

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
PUC-SP

BRUNA NEDER ROSA

**Publicidade e os limites da privacidade: percepção dos usuários a anúncios
personalizados**

Mestrado em Tecnologias da Inteligência e Design Digital

São Paulo
2021

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
PUC-SP

BRUNA NEDER ROSA

**Publicidade e os limites da privacidade: percepção dos usuários a anúncios
personalizados**

Mestrado em Tecnologias da Inteligência e Design Digital

Dissertação apresentada ao PPG
em Tecnologias da Inteligência e
Design Digital, para obtenção do
título de mestre, sob orientação do
Prof. Dr. Diogo Cortiz da Silva.

São Paulo
2021

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

RESUMO

O crescimento da internet modificou aspectos culturais da sociedade com seu alto poder de ampliação da convergência entre campos midiáticos, além de permitir que os usuários se tornassem criadores de conteúdo, alterando seu status de receptor passivo que até então era predominante nos meios de comunicação e também na publicidade. Os usuários passaram a compartilhar diversas informações que são coletadas através de *cookies*, arquivos que armazenam informações dos usuários não somente para tornar uma experiência de navegação mais qualificada, mas que também são comercializadas por *Data Brokers*, empresas que possuem dados de diversas fontes comerciais, governamentais e públicas que coletam e categorizam todos eles, criando públicos para serem utilizados na publicidade com o objetivo de atingir apenas os usuários com maior probabilidade de compra de determinado produto ou serviço. O aumento da vigilância gerou também o aumento de discussões sobre a privacidade de dados, que incentivou a criação de novas leis como a LGPD (Lei Geral de Proteção de Dados), que abordam a transparência dessas empresas na coleta dos dados, mas não em como a sociedade lida com essa vigilância. Por isso, esta dissertação tem como objetivo compreender as percepções dos usuários com relação a anúncios personalizados na publicidade on-line.

Palavras-chave: publicidade on-line, cookies, data brokers, privacidade, anúncios personalizados.

ABSTRACT

The growth of the internet has changed cultural aspects of society with its high power to expand convergence between media fields, in addition to allowing users to become content creators, changing their status as passive receivers, which until then was predominant in the media as a whole and also in advertising. Users started to share various information that are collected through cookies, which are files that store user information not only to make a more qualified browsing experience, but are also marketed by Data Brokers, companies that have data from various commercial, government and public sources that collect and categorize them all, creating audiences to be used in advertising with the aim of reaching users that are most likely to purchase a particular product or service. Increased surveillance has also generated discussions about data privacy and encouraged the creation of new laws such as the LGPD (Lei Geral de Proteção de Dados or General Data Protection Regulation), which addresses the transparency of these companies in data collection, but not how society deals with this surveillance. Therefore, this dissertation aims to understand the perceptions of users regarding personalized ads in online advertising.

Keywords: online advertising, cookies, data brokers, privacy, personalized ads.

LISTA DE FIGURAS

Figura 1: Interesses em configuração de anúncios Google.....	06
Figura 2: Crescimento global da população na internet (2014 - 2020).....	16
Figura 3: Interações por minuto na internet (2020).....	17
Figura 4: Árvore de cookies.....	20
Figura 5: Subgrafo da conexão na cadeia de cookies.....	23
Figura 6: Origem de tráfego - Google Analytics.....	26
Figura 7: Visão Geral de Interesses - Google Analytics.....	27
Figura 8: Dados de Aquisição por Origem/Mídia - Google Analytics	27
Figura 9: Caminhos em Funis Multicanal - Google Analytics	28
Figura 10: Objetivos de campanha - Google Ads	29
Figura 11: Tipos de campanha - Google Ads	30
Figura 12: Busca no Google.....	31
Figura 13: Segmentos Personalizados - Google Ads.....	31
Figura 14: Segmentos de público-alvo - Google Ads.....	32
Figura 15: Detalhe do Segmento de Público-alvo - Google Ads.....	33
Figura 16: Segmentos de Dados - Google Ads.....	34
Figura 17: Objetivos de campanha - Facebook para Empresas.....	35
Figura 18: Segmentações de campanha - Facebook para Empresas.....	35
Figura 19: Público Personalizado - Facebook para Empresas.....	36
Figura 20: Usuários de internet no Brasil.....	41

LISTA DE TABELAS

Tabela 1: Descrição do planejamento da pesquisa.....	11
Tabela 2: Os maiores cookies ghostwriters.....	21
Tabela 3: Maiores empresas envolvidas nas ações de cookies.....	22
Tabela 4: Nós de grau mais alto no grafo por categoria.....	24
Tabela 5: Definição de Gerações.....	40
Tabela 6: Proporção da população por geração, segundo IBGE.....	40
Tabela 7: Quantidade de participantes por geração.....	41
Tabela 8: Pesquisa - Respostas questão 1.....	43
Tabela 9: Pesquisa - Respostas questão 2.....	44
Tabela 10: Pesquisa - Respostas questão 3.....	44
Tabela 11: Pesquisa - Respostas questão 4.....	45
Tabela 12: Pesquisa - Respostas questão 5.....	46
Tabela 13: Pesquisa - Respostas questão 6.....	46
Tabela 14: Pesquisa - Respostas questão 7.....	47
Tabela 15: Pesquisa - Respostas questão 8.....	48
Tabela 16: Pesquisa - Respostas questão 9.....	48
Tabela 17: Pesquisa - Respostas questão 10.....	49
Tabela 18: Pesquisa - Respostas questão 11.....	50
Tabela 19: Pesquisa - Respostas questão 12.....	51
Tabela 20: Pesquisa - Respostas questão 13.....	51
Tabela 21: Pesquisa - Respostas questão 14.....	52
Tabela 22: Pesquisa - Respostas questão 15.....	52
Tabela 23: Pesquisa - Respostas questão 16.....	53
Tabela 24: Pesquisa - Respostas questão 17.....	54
Tabela 25: Pesquisa - Respostas questão 18.....	54
Tabela 26: Quantidade de participantes por geração após realização da pesquisa.....	55
Tabela 27: Casos por gênero.....	55
Tabela 28: Casos por nível de escolaridade.....	55
Tabela 29: Casos por área de formação.....	56

SUMÁRIO

1. INTRODUÇÃO	1
1.1. Questão da pesquisa	3
1.2. Estado da arte	4
1.3.1. Objetivo geral	7
1.3.2. Objetivos específicos	7
1.4. Justificativas	7
1.5. Hipóteses	9
1.6. Metodologia	9
2. FUNDAMENTAÇÃO TEÓRICA	12
2.1. Da Web 1.0 à Web 3.0	12
2.2. Cookies e Data Brokers	17
2.3. Publicidade On-line e o papel de Cookies e Data Brokers	26
2.3.1. Monitoramento	26
2.3.2. Segmentação de campanhas pelo Google Ads	28
2.3.2. Segmentação de campanhas pelo Facebook para Empresas	34
2.4. Uso de dados na publicidade on-line e regulamentações	36
3. MATERIAIS E INSTRUMENTOS DE PESQUISA	39
3.1. Instrumento de pesquisa: questionário	39
3.1.1. Delimitação do universo e da amostra	39
3.1.2. Coleta e tratamento de dados	42
4. RESULTADOS E DISCUSSÕES	43
4.1. Resultados da Pesquisa	43
4.2. Respondendo às Hipóteses	57
4.2.1. Hipótese A	57
4.2.2. Hipótese B	58
4.2.3. Hipótese C	58
4.2.4. Hipótese D	59
4.2.5. Hipótese E	60
5. CONCLUSÕES	62
REFERÊNCIAS	64
ANEXO 1 - QUESTIONÁRIO	69

1. INTRODUÇÃO

Cultura e mídia estão sempre em mutação, e as novas tecnologias digitais permitem uma constante virtualização da informação e comunicação. Segundo Lévy (1999, p. 113), a emergência do ciberespaço tem um efeito tão radical sobre a pragmática das comunicações quanto teve o surgimento da escrita.

A escrita permitiu que as mensagens fossem transmitidas independente da distância, tempo e cultura, ressignificando os papéis dos atores da comunicação, que em sociedades orais possuíam suas posições de emissores e receptores bem delimitadas, com mensagens linguísticas recebidas no tempo e lugar em que eram emitidas. Para que as mensagens escritas pudessem ser interpretadas fora de seu contexto de produção, foram inventadas as artes da interpretação, tradução, dicionários e regras gramaticais. Da parte dos emissores passou a existir um esforço para que as mensagens pudessem circular sem que se perca seu sentido e significado, o que para Lévy (1999, p. 114) corresponde à noção de universalidade.

A linguagem universal iniciada com a escrita é continuada na configuração clássica das mídias em massa: impressa, rádio, cinema e televisão, que possuem mensagens que serão lidas, ouvidas ou vistas por milhões de pessoas, e por isso são configuradas de acordo com um "denominador comum" mental de seus destinatários, e visam exigir o mínimo da capacidade interpretativa de seus receptores, que recebem os conteúdos de forma passiva, sem interação.

As mídias em massa foram o grande destaque do século XX, pois geraram um novo tipo de consciência coletiva, quebrando barreiras sociais com o aumento do acesso à informação, anulando a divisão entre as culturas erudita e popular. A visão humanista de cultura era uma característica forte da época, e considerava apenas alguns segmentos das atividades humanas - como artes visuais, música e literatura - como cultura (SANTAELLA, 2003, p. 33 *apud* Bernard e Spencer 1996, p. 136). Com as mídias em massa, a visão antropológica de cultura começa a ganhar espaço, já que identifica no mundo diferentes tipos de cultura em cada sociedade, cada uma delas valiosa em si mesma, como algo plural e relativista, com base nas ações, ideias e artefatos que os indivíduos numa dada tradição aprendem, compartilham e avaliam.

Já no final do século XX, em meados dos anos 80, com o surgimento de novas mídias como vídeo cassete, CDs e TV a cabo, tornou-se cada vez mais difícil identificar o que era erudito, popular ou massivo. O termo cultura das mídias é dado por Santaella em 1992 para denominar a nova dinâmica de personalização, trânsito e hibridismo dos meios de comunicação.

Com a evolução das mídias, ficou cada vez mais claro que a cultura evolui de forma contínua e cumulativa, e os meios de comunicação se complementam. Assim como a pintura não desapareceu com o surgimento da fotografia, a televisão não desapareceu com o surgimento da internet. São apenas diferentes meios de difusão que ocupam diferentes espaços da cultura.

Nos anos 1990, a internet se propaga como uma fonte de informações, em um formato de uma grande biblioteca. Com seu poder como meio de comunicação, ela logo cresce ao se tornar um catálogo eletrônico que conecta consumidores e empresas, estudantes e universidades, e todo o tipo de informação imaginável. Aos poucos os próprios usuários se uniram e criaram ideias que se propagaram e anularam a distinção entre informação, entretenimento e relacionamento, acabando com a exclusividade de produção das mídias tradicionais, em que o consumidor tinha um espaço de fala limitada (através de cartas para o jornal ou ligação para a rádio), e passaram a ter a sensação de uma conversa fluida na internet. (TORRES, 2009, p.36).

O crescimento da internet marca a transição entre a cultura de massas e a cultura midiática. Enquanto na cultura de massas os três níveis de cultura (conservação, difusão e recepção ou consumo) estavam bem delimitados e as audiências recebiam conteúdos de forma passiva - criando condições favoráveis para o poder de persuasão da publicidade, a cultura midiática está diretamente ligada ao crescimento de tecnologias da comunicação, e é responsável pela expansão de mercados culturais e hábitos de consumo de cultura.

É importante ressaltar que com a internet houve a ampliação da convergência entre diferentes campos midiáticos, como texto, imagem, som e vídeo em uma linguagem universal, tornando possível a tradução, armazenamento, manipulação e distribuição de todas essas informações de forma rápida. Por isso, a cultura das mídias se distingue da cultura de massas, pois possui novas tecnologias segmentadoras, diversificadoras, capazes de uma maior adequação a um público mais individualizado.

O crescimento da multiplicidade de mídias, a multiplicação de suas mensagens e fontes foi dando margem ao surgimento de receptores mais seletivos, individualizados, o que foi, sem dúvida, preparando o terreno para a emergência da cultura digital, na medida em que esta exige receptores atuantes, caçadores em busca de presas informacionais de sua própria escolha. (SANTAELLA, 2003, p. 68)

A comunicação passou por um progresso não apenas tecnológico, mas que também modificou outros aspectos culturais da sociedade como o mercado capitalista, criando novos hábitos de consumo de cultura.

Segundo Jenkins (2008), a convergência das mídias altera a relação entre as tecnologias, indústrias, mercados e públicos, além da forma com que os consumidores processam notícias e entretenimento. Nesse processo, as empresas perceberam as novas oportunidades de lucros e ampliação de mercados, com novas formas de se conectar com seu público com o marketing digital, e em paralelo os consumidores também aprenderam a utilizar as tecnologias para ter um controle maior sobre o fluxo da mídia e compartilhar ideias e conteúdos.

1.1. Questão da pesquisa

As experiências de consumo passaram a transitar entre o digital e analógico com a internet, em um processo que culminou na integração entre os dois ambientes por meio do compartilhamento de informações através da Inteligência Artificial. Segundo Oliveira (2014, p. 126 - 151), o trânsito da informação é feito de forma pervasiva, devido à sua capacidade de propagação e difusão entre diversos meios, sistemas e tecnologias, e também de forma ubíqua, pois é capaz de estar presente em todos os lugares ao mesmo tempo e de forma imperceptível aos usuários, em Smartphones, computadores, e em todos os dispositivos conectados à internet.

Isso significa que as informações que antes estavam limitadas a instituições como bancos, hospitais ou lojas físicas passaram a ser integradas com as informações de interações on-line, tornando possível a determinação de costumes, hábitos e características dos usuários, de forma em que sua privacidade passou a ser vulnerável (TOMIZAWA, 2008).

O aumento de dados disponíveis com a propagação da internet transformou os consumidores em mercadoria, já que suas informações passaram a ser comercializadas para serem utilizadas na publicidade. Um dos principais atores que tornam isso possível são os *cookies*, que são pedaços de código capazes de reter informações de navegação para o monitoramento comportamental dos usuários, o que permitiu a criação de uma comunicação mais segmentada e direcionada por parte dos anunciantes, e as estratégias de marketing digital passaram planejadas com base nos interesses ao bem de consumo anunciado, aumentando a probabilidade de vendas e reduzindo os custos das ações publicitárias (BIONI, 2019, p. 44).

Devido à falta de transparência na coleta e uso de informações, houve um aumento nas discussões em relação à privacidade e vigilância, já que o avanço tecnológico possibilitou o cruzamento de informações de diversas fontes de dados de forma não regulamentada, e dessa forma o seu uso na publicidade facilita a indução ao consumo. Novas leis têm surgido para proteger o direito à privacidade, como a *General Data Protection Regulation* (GDPR)¹, a regulamentação pioneira da Europa que entrou em vigor em maio de 2018, e que também influenciou a aprovação da Lei Geral de Proteção de Dados (LGPD) no Brasil, que entrou em vigor em agosto de 2020. Porém, as novas leis abordam a coleta e transparência no uso dos dados, mas não em como o consumidor lida com a comunicação direcionada. Por isso, a pesquisa busca compreender as percepções dos usuários com relação a anúncios personalizados na publicidade on-line.

1.2. Estado da arte

A sociedade da informação surgiu na era pós-industrial, em que é possível transmitir informações em uma quantidade e velocidade nunca vistas anteriormente: “informação é o (novo) elemento estruturante que (re)organiza a sociedade, tal como o fizeram a terra, as máquinas a vapor e a eletricidade, bem como os serviços, respectivamente, nas sociedades agrícola, industrial e pós-industrial” (BIONI, 2019, p.2). A evolução digital não se resume apenas ao ambiente virtual da internet, mas também é uma ferramenta de grande destaque nesse processo,

¹ GDPR - Para saber mais acesse:
https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en

obtendo influências econômicas através das informações geradas nesse meio e em relações sociais.

Uma das principais mudanças sociais está no uso de dados pessoais, que tornaram o consumidor um produto comercializável, já que na internet a maior parte das informações são gratuitas e a moeda de troca são os dados cedidos pelo usuário, para que seu perfil de interesses e consumo seja identificado.

“Trata-se de um modelo de negócio que é financiado ou suportado predominantemente pela publicidade comportamental. Em um primeiro momento, atrai-se o usuário para que ele usufrua um serviço e/ou produto para, em um segundo momento, coletar seus dados pessoais e, então, viabilizar o direcionamento da mensagem publicitária, que é a sua fonte de rentabilização” (BIONI, 2019, p.12).

Esse tipo de troca de dados por serviços citada por Bioni (2019) pode ser observada pela empresa Meta, por exemplo, que fornece diversos serviços de forma gratuita (como o Facebook, Instagram e WhatsApp), e gerou US\$ 28,2 bilhões² com anúncios no terceiro trimestre de 2021.

A disponibilidade de dados passou a moldar e caracterizar a publicidade on-line, com o grande diferencial de entregar mensagens personalizadas de acordo com o perfil do público. A quantidade e velocidade em que esses dados são compartilhados é tão grande, que seu valor foi perdido com o tempo, já que consideramos algo raro como valioso, mas quando temos em abundância, segundo Shirky (2011, p.48), começamos a tratar como se fosse barato o suficiente para desperdiçar e fazer experiências. Isso significa que os usuários da internet se acostumaram a consumir, compartilhar e gerar dados como se não tivessem valor, e em paralelo seus dados são coletados, classificados e comercializados.

No Google, por exemplo, dentro das configurações de conta do Gmail, é possível acessar os detalhes sobre a personalização dos anúncios e em quais categorias o perfil do usuário foi enquadrado, e que inclui não somente dados pessoais mas também interesses baseados em anunciantes parceiros da plataforma.

² Facebook: mesmo sob crises, rede social cresce em receita e usuários. Tecmundo, 2021. disponível em: <https://www.tecmundo.com.br/mercado/227537-facebook-divulga-balanco-aumento-receita-3-trimestre.html>. Acesso em 15 nov 2021.

Na figura 1 é possível ver algumas das mais de 150 categorias em que um único perfil pode ser classificado.

Figura 1: Interesses em configuração de anúncios Google



Fonte: Google - Configurações de anúncio
<https://adssettings.google.com/authenticated>

A coleta desses dados é feita através de *cookies* de navegador e são tratados e categorizados por empresas chamadas de *Data Brokers* com o objetivo de comercializá-los. Essas empresas também possuem dados de fontes comerciais, governamentais e públicas, o que permite a criação de um perfil detalhado do consumidor, e a segmentação na publicidade pode ser feita desde interesses com base no histórico de navegação até mesmo ao detalhe de renda familiar, cargo ou comportamento de compra.

Com base nessas categorizações, a experiência de navegação se torna mais qualificada, já que o conteúdo que aparece - não só em anúncios, mas também em resultados de buscas ou redes sociais - já são relacionados com os interesses de cada usuário.

1.3. Objetivos

A publicidade on-line se diferencia dos modelos tradicionais como a televisão, pois possibilita a segmentação de públicos personalizados para uma comunicação direcionada. Para empresas, os investimentos em marketing digital se mostram mais vantajosos pois possuem custos mais baixos e são mensuráveis. Porém, diferente dos outros meios de comunicação, na internet o usuário tem um papel ativo e pode mudar a forma como as informações são compartilhadas. Nesse contexto, os objetivos desta pesquisa se desdobram em um objetivo geral e três objetivos específicos.

1.3.1. Objetivo geral

O objetivo geral é analisar as percepções dos usuários em relação a anúncios personalizados de acordo com seus interesses no ambiente on-line.

1.3.2. Objetivos específicos

Analisar como os usuários compreendem a coleta de seus dados pelas plataformas.

Analisar quais ações os usuários estão fazendo para proteger seus dados pessoais.

Analisar as diferenças de comportamento no ambiente on-line entre gerações.

1.4. Justificativas

A evolução da tecnologia de comunicação permitiu que pessoas comuns passassem a ser protagonistas na geração de conteúdos no ambiente on-line, como textos, vídeos, fotos e todo tipo de informações pessoais em comunidades virtuais: “[...] espaço informacional multidirecional que depende da interação do usuário, permite a este o acesso, a manipulação a transformação e o intercâmbio de seus fluxos codificados de informação” (SANTAELLA, 2004, p. 44).

Com essas mudanças, as máquinas também passaram a desempenhar um papel de maior influência sobre as pessoas. Por exemplo, em uma busca feita no

Google Search, os resultados são selecionados pela plataforma, sem que o usuário tenha qualquer controle sobre isso.

Assim, tomar consciência de que somos mediados cada vez mais por agentes computacionais - que controlam os fluxos de informação de nossas vidas e moldam a nossa percepção, e que usamos constantemente para tomar decisões - é um passo extremamente importante para a nossa atuação no mundo de hoje, cada vez mais digital (GABRIEL, 2018, p.61)

Por isso, a utilização das plataformas digitais requer um senso crítico, já que os conteúdos são mostrados com base no que os algoritmos acreditam que o usuário quer ver, e não necessariamente no que ele precisa encontrar. Além disso, também é preciso atenção às informações compartilhadas, por mais difícil que seja a identificação dos diferentes contextos dentro deste ambiente. Quando pensamos em meios físicos *off-line*, está muito clara a diferenciação entre casa, trabalho, igreja, e quais informações serão compartilhadas em cada contexto. Já no ambiente digital, existem diversos contextos existindo simultaneamente, o que dificulta o controle da privacidade. Isso porque os dados gerados são computáveis, registrados e disseminados instantaneamente, tornando um ambiente propício para o compartilhamento de informações.

As informações compartilhadas por pessoas são feitas tanto de forma consciente (como fotos, localização, textos), quanto inconsciente (quando efetuam alguma busca na internet). A obtenção de todos esses dados permite conhecer seus comportamentos, preferências e personalidade, dando poder para quem os detém para auxiliar ou manipular esses indivíduos. Ao mesmo tempo em que permite que o conteúdo que chega à esses usuários seja mais relevante e customizado, dados coletados também podem ser utilizados para fins que não foram acordados - como, por exemplo, se um usuário foi categorizado como diabético com base nas informações coletadas de forma indireta (como através de buscas realizadas), uma empresa de planos de saúde pode escolher cobrar um valor mais alto pelo risco que a pessoa representa. Não importa a finalidade, o formato para a coleta de dados é o mesmo. Assim, a intenção do seu uso determinará se ele será benéfico ou manipulativo. Por isso, “a privacidade é elemento regulador e a transparência, o efeito resultante” (GABRIEL, 2018, p. 59).

1.5. Hipóteses

As hipóteses, segundo Gil (2008, p. 41), são uma suposta resposta ao problema a ser investigado, e que serão aceitas ou rejeitadas após testadas, e sua verificação é o propósito da pesquisa científica. Sendo assim, para responder aos objetivos geral e específicos deste trabalho, foram definidas cinco hipóteses:

- a) Os usuários não têm conhecimento das configurações de publicidade disponíveis nas plataformas.
- b) Os usuários não entendem que o conteúdo visualizado é disponibilizado com base em seus interesses.
- c) Os usuários sabem que seus dados transitam entre plataformas, mas não compreendem como isso acontece.
- d) Acredita-se que as gerações baby boomers e geração X possuem maior preocupação com os dados pessoais no ambiente on-line, em comparação com as gerações Y e Z.
- e) Acredita-se que as gerações Y e Z, por terem crescido já no ambiente digital, possuem mais conhecimento sobre privacidade de dados do que as gerações baby boomers e geração X.

1.6. Metodologia

A contribuição da teoria da midiatização “(...) é prover uma base de análises e criar uma compreensão teórica de como a mídia pode interagir com outros processos culturais e sociais” (HJARVARD, 2013, p.4). Por isso, a primeira etapa do procedimento deste estudo será a pesquisa bibliográfica para compreender o atual cenário da publicidade on-line, conceitos básicos, principais ações de marketing digital e como são criadas e segmentadas as campanhas, entendendo quais tipos de dados podem ser utilizados, sejam nativos das plataformas (Google e Facebook) ou de terceiros, explorando questões sobre privacidade e uso de dados pessoais.

A segunda etapa será a realização de uma pesquisa, planejada com base em uma fundamentação metodológica, definição de objetivos, técnicas de coleta e análise de dados. Assim, será possível compreender melhor e quantificar o perfil dos usuários e qual sua percepção quanto ao uso de seus dados para a personalização em publicidade direcionada, com a quebra geracional para analisar quais as possíveis diferenças de comportamento entre elas.

Ao proporcionar uma visão geral e estudo das características de um grupo (suas opiniões, atitudes e crenças) com o levantamento bibliográfico, o estudo caracteriza-se como exploratório e descritivo, pois segundo Gil (2008, p. 28) "as pesquisas descritivas são, juntamente com as exploratórias, as que habitualmente realizam os pesquisadores sociais preocupados com a atuação prática".

O método hipotético-dedutivo de investigação desenvolvido por Karl Popper (2013) foi escolhido, pois através da observação e identificação do problema são formuladas hipóteses que serão testadas em busca de evidências para derrubá-las para sua corroboração.

Na delimitação do universo, foram selecionados brasileiros com a divisão de amostragem com base nas gerações. Segundo o Pew Research Center³, a geração Baby Boomer é caracterizada por indivíduos nascidos de 1946 a 1964, geração X de 1965 a 1980, geração Y ou millennials de 1981 a 1996 e geração Z de 1997 a 2012. A divisão da amostra por idade será feita segundo fonte IBGE/Diretoria de Pesquisas/Coordenação de População e Indicadores Sociais/Gerência de Estudos e Análises da Dinâmica Demográfica.

A estratégia de pesquisa escolhida foi a quantitativa, pela praticidade na busca de respondentes, podendo seu questionário ser aplicado via plataforma online. O questionário é um grupo de perguntas, onde o entrevistado lê e responde sem a presença do entrevistador, podendo ser recebido via correio, pessoalmente, internet etc. O modelo de questionário que será utilizado é explicado por MATTAR (1996), como estruturado disfarçado, onde o pesquisador busca através da tabulação e informações presentes, presumir a importância e valor do tópico estudado ao respondente.

Como MARCONI E LAKATOS (2002, p.18) afirmam "O investigador, baseando-se em conhecimentos teóricos anteriores, planeja cuidadosamente o método a ser utilizado, formula problema e hipóteses, registra sistematicamente os dados e os analisa com a maior exatidão possível", com base nesta afirmação, o estudo visa utilizar do conhecimento prévio do assunto estudado e encontrar, de forma ampla, afirmações para as hipóteses presentes.

³ Pew Research Center, Defining generations: Where Millennials end and Generation Z begins, disponível em: <https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/>

Tabela 1: Descrição do planejamento da pesquisa

Item	Descrição
Tema	Anúncios personalizados na publicidade on-line
Objetivo geral	Analisar as percepções dos usuários em relação a anúncios personalizados de acordo com seus interesses
Hipóteses	<ul style="list-style-type: none">- Os usuários não têm conhecimento das configurações de publicidade disponíveis nas plataformas.- Os usuários não entendem que o conteúdo visualizado é disponibilizado com base em seus interesses.- Os usuários sabem que seus dados transitam entre plataformas, mas não compreendem como isso acontece.- Acredita-se que as gerações baby boomers e geração X possuem maior preocupação com os dados pessoais no ambiente on-line, em comparação com as gerações Y e Z.- Acredita-se que as gerações Y e Z, por terem crescido já no ambiente digital, possuem mais conhecimento sobre privacidade de dados do que as gerações baby boomers e geração X.
Metodologia de abordagem	Hipotético-dedutivo
Instrumento de pesquisa	Aplicação de questionário
Delimitação do universo	Brasileiros com acesso a internet
Tipos de amostragem	Amostragem Não Probabilística por quotas.
Estratos de amostragem	Grupos geracionais por faixa etária: 18 - 24 anos (Geração Z), 25 - 34 anos (Millennial), 35 - 39 anos (Millennial tardio), 40+ anos (Baby Boomers).
Quantidades	500 participantes, com a seguinte distribuição: 18 - 24 anos (15%): 75 casos 25 - 39 anos (34%): 170 casos 40 - 56 anos (31%): 155 casos 57 - 75 anos (20%): 100 casos
Meios para aplicação do questionário	Painel de respondentes por método digital

Fonte: elaborada pela autora.

2. FUNDAMENTAÇÃO TEÓRICA

Para compreender como as plataformas utilizam dados pessoais, a fundamentação teórica deste estudo inicia-se com um breve resumo sobre o crescimento da internet, e qual o papel de *cookies* e *Data Brokers* para a publicidade. Em seguida, as principais campanhas de publicidade on-line, como são segmentadas, monitoradas e a importância da discussão sobre a privacidade de dados pessoais.

2.1. Da Web 1.0 à Web 3.0

O mundo interconectado partiu de três pontos técnicos principais, sendo o primeiro em 1965 com o conceito de hipertexto, o segundo em 1969 com a implementação da internet e o terceiro em 1990 com a criação da *World Wide Web*.

Como forma de texto não sequencial, o hipertexto possui uma leitura sem caminho determinado em que um usuário de sistema informático seria capaz de desenvolver sequências associativas de informações por meio de links (Nelson 1965, p. 96). Já a internet surgiu durante a Guerra Fria, na busca por uma comunicação em rede capaz de conectar dispositivos com diferentes configurações. No período, já foi possível integrar quatro pontos em comunicação em rede, segundo Leiner *et al.* (2009, p. 23): Universidade da Califórnia em Los Angeles (UCLA), Instituto de Pesquisa de Stanford (SRI), Universidade da Califórnia em Santa Bárbara (UCSB) e Universidade de Utah.

Em 1973, os protocolos básicos da internet foram desenvolvidos pelos pesquisadores Robert Kahn e Vincent Cerf: o Protocolo de Controle de Transmissão (TCP) e o Protocolo de Internet (IP). O primeiro permite o controle de fluxo e recuperação de pacotes de dados transmitidos pela rede, enquanto o segundo permite o esquema de direcionamento e endereçamento destes pacotes aos dispositivos corretos (LEINER *et al.*, 2009, p. 25).

Somente em 1990 surge a *World Wide Web*, quando as funcionalidades de hipertexto e da internet são integradas por Tim Berners-Lee através de regras capazes de traduzir a comunicação entre computadores com sistemas completamente diferentes. Com a implementação dos protocolos de URI (Identificador de Recursos Universal), HTTP (Protocolo de Transferência de

Hipertexto) e HTML (Linguagem de Marcação de Hipertexto) foi possível criar sistemas nominais e universais de identificação de páginas, traduzir solicitações e respostas entre um computador e um servidor de conteúdo on-line e exibir essas informações de forma gráfica e ordenada (BERNERS-LEE, 1999, p. 33-9).

Com o aumento da performance dos equipamentos e a criação de uma rede de servidores que permitiu a comunicação entre os computadores, as informações na internet passaram a ter várias fontes de hospedagem em todo o mundo, sem que houvesse um único dono - o que perpetua até hoje. O acesso a essas informações é feito através de navegadores (ou *browsers*), que são a porta de entrada para a busca dos websites da internet. Para tornar a navegação mais fácil, surgiram as ferramentas de buscas - como o Google, que com seus robôs ou *spiders*, são capazes de acessar todos os sites e procurar as palavras-chave inseridas no buscador e trazem os resultados mais relevantes. Assim, os usuários não precisam conhecer todos os sites que existem para encontrar o que buscam.

O surgimento de sites e grandes portais, como jornais, revistas, bibliotecas, universidades e estabelecimentos comerciais ficou conhecido como Web 1.0. Com a internet, as empresas logo perceberam um novo meio para compartilhar informações sobre seus produtos e serviços, e as URLs (*Uniform Resource Locator*) de cada marca começaram a ser divulgadas em rótulos de produtos para que os consumidores pudessem lembrar e digitar sem erro (TORRES, 2009, p.40).

Nesse primeiro momento da Rede, apenas os usuários com conhecimento em tecnologia ou programação eram capazes de criar seus sites, e por isso o volume de páginas pessoais ainda era pequeno. A segunda geração da internet ficou conhecida como Web 2.0, oficialmente definida em 2004 por Dale Dougherty, vice-presidente da O'Reilly Media, com a característica de ser colaborativa e participativa, e o surgimento dos primeiros *blogs*, em 1997, foram o ponto de transição para que os usuários passassem a participar na produção de conteúdo. A plataforma Blogger ajudou a impulsionar esse movimento, permitindo que qualquer usuário pudesse ter sua própria página para trocar opiniões sobre produtos e criar conversas com outros usuários com interesses em comum. O sucesso dos *blogs* foi tanto, que passaram a ser também uma fonte de renda para os criadores de conteúdo, seja para a disponibilização de espaços no site para *banners*, ou com publicações sobre determinado produto ou serviço.

As redes sociais também tiveram grande destaque nessa fase da Rede, pois através delas as pessoas puderam se encontrar on-line para conversar com amigos e familiares, além do compartilhamento de fotos, vídeos e músicas, se tornando um elemento social importante ao aproximar pessoas independente da distância. Algumas das maiores redes sociais que existem até hoje surgiram nesse período, como o Facebook, que foi criado em 2004 pelo estudante de Harvard e programador Mark Zuckerberg, e em 2020 fechou o ano com 2,8 bilhões de usuários ativos mundialmente na plataforma, segundo dados da revista Exame⁴. Ou também o YouTube, que surgiu em 2005, criado pelos americanos Chad Hurley, Steve Chen e Jawed Karidois, que viram o crescimento do acesso às câmeras digitais como uma oportunidade de criar um novo formato de criação de conteúdos, que antes era limitado apenas a textos e imagens. Seu sucesso foi tanto, que em 2006 o site foi comprado pelo Google pela quantia de U\$ 1,65 bilhão, valor atribuído à quantidade de pessoas cadastradas e volume de conteúdo gerado na plataforma. Em 2021, a plataforma possui mais de 2 bilhões de usuários⁵, e está presente em mais de 100 países com 80 idiomas disponíveis.

Na segunda geração da Web, os geradores de conteúdo passaram a utilizar *tags* (etiquetas) para classificar as informações de forma estrutural em seus sites, com a linguagem XML (*Extensible Markup Language* ou Linguagem de Marcação Extensiva), para que seus *links* pudessem se tornar mais relevantes em ferramentas de buscas. Porém, a escolha da palavra-chave (ou *tag*) é feita pelo usuário e de forma limitada, o que significa que as pesquisas realizadas na web nem sempre traziam os documentos desejados.

A Web 3.0 surge como a Web Semântica, com o objetivo de organizar a enorme quantidade de *links* disponíveis na rede para uma melhor experiência de navegação. Berners-Lee *et al.* (2001) a descreve como a Web do significado, em uma geração em que os computadores compreendem os significados por trás de cada página da rede, baseada na integração, combinação e interpretação de informações. As informações se tornaram mais integradas com o uso do celular como microcomputador, *Cloud Computing* - com acesso à tudo que está na nuvem -

⁴ Facebook fica mais perto de 3 bilhões de usuários ativos e receita cresce em 2020. Por Tamires Vitorio. Acesso em 24/07/2021. Disponível em: <https://exame.com/tecnologia/facebook-fica-mais-perto-de-3-bilhoes-de-usuarios-ativos-e-receita-cresce-em-2020/>

⁵ Usuários do YouTube, disponível em: <https://www.youtube.com/intl/pt-BR/about/press/>

e as máquinas passaram a ser capazes de interpretar dados sem a mediação humana, levando em consideração também o contexto em que o usuário está inserido - como sua região e cultura (ou comportamento), em uma sofisticação dos mecanismos de inteligência artificial que engloba a linguagem semântica, permitindo uma interação com qualidade superior e tornando possível a disponibilização de novos produtos e serviços que até então não estavam presentes no ambiente on-line.

“(…) muitas áreas de consumo que não estavam listadas como candidatas para ser oferecidas pela Web, elas passam a ter essa possibilidade real de prestar serviços via rede, pois agora elas têm as bases teóricas e científicas estabelecidas, mudanças culturais das novas gerações X, Y e quem sabe, geração Z, uma base de consumidores treinada, tecnologias desenvolvidas e acessibilidade quase universal à Web.” (Koo, 2009, p. 121)

Lawrence Koo (2009) já previa as mudanças culturais que a internet trouxe para as gerações, tendo em vista as grandes mudanças tecnológicas que ocorreram desde 1945. As máquinas de calcular evoluíram para máquinas capazes de obter e analisar uma enorme quantidade de dados de forma rápida e interligada, incentivando cada vez mais sociedades hiperconectadas.

As novas interfaces de Inteligência Artificial modificaram o cotidiano e cultura, como o GPS que não apenas guia como também informa rotas mais rápidas, ou o assistente virtual para casas inteligentes que é capaz de tocar músicas, ligar aparelhos conectados à ele, dar notícias ou pedir um Uber⁶. Só no Brasil, já são mais de 134 milhões de usuários de Internet⁷ que compartilham textos, fotos, vídeos e diversos outros dados pessoais.

O estudo Data Never Sleeps 8.0⁸, criado pela empresa DOMO, demonstra a intensificação global da interação digital em 2020 com a pandemia da COVID-19, em que as pessoas passaram a depender mais da internet para interações sociais, trabalho, educação e entretenimento.

⁶ Uber: empresa que oferece serviço de transporte privado.

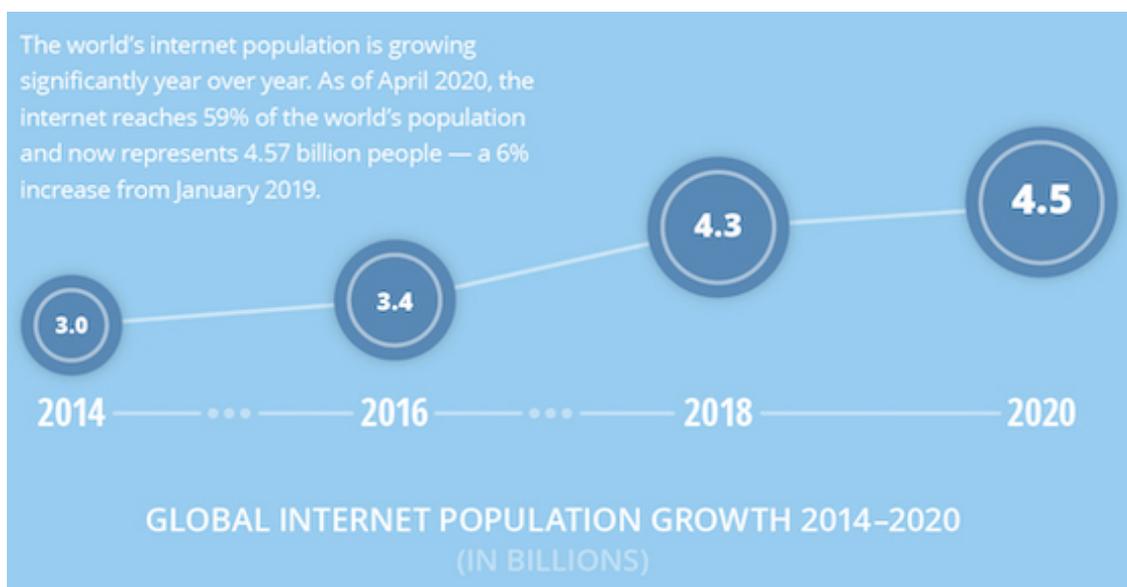
⁷ Dados da Pesquisa TIC Domicílios 2019, disponível em: https://cetic.br/media/analises/tic_domicilios_2019_coletiva_imprensa.pdf

⁸ Data Never Sleeps 8.0, disponível em: <https://www.domo.com/learn/data-never-sleeps-8>

Atualmente vivemos a era da informação, onde há um vasto compartilhamento de informações no ciberespaço, gerando novos modelos de relações e conexões, provocando mudanças em todos os ambientes tradicionais. Para Santaella (2007 apud Ianni 1996), a globalização subverte diversas formas de organização das forças produtivas, envolvendo a produção material e cultural e provocando a desterritorialização e reterritorialização de coisas, gente e ideias.

A *Figura 2* mostra o crescimento da população global na internet, atingindo 59% da população mundial no ano de 2020. São mais de 4 bilhões de pessoas gerando cliques em anúncios, reações nas mídias sociais, compartilhamentos, transações, conteúdo de streaming e diversas outras interações.

Figura 2: Crescimento global da população na internet (2014 - 2020)

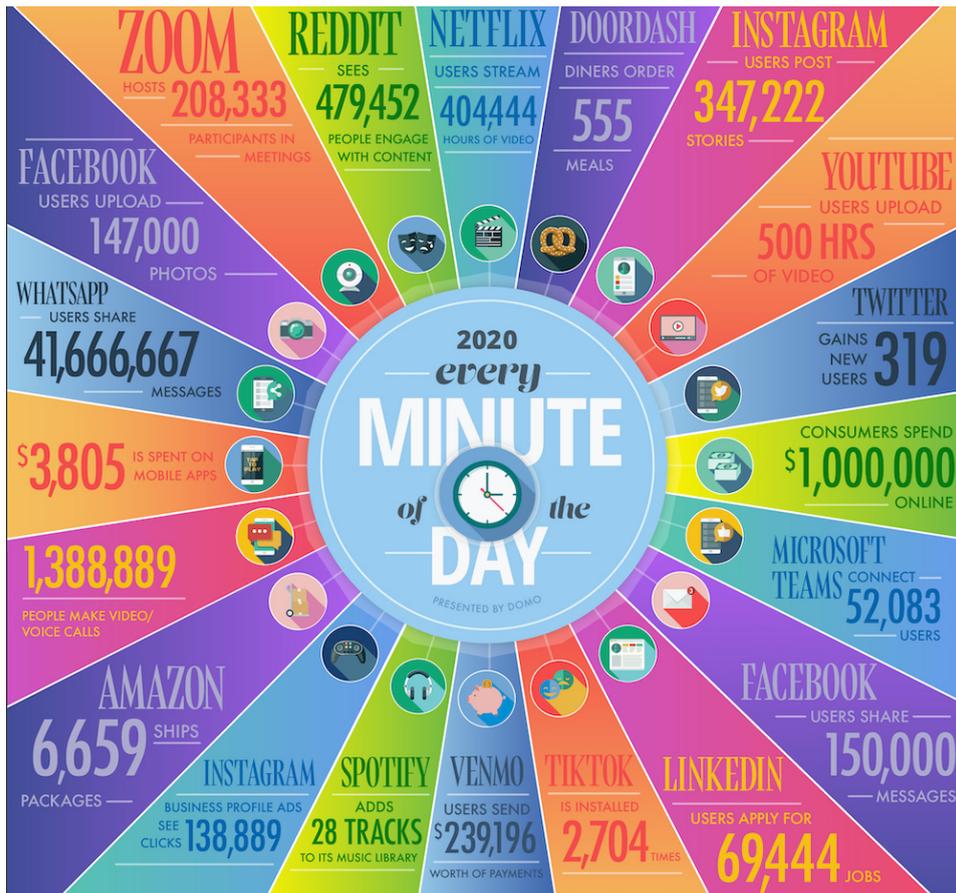


Fonte: DOMO - Data Never Sleeps 8.0

<https://www.domo.com/learn/data-never-sleeps-8>

A pesquisa também traz destaque das interações por minuto das principais plataformas, e nas principais redes sociais é possível identificar o *upload* de 147 mil fotos no Facebook, mais de 41 milhões de mensagens no WhatsApp, 500 horas de vídeos no YouTube e 347 mil posts no Stories do Instagram, como é possível observar na *Figura 3*.

Figura 3: Interações por minuto na internet (2020)



Fonte: DOMO - Data Never Sleeps 8.0

<https://www.domo.com/learn/data-never-sleeps-8>

O papel ativo dos usuários na geração de conteúdo faz com que uma enorme quantidade de dados sobre eles esteja disponível no ambiente *on-line*. O rastreamento da navegação fez com que empresas pudessem capturar, categorizar e comercializar os dados desses usuários para fins publicitários, com o objetivo de criar anúncios personalizados de acordo com os interesses de cada um, sendo os *cookies* e *Data Brokers* os principais responsáveis por esse cenário.

2.2. Cookies e Data Brokers

As interações *on-line*, em blogs, redes sociais ou até mesmo em uma busca na internet, geram dados que podem ser coletados, armazenados e rastreados. O momento da coleta de dados é chamado de *Behavioural Tracking* (ou, em tradução literal, rastreamento comportamental), segundo Castelluccia (2012). Um exemplo

desse rastreamento é quando um usuário tenta fazer uma compra de passagem aérea para São Paulo, faz a busca por voos mas não faz a compra. Em seguida, ao entrar em um website de notícias visualiza um anúncio de passagens aéreas para São Paulo. Apesar de nenhuma informação pessoal ter sido coletada, o interesse foi identificado.

Esse tipo de experiência traz uma navegação mais personalizada, e isso é possível graças aos *cookies* de navegador, que mantêm a comunicação com o servidor de forma contínua enquanto ele navega por uma página na internet (KRISTOL, 2001, p. 152-4).

Segundo Castelluccia (2012 *apud* McKinley 2008), as informações que são utilizadas para construir os perfis de usuários são diferentes fontes de monitoramento de endereços de IP, utilizando técnicas como *cookies*, *Javascripts*, *supercookies* e *evercookies*.

Os *cookies* são pedaços de textos que ficam salvos no navegador dos usuários e vinculados à solicitação ao HTTP, e consiste em pedaços de informações definidas por um servidor. “Desse modo, os *cookies* são responsáveis por manter a comunicação entre cliente e servidor e identificar de forma contínua um usuário enquanto ele navega por uma página na web” (KRISTOL, 2001, p. 152-4), e podem pertencer à própria página visitada (*first-party cookies*) ou *cookies* de terceiros (*third-party cookies*).

Existem dois tipos de *cookies*, de sessão e os persistentes. Os de sessão são utilizados para salvar as preferências dos usuários e são apagados quando os usuários se desconectam (fazem o *log out*). Já os *cookies* persistentes mantêm a sessão autenticada com o servidor e os arquivos permanecem no navegador até que o usuário os delete ou eles expirem, como por exemplo as preferências de idioma e outras configurações, para que os usuários não precisem refazê-las toda vez que acessarem determinado site.

Javascripts são arquivos executáveis baixados por usuários de forma automática para carregar as informações do site. Algumas vezes, esses arquivos junto com os *cookies first-party* enviam informação de volta aos servidores, e as duas informações juntas podem conter dados como endereço de IP, histórico de navegação e endereço de e-mail.

Os *supercookies* funcionam através de *plug-ins*, que são elementos adicionais dos navegadores capazes de gerenciar os dados e controlar o que é

deletado, enquanto os *evercookies* são capazes de permanecer no computador do usuário mesmo após serem deletados, por manipularem as aplicações armazenamento de navegação, como por exemplo o armazenamento temporário de informações conhecido como cache.

No final dos anos 1990, houve um grande aumento no uso de *cookies*, principalmente de terceiros, justamente para garantirem a frequência e personalização de anúncios de acordo com o perfil e histórico de navegação de cada usuário, além de controlarem a frequência com que os anúncios aparecem. Segundo Kristol (2001, p. 159-64), o grande diferencial dos *cookies* de terceiros é que com eles é possível formar bases de perfis mais completos, pois eles não se originam de uma única página e sim por um conjunto delas.

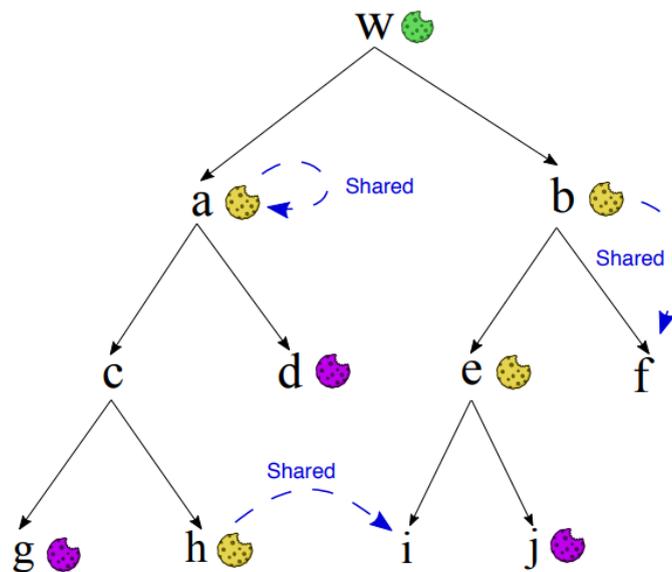
Os *cookies* foram criados inicialmente para melhorar a qualidade da navegação, salvando os sites já visitados, melhorando a experiência de compras on-line, conteúdos visualizados e publicidade direcionada. Porém, a tecnologia evoluiu de forma cada vez mais invasiva na extração e análise de padrões de comportamento on-line, gerando uma enorme quantidade de informações que passaram a ser de grande valor para as empresas.

Cookies passaram a ter um papel fundamental no rastreamento dos usuários, e têm levantado cada vez mais discussões em relação à privacidade de dados pessoais. Porém, segundo Sanchez-Rola *et al.* (2021), as discussões sobre o assunto focam em quais domínios estão associados aos *cookies* criados no navegador e quais domínios compartilham *cookies* com outros domínios, o que segundo os autores não captura a interação sutil de diferentes organizações através dos códigos fonte dos *cookies*. Isso significa que donos de sites muitas vezes não estão cientes de que seus *cookies* - que teoricamente seriam utilizados apenas como *first-party* também são enviados para terceiros.

Na realidade, existem três atores na operação de compartilhamento de *cookies*, sendo o que criou, o que o recupera do navegador e o inclui em uma solicitação *HTTP* e o que recebe o *cookie* no final da cadeia.

Sanchez-rola *et al.* (2021) introduzem o conceito de *cookie ghostwriting*, que está relacionado com *cookies first-party* que acabam equivalente a *third-party*, pois enviam cópia dessas informações para terceiros, tornando possível capturar informações até mesmo dos usuários que só aceitaram *cookies first-party*.

Figura 4: Árvore de cookies



Fonte: Sanchez-rola *et al.* - Eurocom (2021) - <https://www.eurecom.fr/publication/6376>

No exemplo criado pelos autores, a *cookie tree* (ou árvore de *cookies*) demonstra em verde o *cookie first-party* criado por W, os amarelos como *ghosted first-party cookies* e os roxos os *third-party cookies*. O fluxo desses *cookies* pode ser dividido entre: criadores, intermediários, remetentes e receptores.

Criadores: representado por W na figura, são entidades que configuram *cookies*, seja através de *JavaScripts* ou *header HTTP*, além de *cookie third-parties* (d, g e j) e *ghost cookie creators* (b, e e h).

Intermediários: nós internos que podem não criar ou manipular diretamente os *cookies*, mas possuem recursos que os fazem (exemplo: a e c são *cookies* intermediários de h).

Remetentes: recuperam *cookies* definidos por outras entidades e os enviam por meio de uma solicitação HTTP. Podem ser divididos em três categorias: próprios remetentes, que enviam seus próprios *cookies* (ou seja, aqueles criados pelo mesmo nó, como a e b), os remetentes na cadeia, que enviam *cookies* criados por um de seus nós filhos (ou seja, para os quais o remetente faz parte da cadeia de criação) e fora da cadeia, que enviam *cookies* sem fazer parte de sua cadeia de criação (como e enviando *cookie* de h para i).

Receptores: as entidades que recebem explicitamente *cookies* como parte de uma solicitação HTTP. Existem os receptores que obtêm *cookies* criados por eles

próprios (a), receptores em cadeia que obtêm *cookies* criados por seus nós filhos (ou seja, para os quais o receptor faz parte a cadeia de criação), e receptores fora da cadeia que obtêm *cookies* para os quais não faziam parte da cadeia de criação (f e i).

Os autores reforçam também que existem os *ghosted third-party cookies*, que também são considerados no estudo mas não representados na figura de árvores de cookies para facilitar o entendimento.

Para o estudo, foram selecionados os 1 milhão de domínios mais acessados segundo Tranco⁹, coletando 138 milhões de *cookies* em 6.2 milhões de páginas. Dentro dos sites, 76% deles continham pelo menos um *cookie*, sendo o restante deles sites abandonados ou com *links* sem conteúdo ou “quebrados”; 52,45% deles (387.202) possuem pelo menos um *cookie* compartilhável ou evento de colisão (que é quando um novo ator substitui um *cookie* previamente existente). 47% (31.1M) são *cookies third-party* e 28M são *ghosted cookies*, sendo apenas 7.6M de *cookies first-party* que foram criados pela fonte do próprio site.

A figura abaixo reflete os maiores *cookies ghostwriters* de acordo com o número de site em que foram configurados seus *ghosted cookies*. O que a pesquisa destaca é que muitos deles utilizam a maior parte de seus *cookies* ou quase todos como *ghostwriter*, sendo os principais Google e Facebook.

Tabela 2: Os maiores cookies ghostwriters

Actor	Websites	
	as ghostwriter	as third-party
Google	175,395	283,778
Facebook	141,708	6,472
AddThis	47,250	52,648
Yandex	41,089	3,782
Baidu	19,965	1,348
Quantcast	17,902	3,039
AdRoll	10,702	1
Criteo	10,486	68
CNZZ	10,052	0
Shopify	9,657	137
ShareThis	9,486	17,589
Hotjar	9,431	3
HubSpot	8,935	372
Zopim	8,557	3
Amazon	7,883	23,358
StatCounter	7,631	2
Mail.Ru	6,573	1,914
Adobe	6,379	30,918
Tawk	5,890	11
Microsoft	5,782	19,200

Fonte: Sanchez-rola *et al* - Eurocom (2021) - <https://www.eurocom.fr/publication/6376>

⁹ Tranco: classificação de sites orientada para pesquisa contra manipulação. Disponível em: <https://tranco-list.eu/>

A análise da árvore de *cookies* permitiu a identificação de 170 mil organizações envolvidas na criação e compartilhamento de *cookies*, sendo que 43% dos sites estudados possuem pelo menos 3 diferentes organizações envolvidas. Além disso, o experimento também observou 8.97M de compartilhamento de eventos de *cookies* em 378K sites, e identificou que 6% deles (505.926) foram efetuados sem criptografia, e 37,71% foram compartilhados por HTTP quando foram inicialmente criados por HTTPS (sendo o S de Secure, para proteger a integridade e confidencialidade dos dados entre o computador do usuário e o site).

Outro dado importante identificado foi de papéis de remetentes e receptores, que identificou que em 354K dos sites (sendo que 48% deles criam *cookies* identificadores) um ator cria um *ghosted first-party cookie* e o envia para si mesmo. Através dessa prática, é possível que as empresas sincronizem o *third-party cookie* com o *ghosted first-party cookie*, o que significa que mesmo quando um usuário deletar ou recusar o acesso de um *cookie third-party*, é possível que ainda assim ele esteja sendo rastreado.

Sanchez-rola *et al.* (2021) analisaram as maiores 50 empresas envolvidas nas ações de *cookies*, e observaram que elas possuem diferentes papéis dentro do ecossistema, a maior parte dos atores age como remetentes (*sender*), receptores (*receiver*) ou intermediários (*intermediary*), demonstrando que muitos se comportam como intermediários tanto na criação quanto compartilhamento, o que sugere um panorama complexo de conexão entre eles.

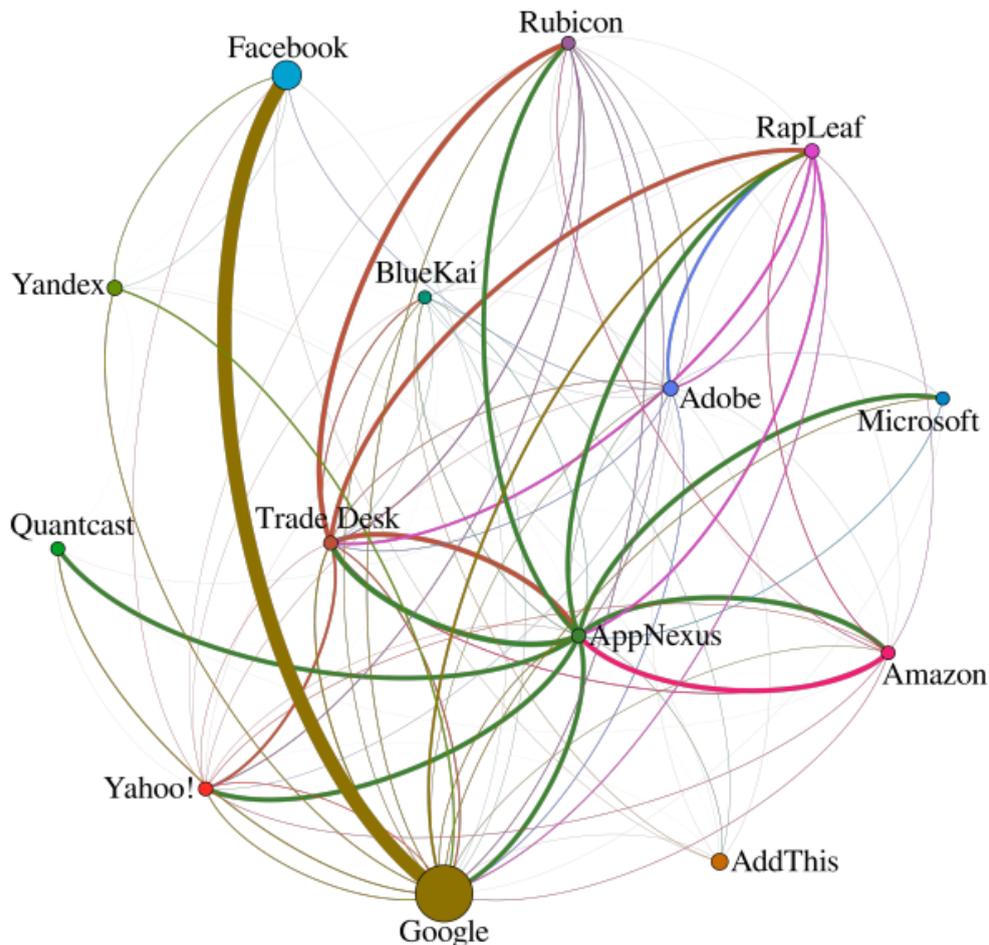
Tabela 3: Maiores empresas envolvidas nas ações de cookies

#	Actor	CC	Websites	Creator	Sender			Receiver (from other / self)			Intermediary			
					<i>own</i>	<i>in-chain</i>	<i>off-chain</i>	<i>own</i>	<i>in-chain</i>	<i>off-chain</i>	Creation	Sharing		
1	Google	US	415,545	93%	–	2%	14%	– / –	2%	1%	7%	10%	29%	32%
2	Facebook	US	171,699	96%	96%	–	20%	– / 96%	– / –	–	3%	20%	–	20%
3	AddThis	US	62,671	84%	80%	–	13%	4% / 80%	– / –	–	14%	9	1%	4%
4	Yandex	RU	49,036	93%	92%	1%	22%	7% / 92%	– / –	–	9%	22%	2%	92%
5	Adobe	US	45,835	71%	27%	2%	34%	2% / 26%	37% / 1%	1%	36%	33%	61%	81%
6	RapLeaf	US	42,698	13%	–	1%	75%	– / –	– / 1%	1%	57%	73%	10%	68%
7	AppNexus	US	39,991	77%	47%	47%	38%	– / 44%	9% / 46%	46%	36%	73%	87%	
8	Trade Desk	US	39,000	76%	14%	12%	4%	8% / 14%	4% / 11%	51%	4%	75%	93%	
9	Yahoo!	US	38,472	48%	20%	19%	55%	9% / 19%	10% / 19%	71%	53%	25%	62%	
10	Quantcast	US	36,662	68%	62%	–	11%	– / 62%	– / –	–	50%	11%	8%	26%
21	Adform	DK	25,627	24%	1%	–	26%	– / 1%	– / –	–	44%	23%	15%	87%
30	Baidu	CN	22,771	99%	12%	5%	5%	1% / 12%	– / 4%	–	5%	1%	11%	
32	WPP	GB	21,943	–	–	–	28%	– / –	– / –	–	86%	23%	2%	84%
34	Bidr	AU	20,918	81%	–	1%	6%	74% / –	–	63%	–	3%	–	90%
35	Horyzon	FR	20,872	89%	1%	11%	9%	1% / –	73% / 11%	20%	8%	93%	84%	
39	Casale	CA	19,595	32%	19%	–	60%	– / –	– / –	–	57%	45%	16%	43%
41	Eyeturn	CA	18,616	90%	–	–	–	– / –	– / –	–	17%	–	–	–
47	Criteo	FR	16,308	79%	64%	–	14%	14% / 64%	– / –	–	15%	12%	5%	39%
48	Mail.ru	RU	16,120	92%	7%	–	36%	2% / 7%	– / –	–	10%	36%	1%	7%
50	Avocet	GB	15,677	51%	3%	–	31%	– / 3%	9% / –	–	20%	31%	50%	86%

Fonte: Sanchez-rola *et al* - Eurocom (2021) - <https://www.eurocom.fr/publication/6376>

Os autores também criam um subgrafo de ligação que inclui organizações que aparecem em pelo menos 30.000 sites rastreados. A espessura da borda é proporcional ao peso: linhas esmaecidas indicam pequenos pesos, já a curvatura em sentido horário indica a direção e a cor identifica a origem.

Figura 5: Subgrafo da conexão na cadeia de cookies



Fonte: Sanchez-rola *et al.* - Eurocom (2021) - <https://www.eurecom.fr/publication/6376>

É possível observar que a grossura da linha entre Google e Facebook é a maior, e está relacionada à criação de *cookies* e inclusão de *tags* em sites (que é o ato de incluir o código dessas plataformas para mensuração do público), e são capazes de criar identificadores no navegador.

Além disso, existe um link entre diversas categorias como por exemplo Google/Alphabet que é dono do buscador google.com, do site de videos youtube.com e do advertiser doubleclick.net. A tabela abaixo mostra as maiores

categorias por grau (definido pela média do grau de entrada e saída), e é possível observar que os traços de *cookies* estão concentrados em algumas categorias, sendo a maior de ads/analytics, que está relacionada com a publicidade.

Tabela 4: Nós de grau mais alto no grafo por categoria

Category	Degree	Indegree	Outdegree
Web Ads/Analytics	31%	41%	22%
Technology/Internet	12%	12%	12%
Business/Economy	10%	9%	10%
Content Servers	7%	6%	9%
Social Networking	4%	7%	2%
Search Engines/Portals	4%	6%	2%
Shopping	3%	2%	5%
Audio/Video Clips	3%	4%	2%
News/Media	2%	1%	3%
Entertainment	2%	1%	3%
Mixed Content	2%	2%	1%
Suspicious	2%	2%	2%
Uncategorized	2%	1%	2%
Education	1%	–	2%
Health	1%	–	2%

Fonte: Sanchez-rola *et al* - Eurocom (2021) - <https://www.eurecom.fr/publication/6376>

O estudo também aponta a incompatibilidade entre os nós de entrada e de saída: os com maior entrada como *ads/analytics*, *social networking* e *search engines* são os que realizam a maior parte do rastreamento; já as categorias de shopping, *news/media* e *entertainment* possuem mais páginas que podem trazer informações aos rastreadores quando visitadas, demonstrando os diferentes papéis de cada um.

É possível observar que existem diversos atores envolvidos e o desafio para donos de sites e desenvolvedores de garantir a privacidade dos usuários, e enquanto não há uma forma definitiva de regulamentação devido a complexidade da rede, empresas continuam a se beneficiar com o uso dos dados pessoais, como os *Data Brokers*.

Data Brokers trabalham com a coleta, cruzamento e comercialização de dados em grande escala, obtidos através de empresas privadas, governo e pela internet, e são capazes de traçar perfis demográficos, econômicos,

comportamentais, de saúde, religião, sexualidade, eventos cotidianos, histórico de registro de imóveis, registro de veículos e registros judiciais, em uma indústria que gera cerca de U\$ 200 bilhões anualmente (Crain, 2016 apud Mott, 2014).

Quando surgem os primeiros *Data Brokers on-line*, no final de 1990, a coleta excessiva e fora da legislação por empresas como *DoubleClick* (do Google) fez com que elas desenvolvessem uma alta capacidade de segmentação para a publicidade, e os modelos de negócio online passaram a ser direcionados através da vigilância dos consumidores. Segundo Crain (2016, p. 96), assim que a *Federal Trade Commission*, órgão regulamentador dos Estados Unidos, descobriu que 85% dos sites coletavam informações sobre os consumidores e apenas 14% deixava clara tais práticas, a Associação de Marketing Direto (*Direct Marketing Association*) criou uma campanha para encorajar empresas a compartilhar suas políticas com os usuários. Quando elas se mostraram incompreensíveis, foram criados templates para uso geral, e que até hoje não possuem completa transparência, o que segundo o autor é uma forma que os *Data Brokers*, *web publishers* e empresas de publicidade on-line utilizam para manter uma auto-regulamentação.

“permita que os usuários saibam o mínimo possível sobre as atividades de coleta de dados, da maneira mais educada, mas complexa o suficiente para que eles não entendam o que está acontecendo, mas possam se sentir bem com isso” (Crain, 2016, p. 96 apud Turow, 2011, p. 83)

A discussão sobre a privacidade começa a aumentar a partir do momento em que os *Data Brokers* começam a cruzar as informações off-line com on-line, como o caso do Serasa Experian, que possui diversas informações pessoais e de análise de crédito. Como comprometimento à transparência na coleta dos dados, empresas como *DoubleClick* criaram portais como *privacychoices.org*, que permitiam os consumidores de excluir seus dados coletados (como a opção *opt-out*), o que se tornou muito comum na publicidade on-line e por consequência também enfraqueceu ações regulamentadoras, já que os consumidores possuem, em teoria, a opção de não compartilhar seus dados.

Importante ressaltar que os *Data Brokers* coletam seus dados de diversas fontes, sendo vendedores e compradores de informações em diferentes mercados,

o que incentiva a vigilância e transforma as pessoas em produtos e não consumidores.

2.3. Publicidade On-line e o papel de Cookies e Data Brokers

2.3.1. Monitoramento

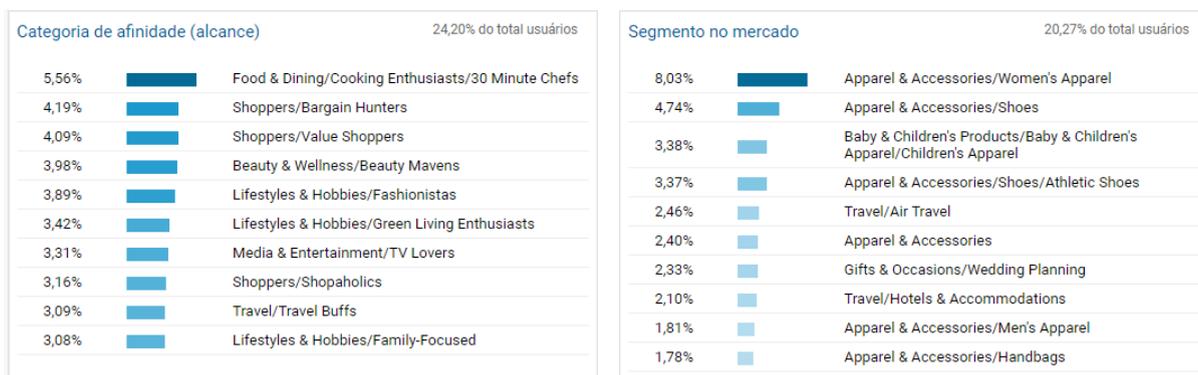
O monitoramento e mensuração em tempo real para a tomada de decisões rápidas e com foco em resultados é um dos principais diferenciais da publicidade *on-line*. Segundo Marques (2018), o site é o meio mais importante no Marketing Digital, pois aparece nos resultados dos buscadores e permite a divulgação de informações do produto ou serviço com maior credibilidade, além de ser onde normalmente ocorrem as vendas. Para o monitoramento de sites, o Google Analytics é uma das principais ferramentas, pois é capaz de capturar dados de público-alvo - como informações demográficas, interesses, páginas acessadas, origem dos acessos, dados de venda (origem, quantidade, produtos e receita), e diversas informações que auxiliam as estratégias de marketing através de uma *Tag* (ou etiqueta, em tradução literal) do Google Analytics que é instalada em todas as páginas do código do site.

Figura 6: Origem de tráfego - Google Analytics



Fonte: Google Analytics - <https://analytics.google.com/>

Figura 7: Visão Geral de Interesses - Google Analytics



Fonte: Google Analytics - <https://analytics.google.com/>

Figura 8: Dados de Aquisição por Origem/Mídia - Google Analytics

Origem/mídia	Aquisição			Comportamento			Conversões - Comércio eletrônico		
	Usuários	Novos usuários	Sessões	Taxa de rejeição	Páginas / sessão	Duração média da sessão	Taxa de conversão do comércio eletrônico	Transações	Receita
	7.094 Porcentagem do total: 100,00% (7.094)	6.449 Porcentagem do total: 100,00% (6.443)	8.707 Porcentagem do total: 100,00% (8.707)	57,59% Média de visualizações: 57,59% (0,00%)	2,78 Média de visualizações: 2,78 (0,00%)	00:01:37 Média de visualizações: 00:01:37 (0,00%)	0,84% Média de visualizações: 0,84% (0,00%)	73 Porcentagem do total: 100,00% (73)	R\$ 177.501,23 Porcentagem do total: 100,00% (R\$ 177.501,23)
1. facebook-instagram / social	2.026 (27,37%)	1.813 (28,11%)	2.121 (24,36%)	72,65%	1,78	00:00:34	0,19%	4 (5,48%)	R\$ 7.280,00 (4,10%)
2. google / cpc	1.838 (24,83%)	1.525 (23,65%)	2.501 (28,72%)	52,06%	3,32	00:02:06	0,92%	23 (31,51%)	R\$ 48.660,00 (27,41%)
3. (direct) / (none)	1.589 (21,47%)	1.520 (23,57%)	1.821 (20,91%)	56,29%	2,81	00:01:33	1,43%	26 (35,62%)	R\$ 59.680,00 (33,62%)
4. google / organic	644 (8,70%)	478 (7,41%)	802 (9,21%)	38,03%	4,08	00:03:06	1,50%	12 (16,44%)	R\$ 37.601,23 (21,18%)
5. 020 - Newsletter DB - BR / email	361 (4,88%)	322 (4,99%)	405 (4,65%)	45,43%	2,99	00:01:50	0,25%	1 (1,37%)	R\$ 410,00 (0,23%)
6. facebook.com / referral	291 (3,93%)	261 (4,05%)	299 (3,43%)	60,20%	2,16	00:01:11	0,00%	0 (0,00%)	R\$ 0,00 (0,00%)
7. m.facebook.com / referral	280 (3,78%)	251 (3,89%)	303 (3,48%)	82,84%	1,40	00:00:29	0,33%	1 (1,37%)	R\$ 2.650,00 (1,49%)
8. instagram.com / referral	109 (1,47%)	87 (1,35%)	115 (1,32%)	63,48%	2,24	00:00:49	0,00%	0 (0,00%)	R\$ 0,00 (0,00%)
9. IGShopping / Social	48 (0,65%)	41 (0,64%)	59 (0,68%)	52,54%	3,88	00:02:33	0,00%	0 (0,00%)	R\$ 0,00 (0,00%)
10. l.facebook.com / referral	42 (0,57%)	28 (0,43%)	44 (0,51%)	63,64%	2,02	00:00:30	0,00%	0 (0,00%)	R\$ 0,00 (0,00%)

Fonte: Google Analytics - <https://analytics.google.com/>

Com as *tags*, é possível monitorar todo o comportamento de navegação dos usuários e obter diversas informações que podem auxiliar nas decisões de marketing. Além de dados em tempo real e de interesses dos usuários, também é possível analisar qual o tempo até a conversão, a origem e impacto de cada canal para a venda final. Na figura abaixo, é possível observar que o principal volume de conversões é proveniente dos usuários que acessam o site duas vezes antes de efetuar uma compra, e a influência da pesquisa paga e redes sociais na jornada do cliente.

Figura 9: Caminhos em Funis Multicanal - Google Analytics

Caminho dos grupos de canais MCF ?	Conversões ?	↓	Valor da conversão ?
1. Direto x 2	84 (10,97%)		R\$ 118.460,00
2. Pesquisa paga → Direto	54 (7,05%)		R\$ 61.765,23
3. Direto x 3	32 (4,18%)		R\$ 22.660,00
4. Pesquisa paga x 2	29 (3,79%)		R\$ 20.300,00
5. Pesquisa orgânica → Direto	28 (3,66%)		R\$ 51.660,00
6. Direto → Pesquisa paga	18 (2,35%)		R\$ 16.790,00
7. Direto x 4	17 (2,22%)		R\$ 26.170,00
8. Pesquisa paga → Direto x 2	12 (1,57%)		R\$ 7.241,60
9. Direto → Rede social	11 (1,44%)		R\$ 6.550,00
10. Pesquisa paga x 2 → Direto	11 (1,44%)		R\$ 16.290,00

Fonte: Google Analytics - <https://analytics.google.com/>

O monitoramento e mensuração também são utilizados em campanhas de diversas plataformas - como as principais Google e Facebook, para otimização de resultados. A instalação das *tags* em todo o site permite que as plataformas possam otimizar os resultados de acordo com o objetivo e meta de cada campanha. Por exemplo, uma campanha de vendas será otimizada para usuários com maior propensão a chegar na página da compra finalizada, enquanto uma campanha de remarketing identifica usuários que acessaram o site e é utilizada para chamar usuários para finalizar a compra. Com todos os dados disponíveis, é possível encontrar o público-alvo na hora e lugar certos, com menor dispersão e foco no maior retorno sobre investimento.

2.3.2. Segmentação de campanhas pelo Google Ads

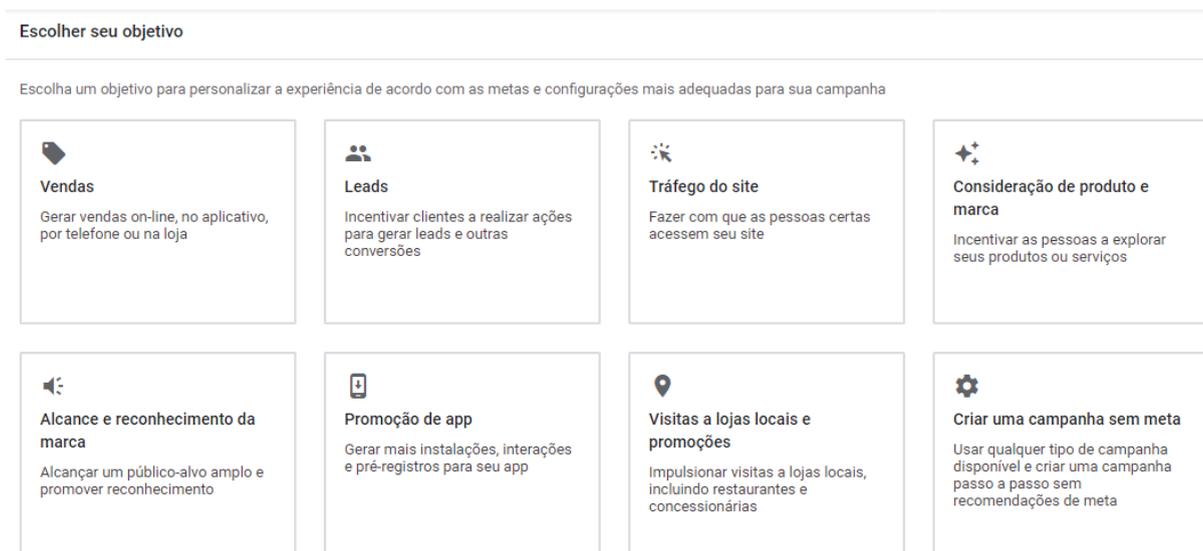
O Google foi fundado em 1998, pelos estudantes de doutorado Larry Page e Sergey Brin na Universidade de Stanford, Estados Unidos, após o sucesso do projeto que descreveu as propriedades técnicas de um buscador capaz de rastrear a web inteira e listar as páginas com base em sua relevância. Para isso, criaram o *PageRank*, uma tecnologia que determinava a relevância do site de acordo com o termo de pesquisa, considerando o número de páginas de determinado domínio e quantas páginas faziam um link para ele. Esse sistema passou a se chamar Google,

em referência à palavra googol, que representa o número 1 seguido pelo número 0 cem vezes, como uma alusão a enorme quantidade de páginas da web.

Apesar de ter se iniciado como um buscador, atualmente o Google possui diversos outros serviços, e entre os mais conhecidos estão: Pesquisa, Mapas, Tradutor, Chrome (navegador), Gmail (e-mail), YouTube e Drive (para armazenamento de arquivos).

Para anunciantes, o Google Ads é a plataforma de criação de campanhas, com diversas opções de acordo com cada objetivo de negócio, capaz de abranger diversas etapas do funil de marketing, que vão desde reconhecimento de marca até a venda final.

Figura 10: Objetivos de campanha - Google Ads

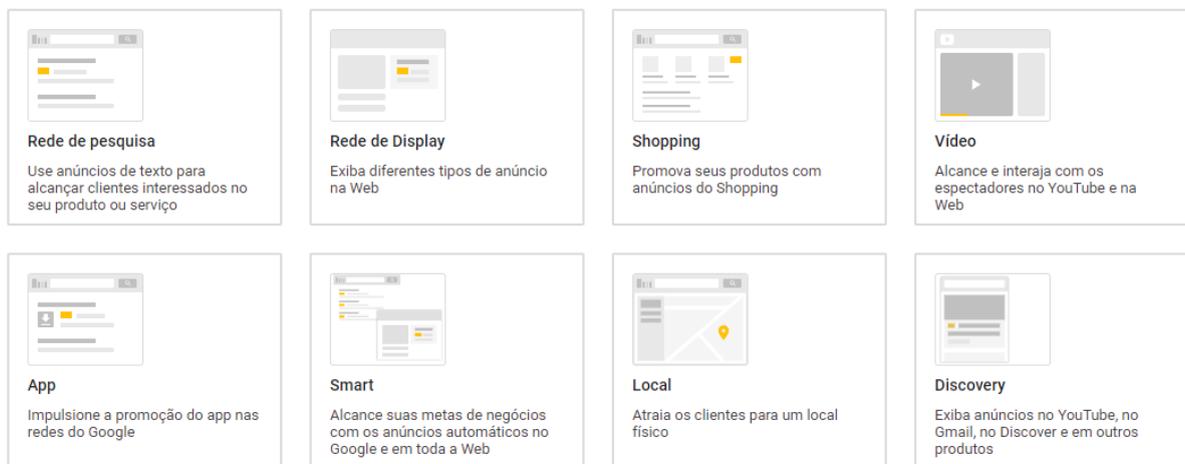


Fonte: Google Ads - <https://ads.google.com/>

Existem oito tipos de campanhas disponíveis, que incluem anúncios em formatos de texto, imagem e vídeo:

Figura 11: Tipos de campanha - Google Ads

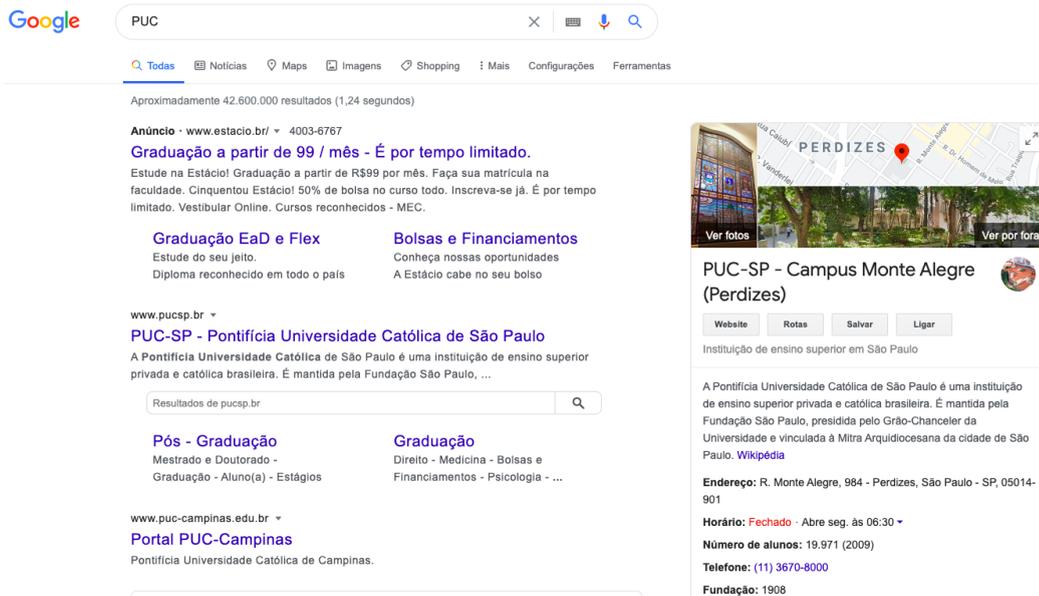
Selecione um tipo de campanha



Fonte: Google Ads - <https://ads.google.com/>

Em todos os tipos de campanhas, os anúncios aparecem de acordo com a segmentação do público-alvo. Na campanha de rede de pesquisa, por exemplo, os anúncios de texto são acionados com base em palavras-chave que são inseridas na pesquisa Google. Esses termos podem ser relacionados ao produto ou serviço, como no exemplo abaixo, em que ao buscar a palavra PUC, aparece em primeiro lugar o anúncio da Estácio. Isso significa que esse concorrente está pagando para aparecer quando um usuário busca pela PUC. Já em segundo lugar, aparece o site da PUC de forma orgânica (ou seja, não é um anúncio), devido à sua relevância e palavras-chave que estão contidas dentro do site e auxiliam o algoritmo do Google a identificar os melhores resultados.

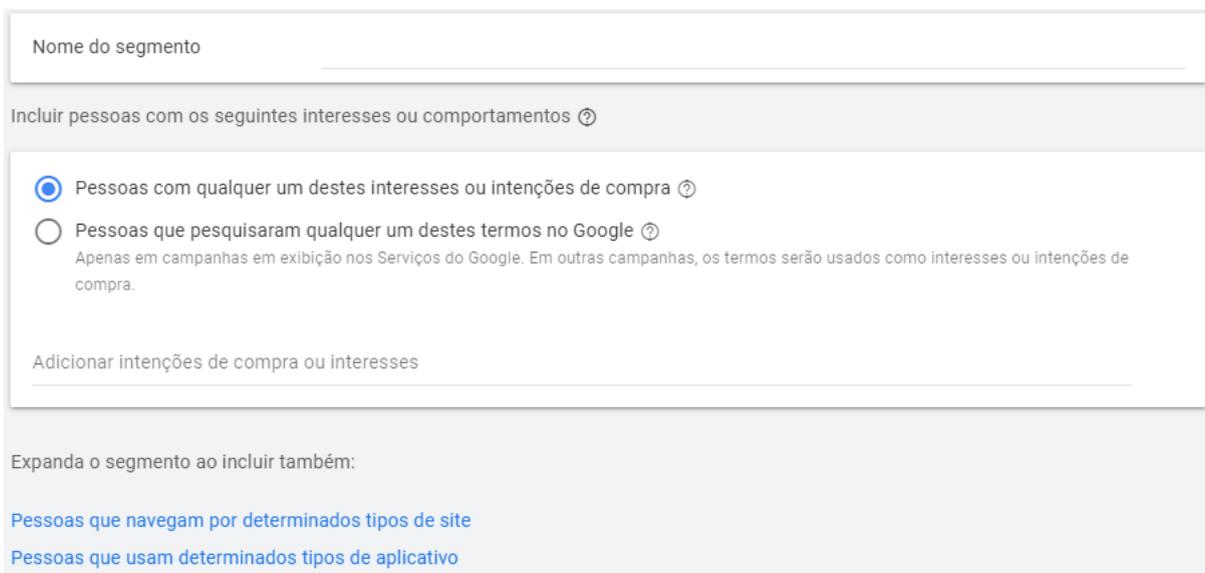
Figura 12: Busca no Google



Fonte: Google.com - <https://www.google.com/search?q=puc>

Nas campanhas de Display, Vídeo ou Discovery, também é possível criar públicos personalizados de acordo com palavras-chave, com a diferença de que os anúncios não irão aparecer na hora exata da busca, e sim para identificar padrões de comportamento dos consumidores para exibir os anúncios. Por exemplo, visualizar um anúncio da PUC no YouTube após fazer a busca pela Universidade no Google.

Figura 13: Segmentos Personalizados - Google Ads



Fonte: Google Ads - <https://ads.google.com/>

Segundo instruções da própria plataforma, a escolha dos interesses é através de palavras-chaves ou frases que representam o cliente ideal, e “os anúncios alcançarão pessoas com possível interesse ou intenção de compra nessas palavras-chave com base no comportamento e atividade delas, como aplicativos que elas usam ou tipo de conteúdo que pesquisam na Internet”.

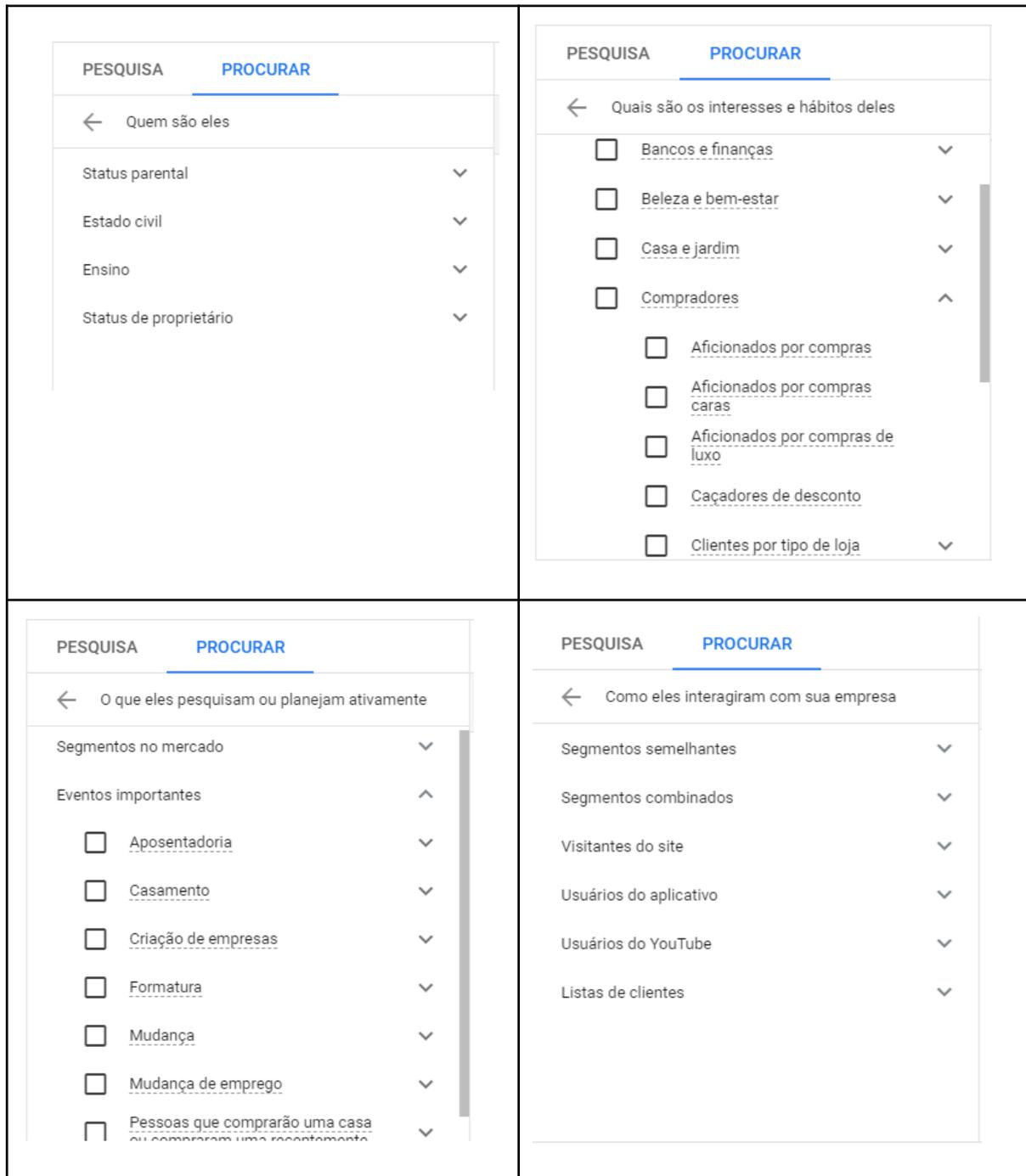
Além da inclusão manual das segmentações, é possível também escolher entre os segmentos de público-alvo já disponíveis na plataforma, que incluem informações demográficas, interesses e hábitos, o que os usuários pesquisam ativamente, interações com a empresa (como usuários que já acessaram o site ou fizeram uma compra) e usuários identificados com comportamentos semelhantes a esses.

Figura 14: Segmentos de público-alvo - Google Ads



Fonte: Google Ads - <https://ads.google.com/>

Figura 15: Detalhe do Segmento de Público-alvo - Google Ads



Fonte: Google Ads - <https://ads.google.com/>

Para segmentos de dados, que são com base em lista de clientes e atividades no site da marca, a própria plataforma disponibiliza o “tamanho” da lista

de *cookies* que estão associados ao público, tornando possível estimar quantas pessoas podem ser alcançadas em cada rede.

Figura 16: Segmentos de Dados - Google Ads

<input type="checkbox"/> Nome do segmento	Tipo ↓	Status da associação	Tamanho: rede de pesquisa	Tamanho: YouTube	Tamanho: Rede de Display	Tamanho: campanha do Gmail
Em uso						
<input type="checkbox"/> Todos os visitantes (Google Ads) Pessoas que acessaram páginas contendo suas tags de remarketing	Visitantes do site Criada automaticamente	Aberta	310.000	310.000	190.000	58.000

Fonte: Google Ads - <https://ads.google.com/>

O Google também disponibiliza a expansão automática do alcance das segmentações, complementando e reforçando os públicos já configurados para melhor atingir a meta da campanha. Como por exemplo, em segmentações de palavras-chave, o Google Ads incluirá outras palavras ou tópicos semanticamente relevantes que podem levar a conversões.

Também há a opção de criar campanhas inteligentes, que não requerem nenhuma forma manual de configuração, exibindo anúncios onde a plataforma identifica que há mais oportunidades de negócio, através do remarketing automático de pessoas que acessaram o site e usuários que ainda estão nas primeiras etapas do processo de compra.

Em todos os tipos de campanhas e segmentações, a similaridade entre elas é a utilização de *cookies* para a identificação do comportamento de navegação do público, independente se isso é feito de forma manual ou automática, o que mostra a grande importância dos *cookies* na publicidade.

2.3.2. Segmentação de campanhas pelo Facebook para Empresas

O Facebook foi criado por quatro alunos da Universidade de Harvard em 2004, época em que os *blogs* eram o grande destaque da internet. Inicialmente, a rede social foi idealizada para ser utilizada apenas por alunos de Harvard, mas logo se expandiu para diversas Universidades, até que em 2006 se tornou aberta ao público até se tornar a maior rede social do mundo, e atualmente possui cerca de 2.5 bilhões de usuários mensais ativos, segundo dados da Adobe¹⁰.

¹⁰ Top Social Media Sites, disponível em: <https://www.adobe.com/express/learn/blog/top-social-media-sites>

Assim como o Google, o Facebook também disponibiliza diversos tipos de campanha de acordo com o objetivo de negócio:

Figura 17: Objetivos de campanha - Facebook para Empresas

Escolha um objetivo da campanha

[Saiba mais](#)

Reconhecimento	Consideração	Conversão
<input type="radio"/> Reconhecimento da marca	<input type="radio"/> Tráfego	<input type="radio"/> Conversões
<input type="radio"/> Alcance	<input type="radio"/> Engajamento	<input type="radio"/> Vendas do catálogo
	<input type="radio"/> Instalações do aplicativo	<input type="radio"/> Tráfego para o estabelecimento
	<input type="radio"/> Visualizações do vídeo	
	<input type="radio"/> Geração de cadastros	
	<input type="radio"/> Mensagens	

Fonte: Facebook Business <https://business.facebook.com/>

No direcionamento detalhado, é possível incluir dados demográficos, de interesses e comportamentos que são obtidos da mesma forma que o Google - através do histórico de navegação dos usuários.

Figura 18: Segmentações de campanha - Facebook para Empresas

Localizações

Localização:

- Brasil

Idade

18 - 65+

Gênero

Todos os gêneros

Direcionamento detalhado

Incluir pessoas com correspondência ⓘ

🔍 Adicione dados demográficos, interesses ou comportament Sugestões Procurar

▼ Dados demográficos ⓘ

- ▶ Formação
- ▼ Financeiro
 - ▶ Renda
- ▶ Acontecimentos
- ▶ País
- ▶ Relacionamento
- ▶ Ocupação
- ▶ Interesses

Fonte: Facebook Business <https://business.facebook.com/>

Também é possível criar públicos personalizados, com pessoas que já demonstraram interesse na empresa ou produto, baseado em atividades no site (através da *tag*), lista de clientes ou interações dos usuários com a marca na rede social.

Figura 19: Público Personalizado - Facebook para Empresas



Fonte: Facebook Business <https://business.facebook.com/>

A grande diferença entre o Google e Facebook, é que o Google além de exibir os anúncios na rede própria (como o buscador ou YouTube) se abrange em diferentes sites parceiros através da rede de Display, enquanto o Facebook está limitado a exibir anúncios dentro de suas plataformas: Facebook, Instagram, Messenger e também através da Audience Network, que permite a monetização de anúncios criados no Facebook através de sua exibição em aplicativos próprios (para sistemas operacionais Android ou iOS) de empresas¹¹.

2.4. Uso de dados na publicidade on-line e regulamentações

Como visto no tópico anterior, o anunciante exibe sua marca, produtos ou serviços em páginas ou aplicativos de terceiros (os chamados *publishers*), que são

¹¹ Sobre o Audience Network - Facebook. Disponível em: <https://pt-br.facebook.com/business/help/452287605232720?id=211412110064838>. Acesso em 24/09/2021.

responsáveis por veicular a mídia (no exemplo citado, o Google e Facebook). Já o consumidor está compartilhando suas informações que serão utilizadas sem seu conhecimento para lhe convencer a comprar produtos que não precisa - ou que precisa menos.

Segundo Calo (2014), a utilização dos dados pessoais na publicidade pode gerar danos econômicos ao diferenciar preços de acordo com indicadores geográficos e financeiros; à privacidade, ao comercializar dados pessoais para diferentes fins, como a coleta de dados que pode determinar que o consumidor está ativamente buscando um produto e cobrar mais caro por ele, ou um usuário que utiliza seu e-mail para fazer login em um site e seu e-mail é vendido para outras empresas sem sua autorização; e danos à autonomia, já que os dados sobre determinado usuário podem ser tão detalhados a ponto de induzir comportamentos de consumo ao identificar o momento e a forma exata para os impactar com anúncios, que tendem a garantir um momento de irracionalidade. Nesse contexto de manipulação no marketing digital, o consumidor não consegue saber se o produto que está sendo oferecido apresenta os mesmos preços do que para outros usuários, ou se a utilização de um navegador diferente mostrará um outro preço.

A Lei Geral de Proteção de Dados (LGPD), que entrou em vigor em agosto de 2020, surge para proteger o direito à privacidade de dados, mas não possui nenhuma regulamentação em relação aos *cookies*. A LGPD possui diversas similaridades com a GDPR (General Data Protection Regulation), porém no caso da lei regulamentadora na União Européia, ela especifica que *cookies* são considerados legítimos para fins de coleta de informações de funcionamento de um site ou para facilitar a publicidade, e por isso as informações acerca de sua utilização precisam estar claras e acessíveis, como seus propósitos e armazenamento no computador do usuário, condicionando a navegação do site à aceitação do *cookie* (OLIVEIRA, 2018, p. 320 *apud* UNIÃO EUROPEIA, 2002), para que o usuário possa avaliar e permitir o acesso aos seus dados.

Com essa nova diretriz, ao acessar um site, o internauta possui duas opções centrais: tentar recusar os *cookies* e sofrer bloqueios de acesso ao servidor ou aceitá-los e ficar exposto aos riscos de privacidade (Oliveira, 2018, p. 322 *apud* Tirtéia *et al.* 2011, p. 12-3). Porém, é importante lembrar que existem também os *supercookies* e *evercookies*, capazes de contornar as preferências dos usuários, além dos *ghostwriters* que por mais que tenham a aprovação do uso de dados

através de *cookies first-party*, não têm sua coleta especificada e aprovada pelo usuário.

A utilização dos *cookies* se mostra essencial para a personalização de anúncios na publicidade on-line, pois agem como potencializadores na identificação e construção de perfil de pessoas físicas, e também que sua complexidade é de difícil compreensão até mesmo aos órgãos regulamentadores. Por isso, a pesquisa busca compreender o quanto os usuários entendem sobre esse fluxo de informações e como lidam com o compartilhamento de seus dados pessoais.

3. MATERIAIS E INSTRUMENTOS DE PESQUISA

Nessa sessão serão apresentados os critérios para a construção do questionário, seleção do público-alvo e instrumentos para coleta e tratamento dos dados.

3.1. Instrumento de pesquisa: questionário

O questionário aplicado será estruturado fechado, com perguntas que trazem espaços para marcação da escolha da resposta, um formato adequado quando a aplicação do questionário é feita sem a presença de um entrevistador e não é possível que o entrevistado faça perguntas adicionais (Andrade, 2010).

A distribuição do questionário foi feita em duas dimensões:

1. Conhecimento sobre a publicidade on-line (hipóteses A e B);
2. Cuidado com o uso dos dados pessoais na internet e sua relação com a publicidade (hipóteses C, D e E).

3.1.1. Delimitação do universo e da amostra

Para aplicação da pesquisa, o primeiro passo é a delimitação do universo, a partir de elementos que possuem características em comum (Lakatos e Marconi, 2021), delimitando-as em um subconjunto de amostras.

Para a seleção da amostragem, foi escolhida a não probabilística por apresentar resultados mais satisfatórios, já que as conclusões se limitam aos participantes da pesquisa, segundo Silva *et al.* (2019).

O tipo de amostragem por quotas é, segundo Gil (2008), o que apresenta maior rigor, e é desenvolvido em três fases: classificação da população em função de propriedades relevantes para o fenômeno estudado; determinação da proporção da população em cada classe e fixação de quotas para que a amostra seja composta de acordo com a proporção das classes consideradas.

Como a pesquisa tem o objetivo de identificar diferentes percepções de acordo com as gerações, a divisão da amostra foi feita segundo dados do Pew Research Center¹², a geração Baby Boomer é caracterizada por indivíduos nascidos

¹² Pew Research Center, Defining generations: Where Millennials end and Generation Z begins, disponível em: <https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/>

de 1946 a 1964, geração X de 1965 a 1980, geração Y ou millennials de 1981 a 1996 e geração Z de 1997 a 2012.

Tabela 5: Definição de Gerações

Geração	Faixa Etária
Geração Z	18 - 24 anos
Geração Y	25 - 39 anos
Geração X	40 - 56 anos
Baby Boomers	57 - 75 anos

Fonte: elaborada pela autora

Apesar de a Pew Research Center considerar a Geração Z os indivíduos nascidos até 2012, por uma questão de ética de pesquisa serão considerados apenas os indivíduos maiores de idade, ou seja, a partir de 18 anos.

Para a definição da proporção por geração, foram utilizados dados do IBGE/Diretoria de Pesquisas - Coordenação de População e Indicadores Sociais, Gerência de Estudos e Análises da Dinâmica Demográfica, considerando a população Brasil e excluindo menores de 18 anos e acima de 75 anos, repercentualizando o total populacional elegível para a pesquisa.

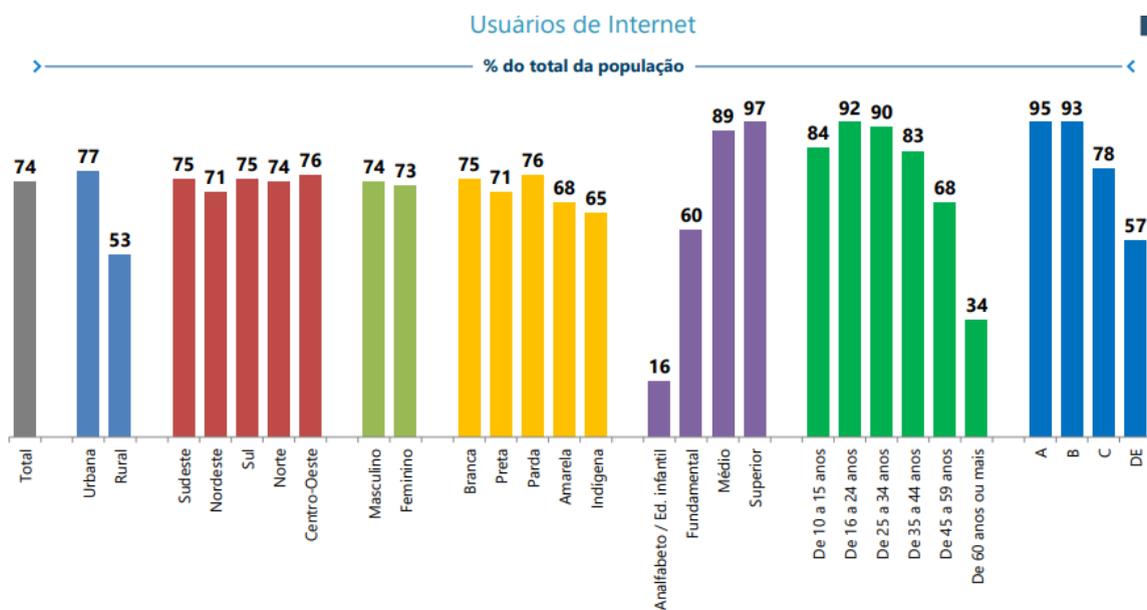
Tabela 6: Proporção da população por geração, segundo IBGE

FAIXA IDADE	GERAÇÃO	POPULAÇÃO	%
18-24	Geração Z	23.354.290	15%
25-39	Geração Y	51.270.587	34%
40-56	Geração X	47.376.374	31%
57-75	Baby Boomers	30.549.652	20%
TOTAL GERAÇÕES		152.550.903	100%

Fonte: elaborada pela autora

Foi considerada também a pesquisa TIC Domicílio 2019, que traz 134 milhões de usuários de internet no Brasil, mas que não possui a mesma quebra geracional considerada na presente pesquisa. Dado o alto número de usuários de internet, optou-se por seguir a quebra proporcional do IBGE.

Figura 20: Usuários de internet no Brasil



Fonte: TIC Domicílios (2019)

https://cetic.br/media/analises/tic_domicilios_2019_coletiva_imprensa.pdf

Sendo assim, foi definida uma amostra com 500 casos, respeitando a proporção de gerações e considerando a otimização de recursos, qualidade para investigação dos objetivos da pesquisa e viabilidade da aplicação do questionário.

Tabela 7: Quantidade de participantes por geração

Geração	Faixa Etária	%	Casos
Geração Z	18 - 24 anos	15%	75
Geração Y	25 - 39 anos	34%	170
Geração X	40 - 56 anos	31%	155
Baby Boomers	57 - 75 anos	20%	100
TOTAL		100%	500

Fonte: elaborada pela autora

3.1.2. Coleta e tratamento de dados

A coleta de dados será feita pela amostragem por voluntários através da empresa MindMiners, com o envio do questionário através da plataforma de pesquisa MeSeems. A escolha dessa empresa foi feita por já possuírem públicos definidos de acordo com perfis demográficos, e o questionário será enviado de acordo com o volume de respondentes necessário.

O tratamento dos dados seguirá conforme os critérios de seleção, codificação e tabulação descritos por Lakatos e Marconi (2021), que incluem, respectivamente: exame dos dados para detectar erros e respostas incompletas ou distorcidas; categorização de dados para facilitar sua tabulação e interpretação e por fim a tabulação para análise estatística e representação gráfica.

A utilização da escala de Likert (Likert, 1932) com 5 níveis será aplicada através das opções: discordo totalmente, discordo parcialmente, indiferente, concordo parcialmente e concordo totalmente, e as respostas indicarão atitudes mais ou menos favoráveis, gerando resultados que destacam de forma mais clara, possibilitando a análise entre as gerações.

4. RESULTADOS E DISCUSSÕES

Nesta seção, serão apresentadas as respostas da pesquisa e sua análise em relação às hipóteses levantadas no trabalho.

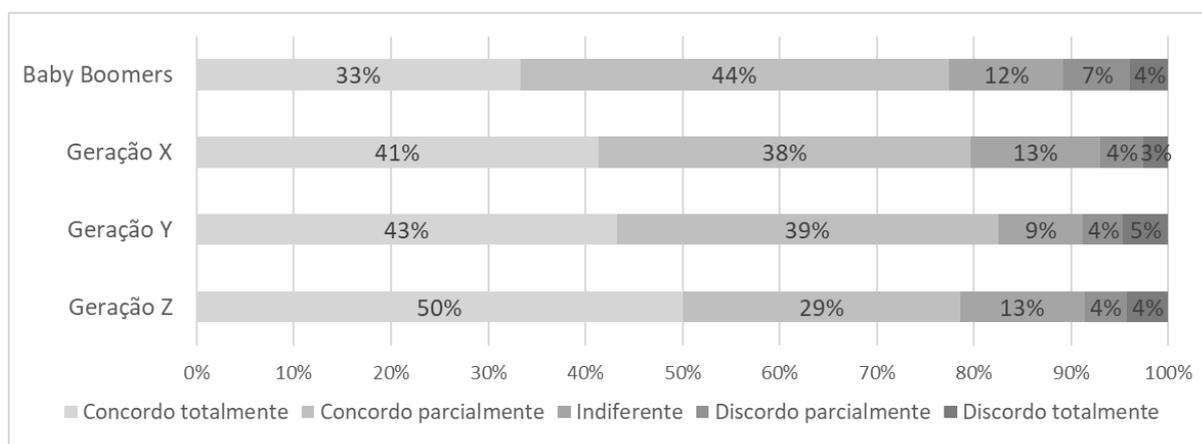
4.1. Resultados da Pesquisa

CONHECIMENTO SOBRE A PUBLICIDADE ON-LINE (HIPÓTESES A E B)

Questão 1 - O seu feed de notícias nas redes sociais é a página principal de acesso a atualizações de amigos, familiares, empresas e fontes de notícias aos quais está conectado. Agora, vamos te apresentar algumas frases e queremos que você responda o quanto concorda ou discorda de cada uma:

Acredito que o conteúdo que aparece no meu feed é personalizado de acordo com meus interesses.

Tabela 8: Pesquisa - Respostas questão 1

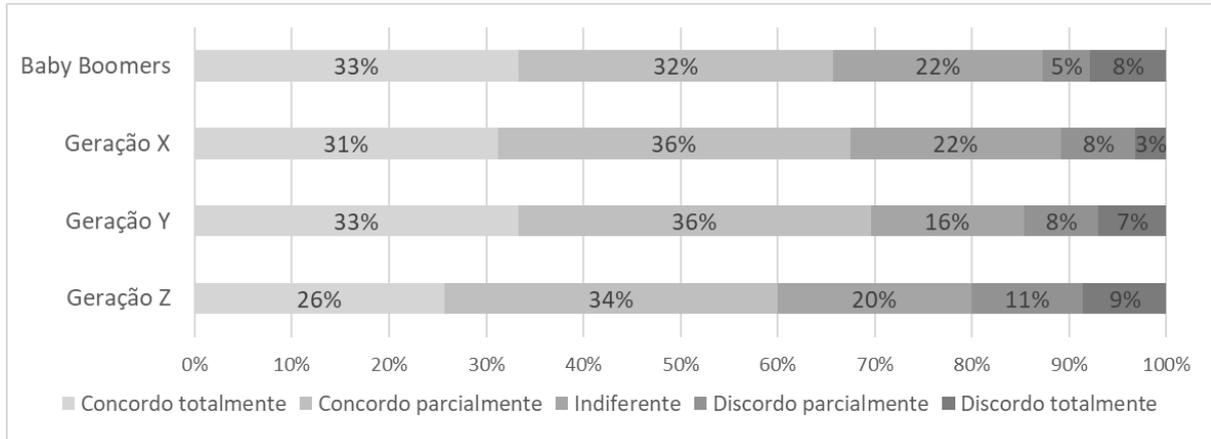


Fonte: elaborada pela autora

A primeira questão busca compreender a percepção dos usuários quanto a seleção dos conteúdos que aparecem em seus feeds em redes sociais. É possível observar que uma média de 80% dos respondentes entre todas as gerações concorda com a afirmação de que o feed é personalizado de acordo com seus interesses.

Questão 2 - Tenho conhecimento de que é possível consultar, adicionar e remover as preferências de propaganda criadas pelas plataformas Google e Facebook.

Tabela 9: Pesquisa - Respostas questão 2

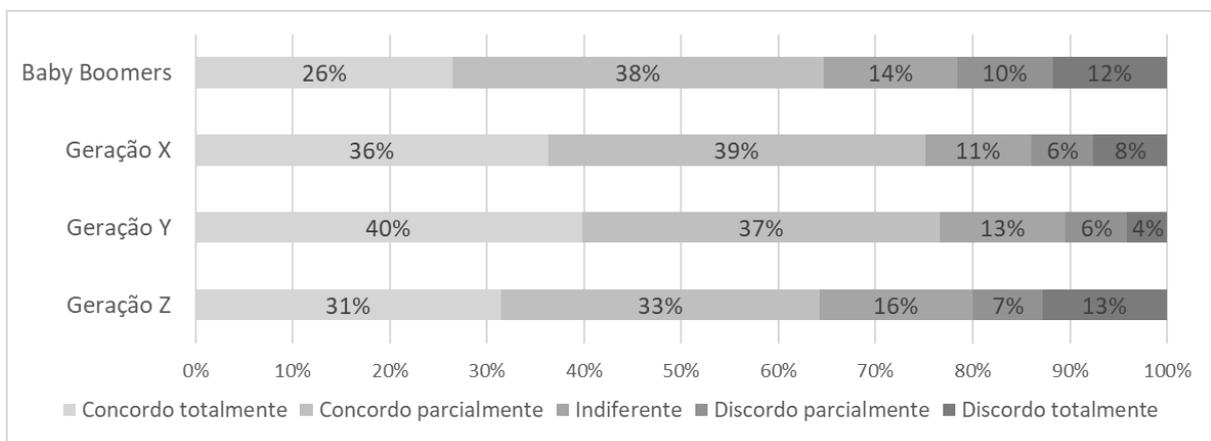


Fonte: elaborada pela autora

Na questão 2, 66% dos respondentes afirmam ter conhecimento sobre as preferências de propaganda nas plataformas, enquanto 19% se mostram indiferentes. Ao fazer o recorte por gerações, é possível observar que a Geração Y possui maior grau de concordância (70%) e a Geração Z possui menor grau de concordância (60%).

Questão 3 - Tenho conhecimento de que meus dados (como idade, gênero, e-mail, localização, cargo e renda) podem ser utilizados para a publicidade direcionada.

Tabela 10: Pesquisa - Respostas questão 3



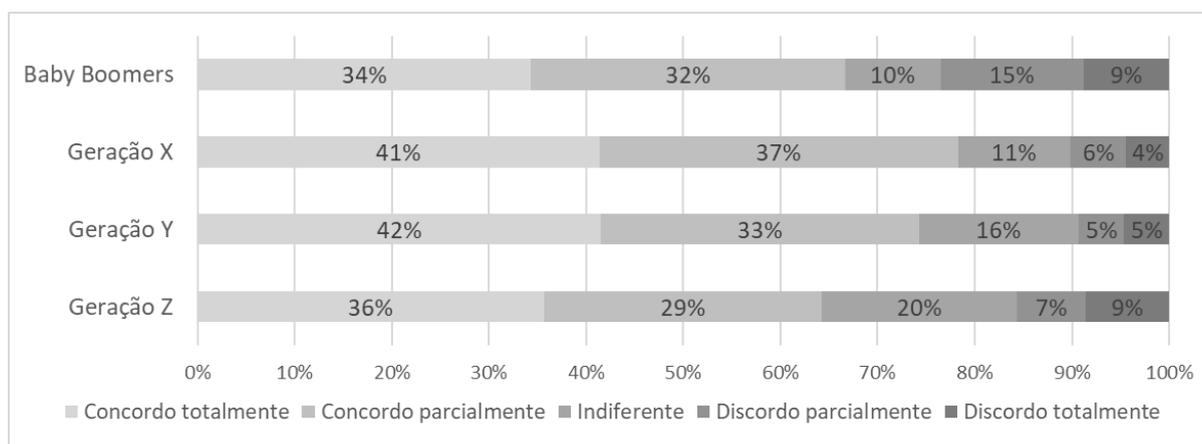
Fonte: elaborada pela autora

É possível observar que 72% dos respondentes concordam que são utilizadas suas informações de idade, gênero, e-mail, localização, cargo e renda para a publicidade on-line.

Quando a quebra é feita por gerações, as Gerações Y e X demonstram mais conhecimento sobre os dados que são utilizados, e respectivamente 77% e 75% concordam com a frase. Já as gerações Baby Boomers e Z também concordam (65% e 64%, respectivamente), mas em menor grau.

Questão 4 - Tenho conhecimento de que o meu histórico de navegação na internet pode ser utilizado para a publicidade direcionada.

Tabela 11: Pesquisa - Respostas questão 4

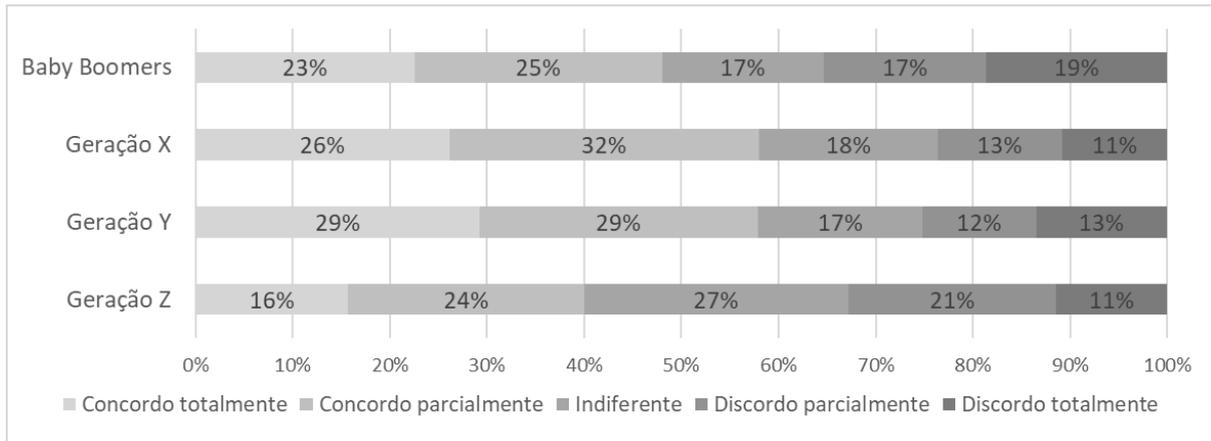


Fonte: elaborada pela autora

A questão 4 complementa a pergunta anterior, mas com foco nos dados de navegação. O grau de concordância com a afirmação se mantém similar - 73%, enquanto na quebra geracional a Geração X demonstra maior grau de concordância (78%) e a Geração Z menor grau (64%).

Questão 5 - Nas plataformas em que tenho perfil cadastrado (como Google e Facebook), já configurei as minhas preferências de anúncios.

Tabela 12: Pesquisa - Respostas questão 5

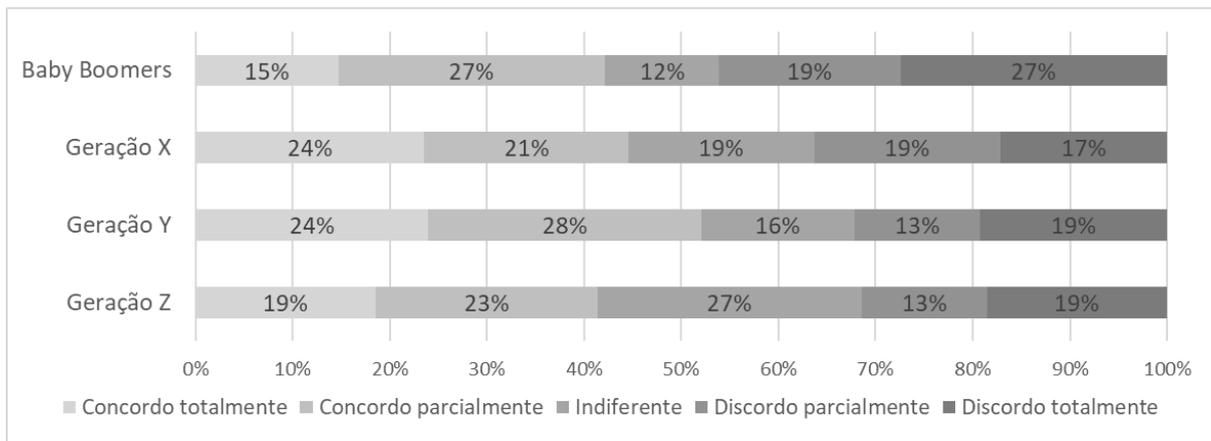


Fonte: elaborada pela autora

Quando o foco da pergunta muda para qual usuário efetivamente fez alguma mudança em suas preferências em suas plataformas, é possível observar uma queda no grau de concordância para 53%. Nas Gerações Y e Z, 58% concordam, enquanto para Baby Boomers 35% discordam da afirmação. A Geração Z possui menor grau de concordância (40%), e o maior volume de respostas indiferentes (27%).

Questão 6 - Consigo ensinar para outras pessoas como configurar suas preferências de anúncios no Google e Facebook.

Tabela 13: Pesquisa - Respostas questão 6



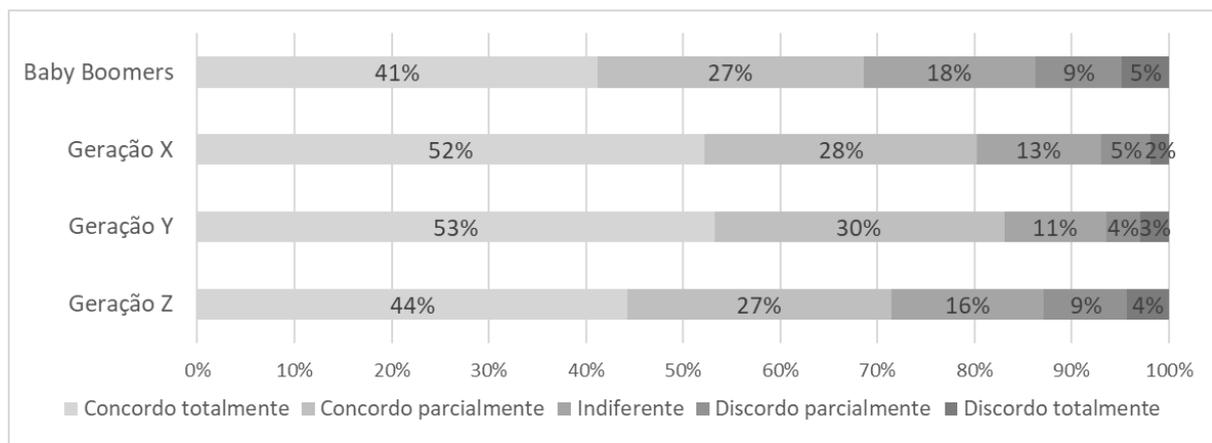
Fonte: elaborada pela autora

Ao questionar os usuários se eles conseguiriam ensinar outras pessoas sobre as configurações disponíveis, observamos uma queda no grau de concordância para 46%. A Geração X demonstra maior queda: de 58% dos que afirmaram já ter feito configurações (questão 5), apenas 45% se diz capaz de ensinar outras pessoas. A menor divergência está na Geração Z, em que o volume vai de 41% na questão anterior para 42% de concordância na questão 6.

CUIDADO COM O USO DE DADOS PESSOAIS NA INTERNET E SUA RELAÇÃO COM A PUBLICIDADE (HIPÓTESES C, D E E)

Questão 7 - Já notei que recebi a mesma propaganda em diferentes sites e redes sociais.

Tabela 14: Pesquisa - Respostas questão 7

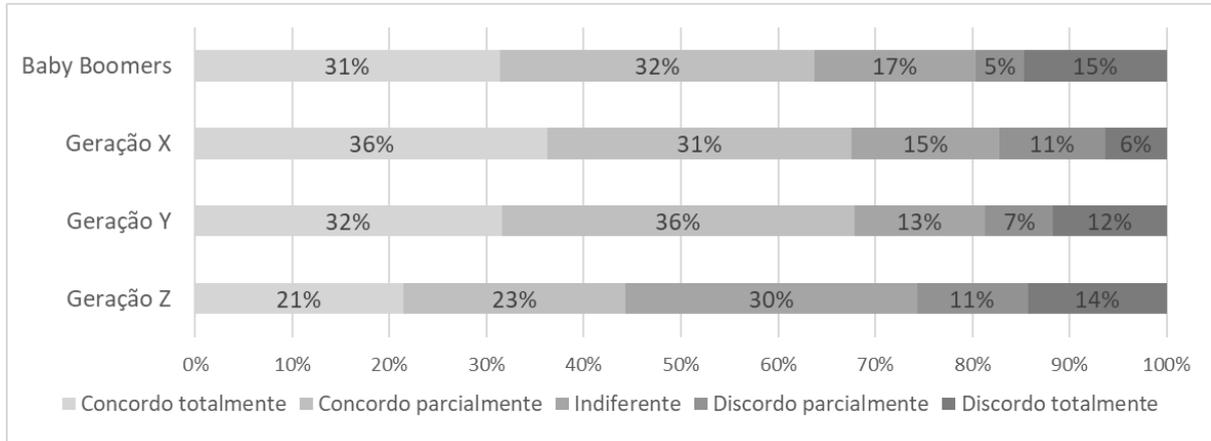


Fonte: elaborada pela autora

Na questão 7, 78% dos respondentes confirmam já ter notado a mesma propaganda em diferentes sites. O volume entre as Gerações X e Y se mantém bem similar (respectivamente 80% e 83%), enquanto para a Geração Z o grau de concordância cai para 71% e para Baby Boomers observamos uma nova queda para 69%.

Questão 8 - Compreendo qual o papel dos cookies de navegador no armazenamento de informações.

Tabela 15: Pesquisa - Respostas questão 8

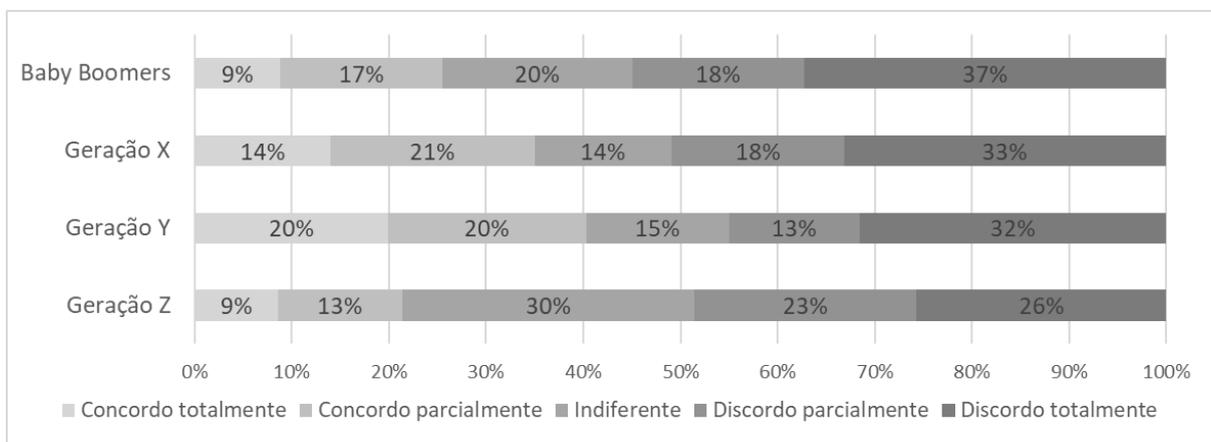


Fonte: elaborada pela autora

A questão 8 busca compreender o conhecimento sobre os *cookies* no papel do armazenamento das informações, e observamos que 64% dos respondentes concordam com a afirmação. O grau de concordância entre as Gerações Y e X está em 80%, enquanto para Baby Boomers o volume cai para 64%. Já para a Geração Z, apenas 44% concordam e 30% se mostram indiferentes.

Questão 9 - Tenho conhecimento para explicar para amigos e colegas de trabalho o que são cookies first-party e third-party.

Tabela 16: Pesquisa - Respostas questão 9

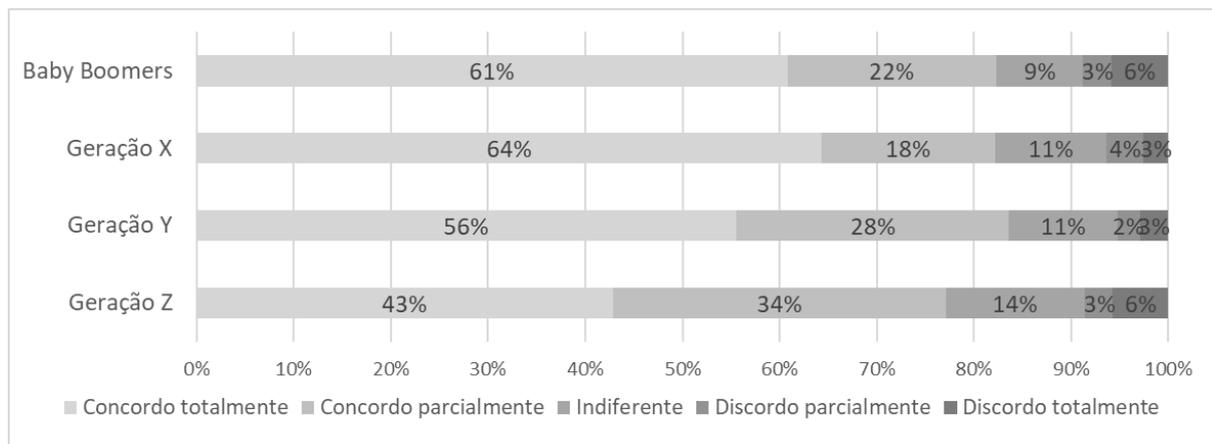


Fonte: elaborada pela autora

Ao detalhar os principais tipos de *cookies* (*first-party* e *third-party*), é possível observar uma mudança expressiva, e o grau de concordância cai para 33%. A Geração Y demonstra maior conhecimento em comparação com as outras, com 40% de grau de concordância. Já a Geração Z demonstra o menor grau, 21%.

Questão 10 - Me preocupo com as informações que compartilho nas redes sociais, como fotos e vídeos pessoais e minhas interações.

Tabela 17: Pesquisa - Respostas questão 10

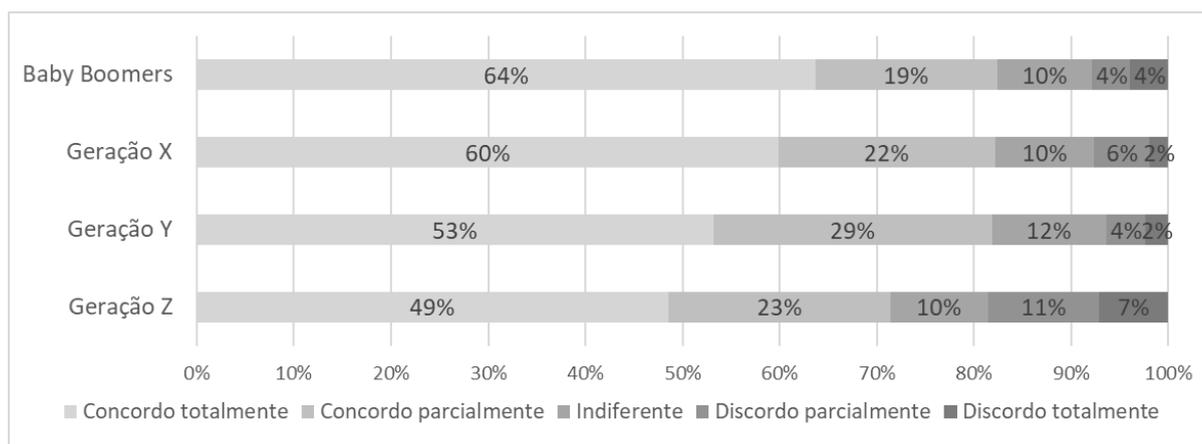


Fonte: elaborada pela autora

No compartilhamento de informações em redes sociais, o grau de concordância fica em 82%, e é possível observar maior preocupação na Geração Y (84%), seguida das gerações X e Baby Boomers (82%), e Geração Z (77%).

Questão 11 - Não cadastro meus dados em qualquer site, somente nos que confio.

Tabela 18: Pesquisa - Respostas questão 11



Fonte: elaborada pela autora

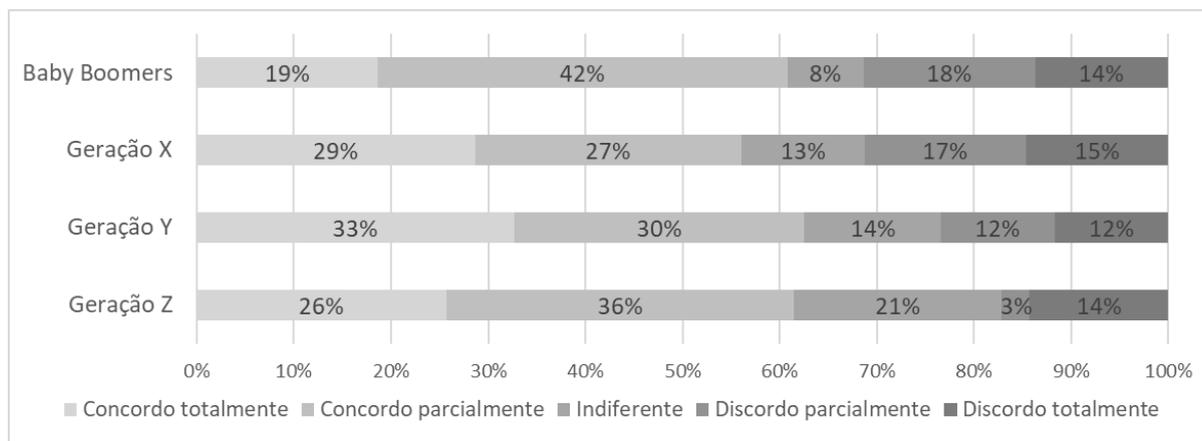
Na questão 11, o grau de concordância entre os respondentes é de 81%. A preocupação se mostra maior nas gerações Y, X e Baby Boomers, em que 82% concordam com a afirmação. Já para a Geração Z, esse volume cai para 71%.

É importante ressaltar que diversos fatores podem tornar um site confiável para a percepção dos consumidores, já que existem riscos no compartilhamento de dados e compras pela internet. Segundo AROEIRA *et al.* (2015, *apud* TURBAN *et al.*, 2007) em estudo sobre o comportamento do consumidor no comércio eletrônico, a principal preocupação é com a privacidade, dividida em dois tipos pelos autores: não transacional e transacional. A primeira se refere ao mau uso das informações pessoais por terceiros, intromissões indesejadas como vírus e recebimento de e-mails não desejados, além do excesso de propaganda. Já a segunda se refere ao mau uso de informações pessoais e fraudes no cartão de crédito.

Os dados são complementados pelo estudo realizado pelos próprios autores (AROEIRA *et al.*, 2015), em que se constata que quanto maior a percepção sobre privacidade pelo consumidor, maior será a confiança pelo site, além da percepção do compromisso com o cliente, que também influencia na confiança e satisfação por parte dos usuários.

Questão 12 - Quando preencho um formulário na internet, costumo utilizar preenchimento automático do meu navegador.

Tabela 19: Pesquisa - Respostas questão 12

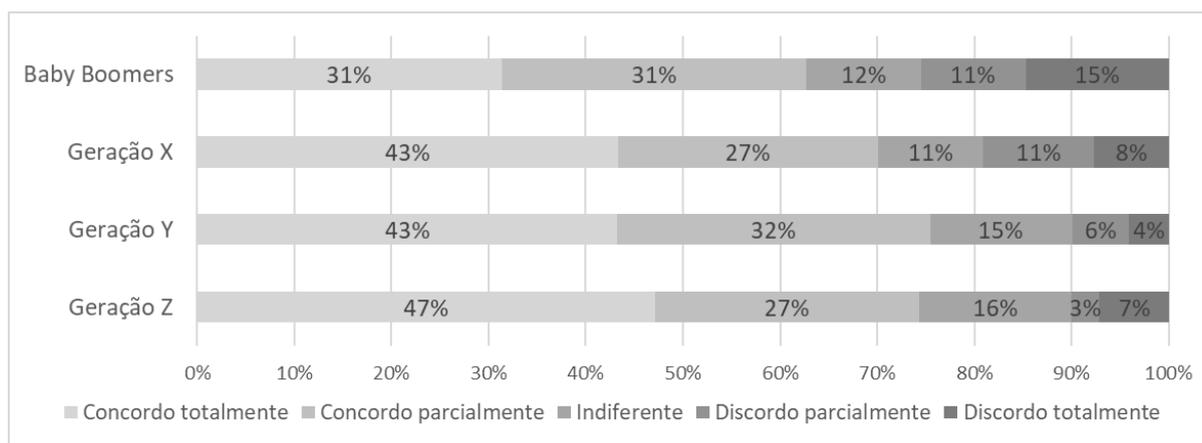


Fonte: elaborada pela autora

A utilização do preenchimento automático através de cookies é realizada por 60% dos respondentes, e se mostra maior entre Geração Y (63%), e Gerações Z e Baby Boomers (61% cada) e menor na Geração X (56%).

Questão 13 - Quando me cadastro em novos sites, dou preferência por fazer login através de contas existentes (como fazer login com minha conta de e-mail ou pelo Facebook), para que não precise preencher meus dados novamente.

Tabela 20: Pesquisa - Respostas questão 13



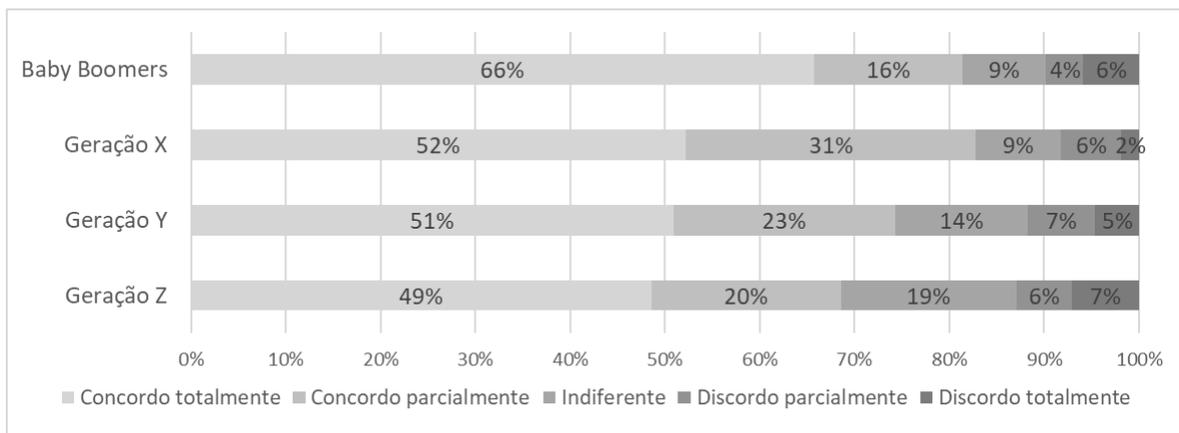
Fonte: elaborada pela autora

O cadastro em novos sites através de contas existentes pode ser feito através APIs (*Application Programming Interface*, ou Interface de Programação de Aplicativo, em

tradução literal), que são capazes de oferecer recursos e funcionalidades como acesso e manipulação de informações para aplicações externas, ou seja, entre sites (XAVIER *et al.*, 2011). 71% dos respondentes afirmam utilizar esse recurso, com maior grau de concordância entre Geração Y (75%) e Z (74%).

Questão 14 - Tenho receio em colocar dados pessoais como CPF, e-mail e telefone quando faço um cadastro em um site.

Tabela 21: Pesquisa - Respostas questão 14

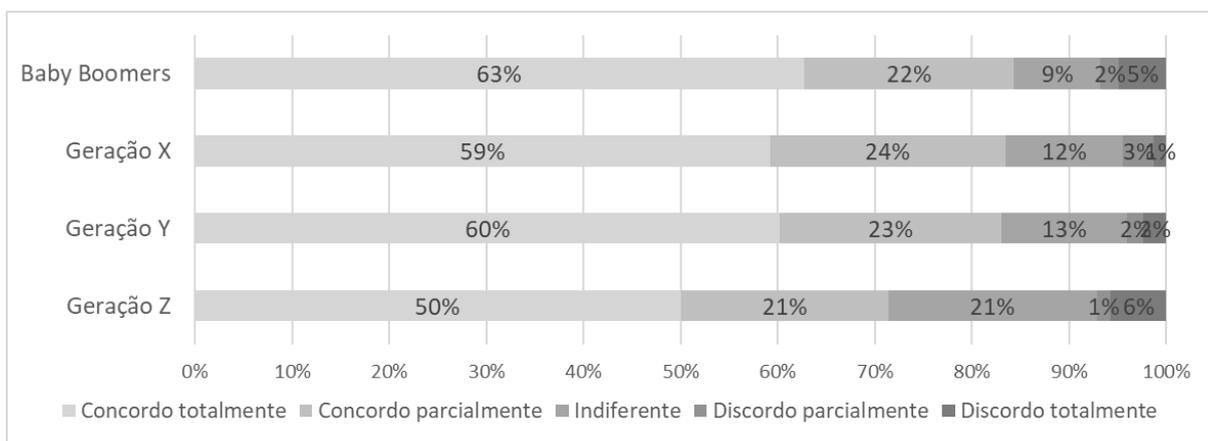


Fonte: elaborada pela autora

Observamos também que no preenchimento dos cadastros, 78% possui preocupação com os dados compartilhados em sites. O maior grau de concordância está na Geração X (83%), em seguida Baby Boomers (81%).

Questão 15 - Se pudesse escolher, gostaria que empresas não tivessem acesso aos meus dados de navegação (histórico de buscas, sites acessados e interesses).

Tabela 22: Pesquisa - Respostas questão 15

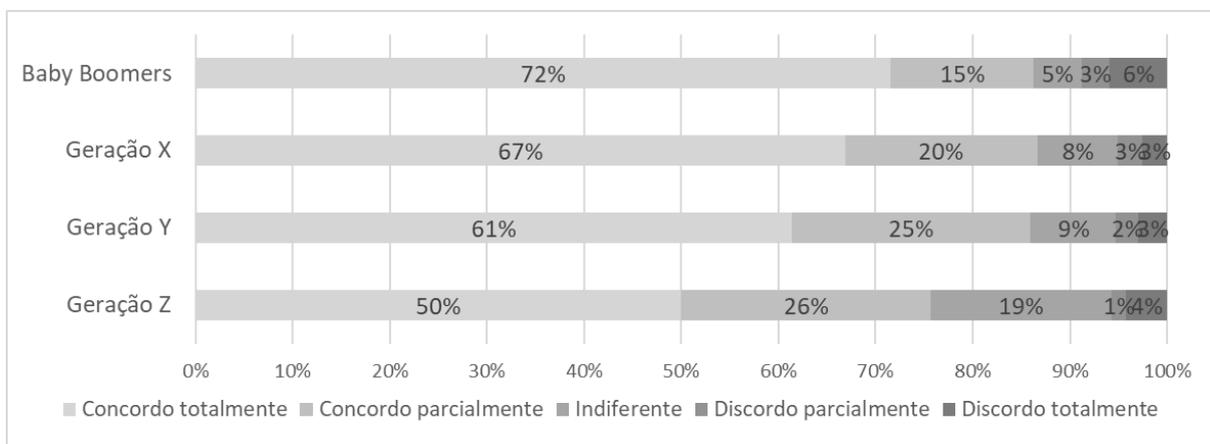


Fonte: elaborada pela autora

Na questão 15, a preferência de 82% dos respondentes seria de que as empresas não tivessem acesso aos dados de navegação. As Gerações Y, X e Baby Boomers mostram um maior grau de concordância, enquanto a Geração Z concorda em menor grau - 71%, e também obteve o maior volume de usuários indiferentes - 21%.

Questão 16 - Me importo com quais empresas têm acesso aos meus dados pessoais (nome, telefone, e-mail e CPF).

Tabela 23: Pesquisa - Respostas questão 16

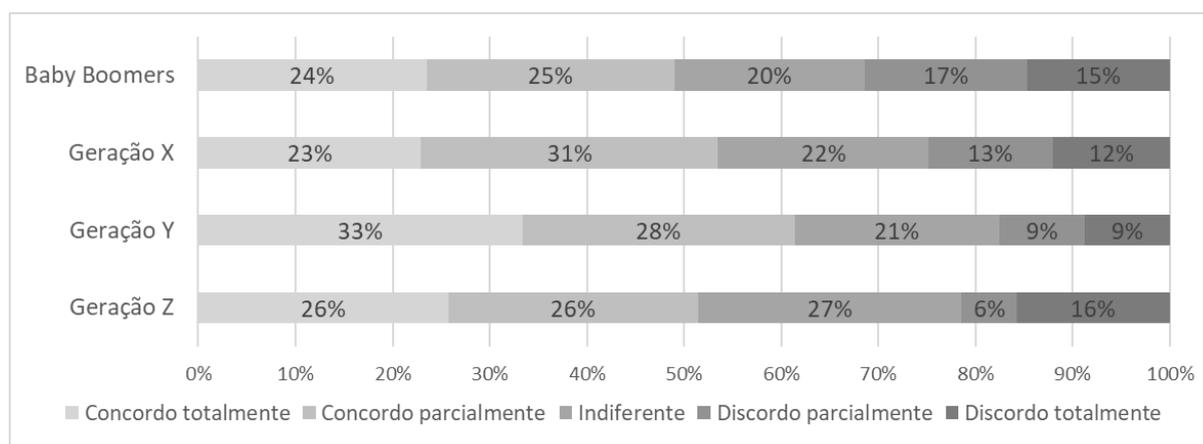


Fonte: elaborada pela autora

O grau de concordância sobre a preocupação com quais empresas possuem dados se mostra menor na Geração Z - 76% afirma que se importa, em detrimento dos 87% da Geração X e 86% das Gerações Y e Baby Boomers.

Questão 17 - Me sinto mais atraído por anúncios na internet do que em outros canais (como televisão, rádio, revista, jornal ou outdoors), pois acredito que representam mais minhas preferências

Tabela 24: Pesquisa - Respostas questão 17

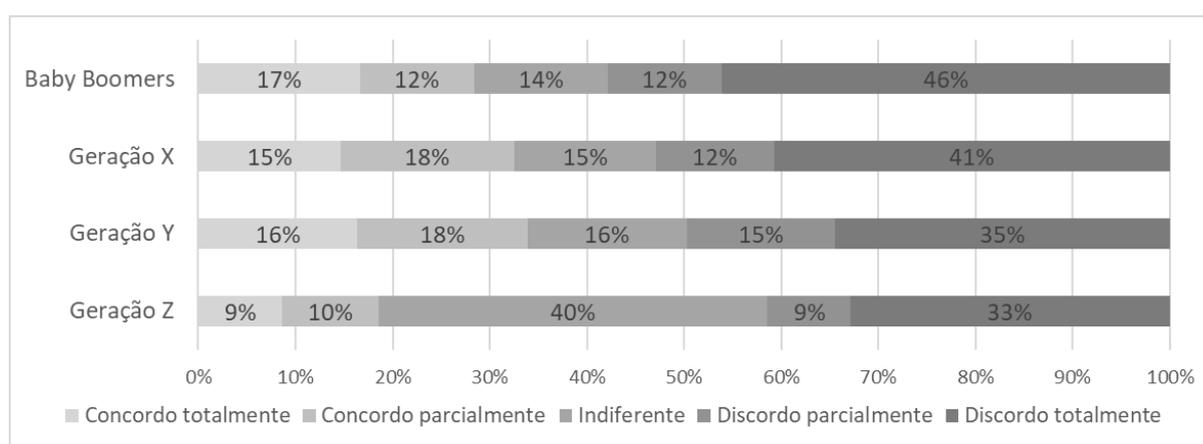


Fonte: elaborada pela autora

Sobre a percepção dos usuários sobre a publicidade on-line, 55% dos entrevistados afirma se sentir mais atraído por anúncios na internet, em comparação a outras mídias. É possível observar uma diferença entre gerações, sendo o maior grau de concordância na Geração Y (61%) e o menor grau na Geração Baby Boomers (49%).

Questão 18 - Se tivesse a opção, venderia meus dados pessoais (nome, telefone, e-mail) e de histórico de navegação em troca de acesso gratuito a conteúdos.

Tabela 25: Pesquisa - Respostas questão 18



Fonte: elaborada pela autora

O interesse na comercialização dos dados pessoais se mostra baixo entre todos os respondentes, e o grau de concordância nessa questão foi de 30%. O maior interesse está

na Geração Y (34%), seguido da Geração X (32%) e Baby Boomers (28%), sendo o menor volume na Geração Z (19%).

DADOS DEMOGRÁFICOS

Questão 19 - Qual a sua faixa etária?

Tabela 26: Quantidade de participantes por geração após realização da pesquisa

Geração	Faixa Etária	%	Casos
Geração Z	18 - 24 anos	14%	70
Geração Y	25 - 39 anos	34%	171
Geração X	40 - 56 anos	31%	157
Baby Boomers	57 - 75 anos	20%	102
TOTAL		100%	500

Fonte: elaborada pela autora

Durante a aplicação do questionário, o volume de casos por geração sofreu pequenas alterações, mas manteve a sua proporção próxima do planejado, como é possível observar na tabela acima.

20 - Com qual gênero você se identifica?

Tabela 27: Casos por gênero

Gênero	Casos	%
Masculino	192	38%
Feminino	308	62%

Fonte: elaborada pela autora

21 - Qual o seu nível de escolaridade?

Tabela 28: Casos por nível de escolaridade

Nível de Escolaridade	Casos	%
Ensino fundamental completo	31	6%
Ensino fundamental incompleto	20	4%

Ensino médio completo	162	32%
Ensino médio incompleto	31	6%
Ensino superior completo	135	27%
Ensino superior incompleto	61	12%
Pós graduação / Mestrado incompleto	21	4%
Pós graduação / Mestrado completo	28	6%
Doutorado completo	2	0%
Doutorado incompleto	9	2%

Fonte: elaborada pela autora

22 - Qual a sua área de formação?

Tabela 29: Casos por área de formação

Área de Formação	Casos	%
Administração	58	23%
Agronomia	1	0,4%
Arqueologia	0	0%
Arquitetura e urbanismo	7	3%
Biomedicina	1	0,4%
Ciência da Computação	8	3%
Ciências Biológicas	3	1%
Comunicação / Publicidade e Propaganda / Jornalismo / Relações Públicas	10	4%
Contabilidade	10	4%
Direito	20	8%
Educação física	4	2%
Enfermagem	8	3%
Engenharia	16	6%
Estatística	2	1%
Farmácia	4	2%
Fisioterapia	1	0,4%
Fonoaudiologia	0	0%
Geografia	2	1%

Informática	13	5%
Letras	15	6%
Odontologia	3	1%
Pedagogia	12	5%
Psicologia	9	4%
Turismo	5	2%
Outro. Qual? - Aberta	44	17%

Fonte: elaborada pela autora

Essa questão foi realizada somente aos usuários que confirmaram ter ensino superior, o que corresponde a 51% do total dos 500 respondentes (256 casos).

4.2. Respondendo às Hipóteses

4.2.1. Hipótese A

Hipótese: Os usuários não têm conhecimento das configurações de publicidade disponíveis nas plataformas.

Durante a pesquisa, 66% dos respondentes afirmam possuir conhecimento sobre as configurações de propaganda disponíveis nas plataformas Google e Facebook. Ao perguntar sobre a realização dessas configurações, 53% afirmam já ter feito mudanças nas preferências de anúncios, sendo o menor grau de concordância na geração Z (40%). Em seguida, ao questionar se esses usuários seriam capazes de ensinar outras pessoas a fazer essas configurações, esse volume cai para 46%, sendo a queda mais expressiva na geração X - que foi de 58% no grau de concordância da aplicação das configurações para 45% capazes de explicar a outras pessoas, enquanto a geração Z manteve grau de concordância semelhante (42%) entre as duas questões.

A partir desses dados é possível concluir que a hipótese A foi corroborada, já que os usuários possuem conhecimento mínimo sobre as configurações de publicidade disponíveis, mas a maior parte demonstra falta de propriedade no assunto para explicar para outras pessoas.

4.2.2. Hipótese B

Hipótese: Os usuários não entendem que o conteúdo visualizado é disponibilizado com base em seus interesses.

Para 80% dos respondentes, o conteúdo de seus feeds de notícias em redes sociais é personalizado de acordo com seus interesses. 72% acreditam que dados pessoais podem ser utilizados para a publicidade direcionada, enquanto 73% concordam que o histórico de navegação também possui influência. Em ambas as questões, a geração Z se destaca com menor grau de concordância (64% em cada). Ao serem questionados sobre a atração de anúncios na internet em detrimento a outras mídias, 55% demonstram maior grau de concordância, sendo a Geração Y (61%) em maior grau e menor grau na geração Baby Boomers (49%).

Pode-se considerar, portanto, que a hipótese B foi refutada, já que a maior parte dos usuários concordam que o conteúdo é disponibilizado com base em seus interesses. Porém, é possível perceber uma resistência à publicidade on-line, pois apenas 55% se dizem mais atraídos por anúncios nesse canal. Isso pode ser um reflexo da publicidade em excesso e com alta frequência, e que tende a saturar o público-alvo. Segundo ALMEIDA *et al.* (2002, *apud* CHEMIN, 2001:41), os pop-ups foram eliminados por falta de critério de uso, e é preciso cautela para que não ocorra o mesmo com outros formatos.

4.2.3. Hipótese C

Hipótese: Os usuários sabem que seus dados transitam entre plataformas, mas não compreendem como isso acontece.

A visualização da mesma propaganda em diferentes sites foi confirmada por 78% dos respondentes - em maior grau de concordância nas Gerações X e Y (respectivamente 80% e 83%), e menor entre geração Z e Baby Boomers (71% e 69%, respectivamente). 64% afirmam compreender o papel dos *cookies* para o armazenamento das informações, sendo um grau de concordância de 80% entre

gerações Y e X, enquanto para Baby Boomers e geração Z observamos novamente uma queda (64% e 44%), sendo 30% da geração Z indiferentes. Quando questionados sobre as diferenças entre os principais tipos de *cookies* (*first-party* e *third-party*), apenas 31% demonstram conhecimento sobre eles, sendo a geração Y com maior grau de concordância (40%) e a geração Z com menor (21%). É possível concluir que a hipótese C foi corroborada, já que os usuários têm a percepção de que seus dados transitam no ambiente on-line, mas a maior parte deles não sabem detalhar como isso acontece.

4.2.4. Hipótese D

Hipótese: Acredita-se que as gerações baby boomers e geração X possuem maior preocupação com os dados pessoais no ambiente on-line, em comparação com as gerações Y e Z.

Durante a pesquisa, 82% dos respondentes se mostraram preocupados com as informações compartilhadas em redes sociais e 81% dos usuários também afirmam que compartilham suas informações pessoais somente nos sites em que confiam. Em ambos os casos, a geração Z se destaca com o menor grau de concordância, de 77% e 71%, respectivamente, enquanto o grau de concordância é menos divergente entre as demais gerações.

Com base nesses dados, são levantadas duas novas hipóteses, a primeira em relação às redes sociais, que permitem o compartilhamento de conteúdos de forma restrita através da configuração de conta privada, o que poderia justificar a menor preocupação da geração Z no âmbito social e não na segurança de dados. A segunda hipótese é em relação a como essa geração navega pela internet e compartilha seus dados, já que o grau de discordância com o cadastro feito somente em sites confiáveis é o mais alto - 19%, enquanto nas demais gerações a média é de 7%.

Sendo assim, podemos considerar a hipótese D refutada, pois o comportamento da geração Y se mostra similar à gerações X e Baby Boomers,

enquanto a geração Z se destaca com opiniões divergentes que demonstram menor preocupação no compartilhamento de dados pessoais.

4.2.5. Hipótese E

Hipótese: Acredita-se que as gerações Y e Z, por terem crescido já no ambiente digital, possuem mais conhecimento sobre privacidade de dados do que as gerações baby boomers e geração X.

Durante a fundamentação teórica, foi demonstrado como a utilização de *cookies* é essencial para o rastreamento dos usuários. Na questão 8, 64% dos respondentes afirmam ter conhecimento sobre o papel dos *cookies*, mas quando é feita a quebra geracional, o grau de concordância entre as gerações X, Y e Baby Boomers é de 66%, enquanto para a geração Z o volume cai para 44%. Já na questão 9, apenas 33% dos respondentes afirmam saber explicar a diferença entre os tipos de *cookies*, sendo que o grau de concordância é menor entre as gerações Baby Boomers (25%) e geração Z (21%).

O uso de *cookies* para preenchimento automático de formulários é utilizado por 60% dos usuários, com grau de concordância maior na geração Y (63%) e menor na geração X (56%). A utilização de APIs para login com contas já existentes se mostra maior nas gerações abaixo de 57 anos: 73% afirmam fazer seu uso para se conectar a novos sites, enquanto apenas 63% dos Baby Boomers afirmam fazer o mesmo. Já a preocupação com o preenchimento de dados pessoais (CPF, e-mail e telefone) em cadastros se mostra maior entre as Gerações X e Baby Boomers (média de 82%), enquanto na Geração Y esse volume cai para 74%, e na Geração Z para 69%.

Isso significa que apesar do receio de inserir seus dados pessoais, a maioria utiliza ferramentas capazes de obter e cruzar dados em diferentes plataformas. É possível observar nas Tabelas 20 e 21 um maior grau de concordância na utilização de *cookies* e APIs entre as gerações Y e Z, e essas mesmas gerações demonstram também menor grau de concordância sobre a preocupação no compartilhamento de seus dados na Tabela 22.

Quanto a preocupação com o acesso aos dados de navegação pelas empresas, para as gerações acima de 24 anos o grau de concordância fica em 84%, e na Geração Z esse volume cai para 71%, enquanto 21% se dizem indiferentes. O cenário é semelhante sobre a preocupação de quais empresas possuem acesso a esses dados, com o grau de concordância de 86% entre as gerações Y, X e Baby Boomers, enquanto para a geração Z esse volume cai para 76%.

Com base nos dados, é possível concluir que a hipótese E foi refutada, já que as gerações Y e Z não possuem conhecimento sobre o funcionamento de como os dados transitam pelas plataformas, e também demonstram menor preocupação com o compartilhamento de suas informações.

5. CONCLUSÕES

Durante a fundamentação teórica deste trabalho foi possível observar como a evolução da internet modificou aspectos culturais da sociedade ao ser pervasiva, ou seja, ao abraçar diferentes domínios (analógicos, digitais e híbridos), diferentes entidades (dados, itens físicos e pessoas) e diferentes meios de comunicação, sendo elas experiências *cross-media* que abrangem diferentes ambientes (OLIVEIRA, 2014, p. 128). Todas essas informações são processadas em torno das pessoas de forma ubíqua, pois estão presentes de diversas formas em atividades diárias e em ambientes cotidianos para prover serviços, informação e comunicação de forma imperceptível.

A sociedade conectada permitiu um fluxo intenso de troca de informações, e a complexidade técnica sobre como os dados transitam gerou oportunidades de lucro para empresas de tecnologia e publicidade. Em destaque foram citados os tipos de *cookies* e suas diferenciações: *cookies first-party*, criados pela própria página visitada e disponível apenas nesse domínio, e que coletam informações sobre os usuários como senha, métodos de pagamento e produtos em um carrinho, salvando suas preferências; *cookies third-party*, criados por um domínio diferente ao da página visitada e disponível em diferentes domínios, capazes de rastrear o usuário em diferentes sites e dispositivos; *supercookies*, que são capazes de controlar quais informações do usuário são deletadas ou não; *evercookies*, que permitem que informações se mantenham salvas mesmo após serem deletadas; e *cookies ghostwriters*, que possuem diferentes agentes em sua cadeia e mesmo sendo instalados por domínios como *cookies first party*, funcionam como *third-party* e são capazes de enviar cópia de informações do site para terceiros.

Na pesquisa, a complexidade dessa cadeia foi confirmada, já que os usuários demonstraram conhecimento superficial sobre o fluxo de suas informações, apesar de estarem conscientes de que os conteúdos são disponibilizados com base em seus interesses. A percepção do valor de seus dados ainda é baixa entre eles, como analisado na questão 18¹³, em que apenas 30% dos respondentes os

¹³ Questão 18: Se tivesse a opção, venderia meus dados pessoais (nome, telefone, e-mail) e de histórico de navegação em troca de acesso gratuito a conteúdos.

venderiam em troca de conteúdos, sendo que seus dados já são comercializados para essa troca.

A geração Z (nascidos após 1997) se destaca com maior divergência de comportamento em relação às outras gerações sobre a coleta e trânsito de suas informações, além da preocupação sobre como elas são utilizadas, apesar de terem crescido em um ambiente já digital. Isso pode ser explicado pelo fato de que esses indivíduos “(...) vêem a tecnologia como apenas mais uma parte de seu ambiente e a assimilam juntamente com as outras coisas.” (TAPSCOTT, 1999, p. 38).

Em estudo realizado pela empresa Criteo (2018) sobre essa geração, é possível observar um perfil de consumidor que possui opiniões fortes moldadas por suas redes sociais e influenciadores digitais, e que busca por estética, design e produtos exclusivos, além de estar atento à diferenciação dos preços e qualidade dos produtos antes da tomada de decisão. A geração Z exige melhor exposição de produtos por ser mais visual e sensorial, e as segmentações disponíveis no marketing digital, que foram detalhadas no subcapítulo 2.3, permitem essas experiências personalizadas.

No geral, é possível concluir que devido à falta de conhecimento sobre como seus dados geram insumos para a publicidade, os consumidores se sentem menos invadidos e incomodados, e as leis de proteção de dados se mostram necessárias para que se crie uma regulamentação sobre sua coleta e comercialização. Ainda não é possível determinar como isso afetará o marketing digital como um todo, mas a limitação do cruzamento de dados proveniente de novas legislações pode reduzir a personalização da comunicação da publicidade on-line, e isso pode ter impacto direto na venda de produtos para a nova geração de consumidores.

REFERÊNCIAS

ALMEIDA, Licinio *et al.*, **A Propaganda na Web: um estudo sobre a percepção e compreensibilidade das mensagens veiculadas pelos anúncios dos Banners Pop-up**. PUC-Rio: 2002. Disponível em: <http://wolverine.ava.ufsc.br/~tearad/repositorio/Fundamentos%20da%20Cor/a%20propaganda%20na%20web.pdf>. Acesso em: 31 dez 2021.

ANDRADE, M. M. **Introdução à metodologia do trabalho científico: elaboração de trabalhos na graduação**. São Paulo, SP: Atlas, 2010.

AROEIRA, Tiago, *et al.*. **Seu Site é Seguro? Privacidade, Compromisso e Confiança em Compras pela Internet**. Simpósio Internacional de Gestão de Projetos, Inovação e Sustentabilidade, IV SINGEP: 2015. Disponível em: http://singep.submissao.com.br/4singep/resultado/an_resumo.asp?cod_trabalho=677. Acesso em: 29 dez 2021.

BERNERS-LEE, Tim. **Weaving the Web: The Past, Present and Future of the World Wide Web by its Inventor**. Londres: Orion Business Books, 1999.

BERNERS-LEE, Tim, *et al.* **The Semantic Web**. Scientific American: 2001. Disponível em: <https://www.scientificamerican.com/article/the-semantic-web/>. Acesso em: 15 set 2021.

BIONI, Bruno. **Proteção de Dados Pessoais: a Função e os Limites do Consentimento**. Rio de Janeiro: Forense, 2019.

CALO, Ryan. **Digital Market Manipulation**. George Washington Law Review: 2014. Disponível em: <https://digitalcommons.law.uw.edu/faculty-articles/25>. Acesso em 12 dez 2021.

CASTELLUCIA, Claude. Behavioral Tracking on the Internet: A Technical Perspective. European Data Protection: In Good Health? (pp.21-33), 2012.

Disponível em: https://link.springer.com/chapter/10.1007%2F978-94-007-2903-2_2. Acesso em: 15 dez 2021.

CRAIN, Matthew. The limits of transparency: Data brokers and commodification. Estados Unidos, 2016. Disponível em: <https://doi.org/10.1177/1461444816657096>. Acesso em: 15 dez 2021.

CRITEO. **Relatório sobre a Geração Z**. Estados Unidos, 2018. Disponível em: < <http://www.criteo.com/br/wp-content/uploads/sites/5/2018/08/18-GenZ-ReportBR.pdf> > Acesso em 15 dez 2021.

Federal Trade Commission. **Data Brokers: A Call for Transparency and Accountability**. Estados Unidos da América, 2014.

GABRIEL, Martha. **Você, Eu e Os Robôs: Pequeno Manual do Mundo Digital**. São Paulo: Atlas, 2018.

GIL, Antônio Carlos. **Como Elaborar Projetos de Pesquisa**. 4. ed. São Paulo: Atlas, 2008.

HJARVARD, Stig. **The Mediatization of Culture and Society**. Abingdon: Routledge, 2013.

IBGE – **Instituto Brasileiro de Geografia e Estatística**. Disponível em <https://www.ibge.gov.br/estatisticas/sociais/populacao/9109-projecao-da-populacao.html?=&t=resultados>. Acesso em: 29 ago. 2021.

JENKINS, Henry. **Cultura da convergência**. São Paulo. Editora Aleph, 2008.

KOO, Lawrence. **O papel da Web 3.0 no Consumo Contemporâneo**. Em: Pensamento & Realidade, Ano XII, Vol 24 - No. 2, 2009.

KRISTOL, David M. **HTTP Cookies: Standards, Privacy, and Politics**. ACM Transactions on Internet Technology, Vol. 1, nº. 02, 2001, pp. 151– 198. Disponível

em: <https://www.cs.stevens.edu/~nicolosi/classes/17fac578/ref4-1.pdf>. Acesso em: 15 dez 2021.

LAKATOS, Eva Maria. MARCONI, Marina de Andrade. **Sociologia Geral**. 8 ed. São Paulo. Atlas. 2019.

LEINER, Barry *et al.* **A brief history of the internet**. ACM SIGCOMM Computer Communication Review, vol. 39, nº 5, 2009, pp 22-31. DOI <<http://dx.doi.org/10.1145/1629607.1629613>>.

LÉVY, Pierre. **Cibercultura**. São Paulo: Editora 34, 1999.

LIKERT, R. **A technique for the measurement of attitudes**. Archives of Psychology. n. 140, p. 44-53, 1932.

SILVA, Karina *et al.* **Metodologia científica**. Porto Alegre. SAGAH. 2019.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 8 ed. Rio de Janeiro. Atlas. 2017.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Metodologia do trabalho científico: projetos de pesquisa / pesquisa bibliográfica/ teses de doutorado, dissertações de mestrado, trabalhos de conclusão de curso**. 8. ed. – São Paulo. Atlas. 2017.

MARCONI, Marina; LAKATOS, Eva. **Técnicas de Pesquisa: Planejamento e Execução de Pesquisas, Amostras e Técnicas de Pesquisas, Elaboração, Análise e Interpretação de Dados**. 5.ed. São Paulo: Atlas, 2002.

MARCONI, Marina de Andrade. LAKATOS, Eva Maria. **Técnicas de pesquisa**. 9. ed. São Paulo. Atlas. 2021.

MARQUES, Vasco. **Marketing Digital 360**. Lisboa: Actual Editora, 2018.

MATTAR, Fauze Najib. **Pesquisa de marketing**: edição compacta. São Paulo: Atlas, 1996.

NELSON, Theodor. **Complex information processing: a file structure for the complex, the changing and the indeterminate**. Proceedings of the 1965 20th national conference, (ACM '65), Lewis Winner (Ed.). ACM, New York, NY, USA, 1965. Disponível em: <https://dl.acm.org/citation.cfm?id=806036>. Acesso em: 15 dez 2021.

Núcleo de Informação e Coordenação do Ponto BR (NIC.br). (2020). **Pesquisa sobre o uso das tecnologias de informação e comunicação**: pesquisa TIC Domicílios, ano 2019. Disponível em: <http://cetic.br/pt/arquivos/domicilios/2019/domicilios/>. Acesso em: 29 ago. 2021.

OLIVEIRA, Henry. **Arquitetura da informação pervasiva: contribuições conceituais**. 2013 [i.e. 2014]. 202 f. Tese (doutorado) - Universidade Estadual Paulista, Faculdade de Filosofia e Ciências de Marília, 2013 [i.e. 2014]. Disponível em: <http://hdl.handle.net/11449/110387>. Acesso em 12 dez 2021.

OLIVEIRA, Henry, *et al.* **Arquitetura da informação pervasiva**. São Paulo: Editora UNESP, 2015. Disponível em: <https://static.scielo.org/scielobooks/6cn9c/pdf/oliveira-9788579836671.pdf>. Acesso em 12 dez. 2021.

OLIVEIRA, Jordan Vinícius de; SILVA, Lorena Abbas da. **Cookies de navegador e história da internet: desafios à lei brasileira de proteção de dados pessoais**. Revista de Estudos Jurídicos UNESP, Franca, ano 22, n. 36, p.307, jul/dez. 2018. Disponível em: <https://ojs.franca.unesp.br/index.php/estudosjuridicosunesp/issue/archive>. Acesso em: 29 ago. 2021.

Pew Research Center, Defining generations: **Where Millennials end and Generation Z begins**. Disponível em:

<https://www.pewresearch.org/fact-tank/2019/01/17/where-millennials-end-and-generation-z-begins/>. Acesso em: 29 ago. 2021.

POPPER, Karl. **A Lógica da Pesquisa Científica**. São Paulo: Cultrix, 2013.

SANCHEZ-ROLA, Iksander et al, **Journey to the center of the cookie ecosystem: unraveling actor's roles and relationships**. 42nd IEEE Symposium on Security & Privacy, São Francisco, Califórnia, 2021. Disponível em: <https://www.eurecom.fr/publication/6376>. Acesso em 24 set. 2021.

SANTAELLA, Lucia. **Cultura das mídias**. 4a. ed. São Paulo: Experimento, 1992 [2003a].

SANTAELLA, Lucia. **Navegar no ciberespaço, o perfil cognitivo do leitor imerso**. São Paulo: Paulus, 2004

SANTAELLA, Lucia. **Linguagens líquidas na era da mobilidade**. São Paulo: Editora Paulus, 2007.

SHIRKY, Clay. **A Cultura da Participação**. Rio de Janeiro: Zahar, 2011.

TAPSCOTT, Don. **Geração digital: a crescente e irreversível ascensão da Geração Net**. São Paulo: Makron Books, 1999.

TOMIZAWA, Guilherme. **A Invasão de Privacidade Através da Internet**. Curitiba: JM Livraria, 2008.

TORRES, Claudio. **A bíblia do marketing digital**. Novatec, 2019.

XAVIER, Otávio *et al.* **Desenvolvimento de Aplicações Sociais a partir de APIs em Redes Sociais Online**. Instituto de Informática Universidade Federal de Goiás, 2011. Disponível em: https://ww2.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_001-11.pdf. Acesso em 29 dez 2021.

ANEXO 1 - QUESTIONÁRIO

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
FACULDADE DE CIÊNCIAS EXATAS E TECNOLOGIA
Programa de Pós-Graduação em Tecnologias da Inteligência e Design Digital

Questionário de Pesquisa Científica

Mestranda: Bruna Neder Rosa

Orientador: Prof. Dr. Diogo Cortiz da Silva

Objetivo do questionário:

Esse questionário faz parte da pesquisa de mestrado **Publicidade e os limites da privacidade: percepção dos usuários a anúncios personalizados** e tem o objetivo de investigar qual a percepção dos usuários sobre a publicidade direcionada de acordo com gerações.

Instruções:

Cada pergunta possui 05 opções de respostas entre **discordo totalmente, discordo parcialmente, indiferente, concordo parcialmente e concordo totalmente**. É preciso escolher somente uma opção de resposta.

Você pode responder o questionário somente uma vez.

INTRODUÇÃO

No ambiente on-line, as informações como dados pessoais e ações realizadas em sites e aplicativos determinam quais tipos de anúncios publicitários - também conhecidos como propaganda - cada usuário irá receber.

Com isso em mente, criamos um questionário para entender melhor como as pessoas lidam com o compartilhamento de informações na internet e se preocupam com a privacidade de dados.

Lembrando que não existe certo ou errado, queremos apenas saber sua opinião.

CONHECIMENTO SOBRE A PUBLICIDADE ON-LINE (HIPÓTESES A E B)

1 - O seu feed de notícias nas redes sociais é a página principal de acesso a atualizações de amigos, familiares, empresas e fontes de notícias aos quais está conectado. Agora, vamos te apresentar algumas frases e queremos que você responda o quanto concorda ou discorda de cada uma:

Acredito que o conteúdo que aparece no meu feed é personalizado de acordo com meus interesses.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

2 - Tenho conhecimento de que é possível consultar, adicionar e remover as preferências de propaganda criadas pelas plataformas Google e Facebook.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

3 - Tenho conhecimento de que meus dados (como idade, gênero, e-mail, localização, cargo e renda) podem ser utilizados para a publicidade direcionada.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

4 - Tenho conhecimento de que o meu histórico de navegação na internet pode ser utilizado para a publicidade direcionada.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

5 - Nas plataformas em que tenho perfil cadastrado (como Google e Facebook), já configurei as minhas preferências de anúncios.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

6 - Consigo ensinar para outras pessoas como configurar suas preferências de anúncios no Google e Facebook.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

CUIDADO COM O USO DE DADOS PESSOAIS NA INTERNET E SUA RELAÇÃO COM A PUBLICIDADE (HIPÓTESES C, D E E)

7 - Já notei que recebi a mesma propaganda em diferentes sites e redes sociais.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

8 - Compreendo qual o papel dos cookies de navegador no armazenamento de informações.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

9 - Tenho conhecimento para explicar para amigos e colegas de trabalho o que são cookies first-party e third-party.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

10 - Me preocupo com as informações que compartilho nas redes sociais, como fotos e vídeos pessoais e minhas interações.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

11 - Não cadastro meus dados em qualquer site, somente nos que confio.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

12 - Quando preencho um formulário na internet, costumo utilizar preenchimento automático do meu navegador.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

13 - Quando me cadastro em novos sites, dou preferência por fazer login através de contas existentes (como fazer login com minha conta de e-mail ou pelo Facebook), para que não precise preencher meus dados novamente.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

14 - Tenho receio em colocar dados pessoais como CPF, e-mail e telefone quando faço um cadastro em um site.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

15 - Se pudesse escolher, gostaria que empresas não tivessem acesso aos meus dados de navegação (histórico de buscas, sites acessados e interesses).

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

16 - Me importo com quais empresas têm acesso aos meus dados pessoais (nome, telefone, e-mail e CPF).

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

17 - Me sinto mais atraído por anúncios na internet do que em outros canais (como televisão, rádio, revista, jornal ou outdoors), pois acredito que representam mais minhas preferências.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

18 - Se tivesse a opção, venderia meus dados pessoais (nome, telefone, e-mail) e de histórico de navegação em troca de acesso gratuito a conteúdos.

- Concordo totalmente
- Concordo parcialmente
- Indiferente
- Discordo parcialmente
- Discordo totalmente

Agora por último queremos conhecer um pouco mais de você. Vamos lá?

DADOS DEMOGRÁFICOS

19 - Qual é a sua faixa etária?

- 18 - 24 anos
- 25 - 39 anos
- 40 - 56 anos
- 57 - 75 anos

20 - Com qual gênero você se identifica?

- Feminino
- Masculino
- Não binário
- Prefiro não declarar

21 - Qual o seu nível de escolaridade?

- Ensino fundamental completo
- Ensino fundamental incompleto
- Ensino médio completo
- Ensino médio incompleto
- Ensino superior completo
- Ensino superior incompleto
- Pós graduação / Mestrado incompleto
- Pós graduação / Mestrado completo
- Doutorado completo
- Doutorado incompleto

22 - Qual a sua área de formação?

- Administração
- Agronomia
- Arqueologia
- Arquitetura e urbanismo
- Biomedicina
- Ciência da Computação
- Ciências Biológicas
- Comunicação / Publicidade e Propaganda / Jornalismo / Relações Públicas
- Contabilidade
- Direito
- Educação física
- Enfermagem
- Engenharia
- Estatística
- Farmácia
- Fisioterapia

- Fonoaudiologia
- Geografia
- Informática
- Letras
- Odontologia
- Pedagogia
- Psicologia
- Turismo
- Outro. Qual?