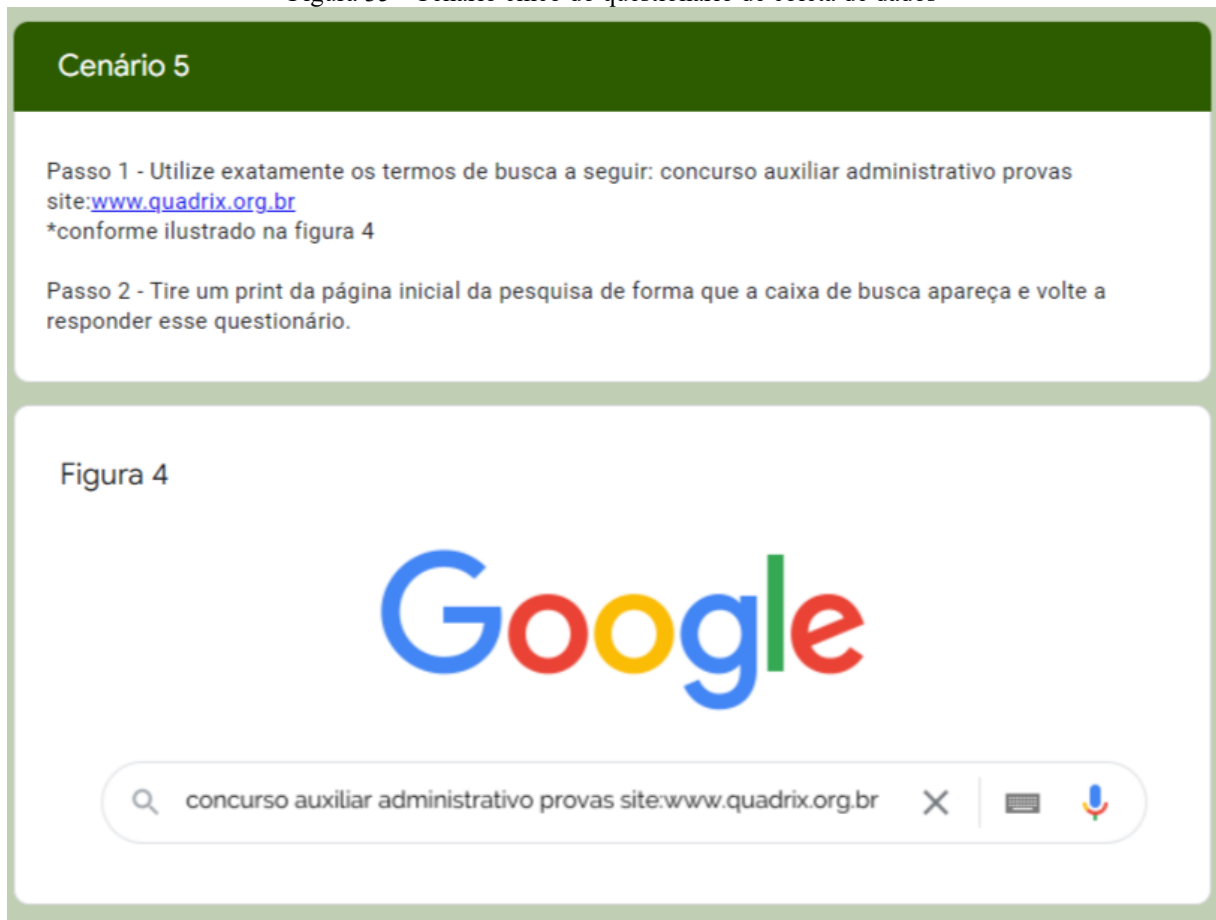


Fonte: Dados da pesquisa (2022)

Ao utilizar o operador de formato filetype: e restringir resultados apenas em PDF, esperava-se uma maior uniformização dos resultados nos diferentes perfis dos usuários participantes do questionário.

Por último, o cenário de número cinco (ver Figura 35), introduziu o operador site: para que a mesma pesquisa reproduzida nos três cenários anteriores estivesse direcionada para resultados que viessem de um único site.

Figura 35 - Cenário cinco do questionário de coleta de dados



Fonte: Dados da pesquisa (2022)

A escolha do site utilizado ocorreu ao levar em consideração a preferência por uma banca de concursos que atuasse com provas para cargos de nível médio, em que o cargo de auxiliar administrativo fosse algo existente. A banca selecionada para o exemplo da expressão de busca com o uso do operador foi o Instituto Quadrix. A organização sem fins lucrativos desenvolve e aplica provas para Conselhos Federais, Conselhos Regionais, Hospitais e Fundações, nos quais o cargo é recorrente, além de estar presente com atuação em todo Brasil, trazendo a abrangência nacional e variedade que o exemplo necessitava (INSTITUTO QUADRIX, 2022).

O cenário cinco apresentou o maior nível de especificidade na aplicação de operadores de pesquisa, restringindo mais ainda os resultados apresentados aos diferentes usuários participantes do questionário.

3.2.3 Coleta de dados

A construção do questionário levou em consideração a necessidade em coletar dados do perfil do respondente (idade e gênero); dados de geolocalização do dispositivo no momento de resposta do questionário (país e estado); cenários de aplicação das expressões de busca; e dados do dispositivo utilizado para responder o questionário (navegador, uso de bloqueadores de anúncios e tipo de dispositivo).

Além dos cenários de busca descritos, as perguntas relacionadas ao perfil e dispositivos dos participantes foram: qual sua idade? Qual seu gênero? De qual país você está respondendo esse questionário? De qual localidade do Brasil você está respondendo esse questionário? Qual navegador você utilizou para realizar as pesquisas indicadas? Você utilizou algum bloqueador de anúncio em seu navegador?; e por último: Você respondeu esse questionário de qual tipo de dispositivo?

O Google Forms foi a plataforma utilizada para coletar os dados. Isto porque é a ferramenta web do próprio Google para gerenciamento de pesquisas a partir do uso de questionários e formulários de registro. Junto a isso, a facilidade do convênio gerado entre o Google e a Universidade Federal de Santa Catarina (UFSC) no segundo semestre de 2020 que permitiu a criação e uso de uma conta institucional com espaço de armazenamento no drive, visto a necessidade de alocar na nuvem as capturas de tela recebidas durante a realização da coleta de dados

O questionário foi aplicado no segundo semestre de 2022 e obteve um total de 76 respostas registradas durante o período. Cada respondente enviou cinco arquivos de imagem resultando num total de 380 capturas de tela coletadas. Destes, 16 não puderam ser analisadas. A exclusão se deu por problemas como: captura da tela não correspondente ao solicitado, expressão de busca formulada incorretamente ou pesquisa realizada em outro buscador diferente do solicitado. Assim sendo, o corpus da pesquisa analisado ficou em 364 capturas de tela.

A Tabela 1 resume a quantidade de capturas de tela que foram excluídas e quantos compuseram o total de cada cenário de busca descrito.

Tabela 1 - Conjunto de dados pré e pós seleção

CENÁRIO DE BUSCA	COLETADOS	EXCLUÍDOS	ANALISADOS
Busca livre	76	00	76
Busca normalizada	76	04	72
Operador booleano AND	76	05	71
Operador filetype:	76	04	72
Operador site:	76	03	73
TOTAL	380	16	364

Fonte: Dados da pesquisa (2022)

A pesquisa de campo foi aprovada pelo Comitê de Ética e Pesquisa com Seres Humanos (CEP) da Universidade Federal de Santa Catarina (UFSC) e aprovada sob o registro CAAE: 61951822.9.0000.0121.

4 ANÁLISE DOS RESULTADOS

Esta seção apresenta e discute os resultados ao examinar o uso de operadores na realização de pesquisas em um buscador. Descreve-se as características nos resultados de pesquisa que se destacaram a partir do uso de operadores avançados na busca por documentos na web.

A discussão considera as diferenças e semelhanças dos resultados obtidos entre os cinco cenários. Estes cenários, dentro da proposta da dissertação podem ser considerados como categorias de análise dos resultados.

4.1 CONTRIBUIÇÕES PARA A CIÊNCIA ABERTA

Para Lecardelli (2020, p. 17) “compartilhar informações científicas é primordial para potencializar os avanços da ciência, poupar de retrabalho, possibilitar reproduzir pesquisas para validação e permitir o reuso de informações por novos interessados para novas descobertas”, portanto, se tem como compromisso neste trabalho, contribuir para a ciência aberta, permitindo a abertura aos dados científicos gerados no desenvolvimento da pesquisa.

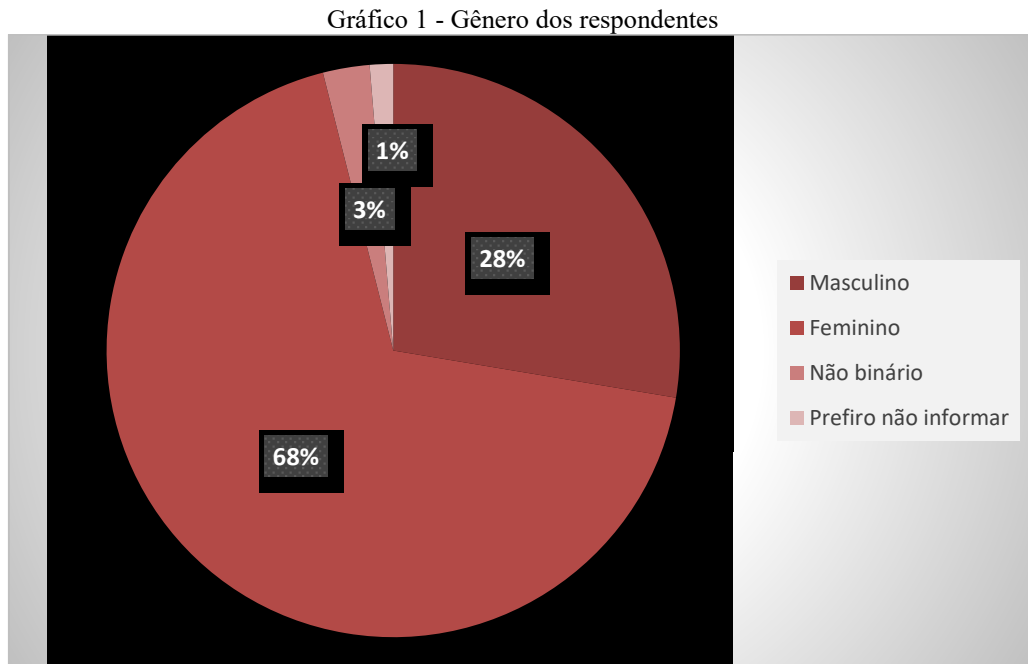
Em conformidade com a Lei Geral de Proteção de Dados (LGPD) em relação ao tratamento de dados em território nacional. (BRASIL, 2018) todas as 380 capturas de tela foram anonimizadas, fotos de perfil, nomes de usuário e outras guias abertas no navegador foram anuladas. Além disso, os arquivos de imagem tiveram seus metadados relacionados a propriedade e informações pessoais foram removidos.

As planilhas criadas a partir das capturas de tela coletadas foram disponibilizadas no repositório de dados científicos abertos Zenodo mantido pela European Organization for Nuclear Research (CERN). A escolha se deu por conta da possibilidade de depósito gratuito dos dados por parte do pesquisador acrescida da atribuição do identificador persistente Digital Object Identifier (DOI) que concede singularidade e permanência a objetos digitais na internet, no caso desta pesquisa, ao conjunto de dados trabalhado.

Dito isso, os dados abertos desta pesquisa encontram-se disponíveis nesse link: [Conjunto de dados da dissertação](#).

4.2 DADOS DEMOGRÁFICOS

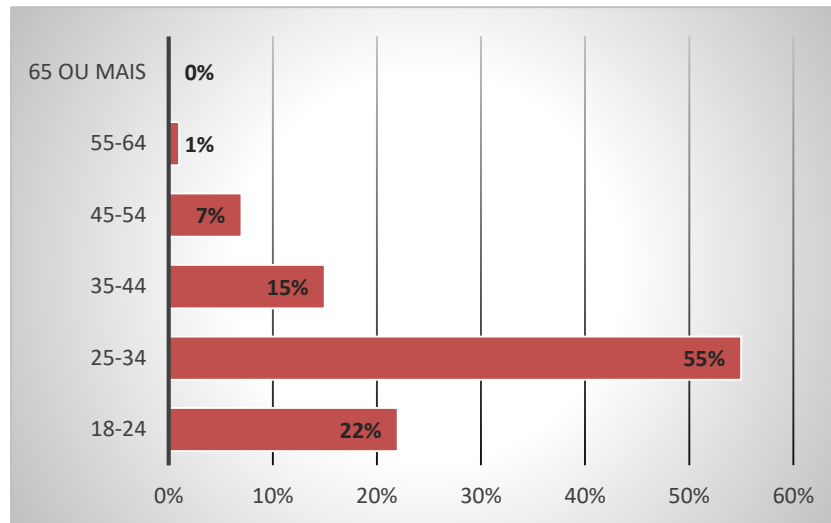
Acerca do perfil, em relação ao gênero, dos 76 participantes 68% declarou pertencer ao gênero feminino, 28% masculino, 3% se declararam não binário e 1% preferiu não informar (ver Gráfico 1).



Fonte: Dados da pesquisa (2022)

O total de participantes foi dividido em um grupo de 6 faixas etárias diferentes, reunindo-os em grupos de 18-24, 25-34, 35-44, 45-54, 55-64 e por último 65 ou mais. A segmentação foi baseada nos grupos dispostos pelo Google (2022a) de acordo com informações do Google Ads Editor, responsável pelos direcionamentos e classificações em anúncios. Sendo assim, a maior parte dos participantes se encontra entre 25 e 34 anos, representando 55% do total. Na sequência o grupo de 18 a 24 anos (22%), 35 a 44 anos (15%), 45 a 54 (7%), 55 a 64 (1%) e nenhum respondente com 65 anos ou mais de idade (ver Gráfico 2).

Gráfico 2 - Faixa etária dos participantes

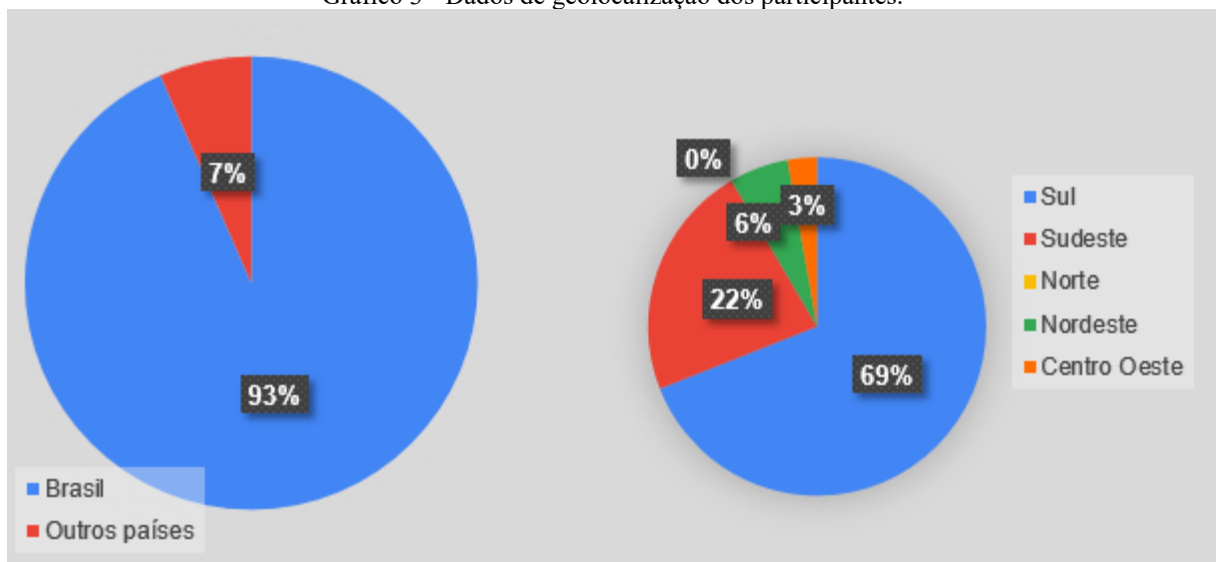


Fonte: Dados da pesquisa (2022)

Os dados geográficos demonstraram que 93% dos participantes estavam em alguma localidade do Brasil no momento de preenchimento do questionário, e 7% estavam em outros países, sendo eles: Portugal, Colômbia, Argentina e Canadá.

Dos que estavam no Brasil, o Sul apresentou maior incidência com 69% do total. Seguido pelo Sudeste com 22%. As demais regiões Nordeste, Centro Oeste e Norte ficaram com 6%, 3% e 0% respectivamente (ver Gráfico 3).

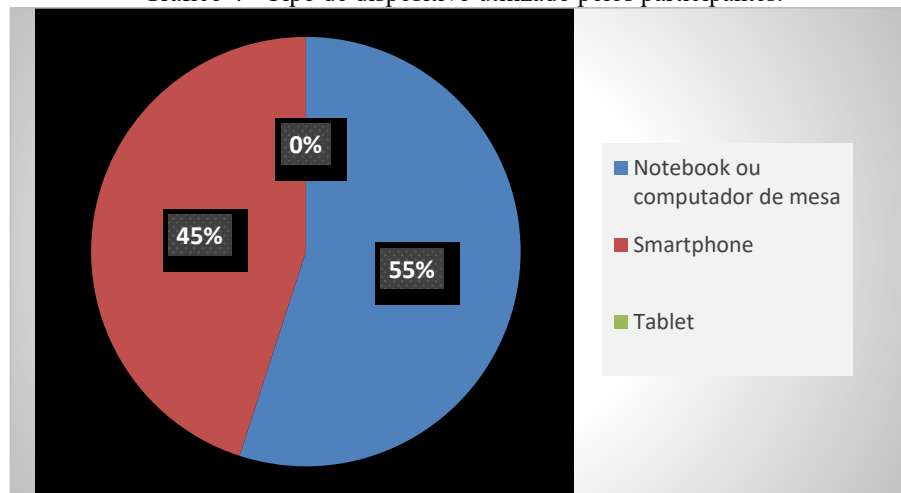
Gráfico 3 - Dados de geolocalização dos participantes.



Fonte: Dados da pesquisa (2022)

Em relação ao tipo de dispositivo que foi utilizado pelos participantes para completar o questionário, a divisão ficou em 55% de um notebook ou computador de mesa e 45% direto de um smartphone (ver Gráfico 4). Nenhum respondente utilizou o tablet para completar a pesquisa.

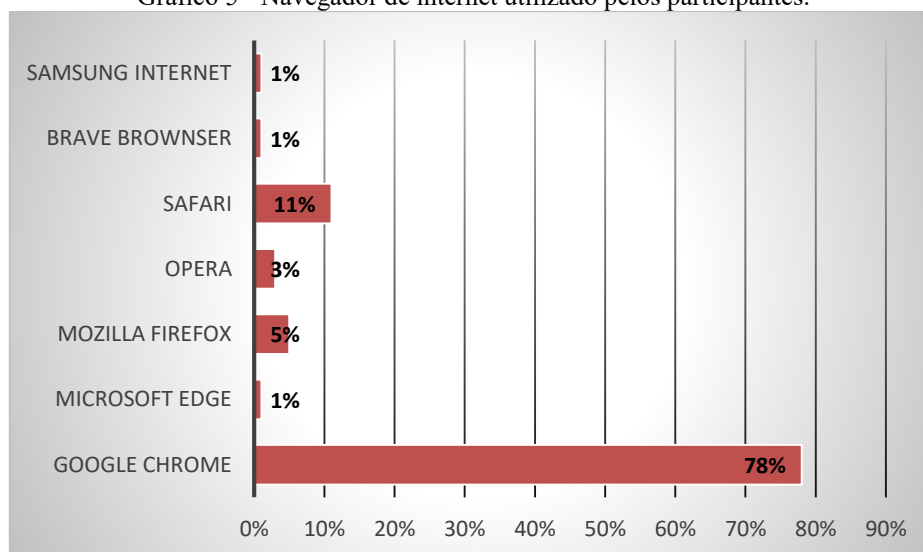
Gráfico 4 - Tipo de dispositivo utilizado pelos participantes.



Fonte: Dados da pesquisa (2022)

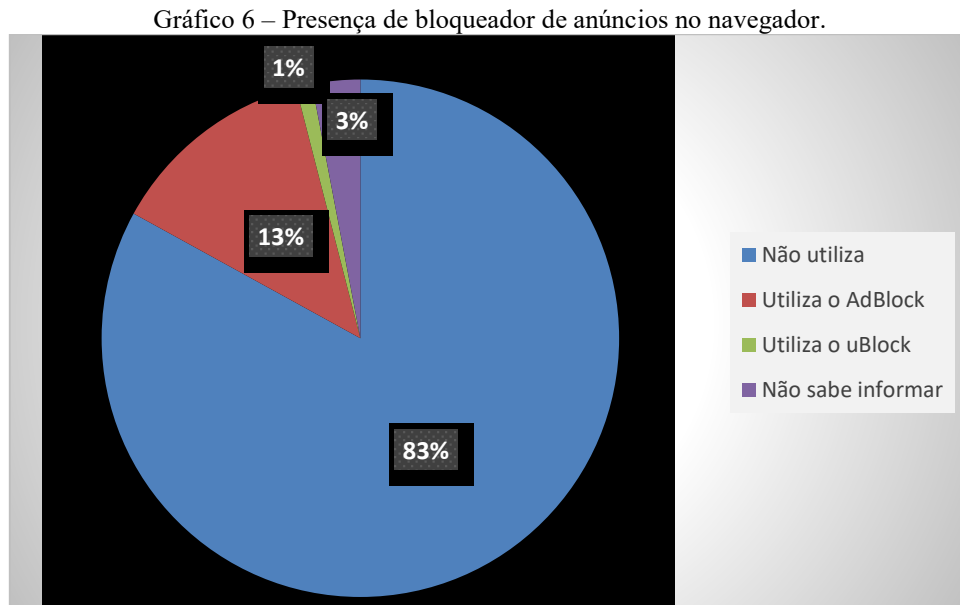
A maioria respondeu ao questionário por meio do navegador Google Chrome, representando 78% do total. 11% pelo Safari, 5% Mozilla Firefox, 3% Opera. Microsoft Edge, Brave Browser e Samsung Internet com 1% cada (ver Gráfico 5).

Gráfico 5 - Navegador de internet utilizado pelos participantes.



Fonte: Dados da pesquisa (2022)

Outro dado coletado foi acerca da utilização ou não de bloqueadores de anúncios nos navegadores utilizados durante o preenchimento do questionário (ver Gráfico 6).



Fonte: Dados da pesquisa (2022)

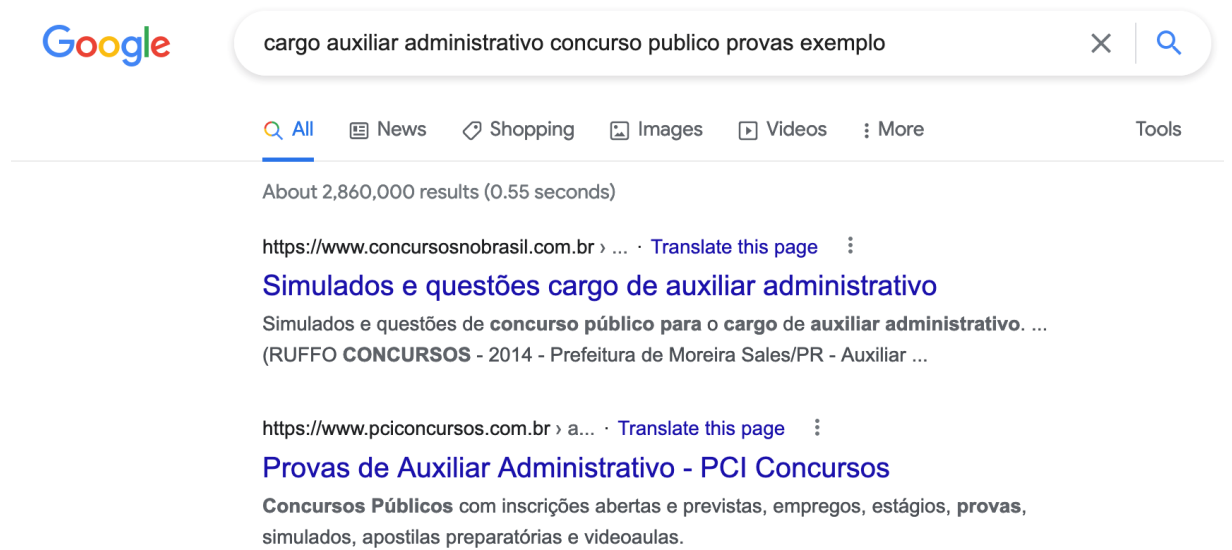
De modo geral, 83% dos participantes afirmaram não utilizar nenhum tipo de bloqueador de anúncio, enquanto 17% que utilizam ou não souberam responder, sendo essa última parcela distribuída em 13% para o uso do AdBlock, 1% para o uBlock e 3% não souberam responder por dúvidas se essa é uma ferramenta que estava embutida em seu navegador de forma automática ou instalada por terceiros sem que eles soubessem.

4.3 BUSCA LIVRE

A busca livre corresponde ao cenário um, dada a narrativa apresentada no questionário para que o usuário realizasse a pesquisa livremente. Portanto, o primeiro resultado diz respeito às escolhas dos termos feitas por cada usuário (ver exemplo Figura 36).

O cenário um solicitava se imaginar na situação de uma pessoa interessada em encontrar **provas de concurso públicos** já realizados, para o **cargo de auxiliar administrativo**.

Figura 36 - Execução da busca livre



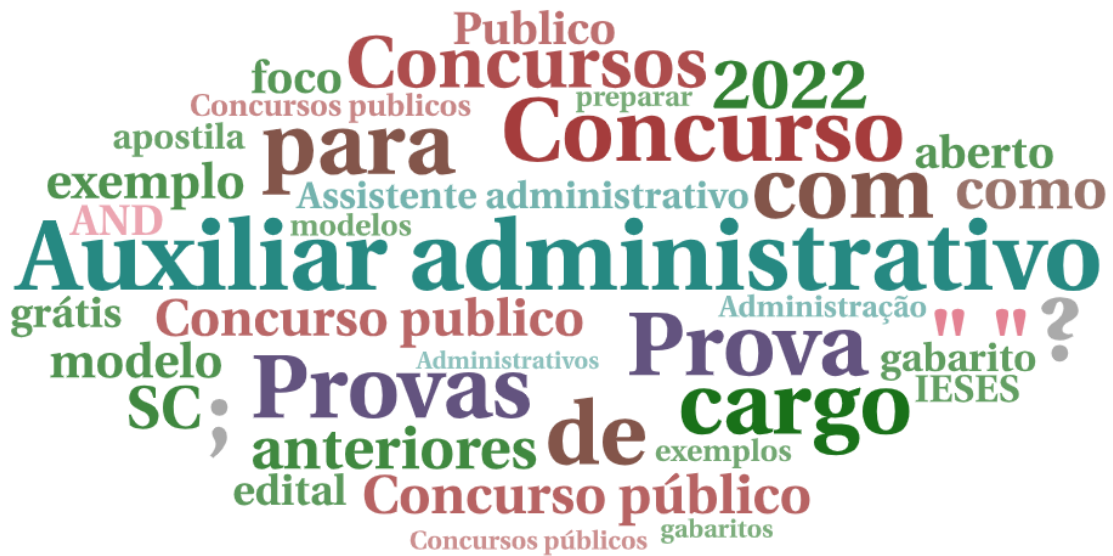
Fonte: Dados da pesquisa – coletado dos participantes (2022)

Os três termos com maior ocorrência na expressão de busca livre criada pelos participantes foram **auxiliar administrativo** com 69 usos (91% dos usuários), seguido de **prova** com 63 (83% dos usuários) e **concurso** com 59 (78% dos usuários). Sem levar em consideração plurais e variações, ainda assim os três termos citados continuam com maior ocorrência dentro das pesquisas, nessa mesma ordem.

Isso mostra que a maioria dos usuários entendeu a busca livre da mesma forma que a busca normalizada. Isto é, grande parte dos participantes utilizou os mesmos termos empregados posteriormente na busca com termos pré-indicados.

Além disso, também é possível notar o acréscimo de outros termos complementares à expressão de busca na nuvem de palavras da Figura 37. Ao todo, 55% dos usuários utilizaram algum termo adicional em relação aos três principais citados. Desses, apenas 12% empregaram algum tipo de operador avançado por conhecimento próprio.

Figura 37 - Nuvem de palavras com os termos da busca livre



Fonte – Dados da pesquisa (2022)

No caso aqui, as *stop words*⁵ não foram ignoradas pois fazem sentido diante do contexto. Mesmo sendo palavras consideradas pela computação como vazias e com uma função semântica praticamente irrelevante para a lógica de indexação e apresentação do conjunto de resultados pelos buscadores, as *stop words* aqui foram mantidas com intuito de representar a necessidade sentida pelos participantes em inseri-las na sua expressão de busca livre.

Contradizendo a questão da total irrelevância dos resultados e tendo por base a observação das capturas de tela coletadas neste cenário, principalmente no que diz respeito a quantidade e ordem dos resultados, toda preposição, conjunção ou até um caractere a mais na expressão de busca, fez diferença (mesmo que não contextualmente) para o buscador. Consequentemente, apresentou resultados diferentes daqueles sem tais acréscimos. Uma busca que utilizou a expressão **provas concurso público** apresentou resultados diferentes de uma pesquisa como **provas de concursos público** e, ainda diferentes de uma pesquisa como **prova de concurso público**, mesmo que a diferença seja apenas um acento agudo.

Os tamanhos representam a frequência com que os termos apareceram. As cores dividem os termos em sete categorias: a cor azul petróleo (■) refere-se aos termos relacionados ao cargo em questão; a cor roxa (■) para termos relacionados ao documento; a cor vermelha (■) para termos relacionados ao evento; a cor marrom (■) para

⁵ Stop words ou palavras-vazias são termos que possuem pouco significado num contexto, como em uma expressão de busca por exemplo. Tais como preposições, artigos, conjunções e outras (IBICT, 2015).

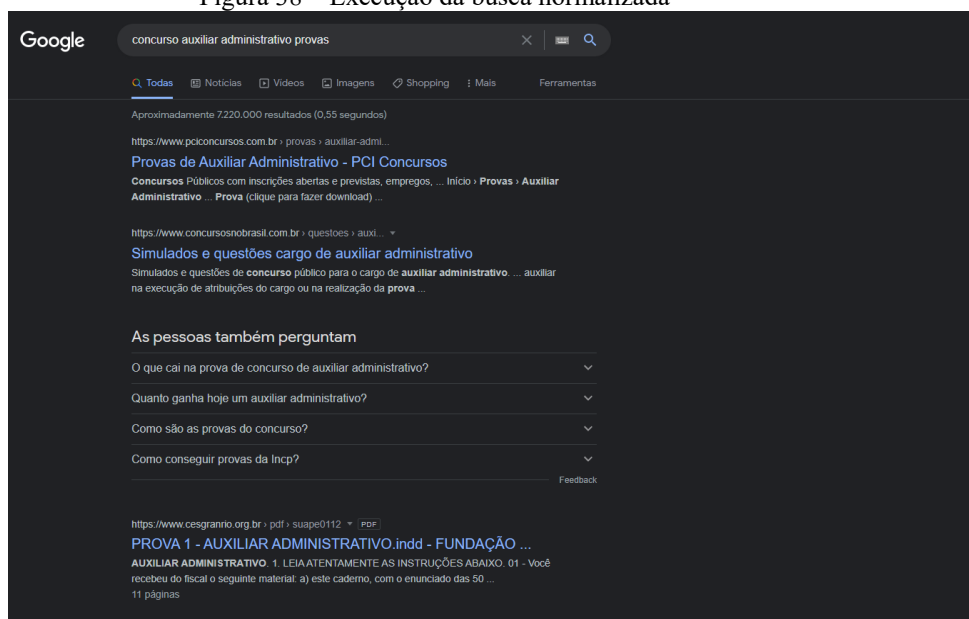
termos de ligação; a cor verde (■) para outros termos relacionados; a cor rosa (■) para operadores de pesquisa avançados e a cor cinza (■) para sinais ortográficos. O grau de opacidade de cada uma delas corresponde à importância dentro da categoria pertencente. Todas as cores foram escolhidas aleatoriamente.

Conseqüentemente, a partir de uma busca livre, os resultados são diversificados, e como já dito, cada acento, plural ou qualquer outro caractere, pode fazer a diferença nos resultados.

4.4 BUSCA NORMALIZADA

O cenário dois apresentou uma busca pronta designada como: **concurso auxiliar administrativo provas** e solicitou que o respondente a replicasse em sua própria caixa de pesquisa (ver exemplo Figura 38).

Figura 38 – Execução da busca normalizada



Fonte: Dados da pesquisa – coletado dos participantes (2022)

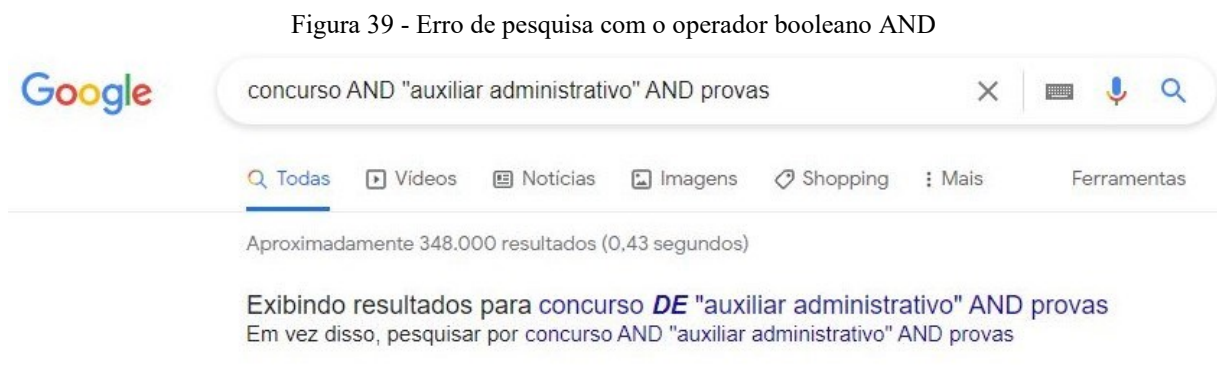
A busca normalizada apresentou resultados mais uniformes que a busca livre. Das 72 capturas de tela válidas analisadas, os resultados em notebooks ou computadores de mesa variaram entre aproximadamente 5.500.000 e 8.000.000 de resultados recuperados. Apenas uma busca teve 34.100.000 links, sendo que o usuário estava no Brasil e informou utilizar um bloqueador de anúncios no navegador Google Chrome.

Em apenas seis capturas de tela nota-se a presença de propagandas no topo dos resultados orgânicos. Esses 8% do total da categoria foram pesquisados utilizando o navegador Google Chrome. Nenhum deles apresentou sugestões de outras pesquisas realizadas no tema, que estavam presentes em outros 83% das capturas de tela, todas sem publicidade. Sugere-se uma relação entre a presença de publicidade e o aparecimento de sugestões de outras pesquisas já realizadas. No entanto, os demais cenários não confirmaram tal hipótese se comparada a pesquisas com o uso de operadores avançados, já que os cenários dois, três e quatro tiveram pesquisas que apresentaram publicidade juntamente com sugestões de pesquisas feitas por outros usuários e não apresentaram dados suficientes para comprovar uma correlação entre esses dois tópicos.

4.5 BUSCA COM OPERADOR BOOLEANO AND

O cenário três foi o primeiro a acrescentar um operador de pesquisa e apresentou uma expressão de busca com o uso do AND. O operador booleano funcionou com apenas seis participantes de um total de 71.

Isto porque para 65 usuários, a expressão de busca foi alterada pelo sistema. A alteração substituiu o primeiro booleano AND pela preposição DE como é possível verificar na Figura 39.



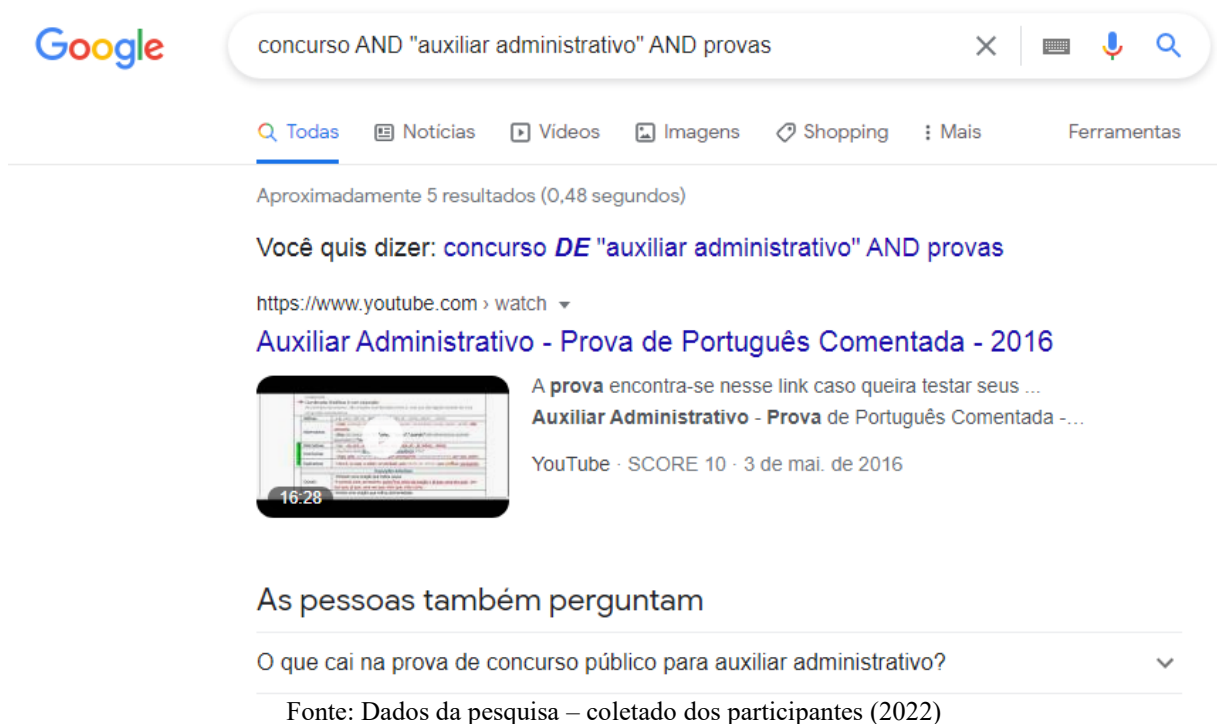
Fonte: Dados da pesquisa – coletado dos participantes (2022)

O buscador considera todas as pesquisas realizadas com mais de um termo como já acrescidas do operador booleanos AND em sua expressão. Portanto, um usuário que pesquisa por 'provas concurso auxiliar administrativo' já tem todos esses quatro termos multiplicados logicamente em decorrência do operador.

Em vista disso, entende-se que nesse caso o sistema realizou uma correção textual ao não interpretar o AND como um operador e sim como mais um termo de busca de valor textual semelhante aos termos ‘concurso’ ou ‘provas’, por exemplo.

As pesquisas em que o booleano foi aceito pelo buscador não apresentaram os avisos da Figura 39 e tiveram resultados exibidos normalmente conforme a Figura 40.

Figura 40 - Pesquisa bem-sucedida com o operador booleano AND



Após análise, ao que parece, somente cinco usuários tiveram suas pesquisas validadas pelo buscador como uma pesquisa que requer ao sistema o uso forçado e explícito do operador booleano AND por duas vezes. As cinco pesquisas não apresentaram nenhum anúncio no topo ou ao lado dos primeiros resultados, sendo que nenhum deles afirmou estar utilizando um bloqueador de anúncios.

Todas as cinco, ainda que o buscador tenha executado a expressão, foram questionadas pelo sistema se a pesquisa de fato era a que foi informada ou se a intenção era a pesquisa: **concurso DE “auxiliar administrativo” AND provas**, executada automaticamente pela grande maioria dos participantes dessa categoria.

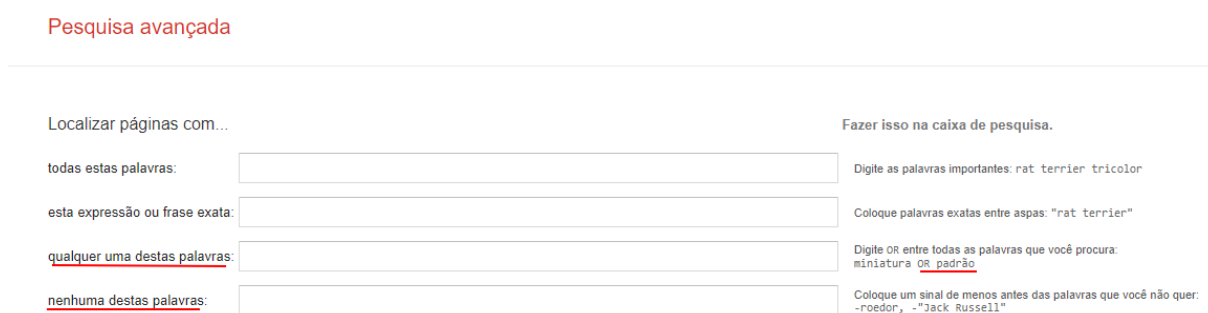
Das seis capturas de tela excluídas da categoria, duas delas, a princípio pareciam ter o operador booleano executado corretamente. No entanto, o número de resultados aproximados

ficou muito aquém do esperado para uma expressão de busca tão específica, ambos acima de 350 mil. Outro ponto percebido foi que duas dessas capturas de tela estavam relacionadas a buscadores configurados em outros idiomas que não o português brasileiro, uma delas estava em inglês e outra em espanhol.

Levando em consideração que das buscas válidas, apenas 7% obtiveram êxito, sugere-se que, para a maioria dos usuários, o operador AND não tende a funcionar obrigatoriamente em buscas com apenas um termo composto acrescido de aspas junto a outros dois termos únicos.

Como se trata de uma função que já ocorre logicamente em toda pesquisa realizada no sistema, ao acessar a aba de pesquisa avançada do Google, somente outros operadores booleanos com o OR e o NOT são disponibilizados aos usuários, conforme a Figura 41.

Figura 41 - Guia de pesquisa avançada do Google



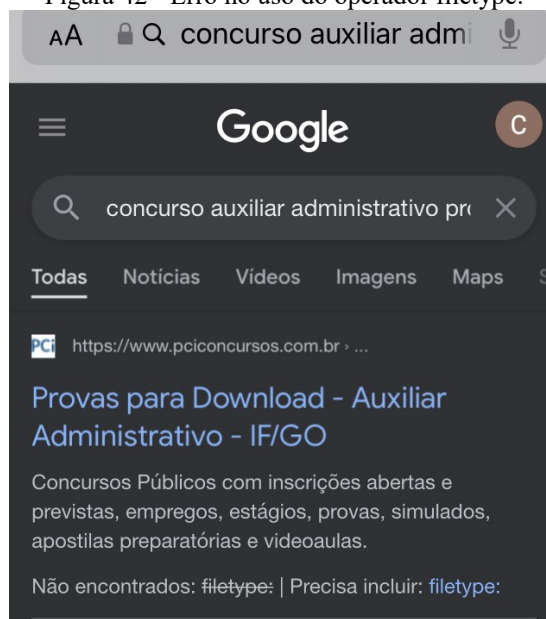
Fonte – Google (2022)

O buscador permite localizar resultados que contenham qualquer uma das palavras, usando o OR e com nenhuma das palavras citadas, utilizando o NOT, na forma de – (sinal de menos).

4.6 BUSCA COM OPERADOR AVANÇADO FILETYPE:

O cenário quatro apresentou o operador avançado filetype: com foco em obter precisão nos formatos dos documentos recuperados. A expressão de busca com o uso de operador filetype: funcionou com 70 participantes do total de 72 válidos. Dos excluídos um deles apresentou a mensagem de que houve resultados que bateram com a expressão pesquisada, no entanto, o operador de pesquisa precisaria ser incluído naquele resultado em si, o que denota que o sistema o interpretou como um termo simples e não um comando de busca avançado como é possível observar na Figura 42.

Figura 42 - Erro no uso do operador filetype:



Fonte: Dados da pesquisa – coletado dos participantes (2022)

Neste caso, o usuário não especificou o formato de arquivo requerido e indicado no questionário, que no caso era PDF. Qualquer outra extensão de arquivo fosse especificada não seria listada pelo buscador como possível inclusão de termo já que ele entenderia como um comando direto mesmo que o resultado fosse zero arquivos recuperados.

Nas 70 demais buscas é possível observar uma etiqueta após a URL de cada resultado, confirmando que o documento hospedado naquele site é na extensão requerida na expressão de busca, no caso, PDF (ver Figura 43).

Figura 43 - Etiqueta com extensão de arquivo



Fonte: Dados da pesquisa – coletado dos participantes (2022)

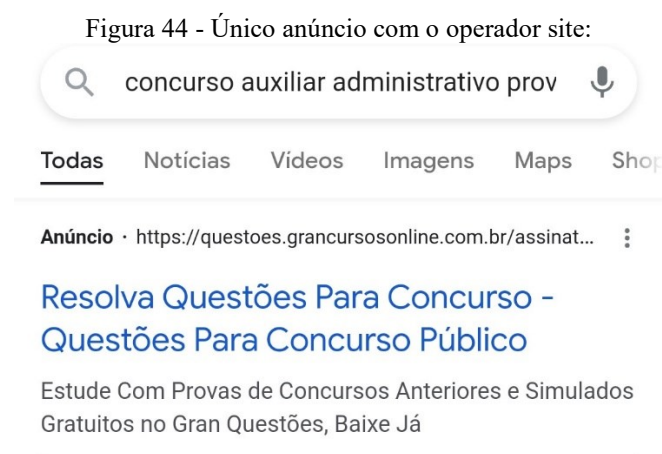
Em geral, também é possível afirmar que para a amostra analisada, o operador não influenciou na aparição ou não de anúncios e suas respectivas posições em tela ou sugestões de pesquisas. Em relação a quantidade aproximada de resultados em pesquisas realizadas em notebooks ou computadores de mesa foi mantida a variação entre 1.000.000 e 1.600.00 de resultados. Apenas dois destoaram desse número, um para mais e outro para menos, o que é justificado pela geolocalização dos dois usuários que estavam fora do Brasil no momento das buscas.

4.7 BUSCA COM OPERADOR AVANÇADO SITE:

O cenário cinco contava com o uso do operador site: a fim de especificar ao buscador a fonte principal de informação dos resultados a serem apresentados. Os resultados mostram que o operador site: funcionou nas 73 capturas de tela válidas analisadas dentro da categoria.

Com o operador avançado site: foi possível observar o maior nível de similaridade entre os primeiros resultados e as perguntas sugeridas pelo próprio buscador como pesquisas similares que outros usuários realizaram.

Em relação aos anúncios, este foi o operador que mais influenciou na presença de publicidade dentre os resultados. De todas as 73, somente uma apresentou um anúncio no topo, o que gerou controvérsias tendo em vista que o usuário em questão afirmou utilizar um bloqueador de anúncios em seu navegador (ver Figura 44).



Fonte: Dados da pesquisa – coletado dos participantes (2022)

Este mesmo respondente também apresentou outra captura de tela com a presença de anúncios na categoria do operador filetype:. Nesse caso, ou a informação cedida por ele em relação ao uso de um bloqueador de anúncios estava incorreta ou o funcionamento deste bloqueador foi comprometido em níveis mais avançados de busca com a presença de operadores como filetype: e site:.

A quantidade aproximada de resultados em pesquisas realizadas em notebooks ou computadores de mesa variou entre 900 e 2300 resultados, sendo, portanto, o operador que mais trouxe exatidão aos resultados, retornando para o usuário um número consideravelmente baixo em comparação aos outros quatro cenários nos quais havia resultados na casa dos milhões. Ocorrência que condiz perfeitamente com a funcionalidade do operador, já que este aponta somente para resultados advindos de uma única fonte, no caso, o site da banca Quadrix. Nesse caso, a quantidade é relativa, mas a fonte de busca é precisa.

4.8 OBSERVAÇÕES GERAIS

Em 10 das 380 capturas de tela, os participantes, por conta própria, rolaram a tela e realizaram a captura. Alguns resultados abaixo da caixa de pesquisa. O que leva a crer que anúncios foram ignorados e desconsiderados por eles como resultados válidos frente à expressão de busca pesquisada.

Com ou sem o uso de operadores avançados, as pesquisas realizadas em smartphones não contêm as quantidades aproximadas de resultados apresentadas pelo mesmo buscador em computadores de mesa e notebooks.

Em oito capturas de tela, mesmo os usuários que informaram utilizar um bloqueador de anúncios, tiveram anúncios exibidos em suas pesquisas. Seis deles ao utilizarem o operador filetype:.

O uso dos bloqueadores não interferiu na quantidade de resultados aproximados informada pelo buscador em comparação aos que não utilizaram.

Os diferentes navegadores, bem como o uso de smartphones ou notebooks e computadores de mesa não manifestaram características que visivelmente interferissem no uso de funcionamento dos operadores propostos na pesquisa.

As características que mais se destacaram nas buscas foram a presença de anúncios, sugestões de outras pesquisas já realizadas que tinham relação com a pesquisa atual e resultados aproximados em notebooks ou computadores de mesa.

Em relação a presença de anúncios no topo dos resultados, calculou-se a porcentagem de pesquisas que apresentaram anúncios dentre os participantes que informaram não utilizar nenhum tipo de bloqueador de anúncios (ver Tabela 2). Um único participante informou utilizar um bloqueador, entretanto, seus resultados apresentavam anúncios. Diante das informações que puderem ser visualizadas nas capturas de tela, ele também foi considerado dentro do grupo que continham anúncios nos resultados.

Tabela 2 - Porcentagem de anúncios nos cenários normalizados

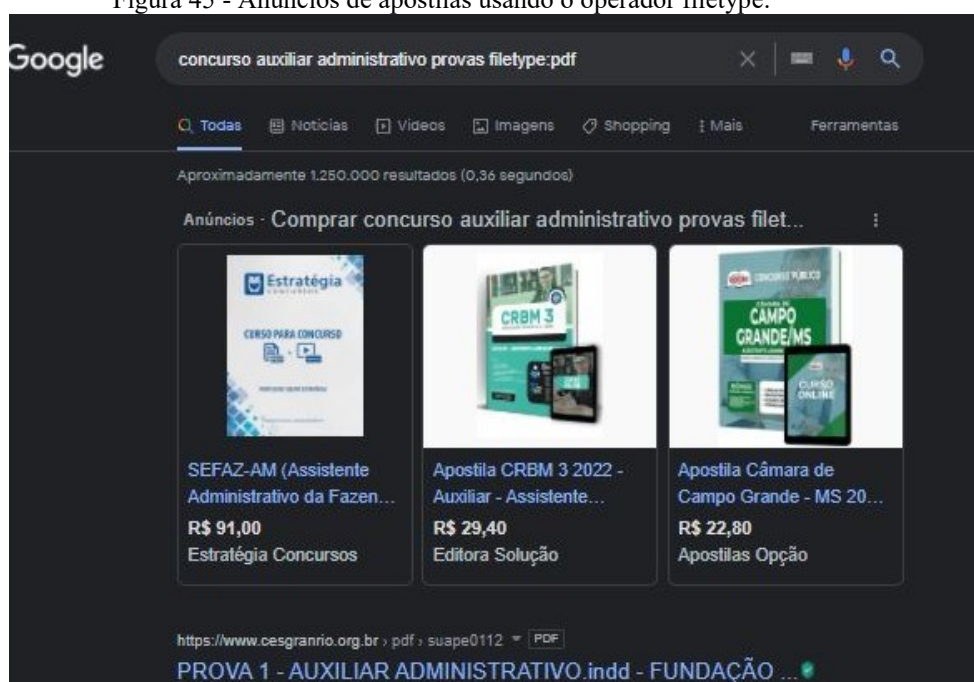
CENÁRIOS DE BUSCA	ANÚNCIOS
Busca normalizada	9%
Busca com o AND e as aspas	0%
Busca com o filetype:	39%
Busca com o site:	2%

Fonte: Dados da pesquisa (2022)

Nota-se que uso de alguns operadores avançados pode favorecer a aparição de anúncios de pesquisa à medida que ou obrigatoriamente o anunciante tenha usado todas as palavras-chave da intersecção resultante do uso do booleano AND acrescido as aspas, ou esteja atrelado ao site indicado como fonte única de pesquisa do operador site:.

Se a busca realizada pelo operador filetype: indicar um formato de arquivo atrelado ao produto anunciado, há chances deste produto se encaixar nos resultados como anúncio de pesquisa (ver Figura 45).

Figura 45 - Anúncios de apostilas usando o operador filetype:



Fonte: Dados da pesquisa – coletado dos participantes (2022)

Nos dados coletados, o formato PDF sempre esteve ligado a anúncios de apostilas, provavelmente comercializada nesse formato de arquivo.

Na sequência, calculou-se a porcentagem das capturas de tela em que as sugestões de pesquisa apareciam dentro dos cenários normalizados (ver Tabela 3).

Tabela 3 - Porcentagem das sugestões de pesquisa nos cenários normalizados

CENÁRIOS DE BUSCA	SUGESTÕES
Busca normalizada	83%
Busca com o AND e as aspas	100%
Busca com o filetype:	76%
Busca com o site:	92%

Fonte: Dados da pesquisa (2022)

Não houve uma grande variação que possa inferir correlação entre o aparecimento ou não de sugestões (ver Figura 46) a partir do uso dos operadores avançados testados.

Figura 46 - Exemplo de sugestões de pesquisas já realizadas

As pessoas também perguntam

O que cai na prova de concurso público para auxiliar administrativo?	▼
Quanto ganha hoje um auxiliar administrativo?	▼
Como são as provas do concurso?	▼
Como conseguir provas da Incp?	▼

Fonte: Dados da pesquisa – coletado dos participantes (2022)

Por fim, calculou-se a mediana entre as pesquisas realizadas em notebooks e computadores de mesas, no que diz respeito à quantidade de resultados aproximados (ver Tabela 4).

Tabela 4 - Mediana dos resultados nos cenários normalizados

CENÁRIOS DE BUSCA	MEDIANA DOS RESULTADOS APROXIMADOS
Busca normalizada	7.400.000
Busca com o AND e as aspas	4
Busca com o filetype:	1.270.000
Busca com o site:	1.930

Fonte: Dados da pesquisa (2022)

Os resultados apontam para uma redução na quantidade total de resultados ao utilizar operadores avançados de pesquisa se comparado a uma busca sem eles. Afinal, quanto mais específica for a expressão, menor é a quantidade de documentos dentro do índice em que o buscador realizará a varredura a procura de resultados precisos.

5 CONSIDERAÇÕES FINAIS

O sucesso de um processo de busca na internet pode ser analisado desde a construção da expressão de busca —com ou sem o uso de operadores—, até o resultado apresentado em tela. Este resultado ainda pode apresentar elementos de feedback que apontam sucesso ou fracasso daquilo que foi solicitado pelo usuário. Ou seja, se a expressão construída de fato foi entendida pelo buscador e funcionou.

Diante disto, e após as observações realizadas nessa dissertação, é possível apresentar duas reflexões. A primeira se refere ao **funcionamento da expressão de busca**. Ao rodar uma expressão de busca com operadores booleanos, e o retorno mostrar nenhum resultado, pode-se imaginar que a expressão de busca não funcionou.

Porém, uma pesquisa retornar zero resultados não quer dizer que a expressão de busca estava incorreta. Ela pode ter funcionado sem nenhum erro de sintaxe e constatado que não há nada ali correspondendo ao que foi pesquisado.

No segundo cenário reflexivo, usa-se uma expressão de busca com operadores booleanos. A pesquisa retornou com 23 resultados.

A segunda reflexão, é sobre a **precisão dos resultados**. Uma pesquisa retornar resultados não quer dizer que a expressão de busca foi precisa. A depender da situação, alguns sistemas podem ignorar operadores ou retornar resultados com partes da expressão. Usuários mais desavisados podem achar que a expressão foi executada exatamente como ele construiu, quando na verdade o sistema usou outros caminhos para apresentar os resultados.

Esse último caso fica explícito nas pesquisas utilizando o operador booleano AND que compuseram esse estudo. No caso, o buscador informou que estava exibindo resultados para uma expressão de busca que não a solicitada. Contudo, nem sempre isso acontece e cabe sempre ao usuário examinar se aqueles resultados que retornaram equivalem a ideia inicial de expressão que ele elaborou e construiu.

Isto é importante e se relaciona com todo o processo de observações realizado neste trabalho à medida que tais operadores avançados não estão presentes na interface inicial da tela de busca e nem são apresentados aos usuários como sugestões. Assim, demandam que os usuários tenham um prévio conhecimento de como funcionam e para que servem, de modo a evitar um uso inadequado e ineficaz dos operadores.

Das características nos resultados de pesquisa que se destacam a partir do uso de operadores avançados na busca por documentos na web estão os anúncios de pesquisa, as

sugestões de pesquisas que já foram realizadas anteriormente por outros usuários e a quantidade aproximada de resultados.

Em relação aos anúncios na página de resultados de pesquisa, nota-se que o uso de operadores avançados de pesquisa pode favorecer seu aparecimento como também os inibir. A criação dos anúncios possivelmente precisará estar alinhada ao objetivo do operador, como por exemplo, um formato de documento compatível com o produto anunciado na web.

No que diz respeito às sugestões de pesquisas já realizadas, os operadores não parecem interferir em seu aparecimento, sendo estas teoricamente associadas ao conteúdo da busca em si a partir somente dos termos de busca selecionados pelo usuário.

Por último, a quantidade de resultados aproximados demonstrou que o operador booleano AND acrescido do uso de aspas “, tem potencial para reduzir os resultados, seguidamente do operador site: e por último o filetype:. Todos os quatro operadores utilizados apresentaram redução na quantidade aproximada de resultados se comparados a mesma pesquisa sem o uso deles. Com isso, sugere-se que o uso de operadores avançados de pesquisa pode diminuir a quantidade de resultados apresentados aos usuários e, portanto, direcionar com maior precisão a expressão de busca para encontrar os documentos desejados.

Os resultados também sugerem que a camada computacional, representada pelos buscadores, é continuamente influenciada pela camada cultural, já que os buscadores mudam a forma como uma expressão de busca é executada a partir da forma como os usuários a executam. Assim, a substituição do operador AND por “DE” ou mesmo sempre considerar um AND entre cada palavra da busca são evidências desta influência.

Os usuários por vezes sentem necessidade de incluir outros termos na expressão de busca que vão além de operadores de pesquisa ou das palavras-chave principais relacionadas a temática de uma pesquisa, termos como preposições, conjunções e demais qualificadores como datas, locais e formatos, reforçando a ideia do introduzida pelo AskJeeves nos anos 90 de que a pesquisa pode ser expressa da forma de diálogo entre humanos, ou seja, com linguagem natural.

Os resultados também sugerem o inverso, ou seja, a influência na camada cultural. Isto porque um motor de busca, na forma em que opera, representa, organiza, disponibiliza e permite que operações como pesquisas de diferentes níveis de complexidade ocorram, é de se esperar que influencie diretamente a camada cultural das mídias digitais, modificando práticas e comportamentos nessa interação humano-computador.

O modo como a informação era organizada por bibliotecas foi remediada nos buscadores no início da década de 90. As caixas de busca implementadas moldaram o comportamento de busca on-line dos usuários, as relações entre conteúdos influenciaram algoritmos de organização e ranqueamento de resultados e, mais recentemente, o volume crescente de dados disponíveis na web tem estimulado novas formas mais específicas de se encontrar uma informação como filtros e operadores avançados de pesquisa.

É fato que tais operadores não são tão populares em pesquisas de rotina e não são todos os usuários que usam ou sabem empregá-los da melhor maneira possível. A interface de pesquisa não os privilegia e nem os sugere, assim sendo, é necessário conhecê-los previamente para utilizá-los.

Dos anos 90 até início dos anos 2000 a caixa de pesquisa dos buscadores saiu de um lugar para assumir destaque dentro da interface, centralizada e indicando ao usuário que a inserção de uma expressão de busca é o fator mais importante da sua página inicial. Em específico no Google, estímulos de filtros aparecem somente após a pesquisa ser executada e operadores avançados ficam restritos a usuários com conhecimento prévio. A depender do desenvolvimento tecnológico para entender as pesquisas de cada usuário somado ao aumento no volume de documentos disponíveis da web, há a possibilidade de novamente a camada computacional influenciar a camada cultural, ao trazer à tona recursos avançados de busca para a interface inicial dos buscadores.

Os algoritmos de pesquisa de fato tornam cada tela de resultado muito particular. Pesquisas bem específicas aparentemente reduzem as diferenças, mas não a ponto de torná-las idênticas. Aliás, esse não é o objetivo, afinal, a personalização de interesses por meio de algoritmos é um dos grandes trunfos do marketing digital, com alto valor.

Quanto às limitações da pesquisa, compreende-se que dados acerca do idioma no qual o buscador utilizado estava configurado, poderiam ter sido coletados para auxiliar na análise dos resultados em relação aos buscadores configurados em português. Sem eles foi possível realizar essa verificação apenas de forma visual e não como um dado fornecido pelo participante em forma de resposta estruturada.

Outro ponto em questão diz respeito à familiaridade dos usuários em relação a operadores de pesquisa avançados. Como alguns erros aconteceram por parte da execução da expressão, em questões de sintaxe, uma breve explicação sobre como eles não funcionam poderia ter sido acrescida ao questionário, de forma a prevenir tal acontecimento.

Sugere-se, portanto, que o design da interface do Google poderia usufruir da remediação para que a oscilação entre a transparência e opacidade dos operadores contribua ativamente para instruir os usuários na sua aplicação. Este design estaria agindo para o movimento contínuo e alternado de influência das camadas culturais e computacionais.

O uso de imagens, ou seja, das capturas de telas como unidade de dados pode gerar adversidades na pesquisa principalmente pela dificuldade em estipular o perímetro da captura, fazendo com que a dimensão destes sejam distintas de dispositivo para dispositivo, o que resulta em mais ou menos informações em cada um deles. Problema esse que poderia ter sido contornado se as páginas tivessem sido salvas de forma completa em formato HTML.

Pesquisas futuras podem abordar uma comparação entre diferentes buscadores e quais são os operadores avançados de pesquisa disponíveis em cada um deles, com foco em descobrir se a interface os promove e incentiva seu uso ou esses são sempre restritos aos usuários com conhecimento prévio em pesquisa avançada.

Outra forma de expandir a pesquisa está em realizar uma discussão a respeito dos mecanismos de busca semânticos baseados em modelos pré-treinados como BERT e GPT. Estes modelos já são realidade e a partir de buscas mais contextuais, podem tornar a experiência do usuário em realizar uma pesquisa online, algo muito próximo a ideia explorada pelo AskJeeves na década de 90 ao prenciar o uso de um buscador com a consulta de um especialista humano.

Ainda a partir deste trabalho, há a possibilidade de focar uma análise em somente um único tipo de operador, como os booleanos por exemplo. Propõe-se também a realização de uma pesquisa que procure compreender como as bibliotecas têm trabalhado a capacitação para pesquisa avançada não somente em buscadores tradicionais, mas também em bases de dados científicas, já que estas são ambientes em que tais operadores podem ser bastante oportunos na construção de expressões de busca.

REFERÊNCIAS

- A HISTÓRIA dos buscadores de internet! – História da Tecnologia. [S.l.: s. n.], 2022. 1 vídeo (13 min.). Publicado pelo canal TecMundo. Disponível em: <https://www.youtube.com/watch?v=zrV3C1DKWiI&t=112s>. Acesso em: 16 abr. 2022.
- ABAR, Celina. Noções de lógica matemática: esboço do desenvolvimento da lógica. PUC-SP. Disponível em: <https://www.pucsp.br/~logica/Desenvolvimento.htm>. Acesso em 22 out. 2021.
- AMERICAN LIBRARY ASSOCIATION – ALA. Presidential Committee on Information Literacy. Chicago, 1989. Disponível em: <http://www.ala.org/ala/publications/whitepapers/presidential>. Acesso em: 09 ago. 2022.
- ARAÚJO, Andréa Nunes de. **A contribuição da indexação no processo de seleção de palavras-chave no Google Adwords**. 2018. Dissertação (Mestrado em Organização, Mediação e Circulação da Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2018. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27163/tde-26122018-165003/pt-br.php>. Acesso em: 24 jan. 2020.
- ASK Jeeves (1999) - Television Commercial, [S. l.: s. n.], 1999. Publicado pelo canal Jim Margle, 1 vídeo (30 seg). Disponível em: <https://www.youtube.com/watch?v=xcJ-xmQqmrA>. Acesso em: 12 abr. 2022.
- BASES DE TESES E DISSERTAÇÕES DO EGC. Disponível em: <http://btd.egc.ufsc.br/?p=3346>. Acesso em: 15 jul. 2022.
- BARNES, Jonathan. **Aristóteles**. 3. ed. São Paulo: Loyola, 2013.
- BERINSTEIN, Paula. Turning visual: image search engines on the Web. **Information Today**, [S. l.], v. 26, n. 1, p. 30–34, 1998. ISSN: 01465422. Disponível em: <https://gogale.ez46.periodicos.capes.gov.br/ps/i.do?p=AONE&sw=w&issn=01465422&v=2.1&it=r&id=GALE%7CA20554340&sid=googleScholar&linkaccess=fulltext>. Acesso em: 24 ago. 2022.
- BERNERS-LEE, Tim.; HENDLER, James.; LASSILA, Ora. The Semantic Web. **Scientific American**. v. 284, n. 5, p. 28-37, 2001. Disponível em: <https://www.jstor.org/stable/26059207>. Acesso em: 10 ago. 2021.
- BIANCHINI, Monica; GORI, Marco; SCARSELLI, Franco. Inside PageRank. **ACM Transactions on Internet Technology (TOIT)**, v. 5, n. 1, p. 92-128, 2005. Disponível em: <https://dl.acm.org/doi/abs/10.1145/1052934.1052938>. Acesso em: 19 abr. 2022.
- BLANCHÉ, Robert; DUBUCS, Jacques. **História da Lógica**. Paris: Edições 70, 1996.
- BRANSKI, Regina Meyer. Recuperação de informações na Web. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 9, n. 1, 2004. Disponível em:

<https://periodicos.ufmg.br/index.php/pci/article/view/23500>. Acesso em: 23 nov. 2022.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais. Brasília, DF: Presidência da República, 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm. Acesso em: 20 abr. de 2021.

BUCHINGER, Diego; CAVALCANTI, Gustavo Andriolli de Siqueira; HOUNSELL, Marcelo da Silva. Mecanismos de busca acadêmica: uma análise quantitativa. **Revista Brasileira de Computação Aplicada**, v. 6, n. 1, p. 108-120, 18 abr. 2014. Disponível em: <https://doi.org/10.5335/rbca.2014.3452>. Acesso em: 29 mar. 2022.

BUCKINGHAM, David. Cultura Digital, Educação Midiática e o Lugar da Escolarização. **Educação & Realidade**, [S. l.], v. 35, n. 3, 2010. Disponível em: <https://seer.ufrgs.br/index.php/educacaoerealidade/article/view/13077>. Acesso em: 29 dez. 2022.

BUSH, Vannevar. **As we may think**. The atlantic monthly, 1945. Disponível em: <https://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>. Acesso em: 20 fev. 2022.

CENDÓN, Beatriz Valadares. Ferramentas de busca na Web. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 39-49, 2001. Disponível em: <https://www.scielo.br/j/ci/a/WdYRz6LmQbBD5ZWnKTfHKkm/abstract/?lang=pt&format=html>. Acesso em: 22 set. 2021.

CHAUÍ, Marilena. **Convite à filosofia**. São Paulo: Editora Ática, 2000.

CHOO, Chun Wei. **A organização do conhecimento**: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. São Paulo: Ed. SENAC, 2003. 425 p.

CLEVELAND, Donald; CLEVELAND, Ana. **Introduction to indexing and abstracting**. 4. ed. Califórnia: Libraries Unlimited, 2013.

COMISSÃO EUROPEIA. **Para uma Europa do conhecimento**: a União Europeia e a sociedade da informação, [S. l.], 2003. Disponível em: <https://op.europa.eu/en/publication-detail/-/publication/3a0f5454-1c91-46a6-aed1-fd3ca5dd8c64/language-pt/format-PDF/source-232609877>. Acesso em 20 set. 2021.

CORREA, Sara Dotta; FARIA, Gustavo Henrique Campos de; FADEL, Luciane Maria; VAZ, Carlos Eduardo Verzola. Digitalização de informações geográficas: a transcodificação no Google Earth. Animus. **Revista Interamericana de Comunicação Midiática**, Santa Maria, v. 21, n. 46, 2022. Disponível em: <https://periodicos.ufsm.br/animus/article/view/44143>. Acesso em: 6 out. 2022.

COVERING, Elizabeth Van. The history of the Internet search engine: navigational media and the traffic commodity. *In*: SPINK, Amanda; ZIMMER, Michael. **Web search**:

- multidisciplinar perspectives. Berlin: Springer, 2008. Disponível em: https://link.springer.com/chapter/10.1007/978-3-540-75829-7_11. Acesso em: 11 abr. 2022.
- CUNHA, Murilo Bastos da; CAVALCANTI, Cordélia Robalinho de Oliveira. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Briquet de Lemos, 2008. Disponível em: <https://repositorio.unb.br/handle/10482/34113>. Acesso em 01. out. 2021.
- D'OTTAVIANO, Ítala Maria Loffredo; FEITOSA, Hércules de Araujo. Sobre a história da lógica, a lógica clássica e o surgimento das lógicas não clássicas. Disponível em: https://arquivos.cruzeirosulvirtual.com.br/materiais/disc_2011/2sem_2011/logicaformal/unII/complementar_II.pdf. Acesso em: 30 mar. 2022.
- DAVIS, Bob. Lycos. *In*: LIVINGSTON, Jessica. **Founder at work: stories of startups' early days**. Nova York: Apress, 2007. [Entrevista cedida a] Jessica Livingston. Disponível em: <https://link.springer.com/content/pdf/10.1007/978-1-4302-0327-8.pdf>. Acesso em: 11 abr. 2022.
- DESCARTES, René. **Discurso do método**. Porto Alegre: L&PM, 2009.
- DOERNHOEFER, Mark. Surfing the net for software engineering notes. **ACM SIGSOFT Software Engineering Notes**, Estados Unidos, v. 35, n. 3, p. 14-23, 2010. Disponível em: <https://dl.acm.org/doi/10.1145/1764810.1764815>. Acesso em: 01 abr. 2022.
- FAGUNDES, Vinícius Bossle. **Proposta de uso de engenharia do conhecimento para revisão sistemática**. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, 2018. Disponível em: <https://repositorio.ufsc.br/handle/123456789/194471>. Acesso em 28 jul. 2021.
- FREIRE, Patrícia de Sá. **Aumente a qualidade e quantidade de suas publicações científicas**: manual para elaboração de projetos e artigos científicos. Curitiba: CRV, 2013. 87 p.
- GABRIEL, Martha Carrer Cruz. **Sem e Seo**: dominando o marketing de busca. 2ª. ed. São Paulo: Novatec, 2012.
- GANDAL, Neil. The dynamics of competition in the internet search engine market. **International Journal of Industrial Organization**, v. 19, n. 7, p. 1103-1117, 2001. Disponível em: [https://doi.org/10.1016/S0167-7187\(01\)00065-0](https://doi.org/10.1016/S0167-7187(01)00065-0). Acesso em: 11 abr. 2022.
- GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed., São Paulo: Atlas, 2008.
- GODOY, Arnaldo Sampaio de Moraes. O julgamento de Sócrates. **Revista Sequência**, Florianópolis, v. 24, n. 46, p. 11-27, 2003. Disponível em: <https://periodicos.ufsc.br/index.php/sequencia/article/view/15289>. Acesso em: 21 out. 2021.
- GOOGLE. 200º aniversário de George Boole. Arquivo de doodles. 2015. Disponível em: <https://www.google.com/doodles/george-booles-200th-birthday>. Acesso em: 13 abr. 2022.

GOOGLE. Ajuda do Google Ads. 2018. Disponível em: <https://support.google.com/google-ads/answer/2567043?hl=pt-BR>. Acesso em: 29 ju. 2022.

GOOGLE. Anúncios e dados. 2022a. Disponível em: https://safety.google/intl/pt-BR_br/privacy/ads-and-data/. Acesso em: 24 ago. 2022.

GOOGLE. Como a busca do Google funciona. [2021a]. Disponível em: <https://www.google.com/search/howsearchworks/how-search-works/>. Acesso em: 23 jan. 2022.

GOOGLE. Como a Pesquisa Google funciona. [2021b]. Disponível em: <https://developers.google.com/search/docs/basics/how-search-works?hl=pt-br>. Acesso em: 28 set. 2021.

GOOGLE. Exibir anúncios na primeira página. [2021c]. Disponível em: <https://support.google.com/google-ads/answer/2472742?hl=pt-BR>. Acesso em: 08 set. 2021.

GOOGLE. Interface de pesquisa, 2023a. Disponível em: <https://www.google.com/>. Acesso em: 02 jan. 2023.

GOOGLE. Mecanismo de pesquisa programável. [2022b]. Disponível em: <https://programmablesearchengine.google.com/about/>. Acesso em: 04 abr. 2022.

GOOGLE. Pesquisa avançada. 2022. Disponível em: https://www.google.com.br/advanced_search. Acesso em: 22 out. 2022.

GOOGLE. Refinar pesquisas na web. [2013]. Disponível em: <https://support.google.com/websearch/answer/2466433?hl=pt>. Acesso em: 05 jan. 2023.

GOOGLE. Vestido Jennifer Lopez Grammy, 2023b. Disponível em: <https://www.google.com/search?q=vestido+jennifer+lopez+grammy>. Acesso em: 02 jan. 2023.

GOOGLE. Visão geral. [2021d]. Disponível em: <https://www.google.com/intl/pt-BR/search/howsearchworks/>. Acesso em: 28 set. 2021.

GOULART, Elias Estevão; HETEM JUNIOR, Annibal. Pesquisas na web: estratégias de busca. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 5, n. 1, p. 53–66, 2007. Disponível em: <https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/2021>. Acesso em: 27 fev. 2022.

GRAMMY AWARDS. **1999 Grammy Winners**: 42nd Annual Grammy Awards. 2022. Disponível em: <https://www.grammy.com/awards/42nd-annual-grammy-awards>. Acesso em: 14 abr. 2022.

IBGE. Educa jovens. **Uso de internet, televisão e celular no Brasil**. 2019. Disponível em: <https://educa.ibge.gov.br/jovens/materias-especiais/20787-uso-de-internet-televisao-e-celular-no-brasil.html#subtitulo-0>. Acesso em: 08 set. 2021.

IBICT. Stop words e sinonimos. 2015. Disponível em: http://wiki.ibict.br/index.php/Stop_words_e_sinonimos. Acesso em: 05 jan. 2023.

INCE, Darrel. **A dictionary of the Internet**. 2 ed. [S. l.]: Oxford University Press, 2009. Disponível em: <https://www.oxfordreference.com/display/10.1093/oi/authority.20110803124909395:jsessionid=CC48DB62943A301E2090CFA2C52C09FD>. Acesso em: 22 abr. 2022.

INSTITUTO QUADRIX. Sobre o instituto. 2022. Disponível em: <http://www.quadrix.org.br/institucional.aspx>. Acesso em: 21 mar. 2022.

INTERNET ARCHIVE. Cadê? 1996. 1 captura de tela em formato PNG. Disponível em: <https://web.archive.org/web/19961225005433/http://cade.com.br/>. Acesso em: 12 abr. 2022.

INTERNET ARCHIVE. Achei. 1997. 1 captura de tela em formato PNG. Disponível em: <https://web.archive.org/web/20000303110712/http://achei.com.br/>. Acesso em: 12 abr. 2022.

INTERNET ARCHIVE. Aonde. 2001. 1 captura de tela em formato PNG. Disponível em: <https://web.archive.org/web/20010401144548/http://aonde.com.br:80/>. Acesso em: 12 abr. 2022.

INTERNET ARCHIVE. Prokura. 1999. 1 captura de tela em formato PNG. Disponível em: <https://web.archive.org/web/19991013050200/http://prokura.com.br/>. Acesso em: 12 abr. 2022.

INTERNET ARCHIVE. Radix. 1999. 1 captura de tela em formato PNG. Disponível em: <https://web.archive.org/web/19991128101010/http://www.radix.com.br/>. Acesso em: 12 abr. 2022.

JANES, Joseph. O mundo Google: ele mudou a internet e a nossa vida. Agora quer revolucionar todo o resto. [Entrevista cedida a] Rafael Kenski. **Super Interessante**. São Paulo, out. 2004. Disponível em: <https://super.abril.com.br/tecnologia/o-mundo-google/>. Acesso em: 24 ago. 2022.

KOEHLER, Wallace. Digital libraries and World Wide Web sites and page persistence. **Information Research**, v. 4, n. 4, 1999. Disponível em: <https://www.sci.brooklyn.cuny.edu/~sklar/teaching/s08/cis20.2/papers/diglib.pdf>. Acesso em 30 set. 2021.

KRAUS, Joe. Excite. In: LIVINGSTON, Jessica. **Founders At Work: stories of startups early days**. Califórnia: Apress, 2007. Disponível em: https://link.springer.com/chapter/10.1007/978-1-4302-0327-8_4. Acesso em: 14 abr. 2022.

KUNY, Terry; CLEVELAND, Gary. The Digital Library: Myths and Challenges. **IFLA Journal**, [S. l.], v. 24, n. 2, p. 107-113, 1998. Disponível em: <http://archive.ifla.org/IV/ifla62/62-kuny.pdf>. Acesso em: 27 set. 2021.

LANCASTER, Frederick Wilfrid. **Indexação e resumos: teoria e prática**. 3. ed. Brasília: Briquet de Lemos, 2003.

LECARDELLI, Jane. **Dados científicos abertos em agências de fomento à pesquisa: cenário dos planos de gestão de dados (PGD) e princípios FAIR**. 2020. 1 recurso on-line (119 p.) Dissertação (Mestrado) - Universidade do Estado de Santa Catarina, Centro de Ciências Humanas e da Educação, Mestrado em Gestão da Informação, Florianópolis, 2020. Disponível em: <http://sistemabu.udesc.br/pergamumweb/vinculos/00008b/00008bdf.pdf>. Acesso em: 08 ago. 2022.

LOWE, Janet. **Google: lições de Sergey Brin e Larry Page, os criadores da empresa mais inovadora de todos os tempos**. Rio de Janeiro: Elsevier, 2009.

MAULDIN, Michael Loren. Lycos: Design choices in an internet search service. **IEEE Expert**, [S. l.], v. 12, n. 1, p. 8-11, 1997. Disponível em: <https://ieeexplore.ieee.org/abstract/document/577466>. Acesso em 02 abr. 2022.

MANOVICH, Lev. **The Language of New Media**. Cambridge: The Mit Press/ Leonardo Books, 2001.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. São Paulo: Atlas, 2017.

MARKEY, Karen. The online library catalog: paradise lost and paradise regained? **D-Lib Magazine**, Estados Unidos, v. 13, n. 1-2, 2007. Disponível em: <http://www.dlib.org/dlib/january07/markey/01markey.html>. Acesso em: 01 out. 2021.

MATOS, José Claudio Morelli. Que valor têm as minhas dúvidas? *In*: MATOS, José Claudio Morelli; BRITO, Evandro Oliveira (org.). **Valores: formação humana e o desenvolvimento da ação**. São José: Centro Universitário Municipal de São José, 2016.

MEME sobre o acesso aos resultados do Google durante uma pesquisa. [S. l.], 2012. Disponível em: <https://expedicaovida.com.br/o-desespero-dos-pos-graduandos/>. Acesso em: 15 set. 2021.

MOREIRA, Ana Gabriela Soares da Costa. **Elementos de história da lógica**. Porto, 2007, 74 p. Dissertação (Mestrado) – Universidade Portucalense Infante D. Henrique, Mestrado em Matemática/Educação. Disponível em: <http://repositorio.uportu.pt/handle/11328/523>. Acesso em 22 out. 2021.

MOREIRA, Vivian Lemes. ROMÃO, Lucília Maria Sousa. Discurso e buscadores na Web: sentidos heterogêneos em movimento. **Verso e Reverso**, São Leopoldo, v. 24, n. 56, p. 112-123, 2010. Disponível em: <http://revistas.unisinos.br/index.php/versoereverso/article/view/95>. Acesso em: 29 mar. 2022.

MORGAN, Gareth. Paradigms, Metaphors, and Puzzle Solving in Organization Theory. **Administrative Science Quarterly**, v. 25, n. 4, p. 605-622, 1980. Disponível em: <https://www.jstor.org/stable/2392283>. Acesso em 17 jul. 2021.

MORTARI, Cezar Augusto. **Introdução à lógica**. São Paulo: UNESP, 2001.

MURRAY, Janet. **Hamlet no Holodeck: o futuro da narrativa no ciberespaço**. Nova York: Free Press, 2016.

NONAKA, Ikujiro.; TAKEUCHI, Hirotaka. **Criação de conhecimento na empresa: como as empresas japonesas geram a dinâmica da inovação**. 10. ed. Rio de Janeiro: Editora Campus, 1997.

NUNES, André José; FREITAS, Filipe Gomes de. O não saber como princípio da sabedoria em Sócrates. **Frontistés Revista Eletrônica do Curso de Filosofia**, [S.l.], v. 12, n. 22, p. 1-12, 2018. Disponível em: <http://revistas.old.fapas.edu.br/index.php/frontistes/article/view/1247>. Acesso em: 24 nov. 2021.

O LIVRO DA FILOSOFIA. Tradução Douglas Kim. Porto Alegre: Globo, 2016.

O LIVRO DA MATEMÁTICA. Tradução Maria da Anunciação Rodrigues. Porto Alegre: Globo, 2020.

PACK, Thomas. Image Searching Across the Web. **Online**, New Jersey, v. 26, n. 1, p. 30–35, 2002. Disponível em: https://rnp-primo.hosted.exlibrisgroup.com/permalink/f/vsvpiv/TN_cdi_proquest_reports_199923429. Acesso em: 24 ago. 2022.

PICALHO, Antonio Carlos; LUCAS, Elaine Rosangela de Oliveira; AMORIM, Igor Soares. Lógica booleana aplicada na construção de expressões de busca. **AtoZ: novas práticas em informação e conhecimento**, Curitiba, v. 11, p. 1 - 12, 2022. Disponível em: <https://revistas.ufpr.br/atoz/article/view/81838>. Acesso em: 29 mar. 2022.

PRODANOV, Cleber Cristiano; FREITAS, Ernani Cesar de. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico**. Rio Grande do Sul: Feevale, 2013.

RENO, Janet. **Attorney General of the United States, et al., Appellants v. AMERICAN CIVIL LIBERTIES UNION et al.** Suprema Corte dos Estados Unidos, 1997. Disponível em: <https://www.law.cornell.edu/supremecourt/text/521/844>. Acesso em 30 set. 2021.

ROSA, Bruna Neder. **Publicidade e os limites da privacidade: percepção dos usuários a anúncios personalizados**. 2022. Dissertação (Mestrado em Tecnologias da Inteligência e Design Digital) – Programa de Pós-Graduação em Tecnologias da Inteligência e Design Digital, Pontifícia Universidade Católica de São Paulo, São Paulo, 2021. Disponível em: <https://tede.pucsp.br/bitstream/handle/25981/1/Bruna%20Neder%20Rosa.pdf>. Acesso em: 24 ago. 2021.

ROWLEY, Jennifer. **A biblioteca eletrônica**. 2. ed. Brasília: Lemos Informação e Comunicação, 2002.

SALTON, Gerard. **Dynamic Information and Library Processing**. Boston: Prentice Hall, 1975.

SANTOS, Luana Carla de Moura dos; GODOY VIERA, Angel Freddy. Avaliação da recuperação da informação em acervos digitais de jornais. **Em Questão**, Porto Alegre, v. 21, n. 2, p. 49–73, 2015. Disponível em: <https://seer.ufrgs.br/index.php/EmQuestao/article/view/52362>. Acesso em: 8 ago. 2022.

SAPO. Hoje o maior motor de busca de Portugal faz anos. E o maior do mundo também. **SAPO**, 2018. 1 imagem em formato PNG. Disponível em: <https://tek.sapo.pt/noticias/internet/artigos/hoje-o-maior-motor-de-busca-de-portugal-faz-anos-e-o-maior-do-mundo-tambem>. Acesso em: 11 abr. 2022.

SAYÃO, Luis Fernando. Afinal, o que é biblioteca digital? **Revista USP**, São Paulo, n. 80, p. 6-17, 2009. Disponível em: <https://www.revistas.usp.br/revusp/article/view/13709>. Acesso em: 22 set. 2021.

SEYMOUR, Tom; FRANTSVOG, Dean; KUMAR, Satheesh. History Of Search Engines. **International Journal of Management & Information Systems (IJMIS)**, [S. l.], v. 15, n. 4, p. 47–58, 2011. Disponível em: <https://clutejournals.com/index.php/IJMIS/article/view/5799>. Acesso em: 31 mar. 2022.

SHERMANN, Will. Are librarians totally obsolete? 33 reasons why libraries and librarians are still extremely important. **Teacher Librarian**, Estados Unidos, v. 35, n. 1, 2007. Disponível em: <https://link.gale.com/apps/doc/A170580781/AONE?u=capes&sid=bookmark-AONE&xid=f48cf00a>. Acesso em 01. out. 2021.

SILVA, Dhiogo Cardoso da. **Uma Arquitetura de business intelligence para processamento analítico baseado em tecnologias semânticas e em linguagem natural**. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, 2011. Disponível em: <https://repositorio.ufsc.br/handle/123456789/95772>. Acesso em 28 jul. 2021.

SIQUEIRA, Ivan Cláudio Pereira. Mecanismos de busca na web: passado, presente e futuro. **PontodeAcesso**, [S. l.], v. 7, n. 2, p. 47–67, 2013. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/6355>. Acesso em: 28 mar. 2022.

SMITH, Ernie. **Jumping-off point**. Tedium, Escócia, 2019. Disponível em: <https://tedium.co/2019/09/05/jumpstation-search-history/>. Acesso em: 02 abr. 2022.

SMITH, Robin. Lógica. *In*: BARNES, Jonathan (org.). **Aristóteles**. Aparecida: Ideias & Letras, 2009. p. 59-101.

SOBRAL, João Bosco Mangueira. **Da computabilidade formal às máquinas programáveis**. Florianópolis, Edição do autor, 2015. Série Pensamento Matemático.

Disponível em: <https://repositorio.ufsc.br/xmlui/handle/123456789/157317>. Acesso em: 30 mar. 2022.

SOUSA, Richard Perassi Luiz de. **Mídia do conhecimento**: ideias sobre mediação e autonomia. Florianópolis: SIGMO/UFSC, 2019. 111 p. Disponível em: <https://sigmo.paginas.ufsc.br/files/2019/05/Midia-do-Conhecimento-LIVRO-FINAL.pdf>. Acesso em: 01 out. 2022.

SOUZA, Giselle Costa de. **Uma reavaliação do pensamento lógico de George Boole à luz da história da matemática**. 2005. 319 f. Dissertação (Mestrado em Educação) - Universidade Federal do Rio Grande do Norte, Natal, 2005. Disponível em: <https://repositorio.ufrn.br/handle/123456789/14224>. Acesso em: 25 out. 2021.

SPUDEIT, Daniela Fernanda Assis de Oliveira. **Diretrizes para o desenvolvimento da competência em informação em bibliotecas públicas com foco nas pessoas em situação de rua**. 423 p. Tese (Doutorado) – Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Programa Pós-Graduação em Ciência da Informação, Florianópolis, 2021. Disponível em: <https://repositorio.ufsc.br/handle/123456789/231102>. Acesso em: 08 ago. 2022.

STATCOUNTER. **Search engine market share worldwide**. Dublin, 2022. Disponível em: <https://gs.statcounter.com/search-engine-market-share/all/brazil#monthly-200901-202107>. Acesso em: 01 set. 2021.

THE SCOTSMAN. **Googling was born in Stirling**. The Newsroom, Escócia, 2009. Disponível em: <https://www.scotsman.com/news/googling-was-born-stirling-2444060>. Acesso em 02 abr. 2020.

TIME INC. Memex. Life, Estados Unidos, v. 19, n. 11, 1945. Disponível em: https://books.google.com.br/books?redir_esc=y&hl=pt-BR&id=uUkEAAAAMBAAJ&q=memex#v=snippet&q=memex&f=false. Acesso em: 29 mar. 2022.

TOTH, Pedro Henrique. **A evolução comunicativa dos mecanismos de busca**: do telégrafo à web semântica. 2017. 162 f. Dissertação (Mestrado em Comunicação Social) – Universidade Metodista de São Paulo, São Bernardo do Campo, 2017. Disponível em: <http://tede.metodista.br/jspui/handle/tede/1672#preview-link0>. Acesso em: 09 nov. 2021.

TOUTAIN, Lídia Maria Batista Brandão. Biblioteca digital: definição de termos: *In*: MARCONDES, Carlos Henrique; KURAMOTO, Hélio; TOUTAIN, Lídia Maria Batista Brandão; SAYÃO, Luis Fernando (org.). **Bibliotecas digitais**: saberes e práticas. Salvador: EDUFBA; Brasília: IBICT, 2005. p. 15-24. Disponível em: <http://livroaberto.ibict.br/handle/1/1013>. Acesso em 01. out. 2021.

UFSC. Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento. Linhas de pesquisa. [2017]. Disponível em: <https://ppgegc.paginas.ufsc.br/linhas-de-pesquisa/>. Acesso em: 23 nov. 2022.

VERA-PÉREZ, Buenaventura Loreto; LUGO-ORTIZ, Sérgio. Matriz de consistencia metodológica. *Ciencia Huasteca Boletín Científico de la Escuela Superior de Huejutla*, v. 4, n. 8, 2016. Disponível em:

https://repository.uaeh.edu.mx/revistas/index.php/huejutla/article/download/318/4703?inline=1&as_qdr=y. Acesso em 28 mar. 2022.

VIDMAR, Dale James. Darwin on the web: the evolution of search tools. *Computers in Libraries*, New Jersey, v. 19, n. 5, p. 1-5, 1999. Disponível em:

<https://link.gale.com/apps/doc/A54720543/AONE?u=capes&sid=bookmark-AONE&xid=19a935e1>. Acesso em: 22 ago. 2022.

VIEIRA, Eduardo. **Os bastidores da Internet no Brasil**: as histórias de sucesso e de fracasso que marcaram a Web brasileira. Barueri: Manole, 2003.

WAZLAWICK, Raul Sidnei. **Metodologia de pesquisa para ciência da computação**. Elsevier, 2021.

WEB DESIGN MUSEUM. Aliweb, 1996. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/gallery/aliweb-1995>. Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. AltaVista. 1999. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/timeline/altavista-1999>. Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. Archie, 1996. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/gallery/archie-query-form-1996>. Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. Ask Jeeves, 1999. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/gallery/ask-jeeves-1999>. Acesso em: 11 abr. 2022.

WEB DESIGN MUSEUM. Excite, 1996. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/gallery/excite-1996>. . Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. Google no início, 1996. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/timeline/google-1998>. Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. Gopher, 1996. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/web-design-history/gopher-1991>. Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. Infoseek, 1997. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/gallery/infoseek-1997>. Acesso em: 11 abr. 2022.

WEB DESIGN MUSEUM. Lycos, 1996. 1 imagem em format PNG. Disponível em:

<https://www.webdesignmuseum.org/gallery/lycos-1996>. Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. Veronica search engine, 1992. 1 imagem em format PNG. Disponível em: <https://www.webdesignmuseum.org/web-design-history/veronica-search-engine-1992>. Acesso em: 01 abr. 2022

WEB DESIGN MUSEUM. W3 Catalog, 1995. 1 imagem em format PNG. Disponível em: <https://www.webdesignmuseum.org/gallery/cui-w3-catalog-1995>. Acesso em: 01 abr. 2022

WEB DESIGN MUSEUM. WebCrawler, 1996. 1 imagem em format PNG. Disponível em: <https://www.webdesignmuseum.org/gallery/webcrawler-1996>. Acesso em: 28 mar. 2022.

WEB DESIGN MUSEUM. Yahoo, 1997. 1 imagem em format PNG. Disponível em: <https://www.webdesignmuseum.org/timeline/yahoo-1997>. Acesso em: 28 mar. 2022.

WU, Xindong; ZHU, Xingquan; WU, Minghui. The evolution of search: three computing paradigms. *ACM Trans. Manage. Inf. Syst.* v. 13, n. 2, 1-20, 2022. Disponível em: <https://doi.org/10.1145/3495214>. Acesso em: 04 set. 2022.

YUS, Francisco. En qué difieren los géneros electrónicos de los analógicos. En el Congreso Leer y Escribir en Español en la Red. Aprendizaje, Renovación y Propuestas. University of Comillas, Cantabria (Espanha), 2011. Disponível em: https://www.researchgate.net/publication/265733107_En_que_difieren_los_generos_electronicos_de_los_analogicos. Acesso em: 29 mar. 2022.

ZHAO, Ying. The Application of Bayesian Learning in the Search Engine. *Wireless Personal Communications*, v. 103, n. 2, p. 1121–1131, 2018. Disponível em: <https://doi.org.ez46.periodicos.capes.gov.br/10.1007/s11277-018-5256-x>. Acesso em: 01 abr. 2022.

ÍNDICE ONOMÁSTICO

- A**BAR, Celina – 33
 AMORIM, Igor Soares – 35, 36
 ARAÚJO, Andréa Nunes de – 68
- B**ARNES, Jonathan – 33
 BERINSTEIN, Paula – 64
 BERNERS-LEE, Tim – 37
 BIANCHINI, Monica – 66
 BLANCHÉ, Robert – 34
 BRANSKI, Regina Meyer – 26
 BUCHINGER, Diego – 45
 BUCKINGHAM, David – 38
 BUSH, Vannevar – 42
- C**AVALCANTI, Cordélia Robalinho de
 Oliveira – 38
 CAVALCANTI, Gustavo Andriolli de
 Siqueira – 45
 CENDÓN, Beatriz Valadares 45, 68
 CHAUI, Marilena – 29, 30, 34
 CHOO, Chun Wei – 21
 CLEVELAND, Ana – 68
 CLEVELAND, Gary – 36
 CLEVELAND, Donald – 68
 CORREA, Sara Dotta – 76
 COUVERING, Elizabeth Van – 54, 55, 61
 CUNHA, Murilo Bastos da – 38
- D**’OTTAVIANO, Ítala Maria Loffredo – 36
 DAVIS, Bob – 53
 DESCARTES, René – 30
 DOERNHOEFER, Mark – 51, 52, 55, 57
 DUBUCS, Jacques – 34
- F**ADEL, Luciane Maria – 76
 FAGUNDES, Vinícius Bossle – 26
 FARIA, Gustavo Henrique Campos de – 76
 FEITOSA, Hércules de Araújo – 36
 FRANTSVOG, Dean – 46, 47, 48, 49, 51,
 54, 55, 56
 FREIRE, Patrícia de Sá – 79
 FREITAS, Ernani Cesar de – 19, 79
 FREITAS, Felipe Gomes de – 30
- G**ABRIEL, Martha Carrer Cruz – 45
 GANDAL, Neil – 61
 GIL, Antônio Carlos - 79
 GODOY, Arnaldo, Sampaio de Moraes – 29,
 30
 GODOY VIERA, Angel Freddy – 21
 GORI, Marco – 66
 GOULART, Elias Estevão – 34
- H**ENDLER, James – 37
 HETEM JUNIOR, Annibal – 36

HOUNSELL, Marcelo da Silva – 45

INCE, Darrel – 49

KOEHLER, Wallace – 36

KRAUS, Joe – 54

KUMAR, Satheesh – 46, 47, 48, 49, 51, 54, 55, 56

KUNY, Terry – 36

LAKATOS, Eva Maria – 80

LANCASTER, Frederick Wilfrid – 67, 68

LASSILA, Ora – 37

LECARDELLI, Jane – 90

LOWE, Janet – 62, 63, 65

LUCAS, Elaine Rosangela de Oliveira – 35, 36

LUGO-ORTIZ, Sérgio – 82

MAULDIN, Michael Loren – 52

MANOVICH, Lev – 19, 22, 75, 76

MARCONI, Marina de Andrade – 80

MARKEY, Karen – 38

MATOS, José Claudio Morelli – 30

MOREIRA, Ana Gabriela Soares da Costa – 34, 35

MOREIRA, Vivian Lemes – 46, 47, 48

MORGAN, Gareth – 80

MORTARI, Cezar Augusto – 33, 34

MURRAY, Janet – 42, 75, 77

NONAKA, Ikujiro – 27

NUNES, André José – 30

PACK, Thomas – 64

PICALHO, Antonio Carlos – 35, 36

PRODANOV, Cleber Cristiano – 19, 79

RENO, Janet – 36

ROMÃO, Lucília Maria Sousa – 46, 47, 48

ROSA, Bruna Neder – 69, 70

ROWLEY, Jennifer – 45

SALTON, Gerard – 43

SANTOS, Luana Carla de Moura dos – 21

SAYÃO, Luis Fernando – 21

SCARSELLI, Franco – 66

SEYMOUR, Tom – 46, 47, 48, 49, 51, 54, 55, 56

SHERMANN, Will – 37

SILVA, Dhiogo Cardoso da – 26

SIQUEIRA, Ivan Cláudio Pereira – 66

SMITH, Ernie – 49

SMITH, Robin – 33

SOBRAL, João Bosco Mangueira – 35

SOUSA, Richard Perassi Luiz de – 41, 75

SOUZA, Giselle Costa de – 32

SPUDEIT, Daniela Fernanda Assis de Oliveira – 38

TAKEUCHI, Hirotaka – 27
TOTH, Pedro Henrique – 35, 43
TOUTAIN, Lídia Maria Batista Brandão –
37

VAZ, Carlos Eduardo Verzola – 76
VERA-PÉREZ, Buenaventura Loreto – 82
VIDMAR, Dale James – 22
VIEIRA, Eduardo – 58

WAZLAWICK, Raul Sidnei – 79
WU, Minghui – 19, 23
WU, Xindong – 19, 23

YUS, Francisco – 67

ZHAO, Ying – 56
ZHU, XINGQUAN – 19, 23