

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO (UFRJ)  
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS (CCJE)  
FACULDADE DE ADMINISTRAÇÃO E CIÊNCIAS CONTÁBEIS (FACC)  
CURSO DE BIBLIOTECONOMIA E GESTÃO DE UNIDADE DE INFORMAÇÃO (CBG)

**MARIANA COSTA VELOZO VITAL**

A INFLUÊNCIA DOS ALGORITMOS DE RECOMENDAÇÃO NAS PESQUISAS  
REALIZADAS NA WEB

Rio de Janeiro

2021

MARIANA COSTA VELOZO VITAL

**A INFLUÊNCIA DOS ALGORITMOS DE RECOMENDAÇÃO NAS PESQUISAS  
REALIZADAS NA WEB**

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia e Gestão de Unidades de Informação da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de bacharel em Biblioteconomia e Gestão de Unidades de Informação.

Orientadora: Profª Nysia Oliveira de Sá

Rio de Janeiro

2021

### Ficha catalográfica

Vital, Mariana Costa Velozo.

A influência dos algoritmos de recomendação nas pesquisas realizadas na web / Mariana Costa Velozo Vital. – Rio de Janeiro, 2021.

39 f.

Trabalho de Conclusão de Curso (graduação) - Universidade Federal do Rio de Janeiro. Curso de Biblioteconomia e Gestão de Unidades de Informação. Rio de Janeiro, 2021.

Orientadora: Nysia Oliveira de Sá.

1. Filtro bolha. 2. Rastros informacionais. 3. Algoritmos de recomendação. I. Sá, Nysia Oliveira de. II. Título

**MARIANA COSTA VELOZO VITAL**

**A INFLUÊNCIA DOS ALGORITMOS DE RECOMENDAÇÃO NAS PESQUISAS  
REALIZADAS NA WEB**

Trabalho de Conclusão de Curso apresentado ao Curso de Biblioteconomia e Gestão de Unidades de Informação da Universidade Federal do Rio de Janeiro, como requisito parcial à obtenção do título de bacharel em Biblioteconomia e Gestão de Unidades de Informação.

Rio de Janeiro, 01 de junho de 2021.

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Nysia Oliveira de Sá – CBG/UFRJ  
Orientadora

---

Prof.<sup>a</sup> Dr.<sup>a</sup> Mariana Zattar – CBG/UFRJ  
Membro interno

---

Prof. Dr.<sup>a</sup> Raimunda Fernanda dos Santos – CBG/UFRJ  
Membro interno

Dedico esse trabalho à Maria da Luz, minha mãe, à minha tia Maria de Jesus, e ao Celso, meu padrasto, que acreditaram em mim quando nem eu mesma acreditei.

## **AGRADECIMENTOS**

Quando comecei a fazer o TCC imaginei que seria fácil escrever os agradecimentos, mas isso se mostrou quase tão difícil quanto começar esse trabalho. Encerrar essa pesquisa é a constatação de que uma das fases mais intensas da minha vida está chegando ao fim, e isso me causa um misto de emoções.

Quando entrei na UFRJ, em 2017, nem em meus pensamentos mais aleatórios e diversos consegui dimensionar tudo o que viveria na graduação: alegrias, tristezas, pensamentos como “meu Deus do céu, me ajuda!” e “calma que vai dar tudo certo” sempre me passavam pela cabeça, e entre trabalhos, provas, estágios, marmitas, mais provas e mais trabalhos, cheguei até o fim (meu Deus, é verdade mesmo!).

Em primeiro lugar agradeço a Deus pela realização desse sonho que é me formar na faculdade que eu sempre sonhei, no curso que eu escolhi e amei cursar, mesmo com todas as dificuldades que o ensino público enfrenta. Sem a força e o sustento d’Ele, e sem as pessoas que Ele colocou em meu caminho para me ajudar, essa trajetória teria sido bem mais difícil.

À minha mãezinha que eu amo tanto, que sempre acreditou em mim mesmo quando eu mesma já não acreditava, dedico toda a intensidade e grandeza do meu amor e da minha gratidão. Ela é a minha base, meu porto seguro, minha maior incentivadora. Agradeço por ela sonhar os meus sonhos comigo, por estar presente mesmo com a distância geográfica que nos separa, pelas ligações que diziam “falta pouco, Mari, a maior parte já passou. Você vai conseguir!”. Te amo infinitamente! Essa conquista também é dela, é por ela, para ela.

À minha tia Maria de Jesus, que é minha segunda mãe, que sempre esteve comigo em todos os momentos e que me ama como filha, dedico meu sincero amor, minha sincera gratidão, e agradeço a Deus por tê-la na minha vida. Ao lado da minha mãe ela sempre lutou por mim, e não há nada no mundo que se compare a tamanho amor.

Ao Celso, meu padrasto, agradeço ao apoio e por sempre estar presente, por me zoar e por acreditar em mim.

À Nysia, minha orientadora, agradeço a imensa paciência que teve comigo, agradeço por cada orientação, por cada sugestão de alteração no texto. Obrigada por me mostrar que o TCC não é um bicho de sete cabeças, que as coisas podem ser mais simples do que a gente imagina, e por ter acreditado na minha capacidade quando eu não fazia a menor ideia do que estava fazendo.

Aos meus amigos que me apoiaram durante toda a graduação, que ouviram minhas reclamações, que me levavam para espairer, meu muito obrigada. Fernando, Débora, Erick, Filipo, e todos que fizeram parte dessa trajetória, que comemoraram comigo e permaneceram junto a mim quando eu estava mal, vocês são incríveis!

À minha turma, 2017.1 da Praia Vermelha, obrigada por cada momento no C.A., por cada conversa no Campinho, por cada lanche no Rio Sul no final do período. Em toda a minha trajetória escolar/acadêmica essa foi a turma com que eu mais me identifiquei.

Em especial agradeço à Thamires Oliveira, à Anna Clara e à Patrícia Pereira por, desde o primeiro período, serem o grupo mais perfeito em que eu poderia estar. Obrigada por me colocarem à par das matérias quando eu estava perdida, das datas de provas e dos trabalhos; obrigada pela paciência comigo quando eu estava com sono e por acreditarem no meu compromisso de sempre entregar a minha parte do trabalho, mesmo começando a fazê-lo domingo à noite. Sem vocês a faculdade não teria sido a mesma.

Em geral, agradeço por cada professor que fez parte da minha formação, que mesmo não tendo sido citado nesse trabalho faz parte dessa realização. Obrigada por cada ensinamento, por cada incentivo e por cada direcionamento. Cada um colaborou para que eu chegasse até aqui, e a eles sou muito grata.

Encerro agradecendo, também, a mim, por não ter desistido e por ter conseguido terminar a faculdade mesmo morando em outro estado, longe da família e num ambiente totalmente novo.

“Cada vez mais, nos sentimos tão seguros envolvidos em nossas bolhas que só aceitamos informações, verídicas ou não, que se enquadram com nossas opiniões, em vez de basear nossas opiniões nas evidências existentes.”

(Barack Obama)



## RESUMO

Apresenta histórico do desenvolvimento tecnológico ao longo dos anos, bem como o desenvolvimento dos buscadores na web e sua evolução. Discorre sobre as influências dos algoritmos de recomendação e tem como objetivo geral identificar se esses algoritmos impactam no resultado das buscas realizadas nos provedores. Objetiva especificamente analisar o resultado dessas buscas; identificar as discrepâncias oriundas das mesmas e propôr recomendações para minimizar os efeitos dos algoritmos de dados. Tem como metodologia a pesquisa bibliográfica de abordagem qualitativa. Os resultados finais apontam que de acordo com o perfil do usuário os retornos de suas pesquisas sofrem influência dos algoritmos de recomendação. Sugere que sejam utilizados outros buscadores nas pesquisas realizadas na web, bem como mecanismos que bloqueiam as coletas de dados realizadas pelos algoritmos presentes na web.

**Palavras-chave:** Filtro bolha; Rastros informacionais; Algoritmos de busca; Algoritmos de recomendação; Mecanismos de busca – Web; Google.

## **ABSTRACT**

Presents a history of technological development over the years, as well as the development of web searches and their evolution. Discusses the influences of recommendation algorithms and has the general objective of identifying whether these algorithms impact the results of searches performed on providers. Specifically aims to analyze the results of these searches; identify discrepancies arising from them and propose recommendations to minimize the effects of data algorithms. Methodology is bibliographic research with a qualitative approach. The final results indicate that, according to the user's profile, the recommendation algorithms influence the returns of their research. Suggests that other search engines be used in searches carried out on the web, as well as mechanisms that block data collections carried out by the algorithms present on the web.

**Keywords:** Bubbler filter; Informational tracks; Recommendation algorithms; Search algorithms; Search engines – Web; Google.

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b> –	Página inicial do Google.....	28
<b>Figura 2</b> –	Print do participante 1.....	30
<b>Figura 3</b> –	Print do participante 2.....	31
<b>Figura 4</b> –	Print do participante 3.....	31
<b>Figura 5</b> –	<i>Print</i> 1 das sugestões de autocompletar dos participantes da pesquisa...	32
<b>Figura 6</b> –	<i>Print</i> 2 das sugestões de autocompletar dos participantes da pesquisa...	33
<b>Figura 7</b> –	<i>Print</i> 3 das sugestões de autocompletar dos participantes da pesquisa...	33

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>13</b>
1.1	PROBLEMA.....	16
1.2	OBJETIVOS.....	16
1.3	JUSTIFICATIVA.....	16
1.4	OBJETO DE PESQUISA.....	18
<b>2</b>	<b>DATAÍSMO.....</b>	<b>19</b>
2.1	COMPETÊNCIA EM INFORMAÇÃO.....	20
2.2	RASTROS INFORMACIONAIS.....	21
2.3	BOLHAS INFORMACIONAIS.....	22
2.4	FONTES DE INFORMAÇÃO.....	24
<b>3</b>	<b>PROCEDIMENTOS METODOLÓGICOS.....</b>	<b>27</b>
3.1	CAMPO DA PESQUISA.....	27
3.2	TÉCNICAS DE COLETA E ANÁLISE DE DADOS.....	28
3.3	POPULAÇÃO/AMOSTRA.....	28
<b>4</b>	<b>ANÁLISE DE DADOS.....</b>	<b>30</b>
<b>5</b>	<b>CONSIDERAÇÕES FINAIS.....</b>	<b>34</b>
	<b>REFERÊNCIAS.....</b>	<b>36</b>

## 1 INTRODUÇÃO

Ao longo da história da humanidade o ser humano sempre foi visto como figura central de todo o desenvolvimento que ocorria. Foram suas ideias, invenções e descobertas que trouxeram a sociedade ao patamar intelectual e tecnológico em que se encontra atualmente. A tecnologia foi criada e desenvolvida, desde sua maneira mais rudimentar até a mais recente, para servir a humanidade, melhorar e possibilitar a qualidade de vida dos indivíduos independente da área em que é aplicada.

É possível acompanhar a importância e a influência da tecnologia na vida do ser humano através das grandes revoluções, como a Revolução Agrícola, que originou o desenvolvimento das técnicas de cultivo, permitindo que o ser humano se fixasse no território e deixasse de ser nômade. A Revolução Industrial, também, foi o evento que desencadeou o aprimoramento e a criação de novas tecnologias que mudaram completamente a vida do ser humano, como o desenvolvimento da energia elétrica e a vapor, o início das indústrias e a produção em massa de produtos industriais.

Apesar de a Revolução Industrial ter ocorrido no século XVIII suas influências perduram hodiernamente, acrescidas de novas tecnologias criadas de maneira cada vez mais veloz, que culminaram na principal revolução que tem mudado, ainda mais, o modo cada pessoa viver e se comunicar: a Revolução Tecnológica.

A Revolução Tecnológica, que teve início na década de 1970 após a Segunda Guerra Mundial e durante a Guerra Fria, irrompeu em inúmeras inovações científicas devido ao número exponencial de pesquisas realizadas naquele período, sendo a mais marcante a difusão da engenharia genética (CASTELLS, 1999).

Outro fator marcante dessa época foi o crescimento acelerado da produção de informação científica, o surgimento dos computadores e a expansão de seu uso durante as décadas subsequentes, além do compartilhamento de informações que, anteriormente, não eram disponibilizadas para a população. Todos esses agentes, principalmente o advento da Internet em 1969, resultaram na sociedade da informação e deram início ao modo de vida atual, que está intrinsecamente ligado às mídias digitais e redes sociais via Internet.

Segundo Castells, “a criação e o desenvolvimento da Internet nas três últimas décadas do século XX foram consequência de uma fusão singular de estratégia militar, grande cooperação científica, iniciativa tecnológica e inovação contracultural”(1999, p.82). A Internet passou por diversas mudanças e aprimoramentos até chegar ao formato pela qual é

conhecida, tornando-se global após a criação da teia mundial WWW (*world wide web*), oferecendo ao usuário um sistema fácil de pesquisas para encontrar a informação desejada (CASTELLS, 1999).

A partir da década de 1980 o uso da Internet cresceu e se popularizou, passando por transformações tecnológicas “que converteram o computador pessoal em computador coletivo, conectado ao ciberespaço” (CALAZANS; LIMA, 2013, p. 2). O desenvolvimento de aparelhos móveis que permitem ao indivíduo permanecer conectado mesmo sem estar presos a fios e tomadas “levou as sociedades a um processo de conexão cada vez mais generalizada, de maneira que estar conectado passou a ser um fator preponderante do nosso cotidiano” (CALAZANS; LIMA, 2013, p. 2). Ainda nessa perspectiva, Jenkins (2008) reflete que

a convergência dos meios de comunicação impacta o modo como consumimos esses meios. Um adolescente fazendo a lição de casa pode trabalhar ao mesmo tempo em quatro ou cinco janelas no computador: navegar na internet, ouvir e baixar arquivos MP3, bater papo com amigos, digitar um trabalho e responder e-mails, alternando rapidamente as tarefas (JENKINS, 2008, p. 42).

Contudo, apenas em 1994 a Internet passou a ser comercializada, e com ela surgiu o conceito que, posteriormente, foi categorizado como Web 1.0 “isto é, quando se usava a internet basicamente como fonte de consulta para encontrar dados disponibilizados por sites de empresas e instituições, e também para a comunicação via e-mail” (GOMES; PASSARELLI, 2020, p. 257). Nesse período o usuário era apenas um “mero espectador da informação, sem nenhum tipo de interferência no fluxo informacional disponibilizado em rede ou qualquer tipo de participação ativa na produção de conteúdos online” (GOMES; PASSARELLI, 2020, p. 257). Assim, as informações consumidas pelos usuários eram as publicações eletrônicas de especialistas da área de tecnologia da informação.

Com o passar dos anos surgiu a necessidade, por parte do usuário, de interagir com as informações disponíveis em rede, iniciando assim a Web 2.0, também conhecida como Web social, onde o indivíduo deixa de ser um espectador da informação e passa a produzi-la, “oportunizando a fase de criação de conteúdo livre” (GOMES; PASSARELLI, 2020, p. 259). Um dos fatores marcantes da Web 2.0 foi o surgimento das mídias sociais, que são locais no ciberespaço onde o usuário, agora ativo, aloca a informação por ele produzida. O ciberespaço, segundo Lévy é:

O espaço de comunicação aberto pela interconexão mundial dos computadores e das memórias dos computadores. Essa definição inclui o conjunto dos sistemas de comunicação eletrônicos (aí incluídos os conjuntos de redes hertzianas e telefônicas clássicas), na medida em que transmitem informações. Consiste de uma realidade multidirecional, artificial ou virtual incorporada a uma rede global, sustentada por computadores que funcionam como meios de geração de acesso (LÉVY, 2009, p. 92).

Com a contínua evolução tecnológica da Internet e a convergência dos meios de comunicação surge a Web 3.0 também conhecida como Web semântica, com mudanças ainda mais significativas na maneira dos indivíduos se comunicarem, pois organiza e sistematiza semanticamente os ambientes digitais, tornando mais precisos os resultados das buscas dos recursos informacionais nesse ambiente.

Importante destacar aqui que ainda não se vive a era da Web 3.0, os profissionais que atuam em ambientes digitais e os pesquisadores têm unido esforços para desenvolver estudos, instrumentos, processos, plataformas e sistemas inteligentes, dotados de semântica.

O autor Marcondes (2012), por exemplo, evidencia que, apesar de a cobertura dos atuais mecanismos de busca seja ampla, esses sistemas ainda apresentam ineficácia no que concerne à precisão e à semântica.

De acordo com Gomes e Passarelli, após essa mudança

surge, então, o conceito da web semântica, que inaugura um processo mais complexo na utilização da ferramenta. A World Wide Web (rede mundial) se tornaria a World Wide Database (base de dados mundial), isto é, um mar de dados em que programas entendam como fazer o melhor uso desses dados por meio da inteligência artificial (IA) e do machine learning (máquinas que aprendem). Do ponto de vista dos internautas, a organização dos conteúdos e informações por IA significa que os programas serão capazes de melhor entender as necessidades individuais, fornecendo um conteúdo personalizado, seja ele uma notícia, um diagnóstico, um site de pesquisa, uma loja ou até mesmo uma propaganda (2020, p. 265).

Contudo, tal avanço chegou a um ponto no qual a tecnologia criada está começando a ultrapassar as fronteiras do artificial, tomando uma espécie de consciência, ainda que de maneira primária, das vontades e desejos de todas as pessoas que têm acesso à Internet.

Além disso, segundo Castells (1999 apud NEGROPONTE, 1995):

vivemos em um mundo que se tornou digital, pois o processo atual de transformação tecnológica expande-se exponencialmente em razão de sua capacidade de criar uma interface entre campos tecnológicos mediante uma linguagem digital comum na qual a informação é gerada, armazenada, recuperada, processada e transmitida.

Dessa maneira, novas correntes científicas irrompem para entender e explicar o fenômeno da interação cada vez mais próxima entre humano e máquina, e como a segunda influencia na primeira. Segundo Harari, “como a inteligência está se desacoplando da consciência, e como a inteligência não consciente está se desenvolvendo a uma velocidade vertiginosa” (2016, p. 355), os seres humanos correm o risco de ficarem ultrapassados diante da “mente tecnológica” que os tecno-humanistas almejam alcançar, excluindo toda influência humana e deixando apenas que a inteligência algorítmica comande o futuro e as decisões tomadas.

Os algoritmos se configuram como uma sequência lógica finita de tarefas executáveis que visam obter uma solução para um determinado problema. Se constituem como a base para o processo de desenvolvimento de *software*, e o seu uso em todos os meios virtuais acessados faz com que essa sequência numérica saiba mais de nós do que nós mesmos (GILLESPIE, 2018), e quanto mais um indivíduo lança seus desejos e opiniões nas redes, mais informações os algoritmos armazenam sobre cada um.

Dito isto, a autora trará a abordagem presente na literatura acerca desse assunto e fará a análise comparativa dos resultados das buscas solicitadas aos participantes desta pesquisa para apresentar os resultados encontrados, ou seja, mostrar se o retorno de nossas buscas na Internet sofrem influências de todo o nosso histórico de pesquisas realizadas anteriormente.

## 1.1 PROBLEMA

A utilização de mecanismos que criam uma internet personalizada para cada usuário produz um ambiente alienante, pois em vez de apresentarem a pluralidade de assuntos referentes ao resultado de uma busca, mostram apenas temas relacionados às buscas anteriormente realizadas pelos usuários, dificultando o aumento de seu conhecimento sobre o assunto pesquisado.

Devido a este fato surgiu o seguinte questionamento: Os algoritmos de busca impactam no resultado das pesquisas dos usuários?

## 1.2 OBJETIVOS



Objetivo geral: identificar se os algoritmos impactam no resultado das pesquisas realizadas pelos usuários em mecanismos de busca como o Google.

Objetivos específicos:

- Analisar o resultado das buscas realizadas no Google;
- Identificar as possíveis discrepâncias oriundas das buscas;
- Propor recomendações para minimizar os efeitos dos algoritmos de buscas.

### 1.3 JUSTIFICATIVA

O tema da pesquisa surgiu após a autora ter participado do IV Fórum Sobre Competência em Informação: pesquisas e práticas no Rio de Janeiro, evento coordenado por professoras do Curso de Biblioteconomia e Gestão de Unidades de Informação da Universidade Federal do Rio de Janeiro (UFRJ) para promoção da Competência em Informação (CoInfo). Após a exposição sobre o Dataísmo, surgiu na autora o questionamento de como o ser humano pode ser considerado um conjunto de algoritmos, um “divíduo” e não um indivíduo, visto que ele é, segundo essa nova corrente científica, a montagem de muitos algoritmos diferentes. Essa premissa justifica o fato de um algoritmo externo poder, teoricamente, conhecer uma pessoa muito melhor do que ela mesma (informação verbal).<sup>1</sup>

Além disso, gerou-lhe incômodo o resultado do questionamento das ciências biológicas às premissas do liberalismo, da liberdade humana, uma vez que argumenta que os algoritmos que formam o ser humano não são livres, mas sim configurados por genes e pressões ambientais, tomando decisões aleatoriamente, mas não livremente (informação verbal).<sup>2</sup>

Dito isto, manifestou-se na autora a necessidade de entender, através do uso dos indivíduos às redes sociais e mídias digitais, como os algoritmos não-humanos influenciam, na Internet, as informações que os “algoritmos humanos” acessam diariamente.

Assim, tornou-se importante a necessidade de conscientização dos indivíduos com acesso à Internet sobre a influência dos algoritmos, na hora de escolher uma informação, nas decisões tomadas, tanto nas individuais, do dia a dia, quanto nas que definem o futuro da sociedade como um todo, à exemplo do resultado da eleição do ex-presidente dos Estados Unidos, Donald Trump, contra sua rival Hillary Clinton, em 2016.

---

<sup>1</sup> Informação concedida no IV Fórum Sobre Competência em Informação: pesquisas e práticas no Rio de Janeiro, no Rio de Janeiro, em outubro de 2018.

<sup>2</sup> Informação concedida no IV Fórum Sobre Competência em Informação: pesquisas e práticas no Rio de Janeiro, no Rio de Janeiro, em outubro de 2018.

À época, segundo a Revista Exame, a empresa Cambridge Analytica - responsável pela campanha digital de Trump – coletou do Facebook, sem o consentimento dos usuários, dados privados de 87 (oitenta e sete) milhões de pessoas e, posteriormente, utilizou estes dados para enviar aos usuários publicidade política especialmente adaptada para cada indivíduo, de acordo com seus interesses expostos nessa rede social, e elaborar informes detalhados para ajudar Trump a ganhar a eleição.

Dados da pesquisa TIC Domicílios (2019), mais importante levantamento sobre acesso às tecnologias da informação e comunicação, realizada pelo Centro Regional para o Desenvolvimento de Estudos sobre a Sociedade da Informação, mostram que em 2019, 134 (cento e trinta e quatro) milhões de brasileiros já possuíam acesso à internet, e a maioria dessas pessoas, ou seja, 99%, usa *smartphones* para se conectarem à web.

Ademais, é sabido que a tecnologia evoluiu de tal forma que não há maneiras de retroceder (PASSARELLI, 2018, apud LATOUR, 1994) e evitar que soframos influências de sistemas especializados em captar dados deixados ao navegar pela internet, por isso torna-se essencial que o usuário tenha noção de que, em linhas gerais, o resultado de suas buscas são determinados pelo que os algoritmos entendem como útil para ele.

Um dos principais papéis do(a) bibliotecário(a) é auxiliar e orientar o usuário na busca da informação pertinente (ACCART, 2012), dando-lhe autonomia para encontrar essa informação e escolhê-la de acordo com sua necessidade informacional. Assim, o indivíduo torna-se competente para decidir o que é melhor para ele, o(a) profissional, diante dessa nova realidade da influência dos algoritmos nas pesquisas dos usuários, tem a oportunidade de incentivá-lo a buscar novos mecanismos para acessar essas informações sem expor excessivamente seus dados aos algoritmos presentes na rede.

Dessa forma, faz-se importante para o campo um novo olhar para a busca e recuperação da informação e para os buscadores utilizados pelos usuários sob a perspectiva da influência dos algoritmos no retorno de suas pesquisas. Assim, esse projeto terá como instrumento de pesquisa o Google, ferramenta de busca mais utilizada ao redor do mundo, segundo o site TechTudo.<sup>3</sup>

Portanto, essa nova realidade exigirá do profissional bibliotecário, e do usuário, maior atenção aos resultados de suas pesquisas, tendo sempre a consciência de que seus retornos são frutos de uma internet personalizada exclusivamente para eles.

Além disso, esse fato também pode tirar o usuário de sua zona de conforto e fazê-lo

---

<sup>3</sup> Disponível em <https://www.techtudo.com.br/listas/2019/02/melhor-que-google-veja-nove-sites-de-busca-rivais-com-funcoes-diferentes.ghtml>. Acesso em 25 de fevereiro de 2021.

procurar meios mais seguros, para a sua privacidade, de obter informação.

#### 1.4 OBJETO DE PESQUISA

O objeto de pesquisa do presente trabalho são os motores de busca na Web, especificamente o Google.

## 2 DATAÍSMO

O Dataísmo se origina de duas ciências: a Biologia e a Ciência da Computação (HARARI, 2015). Ele classifica o ser humano como um algoritmo bioquímico e o define como uma sequência de algoritmos, que tendo sua ordem modificada, pode se assemelhar a outras coisas existentes. É uma forma de desumanização para que apenas a inteligência pura e “perfeita” dos algoritmos seja considerada.

Segundo a definição de Harari (2015, p.370), o Dataísmo "consiste num fluxo de dados e o valor de qualquer fenômeno ou entidade é determinado por sua contribuição ao processamento de dados". Essa nova “religião” dos dados estabelecida por Harari surge num contexto em que o número de informações disponíveis ultrapassou exponencialmente a capacidade humana de processá-las, tornando impossível deixar essa função sob a responsabilidade de seres biologicamente intelectuais e colocando esses processamentos nas mãos de mecanismos de construção de inteligências artificiais: os algoritmos computacionais.

Segundo Gillespie, “os algoritmos não são necessariamente softwares: em seu sentido mais amplo, são procedimentos codificados que, com base em cálculos específicos, transformam dados em resultados desejados” (2018, p. 97). Na era atual, os dados têm mais valor do que qualquer outra substância, como o petróleo, por exemplo, e a sua utilização para todo tipo de tomada de decisão objetiva alcançar maior exatidão nos resultados, economizar o tempo para obter uma resposta e economizar recursos humanos, tecnológicos e financeiros. E esses dados servem como base para que os algoritmos possam aprender automaticamente a partir dos mesmos.

Dessa maneira, com a visão de que as leis matemáticas se aplicam tanto aos organismos bioquímicos quanto aos artificiais, Harari afirma que o Dataísmo tira a barreira que havia entre animais e máquinas (HARARI, 2015, p.370), tornando as pessoas e os objetos tecnológicos uma mera sequência de algoritmos.

Em suma, os dataístas são céticos no que diz respeito ao conhecimento e à sabedoria humanos e preferem acreditar em megadados e algoritmos computacionais a acreditar na capacidade humana para processar, avaliar e escolher, dentre as milhares de informações disponíveis, qual é relevante para uma determinada situação.

Independente de a inteligência humana ser capaz, ou não, de analisar e processar os dados disponíveis, é importante que os indivíduos sejam habilitados para discernir as informações que recebem dos diversos transmissores que os cercam, sejam eles digitais ou analógicos, para que não sejam influenciados por informações tendenciosas provenientes de fontes não confiáveis e, também, por algoritmos enviesados, haja vista que serviços baseados em Inteligência Artificial são amplamente utilizados em nosso cotidiano e tendem a ter esses vieses.

## 2.1 COMPETÊNCIA EM INFORMAÇÃO

O termo competência em informação (do inglês *information literacy*) surgiu em 1974, no relatório *The information service environment relationships and priorities*, do bibliotecário americano Paul Zurkowski. A expressão, que no início “sugeriu que os recursos informacionais deveriam ser aplicados às situações de trabalho, na resolução de problemas, por meio do aprendizado de técnicas e habilidades no uso de ferramentas de acesso à informação” (DUDZIAK, 2003, p.24) sofreu modificações ao longo das décadas.

A Competência em Informação já foi definida como

o processo contínuo de internalização de fundamentos conceituais, atitudinais e de habilidades necessário à compreensão e interação permanente com o universo informacional e sua dinâmica, de modo a proporcionar um aprendizado ao longo da vida (DUDZIAK, 2003, p.28).

Dessa maneira, esperava-se que o indivíduo competente em informação fosse capaz de

identificar sua real necessidade informacional, identificar e manusear fontes potenciais de informação de forma efetiva e eficaz, avaliar criticamente a informação segundo critérios de relevância, objetividade, pertinência, lógica, ética, incorporando as informações selecionadas ao seu próprio sistema de valores e conhecimentos, além de usar e comunicar a informação (DUDZIAK, p.29, 2003).

No entanto, o conceito de competência em informação evoluiu e se atualizou com o passar dos anos e, dentre as inúmeras definições para esse termo, no presente trabalho

destaca-se o termo defendido por Zattar, (2017 apud ASSOCIATION OF COLLEGE AND RESEARCH LIBRARIES, 2015), que diz que a competência em informação se constitui em um

conjunto de habilidades integradas que abrangem a descoberta reflexiva da informação, a compreensão de como a informação é produzida e valorizada, e o uso da informação na criação de novos conhecimentos para atuação de forma ética em comunidades de aprendizagem.

Desse modo vemos que o indivíduo competente em informação tem ciência da importância da informação e a usa não apenas para si mesmo, mas também em benefício da comunidade da qual faz parte.

O indivíduo competente em informação, através de suas habilidades de identificar todos os aspectos que envolve uma informação, como fonte ou meio em que está localizada, não fica completamente vulnerável aos mecanismos de coleta de dados que atuam em sites na web, pois mesmo que seja difícil identificá-los, tem consciência de que precisa se atentar ao que é retornado em suas buscas e, também, ao que aparece como sugestão em anúncios, devido a influência dos algoritmos sobre o que é pesquisado.

Toda pesquisa realizada na internet deixa uma espécie de “pegada digital”, que é um rastro informacional do que foi buscado, quantas vezes aquela busca foi realizada e com qual frequência, formando um perfil dos hábitos e preferências daquele indivíduo (CALAZANS; LIMA, 2013).

## 2.2 RASTROS INFORMACIONAIS

Ao acessar a internet, seja para uma pesquisa ou com a finalidade de entrar em uma rede social, rastros dos interesses, desejos e vontades de cada pessoa são deixados nos sites e recolhidos por algoritmos que, em posse dessas informações, podem induzir o comportamento de uma pessoa.

Um exemplo claro é que, provavelmente, já deve ter acontecido com muitas pessoas, ocorre quando realiza-se uma pesquisa sobre determinado produto no site de uma loja específica e, “de repente”, anúncios desse mesmo produto começam a aparecer em outras páginas acessadas pelo usuário e que não tinham relações com o primeiro endereço web pesquisado, induzindo à compra, visto que os mecanismos de coletas de dados possuem o perfil dos desejos das pessoas, aumentando a possibilidade de concretizar esse interesse em compra (CALAZANS; LIMA, 2013).

Outro exemplo desse fato acontece nas redes sociais quando são visualizadas muitas publicações sobre um assunto e, a partir daí, todo o conteúdo oferecido para o usuário daquela conta é relacionado apenas ao universo pesquisado.

É necessário dizer que, no cenário atual, não há como retroceder a uma era na qual os dados e a tecnologia não influenciam a vida e as decisões das pessoas, todavia é importante que cada indivíduo saiba que existe um algoritmo de recomendação por trás de cada resultado de pesquisa que lhe mostra apenas aquilo que é do seu maior interesse (GILLESPIE, 2018). Cada vez mais torna-se necessária a conscientização de que nem tudo o que retorna para um usuário das redes é o que ele quer, mas sim o que os algoritmos entendem como sua necessidade informacional.

Essas sequências matemáticas estão ficando mais cientes dos gostos e desejos de uma pessoa do que ela mesma, e se tornam capazes de prever ações baseadas nas informações disponibilizadas, inconscientemente, pelo usuário. Assim sendo, fica cada vez mais difícil saber se a informação que uma pessoa consome é algo que ela quer ou se é uma sugestão dos algoritmos (CALAZANS; LIMA, 2013).

Outro problema que surge na questão da coleta de dados realizada pelos algoritmos é a invasão da privacidade do indivíduo e a questão ética, pois as informações disponibilizadas em uma pesquisa não são deliberadamente entregues a esses mecanismos de busca, mas sim recolhidas sem a permissão do usuário.

Gillespie (2018, p. 102) afirma que

os provedores também se beneficiam do caráter cada vez mais participativo da web, em que usuários são encorajados a oferecer todo tipo de informações sobre si mesmos e a se sentirem empoderados ao fazê-lo. À medida que nossas micro-práticas migram cada vez mais para essas plataformas, torna-se sedutor para os provedores de informação (apesar de não ser compulsório) rastrear essas atividades e transformá-las em mercadoria de várias formas.

Contudo, quando bem utilizados, esses dados têm a capacidade de entregar resultados mais certos, economizando recursos importantes como o tempo do usuário, e possibilitando um rápido acesso, em meio a explosão informacional, às informações relevantes. Se empregados com o objetivo de trazer melhorias ao interesse comum, processos que são demorados seriam resolvidos de maneira rápida e prática.

O problema acontece quando grandes empresas usam indevidamente os dados pessoais de cada pessoa para seus próprios interesses, como no caso da Cambridge Analytica, que usou

mineração de dados para influenciar nas eleições americanas de 2016 e no processo de separação do Reino Unido da União Europeia.

Além disso, de acordo com Gillespie (2018, p. 101) “os algoritmos [...] são os padrões de inclusão nos bancos de dados que pré-determinam o que vai ou não aparecer entre seus resultados”. Esse fato mostra que, além de invadir a privacidade dos usuários, os algoritmos excluem o acesso de informações que julgam não serem interessantes para eles, criando bolhas informacionais que isolam os indivíduos a apenas o que é de seu agrado, de acordo com seu histórico de pesquisas.

### 2.3 BOLHAS INFORMACIONAIS

“Bolha informacional” é o termo criado pelo ativista da internet Eli Pariser, em 2010, para identificar o estado de isolamento intelectual resultante de pesquisas personalizadas, quando um algoritmo de site adivinha seletivamente quais informações um usuário gostaria de ver com base nas informações sobre o mesmo, tal como localização, comportamento de cliques anteriores e histórico de pesquisa.

Devido ao fato de os algoritmos retornarem, numa busca, apenas o que eles acreditam ser mais relevante para cada pessoa, criam-se bolhas informacionais. Essas bolhas são preenchidas apenas com os interesses do indivíduo, de acordo com as informações que ele deixou em rede, e acaba se tornando um círculo no qual nada que é desagradável entra, mas apenas informações que confirmam o que ele já acredita (GILLESPIE, 2018). A pessoa não tem controle do que entra na bolha e, também, não enxerga o que fica do lado de fora.

Essa situação não é saudável, pois gera uma alienação na qual o indivíduo só recebe informação daquilo que é do seu interesse, e para uma pessoa se tornar competente em informação é necessário saber de tudo o que acontece, até mesmo o que não lhe é agradável.

Gillespie (2018, p.114) aponta que “somos levados - por algoritmos e por nossa própria preferência pelos que pensam de forma semelhante - para dentro de ‘filtros bolha’, onde encontramos apenas as notícias que esperamos encontrar e as perspectivas políticas que já nos são caras”.

Ter acesso a opiniões diferentes é fundamental para o crescimento intelectual e social de uma pessoa, pois a troca de conhecimento enriquece o saber dos que tem a mente aberta para respeitar e entender o que lhe é diferente. É saudável uma introdução às novas ideias, pessoas e perspectivas, e tal ação ajuda a diminuir discursos de ódio e situações violentas que a sociedade atual vive devido as diferenças de cada ser.

O ideal seria que houvesse um equilíbrio nos resultados das pesquisas, onde o que é visto com mais frequência pelo usuário e os assuntos menos pesquisados por ele fossem retornados com a mesma relevância.

O controle da informação sempre existiu, desde quando essa se encontrava, majoritariamente, em mídias impressas. A questão levantada é a ética por trás desse controle para evitar filtros gerados por algoritmos tendenciosos manipularem as decisões dos usuários. Os algoritmos, pelo menos por enquanto, ainda não possuem senso de responsabilidade moral.

Somando-se a isso, o resultado de pesquisa de duas pessoas distintas, buscando o mesmo termo, pode ser completamente diferente, pois o resultado trazido é de acordo com cada pesquisa realizada anterior e individualmente, de acordo com Gillespie

com as ferramentas de busca contemporâneas, os resultados que dois usuários recebem para a mesma busca podem ser bem diferentes; em um serviço de notícias ou em uma rede social, as informações oferecidas podem ser adaptadas especificamente às preferências do usuário (pelo usuário ou o pelo provedor), de modo que, na prática, as histórias apresentadas, como as mais importantes, podem ser tão distintas de usuário para usuário que não exista sequer um objeto comum de diálogo entre elas (2018, p. 114).

Com o passar dos anos e com o avanço da tecnologia, o modo de buscar uma informação na Web passou por diversas transformações até chegar ao modo pelo qual é atualmente conhecido, com buscadores online que retornam para o usuário milhares de resultados, em poucos segundos, para a sua pesquisa.

Sendo assim, surgiram cada vez mais buscadores que, posteriormente, foram denominados como fontes de informação, e tais fontes podem ser classificadas de diversas maneiras de acordo com o conteúdo informacional que detém, sendo todas elas importantes para o acesso à informação.

## 2.4 FONTES DE INFORMAÇÃO

Fontes de informação são qualquer documento que forneça informações solicitadas por um usuário ou pessoa que esteja pesquisando num catálogo online ou num banco de dados bibliográficos (ALVES; SANTOS, 2018, p. 38) e, além disso, se refere a um documento que fornece informações copiadas ou reproduzidas em outro documento, por exemplo, uma



citação ou um trecho. Ainda, as fontes de informação são classificadas em quatro tipos: primárias, secundárias, terciárias e quaternárias.

É importante ressaltar que apesar de haver essa divisão na classificação das fontes, elas, muitas vezes, se misturam e se apresentam tanto em sua definição própria quanto em outra que, supostamente, não teria a ver com ela, havendo diferentes definições, de diferentes autores, acerca da designação dessas fontes.

Posteriormente, com a evolução da tecnologia, surgiram novas fontes que produziram uma grande transformação na produção e no acesso ao conhecimento (ALVES; SANTOS, 2018, p. 40), fontes estas que, além do formato analógico, também, começaram a ser disponibilizadas em formato digital, como os mecanismos de busca, facilitando o acesso e a consulta pelo usuário.

O primeiro mecanismo de busca surgiu junto com a Internet, “possibilitando o usuário dessa imensa rede acessar milhares de informações dispersas em páginas iniciais (*homepages*)” (CUNHA, 2001, p. 98).

Daí surgiu o termo “surfar” na rede, pois para encontrar alguma informação o usuário fazia “uma busca sem estrutura e caminhos definidos. O usuário acessava uma determinada página conhecida e, a partir de hipervínculos (*links*), ia ‘navegando’ até encontrar a informação desejada” (CUNHA, 2001, p. 98).

Dessa forma,

o surgimento dos diretórios eletrônicos, também conhecidos como mecanismos ou motores de busca (*search engines*), abriu uma nova perspectiva para melhorar a qualidade da informação recuperada na Internet. Eles funcionam como um repositório eletrônico de informações (CUNHA, 2001, p.98).

Em virtude disso, ao longo dos anos foram surgindo mais e mais buscadores, sofrendo modificações e aperfeiçoamentos até que chegassem ao modo como são conhecidos atualmente, sendo o Google o principal e mais utilizado a nível mundial.

De acordo com Batti (2015, p. 19) a palavra “Google” é proveniente de “um trocadilho feito com o termo matemático “Googol”: número representado pelo dígito um seguido de cem dígitos zeros. A utilização da expressão traduz a vontade de Larry Page e Sergey Brin, os fundadores do Google, de organizar uma volume muito grande de dados na internet” .

Atualmente o Google é o principal buscador utilizado pelos usuários da web, e usa ferramentas de *machine learning* (aprendizado de máquina) que conferem aos códigos de um conjunto de algoritmos a capacidade de autoajuste e autocorreção, de modo a se adaptarem e revisarem seus resultados a partir de ações anteriores. O conjunto de algoritmos do Page Rank

do serviço de buscas do Google faz com que em cada interação do usuário na plataforma ou no buscador, seus algoritmos aprendem quais conteúdos seriam de maior relevância ou interesse dos usuários (BENTES; BRUNO; FALTAY, 2019).

Em relação à privacidade do usuário a plataforma já recebeu da *Privacy International*, em 2007, a classificação de “hostil à privacidade”, devido à grande quantidade de dados que coleta sobre os usuários e suas atividades, segundo a BBC News.

Atualmente, de acordo com as políticas de privacidade do Google, a plataforma recolhe de seus usuários as seguintes informações: termos pesquisados, vídeos assistidos, visualizações e interações com conteúdo e anúncios, informações de voz e áudio quando tal recurso é utilizado, atividade de compra, pessoas com quem você se comunica ou compartilha conteúdo, atividades em sites e *apps* de terceiros que usam serviços do Google, histórico de navegação do Chrome sincronizado com a Conta do Google, sendo que as informações coletadas pelo Google e como essas informações são usadas dependem de como o serviço é utilizado pelo indivíduo e de como ele gerencia os controles de privacidade<sup>4</sup>. A plataforma justifica a coleta de todas essas informações como meio de oferecer serviços personalizados para cada usuário.

De acordo com Butti (2015), quando fazemos uma pesquisa na internet usando o motor de busca Google, não buscamos em toda a rede, e sim, nas páginas “pré-cadastradas” pelo Google que estão armazenadas em seu banco de dados.

Complementarmente, Oliveira (2014, p. 52, grifo nosso) ressalta que “o Google utiliza dois critérios: “On Page” e “Off Page”. O “On Page” basicamente **usa as informações que a página disponibiliza para classificar a sua relevância, verificando quais palavras e/ou expressões da busca estão contidas nela**. Já o critério “Off Page” usa informações fora da página para classificar sua relevância, que é medida através das **indicações de outras páginas sobre a página em questão**. A importância de uma página é medida pelo “PageRank”, que é um tipo de algoritmo de busca que vasculha e classifica as páginas na rede.

Conclui-se, assim, que o Google possui informações suficientes de seus usuários para lhes oferecer conteúdos “sob medida” para cada um, mediante seus rastros informacionais deixados ao navegar pela web, criando espaços – bolhas – onde apenas as informações que são do agrado do usuário podem entrar.

---

<sup>4</sup> Informação retirada da Política de Privacidade do Google.

### 3 PROCEDIMENTOS METODOLÓGICOS

Para ser realizada, uma pesquisa científica precisa de técnicas, abordagens e métodos para ser bem desenvolvida. São os métodos científicos que garantem a veracidade e a qualidade da pesquisa para que tenha credibilidade. Nessa seção serão apresentados os procedimentos metodológicos para a realização do projeto de pesquisa proposto.

O estudo caracteriza-se como pesquisa exploratória, documental e de abordagem qualitativa e de campo bibliográfico, visando esclarecer o problema de pesquisa citado anteriormente.

Segundo Gil (2008), a pesquisa exploratória visa “desenvolver, esclarecer e modificar conceitos e ideias” e é realizada quando o tema ainda é pouco explorado, como no caso do Dataísmo, que é estudado nesse trabalho.

Já a pesquisa qualitativa segundo Minayo (2009, p. 21):

[...] responde a questões muito particulares. Ela se ocupa, nas Ciências Sociais, com um nível de realidade que não pode ou não deveria ser quantificado. Ou seja, ela trabalha com o universo dos significados, dos motivos, das aspirações, das crenças, dos valores e das atitudes. Esse conjunto de fenômenos humanos é entendido aqui como parte da realidade social, pois o ser humano se distingue não só por agir, mas por pensar sobre o que faz e por interpretar suas ações dentro e a partir da realidade vivida e partilhada com seus semelhantes.

O campo da pesquisa é bibliográfico, ou seja, foi desenvolvido através de fontes bibliográficas secundárias, nesse caso, artigos e livros. As buscas foram realizadas em fontes nacionais e internacionais, sendo elas Brapci, Scielo, Google Acadêmico, Peri, Re3Data e Portal Capes, entre os meses de dezembro de 2020 e março de 2021. Além disso, os slides disponibilizados pelos professores ao longo do curso também foram utilizados como fonte de consulta.

As palavras-chave pesquisadas foram: “dataísmo”, “algoritmos + sistemas de informação”, “pegadas digitais”, “competência em informação”, “filtro bolha” e “tecnologia da informação e comunicação”, sendo que a maioria dos resultados foi encontrada no Google Acadêmico.

É importante dizer que devido ao assunto ser novo há escassez na literatura nacional acerca do Dataísmo, com poucos artigos publicados ou trabalhos acadêmicos realizados tendo esse tema como abordagem principal. Por isso, o levantamento bibliográfico do referencial teórico foi realizado a partir de temas que se relacionam com o assunto e se encaixam com o

objetivo da pesquisa, sem fugir do foco, que é a influência dos algoritmos nas buscas dos indivíduos na Internet.

Portanto, a presente pesquisa baseia-se nos artigos encontrados, em livros e na leitura de partes do livro *Homo Deus: uma breve história do amanhã*, cujo o autor, Harari, discorre sobre o passado da humanidade até sua chegada à atualidade.

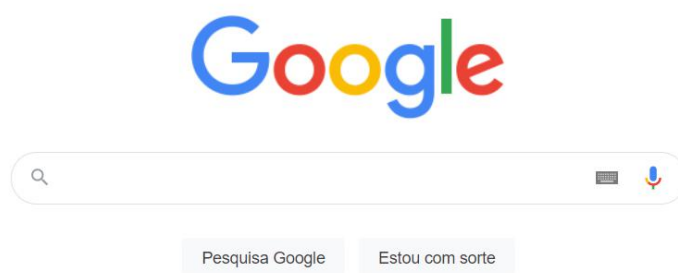
Além disso, foram utilizados artigos encontrados em sites que abordam temas como tecnologia e inovação, e que buscam entender e demonstrar as mudanças que o Dataísmo ocasionará no modo de vida das pessoas.

### 3.1 CAMPO DA PESQUISA

O campo da pesquisa são os mecanismos de busca na web, e o Google foi o buscador escolhido, conforme explicado anteriormente, por ser a ferramenta de pesquisa mais conhecida e utilizada mundialmente, segundo o site de tecnologia TechTudo, permitindo a recuperação das informações buscadas pelos usuários de maneira intuitiva, rápida e prática.

O Google foi fundado como empresa privada em 4 de setembro de 1998, com a missão de "organizar a informação mundial e torná-la universalmente acessível e útil". Apesar de o objetivo inicial da plataforma ter sido funcionar como motor de busca, seu rápido crescimento culminou em uma cadeia de outros produtos.<sup>5</sup>

**Figura 1** – Página inicial do Google



Fonte: Google.

---

<sup>5</sup> Informação retirada da Wikipédia, referência no final do trabalho.

Os fundadores do Google, Larry Page e Sergey Brin, criaram em 1998, enquanto estavam na Universidade de Stanford, Califórnia, o algoritmo PageRank, utilizado pela ferramenta de busca Google para posicionar websites entre os resultados de suas buscas, e tem a função de “medir a importância de uma página contabilizando a quantidade e qualidade de links apontando para ela”, e tal fato faz com que algumas páginas específicas apareçam sempre no topo da lista de sites pesquisados.

### 3.2 TÉCNICA DE COLETA E ANÁLISE DE DADOS

O projeto foi realizado da seguinte maneira: a autora escolheu uma expressão de busca e pediu, devido a necessidade de isolamento social e seguindo os protocolos de biosegurança emitidos pelas OMS devido a pandemia do COVID – 19, via WhatsApp, para que seus contatos mais próximos, seus amigos, buscassem o termo no Google.

Após a realização da procura da expressão de busca no Google, utilizando o campo “busca simples” e logados em sua conta do Gmail, a autora pediu para as pessoas tirarem *prints* de suas telas contendo o resultado de suas buscas, e encaminharem, via Google Forms, esses *prints* para ela, que fez análise comparativa dos *prints* e expôs se o retorno das pesquisas feitas com o mesmo termo traz resultados iguais ou diferentes.

A expressão de busca utilizada pelos participantes da pesquisa foi “governo brasileiro e aquecimento global”, e o questionário – que teve a utilidade de saber a idade dos respondentes – ficou disponível durante o período de 28 de abril de 2021 até 08 de maio de 2021.

A solicitação foi para que apenas estudantes da graduação respondessem a pesquisa, com a faixa etária de jovens e adultos de acordo com o IBGE, e quinze pessoas responderam a pesquisa.

### 3.3 POPULAÇÃO/AMOSTRA

A população escolhida como público-alvo do projeto foram jovens e adultos estudantes de graduação de diversos cursos, e o motivo dessa escolha foi o fato de a autora, por estar concluindo o ensino superior, ter mais contato com esse grupo de pessoas e se enquadrar na definição de jovem que, de acordo com o IBGE (Instituto Brasileiro de

Geografia e Estatística), inclui a população com a faixa etária entre 15 anos e 24 anos, enquanto a população entre 25 anos e 59 anos é considerada adulta.<sup>6</sup>

Partindo do pressuposto que jovens e adultos usam mais a Internet e, conseqüentemente, utilizam o Google como principal buscador, esse público foi escolhido por se encaixar no grupo alvo da pesquisa.

Além disso, escolheu-se, também, estudantes de graduação para participarem da pesquisa devido sua demanda acadêmica, pressupondo que seus rastros informacionais têm relação com o curso da graduação e, também, com seus interesses pessoais.

---

<sup>6</sup> Disponível em <https://www.ibge.gov.br/estatisticas/sociais/populacao/9292-populacao-jovem-no-brasil.html?=&t=o-que-e>. Acesso em 24 de abril de 2021.

## 4 ANÁLISE DE DADOS

Após o público-alvo ter respondido a pesquisa e enviado o *print* com o resultado de sua busca no Google com a expressão “governo brasileiro e aquecimento global”, foi feita a análise de dados para verificar se os resultados seriam discrepantes ou semelhantes.

O perfil dos respondentes é composto por pessoas entre 18 e 30 anos, todas cursantes do ensino superior e acostumadas a utilizar recursos eletrônicos para realizar suas pesquisas. Para esta pesquisa todos os participantes utilizaram o Google como buscador, conforme exemplificado no referencial teórico.

As seguintes notícias retornaram nas pesquisas dos respondentes: o negacionismo do aquecimento global no governo brasileiro (12 retornos); governo brasileiro participa de reunião com negacionistas do aquecimento global (8 retornos); Ministério do Meio Ambiente quase zera verba de combate à mudança climática (8 retornos); enfrentando o aquecimento global e as mudanças climáticas (6 retornos); Brasil chega à cúpula do clima sob cobrança e descrédito (4 retornos); como o Brasil vai contribuir para conter a crise do clima (2 retornos); o que dizem os bolsonaristas que negam o aquecimento global (1 retorno); a difícil missão do governo brasileiro na cúpula do clima (1 retorno).

Após análise dos *prints* verificou-se que não houve discrepâncias entre os resultados, pois todas as notícias acima retornaram de maneira recorrente nas pesquisas dos respondentes, com exceção da última e da penúltima, que retornaram apenas uma vez. A única diferença entre as telas dos participantes está na ordem em que as notícias apareceram, uma vez que elas não seguiram a mesma ordem de apresentação, como exemplificado nas imagens abaixo.

**Figura 1 – Print do participante 1**



Fonte: retorno da pesquisa feita com os participantes.

**Figura 2 – Print do participante 2**



Fonte: retorno da pesquisa feita com os participantes.

**Figura 3 – Print do participante 3**



Fonte: retorno da pesquisa feita com os participantes.



Depreende-se que os resultados foram semelhantes devido ao fato de o perfil do grupo analisado ser muito próximo – estudantes de graduação de instituição federal com faixa etária aproximada –, confirmando as bolhas informacionais que envolvem usuários com opiniões e vivências parecidas, além da relação com a lógica adotada no *PageRank*, uma vez que ele trabalha com o ranqueamento das páginas de busca.

Além disso, após receber *prints* com as sugestões de autocompletar que o Google indicou aos usuários, a autora percebeu que algumas dessas indicações foram diferentes para cada usuário, conforme as imagens abaixo.

**Figura 4** – *Print* 1 das sugestões de autocompletar dos participantes da pesquisa.



Fonte: retorno da pesquisa feita com os participantes.

**Figura 5** - *Print* 2 das sugestões de autocompletar dos participantes da pesquisa.



Fonte: retorno da pesquisa feita com os participantes.

**Figura 6 - Print 3** das sugestões de autocompletar dos participantes da pesquisa.



Fonte: retorno da pesquisa feita com os participantes.

Essa brusca alteração de tela para tela dos usuários, em relação ao que o Google sugere para eles, confirma a atuação dos algoritmos de busca e de recomendação, indicando uma Internet com resultados personalizados para cada usuário de acordo com suas pesquisas anteriores.

## 5 CONSIDERAÇÕES FINAIS

O trabalho teve o intuito de apresentar como os algoritmos de coleta de dados e os algoritmos de recomendação influenciam as futuras pesquisas dos usuários da web, tendo o Google como buscador utilizado para elaborar a pesquisa.

O panorama tecnológico atual mostra como os buscadores online evoluíram, permitindo que milhares de informações sejam recuperadas em poucos segundos, possibilitando uma ampla gama de fontes de pesquisa para seus usuários. Devido a essa evolução, novos mecanismos de aprimoramento da pesquisa foram desenvolvidos até chegar aos algoritmos de coleta de dados que personalizam os resultados das buscas de cada indivíduo.

Os resultados obtidos através da análise dos *prints* enviados pelos participantes da pesquisa mostraram que, de acordo com a similaridade do perfil (faixa etária, escolaridade, local onde mora, ideias) de quem utiliza o Google como buscador, os retornos de suas pesquisas são semelhantes, pois seus rastros informacionais deixados na web montam resultados personalizados para cada um.

A pesquisa teve como objetivo geral identificar se os algoritmos de coleta de dados impactam no resultado de busca realizadas nos provedores, fato que se comprovou verídico após a análise dos resultados enviados pelos respondentes. A similaridade entre os perfis dos entrevistados impossibilitou um resultado discrepante na pesquisa devido a expressão de busca utilizada, porém as sugestões de autocompletar registradas indicam que, de acordo com cada indivíduo, o Google apresenta diferentes sugestões para seus usuários.

Os objetivos específicos também foram alcançados, pois as discrepâncias oriundas das buscas foram identificadas e as recomendações para minimizar o efeito dos algoritmos de dados foram propostos, como utilização de outros buscadores online, como o DuckDuckGo, entre outros. Além disso, pode-se utilizar, também, programas que possuem bloqueadores de algoritmos de coleta de dados e bloqueadores de anúncios.

Além disso, o papel do(a) bibliotecário(a) mostra-se fundamental na formação em competência informacional do indivíduo, pois através dela torna-se possível discernir sobre a informação recuperada, analisando criticamente as características que envolvem o conteúdo informacional, como a fonte na qual está inserida. Ainda, surgem novos desafios a serem enfrentados pelo profissional frente aos ambientes algoritmicamente modelados, tal como reconhecer os mecanismos por trás dos sistemas de recuperação da informação e, também, conscientizar as pessoas do problema que é ter seus dados expostos para terceiros.

A falta dessa competência em informação na formação de um indivíduo o deixa vulnerável em meio a diversas informações falsas e conteúdos tendenciosos presentes na web, tornando-o propenso a ser influenciado de forma demasiada (e inconscientemente) pelos filtros-bolhas presentes na Internet.

Sendo assim, conclui-se que a pesquisa atingiu seu objetivos e, também, sugere que o desenvolvimento dessa temática seja utilizado em futuros trabalhos acadêmicos, pois o tema é muito amplo e abre possibilidades para inúmeras pesquisas com diferentes abordagens, como uma comparação entre os resultados de pesquisas entre dois ou mais buscadores diferentes utilizando a mesma expressão de busca, ou uma pesquisa realizada com participantes com perfis mais divergentes.

## REFERÊNCIAS

- ALVES, Fernanda Maria Melo; SANTOS, Bruno Almeida dos. **Fontes e recursos de informação tradicionais e digitais**: propostas internacionais de classificação. Biblios [online]. 2018, n.72, pp.35-50. ISSN 1562-4730. . Disponível em: [http://www.scielo.org.pe/scielo.php?script=sci\\_arttext&pid=S156247302018000300003&lng=en&nr=iso](http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S156247302018000300003&lng=en&nr=iso). Acesso em 18 mar. 2021
- BATTI, João Carlos Bez. **Um pouco da matemática por trás do algoritmo PageRank do Google**. 2015. Dissertação (Mestrado em matemática)- Universidade Federal de Santa Catarina, 2015.
- BRUNO, F. G.; BENTES, A. C. F.; FALTAY, P. Economia psíquica dos algoritmos e laboratório de plataforma: mercado, ciência e modulação do comportamento. **Revista FAMECOS**, v. 26, n. 3, p. e33095, 27 dez. 2019. Disponível em: <https://doi.org/10.15448/1980-3729.2019.3.33095>. Acesso em: 20 mar. 2021.
- CAMBRIDGE Analytica se declara culpada por uso de dados do Facebook: Rede social informou que assessoria política da campanha digital de Trump usou aplicativo para coletar informações de 87 milhões de usuários. [S. l.], 9 jan. 2019. Disponível em: <https://exame.com/tecnologia/cambridge-analytica-se-declara-culpada-por-uso-de-dados-do-facebook/>. Acesso em: 20 fev. 2021.
- CASTELLS, Manuel. **A sociedade em rede**. 8. ed. v. 1. São Paulo: Paz e Terra, 1999.
- COSTA, Carlos. **Dataísmo, a religião do século XXI**: Google, Facebook, LinkedIn, Wikipedia e até os sites de meteorologia e o Youtube começam a ser de maior confiança e rapidez do que os antigos métodos de consulta e análise. [S. l.], 1 ago. 2018. Disponível em: <https://cio.com.br/tendencias/dataismo-a-religiao-do-seculo-xxi/>. Acesso em: 11 fev. 2021.
- CUNHA, Murilo Bastos da. **Para saber mais**: fontes de informação em ciência e tecnologia. Brasília: Briquet de Lemos, 2001.
- CUNHA, M.B.; CAVALCANTI, C.R. de O. **Dicionário de Biblioteconomia e Arquivologia**. Brasília: Briquet de Lemos, 2008.
- DUDZIAK, Elisabeth Adriana. **Information literacy**: princípios, filosofia e prática. Ciência da Informação, Brasília, v.32, n.1, p. 23-35, jan./abr. 2003
- FILTRO bolha. [S. l.], 12 fev. 2021. Disponível em: [https://en.wikipedia.org/wiki/Filter\\_bubble](https://en.wikipedia.org/wiki/Filter_bubble). Acesso em: 17 fev. 2021.
- GIL, Antônio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.
- GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2006.
- GILLESPIE, Tarleton. **The relevance of algorithms**. In: Media Technologies: Essays on Communication, Materiality, and Society. MIT Press, 2014. Disponível em: [https://www.researchgate.net/publication/281562384\\_The\\_Relevance\\_of\\_Algorithms](https://www.researchgate.net/publication/281562384_The_Relevance_of_Algorithms). Acesso em: 20 abr. 2021.
- GOOGLE. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2021. Disponível em: <<https://pt.wikipedia.org/w/index.php?title=Google&oldid=60941384>>. Acesso em: 18 abr. 2021.

GOOGLE classificado como o pior em privacidade: O Google tem a pior política de privacidade de empresas populares na Internet, diz um relatório. [S. l.], 11 jul. 2007. Disponível em: <http://news.bbc.co.uk/2/hi/technology/6740075.stm>. Acesso em: 21 mar. 2021.

HARARI, Yuval Noah. **Homo Deus: uma breve história do amanhã**. São Paulo: Companhia das Letras, 2016.

HOMO deus, de Yuval Noah Harari: como a informação vai destruir a liberdade humana. [S. l.], 17 out. 2016. Disponível em: <https://www.blogdacompanhia.com.br/conteudos/visualizar/Homo-deus-de-Yuval-Noah-Harari-como-a-informacao-vai-destruir-a-liberdade-humana#:~:text=No%20%C3%A2mago%20desse%20livro%20fascinante,o%20que%20se%20chama%20consci%C3%A2ncia>. Acesso em: 11 fev. 2021.

JENKINS, Henry. **Cultura da Convergência**. São Paulo: Aleph, 2008.

LIMA, Cecília Almeida de Lima; CALAZANS, Janaina de Holanda Costa. **Pegadas digitais: “Big Data” e informação estratégica sobre o consumidor**. Simpósio em Tecnologias Digitais e Sociabilidade. Salvador, 2013.

MARCONDES, Carlos Henrique. Em busca de uma semântica do digital, ou “as they may think”. **PontodeAcesso**, Salvador, v. 6, n. 2, p. 35-73, dez. 2012.

MATTEU, Douglas de. **O QUE É DATAÍSMO?**. [S. l.], 10 fev. 2020. Disponível em: <https://gennegociosegestao.com.br/o-que-e-dataismo/>. Acesso em: 11 fev. 2021.

MINAYO, M. C. de S. O desafio da pesquisa social. In: MINAYO, M. C. de S. (Org.). **Pesquisa social: teoria, método e criatividade**. 28. ed. Petrópolis, RJ: Vozes, 2009.

OLIVEIRA, José Carlos Francisco de. Noções de **Grafos dirigidos, cadeias de Markov e as buscas do Google**. 2014. Dissertação (Mestrado em Matemática profissional)- Universidade Federal de Sergipe, São Cristóvão, 2014.

RIBEIRO, Daniel. **Melhor que Google? Veja nove sites de busca rivais com funções diferentes**: Conheça um pouco mais sobre alguns dos principais concorrentes do famoso mecanismo de busca. [S. l.], 6 fev. 2019. Disponível em: <https://www.techtudo.com.br/listas/2019/02/melhor-que-google-veja-nove-sites-de-busca-rivais-com-funcoes-diferentes.ghtml>. Acesso em: 22 fev. 2021.

ROJO, Verónica López. Dominados por un ejército de algoritmos. **La razón histórica**, volume, número 39, p. 28-32, ano 2018. Disponível em: [www.revistalarazonhistorica.com](http://www.revistalarazonhistorica.com). Acesso em: 10 fev. 2021.

SÁ, Nysia Oliveira de. **Recursos informacionais**. 2018. 19 slides.

SILVA, I. B. D. S.; CALHEIRA, F. J. S.; BULCÃO, M. C. **Sistemas de informação e o planejamento do acesso e recuperação da informação: uma reflexão acerca dos usuários na concepção do dataísmo**. Ponto de Acesso, v. 13, n. 3, p. 156-166, 2019. DOI: 10.9771/rpa.v13i3.34900. Acesso em: 10 fev. 2021.

VALENTE, Jonas. **Brasil tem 134 milhões de usuários de internet, aponta pesquisa**: A maioria acessa a internet pelo celular. [S. l.], 26 maio 2020. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2020-05/brasil-tem-134-milhoes-de-usuarios-de-internet-aponta-pesquisa>. Acesso em: 20 fev. 2021.

ZATTAR, M. Competência em informação e desinformação: critérios de avaliação do conteúdo das fontes de informação. **Liinc em revista**, v. 13, n. 2, 2017. DOI: 10.18617/liinc.v13i2.4075 Acesso em: 24 maio 2021.