UNIVERSIDADE FEDERAL FLUMINENSE

ANTONIO CASA NOVA RAYMUNDO

COOKIES E O FUTURO DA PROPAGANDA NA INTERNET

NITERÓI

2022

ANTONIO CASA NOVA RAYMUNDO

COOKIES E O FUTURO DA PROPAGANDA NA INTERNET

Trabalho de Conclusão de Curso submetido ao Curso de Tecnologia em Sistemas de Computação da Universidade Federal Fluminense como requisito parcial para obtenção do título de Tecnólogo em Sistemas de Computação.

Orientador: Professor Nilson Luís Damasceno

NITERÓI

2022

Ficha catalográfica automática - SDC/BEE
Gerada com informações fornecidas pelo autor

R267c Raymundo, Antonio Casa Nova
COOKIES E O FUTURO DA PROPAGANDA NA INTERNET / Antonio Casa Nova Raymundo. - 2022.
60 f.: il.

Orientador: Nilson Luís Damasceno.
Trabalho de Conclusão de Curso (graduação)-Universidade Federal Fluminense, Instituto de Computação, Niterói, 2022.

1. Cookies HTTP. 2. Segurança de dados on-line. 3. Publicidade Digital. 4. Produção intelectual. I. Damasceno, Nilson Luís, orientador. II. Universidade Federal Fluminense. Instituto de Computação. III. Título.

CDD - XXX

Bibliotecário responsável: Debora do Nascimento - CRB7/6368

Dedico este trabalho a minha família e especialmente a minha mãe e minha tia, pois é graças a elas que eu me tornei a pessoa que sou hoje.

AGRADECIMENTOS

Agradeço aos meus familiares e amigos por terem me apoiado durante o desenvolvimento deste trabalho. Principalmente minha mãe e minha tia, que tiveram que me aturar várias horas por dia enquanto eu comentava sobre toda a progressão deste trabalho. Também agradeço ao meu orientador, o Professor Nilson Damasceno por ter me acompanhado neste trabalho. A sua dedicação foi essencial para me guiar e me acalmar durante todo o processo. E todos os seus conselhos foram de grande importância para mim.

"O homem é tão bom quanto o seu desenvolvimento tecnológico o permite ser."

George Orwell

RESUMO

No início da Internet, houve a necessidade de criar uma tecnologia capaz de manter o estado da comunicação entre o cliente e o servidor: o Cookie HTTP. Com a evolução da Internet, o Cookie também evoluiu, porém, seu uso não trouxe apenas inovações, mas também novos problemas, principalmente na área da privacidade. Este trabalho tem como objetivo descrever o funcionamento dos Cookies HTTP e dos Cookies de Terceiros; relacionar o uso dos Cookies de Terceiros na área da publicidade digital com os problemas de privacidade do usuário na Internet; apontar como a crescente busca dos usuários por mais privacidade está levando à decadência dos Cookies de Terceiros; por fim, analisar quais são as tecnologias que podem sucedê-los.

Palavras-chave: Cookies, Cookies de terceiros, privacidade, publicidade digital, FLoC, Topics, FLEDGE.

LISTA DE FIGURAS

LISTA DE QUADROS

LISTA DE ABREVIATURAS E SIGLAS

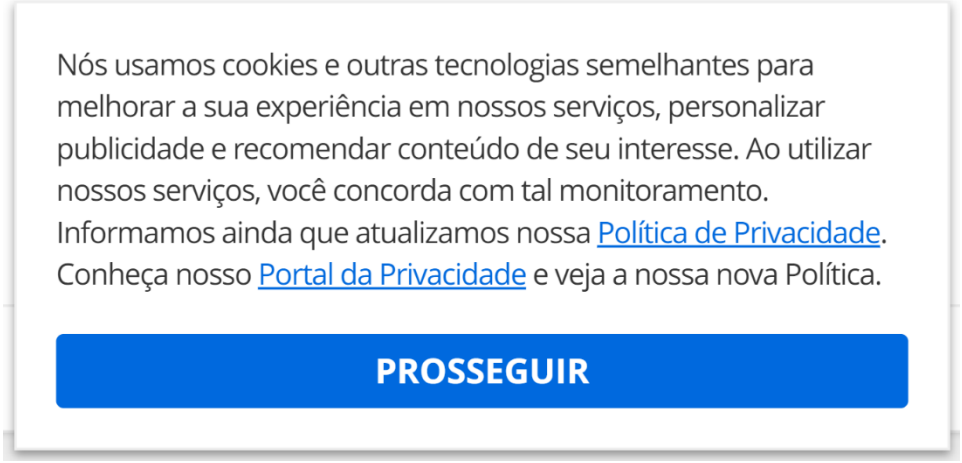
CRLF	*Carriage Return Line Feed*
CSRF	*Cross-Site Request Forgery*
EPD	*ePrivacy Directive*
FLEDGE	*First "Locally-Executed Decision over Groups" Experiment*
FLoC	*Federated Learning of Cohorts*
GDRP	*General Data Protection Regulation*
HTML	*HyperText Markup Language*
HTTP	*Hypertext Transfer Protocol*
HTTPS	*Hypertext Transfer Protocol Secure*
IEFT	*Internet Engineering Task Force*
IP	*Internet Protocol*
LGPD	Lei Geral de Proteção de Dados
RFC	*Request for Comments*
SMTP	*Simple Mail Transfer Protocol*
TCP	*Transmission Control Protocol*
TLS	*Transport Layer Security*
VoIP	*Voice over Internet Protocol*
WWW	*World Wide Web*
XSS	*Cross Site Scripting*

SUMÁRIO

1 INTRODUÇÃO

Nos últimos anos, sites ao redor do mundo começaram a exibir notificações sobre o uso de *Cookie*s e Cookies de Terceiros para os seus usuários, como ilustrado na Figura 1. Essa prática se tornou mais frequente após a implementação de regulamentos com o intuito de proteger a privacidade do usuário na Internet, como a Lei Geral de Proteção de Dados Pessoais (LGPD) [1], no Brasil, e o Regulamento Geral sobre a Proteção de Dados (GDPR), na União Europeia [1], [2]. Porém, mesmo que o termo "*Cookie*" esteja presente nessas notificações em vários sites Web, poucos usuários entendem como funcionam antes de tomar qualquer decisão sobre o seu uso [3].

Figura 1 – Notificação informando sobre o uso de Cookies em sites.



Fonte: Elaboração própria.

O Cookie HTTP, ou simplesmente Cookie, é uma tecnologia importante para o funcionamento da Internet como conhecemos, pois permite o funcionamento de diversos sites e serviços da Web utilizados diariamente por todos os usuários. Porém, uma de suas vertentes, conhecida como Cookies de Terceiros, não é tão bem vista pelos especialistas [4]. Esses *Cookies de Terceiros* são principalmente usados pela indústria de publicidade digital para rastrear os gostos do usuário entre os sites que ele acessa e oferecer propagandas mais "relevantes" [4]. Entretanto, a partir de 2020, as pesquisas na Internet referentes a privacidade de dados do usuário tiveram um aumento [5]. O que mostra que o usuário comum está mais preocupado com a sua privacidade na Internet. O aumento do interesse dos usuários pela sua privacidade

online está mudando o modo com que os produtos e sites funcionam. Diversos produtos estão abandonando o uso dos *Cookies de Terceiros*, visando um futuro sem a necessidade deles. Isso inclui, por exemplo, navegadores que permitem que seus usuários tenham um maior controle de como os Cookies de Terceiros funcionam [6].

O objetivo deste trabalho é descrever a tecnologia dos Cookies HTTP e dos Cookies de Terceiros em particular; detalhar como os *Cookies de Terceiros* se tornaram uma ferramenta tão relevante para a publicidade digital; apontar como os seus problemas envolvendo a privacidade podem levar ao seu desuso; e também estudar o funcionamento das tecnologias que estão sendo desenvolvidas como alternativa aos Cookies de Terceiros, levando em consideração à privacidade do usuário.

Este trabalho está organizado como descrito a seguir. O Capítulo 2 apresenta a Fundamentação Teórica para esclarecer alguns termos que são utilizados durante este trabalho. No Capítulo 3 é descrita a metodologia utilizada para a elaboração deste trabalho. No Capítulo 4 é explorada a origem dos Cookies HTTP e seu funcionamento nos navegadores. O Capítulo 5 é dedicado aos *Cookies de Terceiros*, sua utilização na publicidade direcionada e seus problemas de privacidade. No Capítulo 6 são apresentadas as tecnologias que estão sendo desenvolvidas para substituir os *Cookies de Terceiros*. O Capítulo 7 contém uma análise comparativa entre os Cookies de Terceiros e as demais tecnologias. E, por fim, no Capítulo 8 é apresentada a conclusão do estudo e quais temas podem ser explorados em trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os conceitos fundamentais para o bom entendimento do trabalho. Inicialmente, o capítulo define o conceito de protocolo e descreve a fundamentação da comunicação na Internet através da pilha de protocolos TCP/IP, com ênfase, na camada de aplicação. Em seguida, é apresentado o protocolo HTTP, considerado a base da World Wide Web (WWW), ou simplesmente Web.

2.1 DEFINIÇÃO DE PROTOCOLO

Um protocolo é um conjunto de regras responsável por padronizar e facilitar a comunicação entre pares de entidades na rede. Durante a comunicação, os protocolos possuem padrões que permitem que entidades distintas se comuniquem, desde que essas entidades utilizem as regras definidas nos protocolos utilizados [7]. A implementação dos protocolos pode ser feita por meio de software, hardware ou por uma mescla dos dois.

Os protocolos podem ser utilizados por diferentes entidades, que possuem arquiteturas, linguagens ou tecnologias diferentes. Para isso, esses protocolos devem definir um padrão que seja de conhecimento de todos. A regularização dos protocolos de comunicação da Internet é feita pela *Internet Engineering Task Force* (IEFT) [8], que documenta os protocolos em publicações chamadas de *Request for Comments* (RFC).

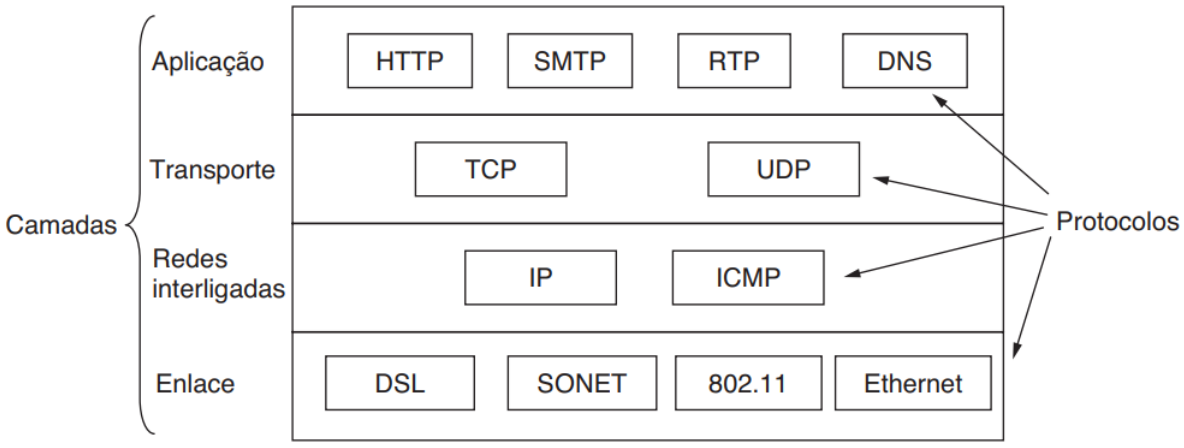
2.2 PILHA DE PROTOCOLOS TCP/IP

Para reduzir a complexidade dos projetos [9], os protocolos de rede, juntamente com os softwares e hardwares que os executam, são organizados em camadas superpostas de acordo com suas funcionalidades [7]. Cada camada possui o objetivo de oferecer às camadas superiores determinados serviços de acordo com suas funções [9]. Essa superposição de camadas de protocolos leva ao conceito de uma pilha de protocolos.

O Protocolo de Internet (IP) e o Protocolo de Controle de Transmissão (TCP) são dois dos mais importantes protocolos da Internet [7]. O Protocolo IP, definido pelo RFC 791 [10], permite encaminhamento de pacotes contendo dados, chamados de datagramas, entre hospedeiros de origem e de destino conectados através de uma ou mais redes de computadores, similares ou não. Esses hospedeiros possuem um endereço numérico único de tamanho fixo definido pelo protocolo IP, responsável por identificá-los na Web, comumente chamado de Endereço IP. O Protocolo IP permite que os datagramas muito longos sejam fragmentados em pacotes menores para a transmissão. Os datagramas necessitam de uma conexão para serem transmitidos e o Protocolo TCP é o responsável por isso. Descrito atualmente pelo RCF 9293 [11], o Protocolo TCP é responsável por estabelecer uma conexão confiável entre os hospedeiros de origem e destino, antes do envio de qualquer dado. O protocolo TCP garante a integridade dos dados, pois confirma a sua entrega. Em uma analogia com mundo real, o Protocolo IP pode ser considerado o envelope de uma carta, contendo os endereços de destinatário e remetente. Já o Protocolo TCP pode ser visto como o serviço dos Correios, responsável por entregar com segurança a correspondência do remetente ao destinatário.

Devido à frequência com que estes protocolos são usados, o conjunto dos protocolos da Internet recebe o nome de “Pilha de Protocolos TCP/IP”. A arquitetura TCP/IP é dividida em quatro camadas: Camada de Enlace, Camada de Rede, Camada de Transporte e Camada de Aplicação [9]. Cada uma dessas camadas exerce funções no processo de comunicação e possui seu próprio conjunto de protocolos.

Figura 2 – Modelo TCP/IP e alguns dos seus protocolos.



Fonte: Rede De Computadores, 5ª ed [9]

2.2.1 Camada de Aplicação

A Camada de Aplicação está no topo da pilha de protocolos TCP/IP. Ela é usada por programas ligados ao usuário final, como os navegadores Web (Chrome, Firefox) ou aplicativos de comunicação por voz (Google Meets, Discord). Diferente das outras camadas, a Camada de Aplicação de fato executa alguma tarefa direcionada ao usuário [9].

Nessa camada estão reunidos os protocolos de mais alto nível utilizados na comunicação entre as aplicações. Cada um desses protocolos descreve detalhes específicos da comunicação conforme o objetivo proposto. No caso de aplicativos de comunicação por voz, o protocolo utilizado pode ser o Voice over Internet Protocol (VoIP) [12]; em serviços de e-mail o protocolo responsável pode ser o Simple Mail Transfer Protocol (SMTP) [13]; e na comunicação entre o navegador e o servidor o protocolo responsável é o Hypertext Transfer Protocol (HTTP) [14], [15]. Quando existe uma conexão entre dois hospedeiros, um assume a função de cliente, que inicia a conexão e que requisita por recursos e serviços, e outro a de servidor, que dispõe seus recursos e serviços para o uso de terceiros.

2.3 PROTOCOLO HTTP

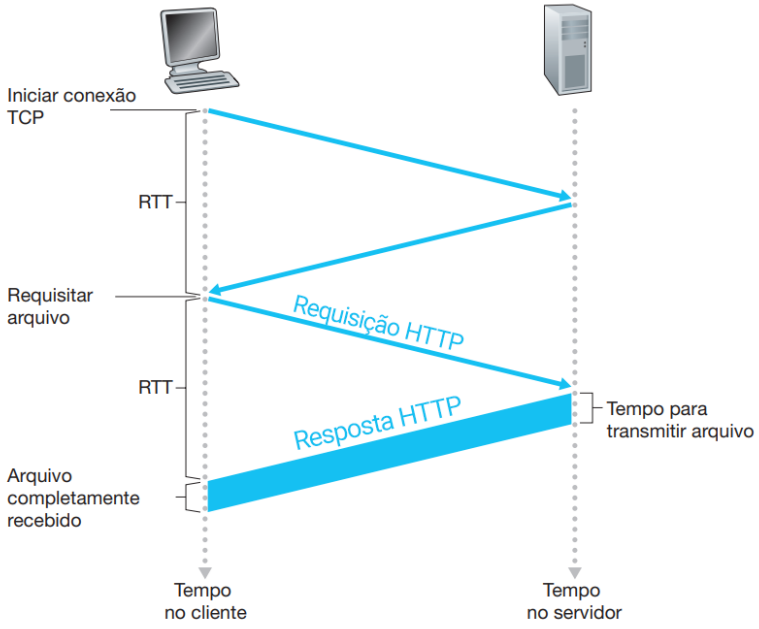
Esta seção é responsável por descrever como o protocolo HTTP está estruturado e quais são suas características. Além disso, esta seção também detalha o formato das mensagens utilizadas durante o funcionamento desse protocolo. Por fim, é apresentada a razão do HTTP ser considerado um protocolo sem estado e quais são suas relações com a tecnologia de Cookies.

2.3.1 Arquitetura do Protocolo HTTP

Dentre os protocolos da Camada de Aplicação, está o Hypertext Transfer Protocol (HTTP - Protocolo de Transferência de Hipertexto), que pode ser considerado o coração da Web. O protocolo HTTP deve ser executado tanto do lado do cliente (navegador do usuário) quanto do lado do servidor. Ele é responsável por definir a estrutura das mensagens trocadas durante uma eventual comunicação cliente-servidor. Foi inicialmente especificado pelo RFC 2616 [15] e, atualmente, é mantido pelo RFC 9110 [16].

Quando o usuário deseja acessar uma página da Web, o seu navegador executa o HTTP do lado do cliente, estabelecendo uma conexão com o servidor através do protocolo TCP. Após o êxito da conexão, o cliente e o servidor podem começar a se comunicar através do protocolo HTTP. Então o cliente envia uma mensagem de requisição ao servidor para cada um dos objetos que compõem a página acessada, como o arquivo HTML e todas as imagens e scripts associados a ele. Cada objeto que compõe a página está associado a um Identificador Uniforme de Recurso (URI), no qual é basicamente responsável por identificar um objeto no servidor. O servidor, por sua vez, responde enviando os arquivos. Após a transferência dos dados, a conexão pode ser encerrada, mas dependendo da versão do protocolo, pode ser reutilizada várias vezes.

Figura 3 – Comunicação entre o Cliente e o Servidor através do protocolo HTTP.



Fonte: Adaptado de COMPUTER NETWORKING A Top-Down Approach [7].

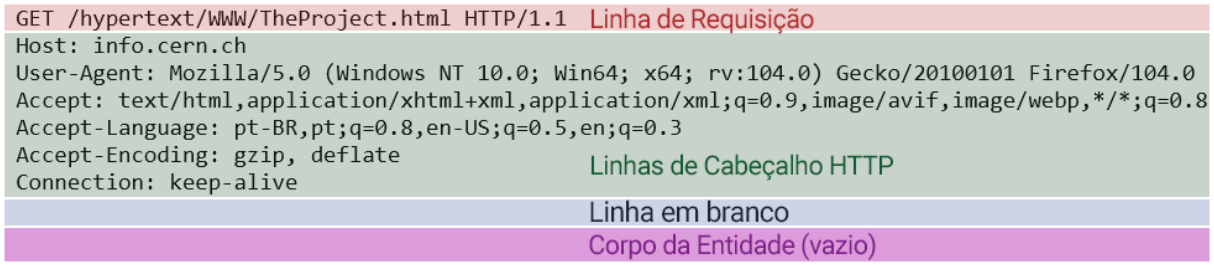
2.3.2 Mensagens HTTP

As requisições enviadas pelo cliente e respostas enviadas pelo servidor possuem um formato de mensagem semelhante. Ambas as mensagens começam com uma linha de texto seguida por nenhuma ou várias linhas de cabeçalho. O fim das linhas de cabeçalho é indicado por uma nova linha de texto em branco. Cada uma das linhas de texto é separada da anterior por uma nova linha, também chamada de quebra de linha ou, em computação, Carriage Return Line Feed (CRLF). Por fim, a mensagem pode conter um último campo chamado corpo da entidade, que contém algum tipo de arquivo [7], [15].

2.3.2.1 Mensagem de Requisição

A primeira linha da mensagem de requisição recebe o nome de linha de requisição [15]. A linha de requisição é composta pelo método que será executado, pelo caminho onde a requisição será aplicada e pela versão do protocolo HTTP.

Figura 4 – Mensagem de requisição HTTP dividida em seus campos.



Fonte: Elaboração própria.

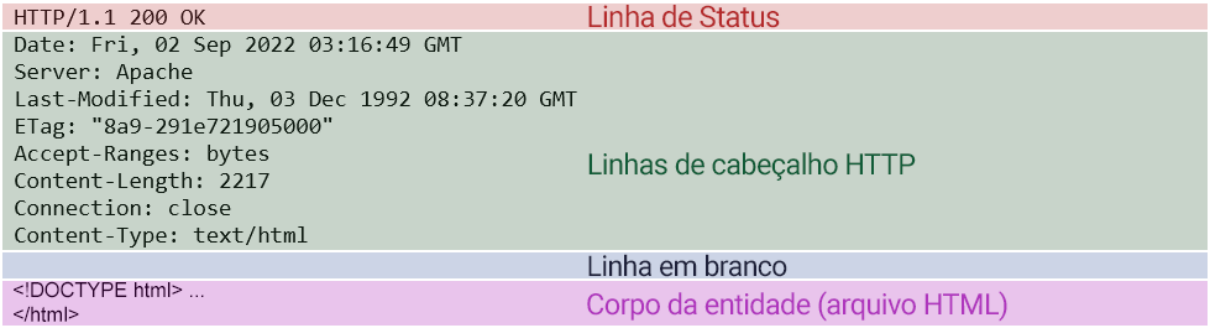
As linhas de cabeçalho têm a função de passar informações adicionais na mensagem. Elas são compostas pelo nome do campo de cabeçalho e pelo valor atribuído ao campo. Os cabeçalhos de requisição estão divididos em três grupos: os cabeçalhos gerais, que podem ser usados nas requisições e nas respostas passando alguma informação sobre a mensagem; os cabeçalhos de requisição, que permite o envio de informações relacionadas apenas a requisição; e os cabeçalhos de entidade, que contém informações sobre o corpo da entidade.

O último campo que a requisição pode ter é o corpo da entidade. O corpo da entidade só existirá caso o método chamado pela requisição necessite enviar algum arquivo para o servidor. No exemplo mostrado na Figura 4, o corpo da entidade está vazio, pois o método utilizado é o `GET`, que não necessita enviar arquivos ao servidor.

2.3.2.2 Mensagem de Resposta

Numa mensagem de resposta, a primeira linha recebe o nome de linha de status [15]. Essa linha possui a versão do protocolo HTTP, um valor numérico de três dígitos indicando o código de status e uma mensagem correspondente ao status. A Figura 5 apresenta um exemplo da linha de status.

Figura 5 – Mensagem de resposta HTTP dividida em seus campos



Fonte: Elaboração própria.

A resposta pode conter também linhas de cabeçalhos de resposta, exclusivos das respostas do servidor, além dos cabeçalhos gerais e de entidade também vistos nas mensagens de requisição. Os cabeçalhos de resposta são responsáveis por passar informações referentes ao servidor e à própria resposta.

Semelhante à mensagem de requisição, após o fim dos cabeçalhos e da linha em branco, está o corpo da entidade. O corpo da entidade são os objetos requisitados pelo cliente, como arquivos HTML, scripts ou imagens. O corpo da entidade também pode estar vazio na requisição, isso ocorre caso a mensagem de status seja suficiente entregar a informação. Na Figura 5, há uma resposta a requisição mostrada da Figura 4. Essa resposta contém no corpo da entidade um arquivo HTML.

2.3.3 Cookies e Protocolo sem Estado

Durante toda a troca de mensagens, o servidor não mantém nenhum conhecimento sobre o que foi requisitado em conexões anteriores [7]. Em condições usuais, o servidor não consegue interligar duas requisições que foram feitas. Para ele, cada requisição é tratada como única. Por essa característica, o HTTP é considerado um protocolo sem estado, pois não mantém memória das conexões e trocas de informação anteriores.

Por um lado, o fato de o protocolo HTTP não possuir estado, é interessante para o planejamento e o desenvolvimento da arquitetura dos servidores. Mas, por outro lado, para algumas aplicações é interessante que o servidor reconheça a existência de comunicações oriundas do usuário, como é o que ocorre com os carrinhos de compras de uma loja virtual, onde não é interessante perder os itens do pedido toda vez que o usuário sai da página. A tecnologia de *Cookie HTTP*, então, foi desenvolvida visando contornar a falta de estado do protocolo HTTP, permitindo, dentre outas facilidades, a identificação dos usuários através de diferentes conexões. A estratégia de manter alguma memória entre conexões HTTP é descrita com mais profundidade no Capítulo 4.

3 METODOLOGIA

Para a elaboração deste trabalho foi feita uma pesquisa exploratória sobre os Cookies HTTP e sobre as tecnologias em desenvolvimento: FLoC, Topics e FLEDGE. Para se obter as informações sobre os Cookies HTTP e Cookies de terceiros foram utilizados artigos presentes na base de dados bibliográficos Scopus [17], livros sobre o assunto, sites voltados para o desenvolvimento e Requests for Comments (RFC). Em agosto de 2022 foi feita uma busca pelos termos “Cookies HTTP” e “Third Party Cookies” na base Scopus, onde quase 300 artigos foram encontrados. Após uma análise dos resultados da busca, cinco desses artigos foram efetivamente utilizados na elaboração do texto. Os livros utilizados para pesquisa foram Computer Network: A Top-Down Approach [7] e Rede de Computadores [9].

A partir das informações contidas nos RFCs e nos artigos acadêmicos, foi possível identificar os problemas de segurança e privacidade enfrentados pelos Cookies de Terceiros. Com o intuito de enriquecer o trabalho, foi feita uma busca no Google pelo termo “privacidade”, a qual resultou em artigos de Filosofia e de Direito sobre o tema [18], [19]. A partir do resultado dessa busca foi possível apresentar no trabalho uma definição mais completa do conceito de privacidade. Além disso, foi possível identificar a razão da privacidade ser um direito tão relevante para os seres humanos, como expresso na Constituição brasileira [20].

Com o objetivo de localizar informações referentes às tecnologias FLoC, Topics e FLEDGE, foram utilizados os seus repositórios no GitHub [21] e blogs de desenvolvimento. Além disso, foram realizadas buscas por artigos científicos sobre essas tecnologias no Scopus. Porém, ao buscar por seus nomes na base Scopus, foi retornado apenas um artigo. O número reduzido de citações na Scopus demostra que essas tecnologias ainda não possuem um estudo tão aprofundado no meio acadêmico por estarem ainda no meio de seus desenvolvimentos.

Com o entendimento sobre os Cookies HTTP, Cookies de Terceiros e tecnologias em desenvolvimento já definidos, foram feitas novas pesquisas através do motor de busca do Google. Os termos pesquisados foram “*most popular websites*” e “*google annual revenue*”, com o objetivo de estudar o tamanho do Google como empresa. Um dos resultados encontrados apresentou uma lista dos sites mais populares do mundo [22], incluindo os serviços do Google. Outro documento utilizado

foi um relatório da Statcount [23] que analisa anualmente a participação de cada navegador, incluindo o Google Chrome, no mercado mundial. A partir desses documentos, foi possível deduzir como o Google se tornou um chamariz para as marcas que desejam anunciar seus produtos.

Em uma nova pesquisa foi investigado o que torna o Google tão interessado em criar uma tecnologia para substituir os Cookies de Terceiros. Para isso, foram analisados os relatórios de ganhos da Alphabet Inc. [24], conglomerado dono do Google. Esse relatório é exigido anualmente pela Comissão de Valores Mobiliários dos Estado Unidos a todas as companhias nacionais [25]. Através desse relatório, foi possível identificar a importância financeira das propagandas para o faturamento anual do Google, por meio de seus diversos produtos e serviços.

4 COOKIES HTTP

Este capítulo traz informações essenciais para o entendimento da tecnologia dos Cookies HTTP. São apresentadas a origem e a descrição sobre o seu funcionamento. E, por fim, questões relacionadas à segurança dessa tecnologia são levantadas.

4.1 ORIGEM DOS COOKIES

Os *Cookies HTTP*, também chamados de *Cookies* Primários *ou Cookies* Web, são pequenos fragmentos de dados criados por um servidor e armazenados no navegador do usuário durante a comunicação. O programador e engenheiro-fundador da Netscape, Lou Montulli, foi o responsável por desenvolver essa tecnologia [4]. Foi inicialmente apelidado de *magic Cookie* – biscoito mágico – em referência ao “biscoito da sorte chinês”, que contém um bilhete dentro [26]. Os Cookies tiveram sua implementação na primeira versão pública do Netscape Navigator, em 1994 [4].

Embora nessa época o uso dos Cookies não era tão difundido, o desenvolvimento dessa tecnologia foi um pedido direto dos clientes da Netscape, que desejavam uma maneira de gerenciar o estado da conexão entre diferentes requisições do protocolo HTTP. Montulli escreveu a especificação dos Cookies, os tratando como parte do protocolo HTTP e não como parte do conteúdo, o que permitiu a criação de uma tecnologia mais robusta do que outras alternativas [4]. O Cookie HTTP é especificado pela primeira vez no RFC 2109 [27], onde recebe o nome técnico de Mecanismo Gerenciador de Estado HTTP, e é atualmente especificado pelo RFC 6265 [28].

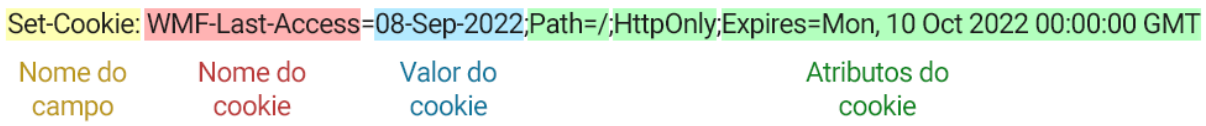
4.2 FUNCIONAMENTO DOS COOKIES

Resumidamente, quando um usuário acessa um site, é iniciada uma série de trocas de requisições e respostas entre o cliente e o servidor, através do protocolo HTTP. Durante essa troca de mensagens, caso o servidor queira salvar o estado da

comunicação com o cliente, ele pode criar um Cookie e armazená-lo no navegador do usuário. Para isso, o servidor inclui uma linha de cabeçalho na resposta, para cada Cookie que deseja criar, com o nome de: `Set-Cookie` [28].

O cabeçalho Set-Cookie é composto pelo nome do Cookie, um sinal de igual e o valor do Cookie criado. O nome e o valor do Cookie são de escolha do servidor, dependendo de quais são suas necessidades. O Set-Cookie também pode conter vários atributos opcionais com informações adicionais dos Cookies. Como a Figura 6 mostra, durante o acesso a uma página da Web, nesse caso à wikipedia.org, um Cookie foi criado e retornado pelo servidor. Esse Cookie recebeu o nome de `WMF-Last-Access` e o valor de `08-Sep-2022`. Além do nome e do valor, esse Cookie possui outros três atributos: o atributo `Path`, que indica o caminho onde esse Cookie deve ser utilizado; o atributo `HttpOnly`, que não permite que o Cookie seja acessível por nenhum script do lado do cliente; e o atributo `Expires`, que contém a data de validade do Cookie [29]

Figura 6 – Linha de cabeçalho da resposta responsável por criar um Cookie



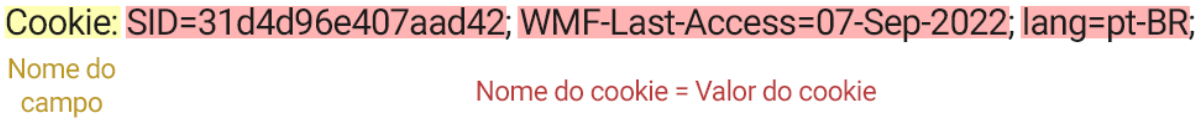
Fonte: Elaboração própria.

Em uma mesma resposta podem existir várias linhas de Set-Cookie o que permite a criação de múltiplos Cookies, porém cada Cookie deve possuir um nome distinto para evitar qualquer dependência da ordem de envio dos Cookies. O navegador é responsável pelo armazenamento e gerenciamento dos Cookies recebidos [7]. O número máximo de Cookies armazenados varia conforme o navegador, mas segundo o RFC 6265 [28], responsável pela padronização dos Cookies, o número mínimo é de 3000 Cookies por navegador e 50 por domínio.

Os Cookies armazenados no navegador são enviados de volta ao servidor de origem quando o usuário acessa aquela página novamente. O envio é feito através de uma das linhas do cabeçalho de requisição, esse campo do cabeçalho recebe o nome de `Cookie`. Em cada requisição deve existir apenas uma linha responsável pelo envio dos Cookies. Caso exista mais de um Cookie por requisição, eles são enviados na mesma linha sendo separados pelo sinal de ponto e vírgula [28], como visto na Figura 7. Para cada Cookie são enviadas duas informações: seu nome e seu valor. As

informações contidas nos Cookies são usadas pelo servidor de acordo com suas necessidades.

Figura 7 – Linha de cabeçalho da requisição enviando três Cookies



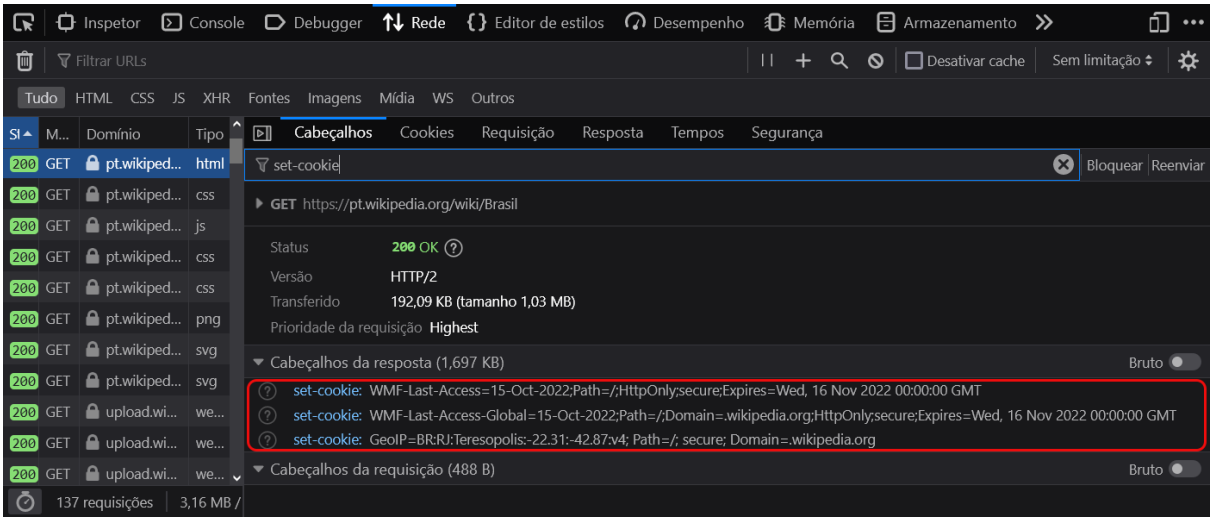
Fonte: Elaboração própria.

4.3 VISUALIZANDO EXEMPLOS REAIS

Qualquer usuário é capaz de visualizar os Cookies criados e enviados pelos sites que acessa. Utilizando as *DevTools* (Ferramentas de desenvolvimento) do navegador, o usuário pode observar as mensagens trocadas entre o servidor e o cliente (navegador), inclusive as linhas de cabeçalho `Cookie` e `Set-Cookie`. As *DevTools* estão disponíveis na maioria dos navegadores modernos e podem ser acessadas através do atalho Ctrl + Shift + i. Através da aba de Redes é possível visualizar todos os objetos requisitados que constituem a página, com suas mensagens de requisição.

Durante o acesso a um site, como a Wikipedia, caso a aba de Redes esteja aberta, ela é capaz de registrar o tráfego de requisições HTTP. A aba Redes exibe a esquerda todos os objetos que compõem a página e a direita as mensagens trocadas para a requisição do objeto selecionado. Como ilustrado na Figura 8, o objeto selecionado, neste exemplo, é o documento HTML. Na requisição estão os Cookies enviados para o servidor e na resposta os Cookies criados ou atualizados. É importante ressaltar que para cada objeto da página, podem existir diferentes Cookies sendo criados, atualizados e enviados.

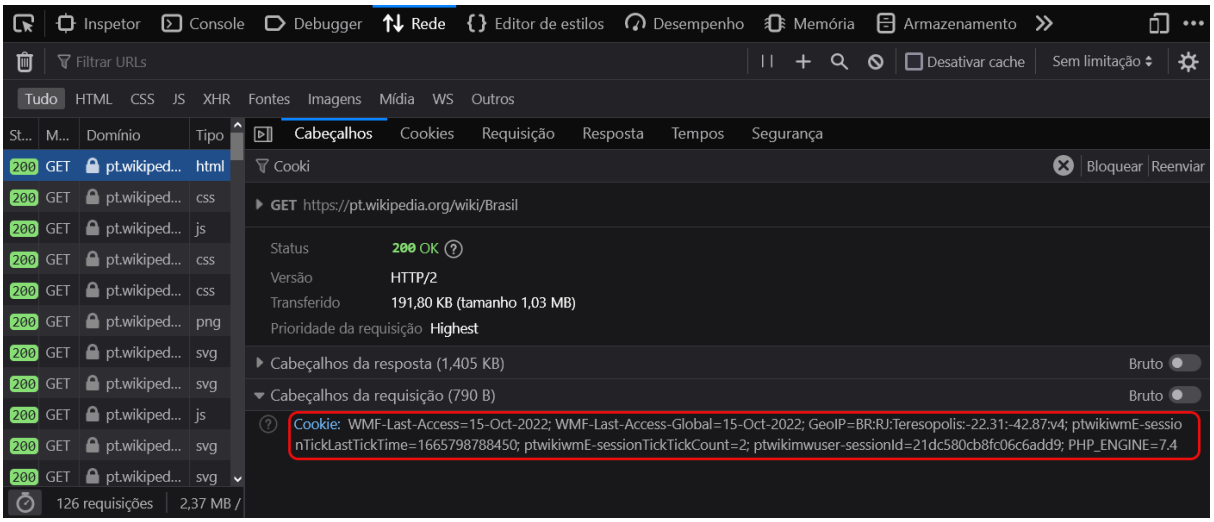
Figura 8 – Cookies criados exibidos através das *DevTools*



Fonte: Elaboração própria.

Neste exemplo, não existiam Cookies armazenados na máquina antes do acesso. Por isso, durante a comunicação com o servidor, os novos Cookies criados pelo servidor foram armazenados no navegador. Porém, quando essa página é revisitada, esses Cookies, que agora estão armazenados no navegador, são apresentados para o servidor através do cabeçalho da requisição, como indicado na Figura 9.

Figura 9 – Cookies enviados exibidos através das *DevTools*



Fonte: Elaboração própria.

4.4 SEGURANÇA E GERENCIAMENTO DE COOKIES

Esta seção apresenta informações relacionadas a segurança dos Cookies. Na primeira parte são apresentadas problemas de segurança e recomendações direcionadas ao servidor para a criação de Cookies mais seguros. A segunda parte descreve o uso de cookies como identificadores de sessão, uma forma de não armazenar informações sensíveis diretamente no Cookie. Por último são demonstradas as ferramentas do navegador que permitem um maior controle do usuário sobre os Cookies de seu navegado.

4.4.1 Protegendo os Cookies

Os servidores devem ser responsáveis por proteger as informações armazenados nos Cookies [28]. O RFC 6265 [28] recomenda o servidor de não salvar dados sensíveis nos Cookies, como senhas, nomes de usuário ou dados de cartão de crédito. Caso ocorra uma invasão na conexão ou na máquina do cliente, os Cookies podem ser roubados ou terem seus valores alterados. Por isso é interessante que as informações não estejam armazenadas nos Cookies em texto simples, mas sim protegidas através de algum tipo de criptografia escolhida pelo servidor.

É interessante, que durante a criação de um Cookie pelo servidor, sejam definidos certos atributos para aumentar a sua confiabilidade. Um desses atributos é o `HttpOnly`, que impede a manipulação dos Cookies por meio de scripts, limitando seu acesso apenas às mensagens HTTP. O `HttpOnly` ajuda a amenizar o impacto de ataques maliciosos como o Cross-Site Scripting (XSS) [30], [31]. O XSS acontece quando um indivíduo mal-intencionado consegue invadir ou injetar scripts maliciosos em um site, que até o momento era considerado seguro. Quando o usuário acessa esse site, os scripts são capazes de obter e modificar os Cookies enviados pelo navegador, podendo roubar informações do usuário. Quando o `HttpOnly` é utilizado, esses scripts se tornam incapazes de manipular os Cookies do usuário, aumentando a segurança pelo lado do cliente.

O atributo `Secure` é outro atributo responsável por aumentar a segurança dos Cookies. Com esse atributo, os Cookies só podem ser criados ou alterados através

do protocolo HTTP sobre uma camada adicional de proteção fornecida pelo protocolo Transport Layer Security (TLS), que torna o HTTP seguro (HTTPS) [32]. O HTTPS estabelece que todos os dados da conexão, incluindo o conteúdo dos Cookies, são criptografados por chaves conhecidas apenas pelo cliente e pelo servidor [9]. Isso não impede que o Cookie seja interceptado, mas impede que seu conteúdo seja decifrado e exposto ao invasor.

Quando o atributo `Domain` é omitido durante a criação de um Cookie, esse Cookie só pode ser lido pelo domínio que o criou [33]. Em um cenário onde o domínio *exemplo.com* criou o Cookie, o domínio *app.exemplo.com* não tem acesso ao Cookie. Caso exista um interesse desse Cookie ser utilizado pelos subdomínios, o atributo `Domain` deve receber o endereço base do subdomínio. Por exemplo, se `Domain=exemplo.com`, qualquer subdomínio **.exemplo.com* terá acesso ao Cookie. Isso é pode ser útil para certas funcionalidades do site, porém, caso os subdomínios não sejam confiáveis, podem ser explorados por invasores através de Cookie Tossing e Cookie Jar Overflow [34]. O Cookie Tossing se aproveita do acesso do subdomínio ao Cookie para criar um Cookie com o mesmo nome e que não possua atributos de segurança, como o `HttpOnly` e o `Secure`. Com isso, o servidor não consegue identificar qual Cookie foi criado por ele e qual foi criado pelo invasor no subdomínio, podendo utilizar de maneira arbitrária qualquer um dos dois. Caso o Cookie criado pelo invasor seja o escolhido pelo servidor, o invasor pode ter acesso ao conteúdo do Cookie por meio de scripts. A outra forma de se aproveitar de um subdomínio é através do Cookie Jar Overflow [35]. Como citado no 4.2, o navegador limita por domínio a quantidade de Cookies armazenados. Portanto, se um domínio for inundado por novos Cookies, os mais antigos são excluídos e podem ser substituídos por Cookies não seguros de mesmo nome vindos de um subdomínio infectado.

Para prevenir a substituição de Cookies por meio de Cookies Tossing e Cookie Jar Overflow, os prefixos de Cookies podem ser utilizados. Os prefixos de Cookies são termos especiais que vão ao início do nome do Cookie [33]. Como o servidor não verifica os atributos de um Cookie quando ele é chamado, os prefixos de Cookie obrigam que certos atributos estejam presentes durante a criação do Cookie. Um Cookie criado com o prefixo `__Host-`, por exemplo, só será aceito pelo servidor caso tenha o atributo `Secure`, o atributo `Path=/` e não possua o atributo `Domain`. Dessa

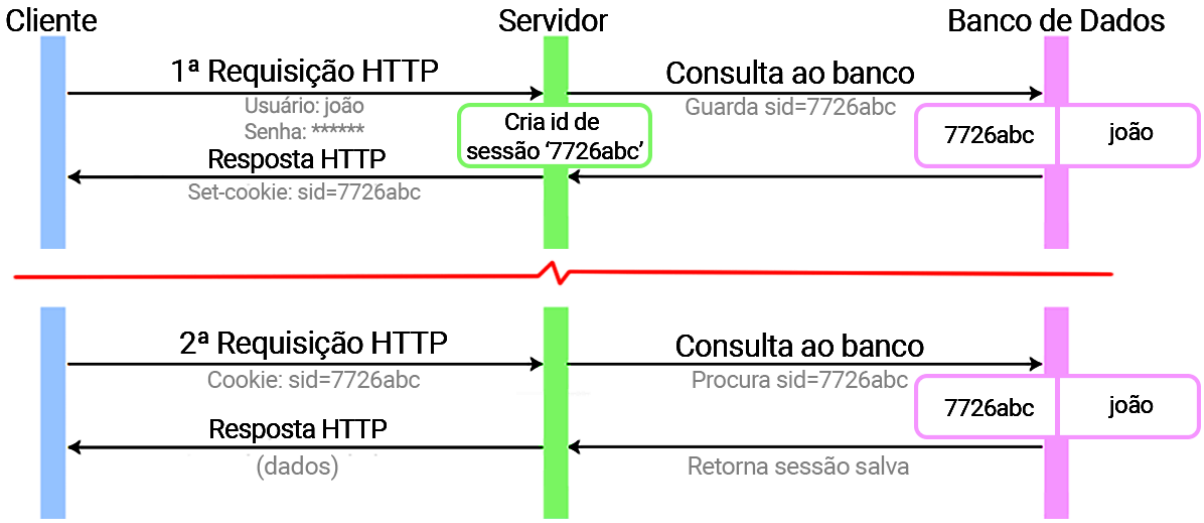
forma, esse Cookie só pode ser usado no domínio em que foi criado, não permitindo que ele substitua Cookies criados em outros domínios.

4.4.2 HTTP com estado: Identificadores de Sessão

Para evitar o armazenamento de informações sensíveis diretamente no Cookie, o servidor pode utilizar outra abordagem: o uso de identificadores de sessão [28]. Com esse método, a informação guardada no Cookie é um código de identificação, de preferência criptografado, que está relacionado a uma entrada no banco de dados do servidor. O identificador que fica salvo no Cookie possui apenas um uso e expira depois de um período.

Assim, caso o usuário acesse um site pela primeira vez, um Cookie com o identificador da sessão é criado e o estado da sessão é salvo num banco de dados auxiliar ao servidor. Num segundo acesso ao site, o navegador envia esse Cookie ao servidor, que o relaciona com a entrada no banco de dados, recuperando os dados da sessão salva. Como ilustrado na Figura 10, as informações da sessão deixam de ser salvas no Cookie armazenado no navegador do usuário e se tornam responsabilidade do servidor.

Figura 10 – Interação entre cliente e servidor através de um Cookie de sessão



Fonte: Elaboração própria.

4.4.2.1 Cross-Site Request Forgery

Identificadores de sessão que utilizam Cookies podem sofrer um ataque chamado Cross-Site Request Forgery (CSRF) [36]. O ataque CSRF utiliza os identificadores de sessão armazenados nos Cookies para executar requisições maliciosas em nome do usuário conectado. Como o protocolo HTTP não possui estado e os navegadores sempre enviam os Cookies armazenados nas requisições, não há como o servidor diferenciar uma requisição gerada pelo usuário de uma requisição forjada por um invasor. Portanto, o ataque CSRF se aproveita dessa brecha para induzir o usuário a executar requisições sem o seu conhecimento, desde que exista uma sessão ativa no site explorado.

O ataque CSRF é iniciado quando a vítima acessa um link malicioso enviado por um invasor [36]. Esse link pertence a um site no qual a vítima possui uma conta, como, por exemplo, uma rede social. O link malicioso também pode estar escondido em sites de fachada, como em um formulário HTML invisível, que é enviado automaticamente quando o site é acessado, ou dentro de uma imagem falsa, que requisita pelo link malicioso. Quando a vítima acessa esse link, sua conta é conectada ao site real devido aos Cookies de sessão, como acontece normalmente. Porém, esse link possui instruções para alguma requisição de interesse do invasor, como, por exemplo, trocar o e-mail da conta para o e-mail do invasor. Como o servidor acha que essa requisição veio do usuário, ela é executada. A partir desse momento, a conta passa a estar conectada a outro e-mail e pode ter sua senha redefinida pelo invasor, que controla por completo a conta.

Para impedir a ocorrência do CSRF, o servidor deve implementar algumas medidas de segurança. Uma das medidas mais populares são os Tokens CSRF sincronizados [37]. Um Token CSRF é um valor único, secreto e muito difícil de ser adivinhado. Ele é gerado pelo servidor e deve ser armazenado junto aos dados da sessão no banco de dados do servidor [38]. O Token CSRF deve ser transmitido ao cliente de maneira segura, como dentro de um formulário HTML oculto. Porém, os Tokens CSRF não devem ser enviados dentro de Cookies. No momento em que uma requisição é realizada, o Token CSRF é enviado com o formulário e validado pelo servidor. Caso o servidor confirme que o Token enviado é igual ao Token armazenado

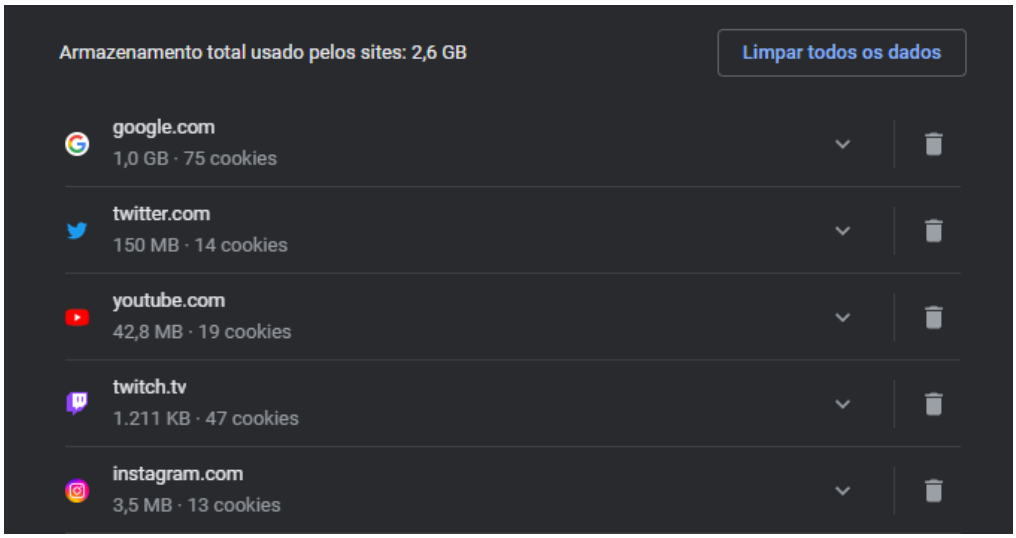
no seu banco de dados, a requisição é executada. Dessa forma, o invasor fica impossibilitado de criar qualquer requisição sem saber o Token CSRF.

Juntamente ao Token CSRF, os Cookies podem receber o atributo `SameSite` quando são criados, para aumentar sua segurança [37]. Quando esse atributo recebe o valor `Strict`, ele impede que o Cookie criado seja requisitado por qualquer outro site, mesmo que seja um link normal, podendo influenciar um pouco na experiência do usuário. Com o valor `Lax`, a requisição é menos restrita, oferecendo certo nível de proteção contra ataques CSRF, sem comprometer tanto a experiencia do usuário.

4.4.3 Gerenciamento de Cookies

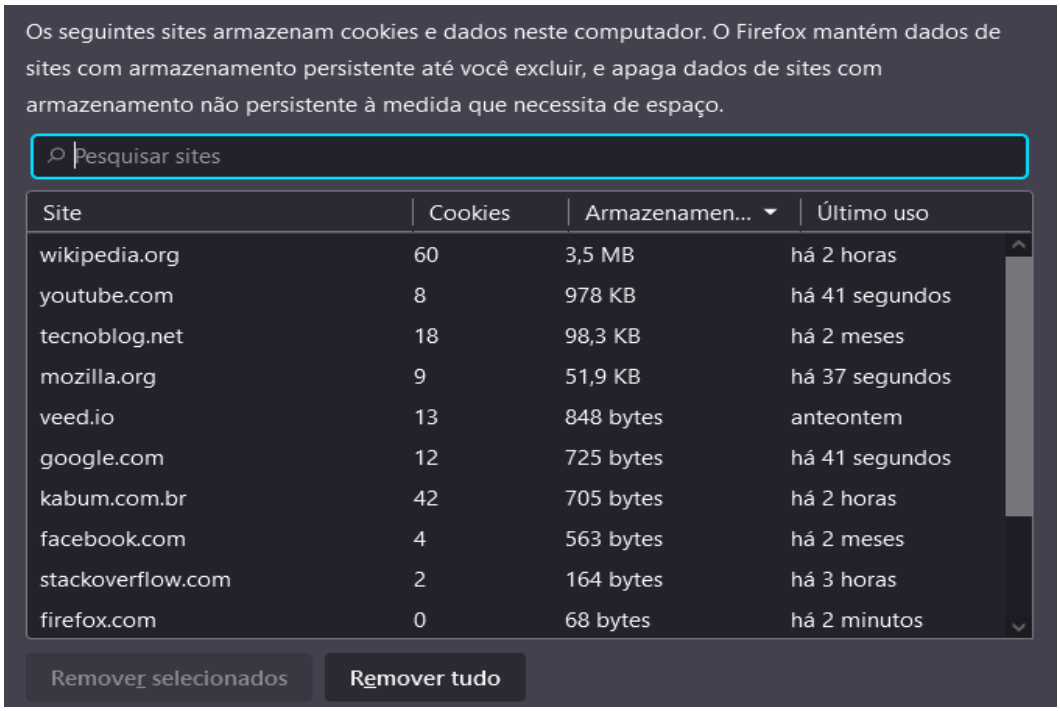
Segundo o RFC 6265 [28], que documenta a tecnologia de Cookie, os navegadores são responsáveis por oferecer meios aos usuários de gerenciar e bloquear o uso dos Cookies, como mostra a Figura 11 e Figura 12. O usuário deve ser capaz de ver todos os Cookies armazenados no seu navegador e excluir aqueles que desejar. O bloqueio total e parcial dos Cookies pelo usuário também deve ser possível. Porém, não é recomendado que todos os Cookies sejam bloqueados, uma vez que possa prejudicar a experiência do usuário em sites onde o uso dos Cookies seja imprescindível.

Figura 11 – Página de gerenciamento de Cookies no navegador Google Chrome



Fonte: Chrome/Elaboração própria.

Figura 12 – Página de gerenciamento de Cookies no navegador Mozilla Firefox



Fonte: Firefox/Elaboração própria.

5 COOKIES DE TERCEIROS

Os Cookies de Terceiros possuem um funcionamento bem semelhante aos Cookies Primários, porém eles são utilizados com finalidades diferentes. Este capítulo tem como propósito descrever o que são os Cookies de Terceiros, como seus usos para a publicidade os diferem dos Cookies Primários e como os seus problemas com relação a privacidade dos usuários estão levando ao seu declínio. Além disso, este capítulo aborda uma breve definição do conceito de privacidade.

5.1 O QUE SÃO OS COOKIES DE TERCEIROS?

Quando um usuário acessa uma página Web e o navegador inicia a comunicação com o servidor, são executadas requisições para a obtenção dos objetos que compõem essa página. Alguns desses objetos podem estar hospedados em outros servidores além do servidor principal. Esses outros servidores recebem o nome de servidores de terceiros, pois não pertencem à página que foi acessada. Os objetos requisitados em servidores de terceiros, são normalmente ligados a serviços utilizados pelo site, como redes de anúncios e propagandas ou ferramentas de integração com as redes sociais. Como os servidores de terceiros também trocam mensagens com o navegador, eles são capazes de depositar Cookies na máquina do usuário e, num momento futuro, interpretar os dados desses Cookies segundo as suas necessidades. Esses Cookies recebem o nome conforme a sua origem: Cookies de Terceiros [28].

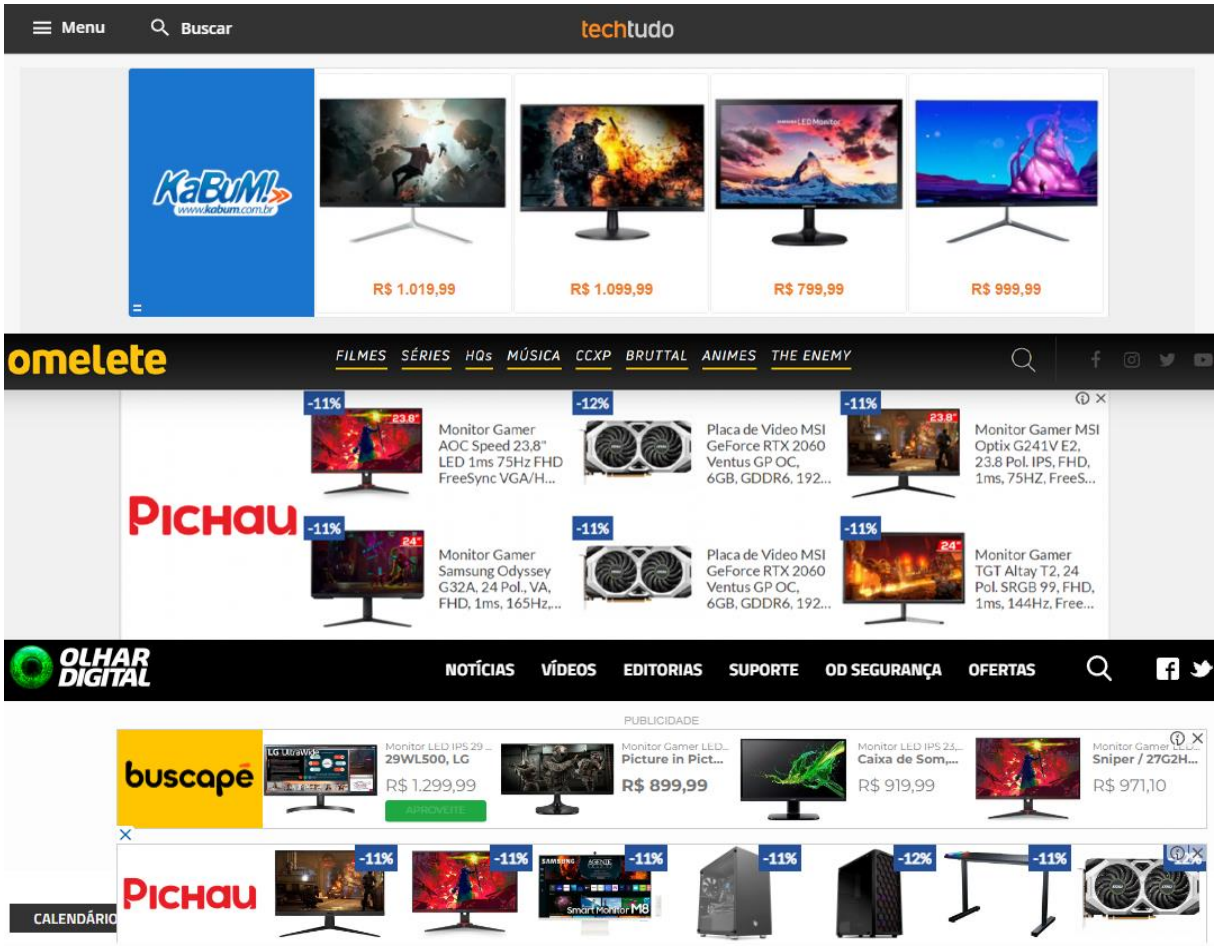
5.2 PUBLICIDADE, PROPAGANDA DIRECIONAL E REMARKETING

A partir de Cookies de Terceiros criados por redes de anúncios, informações dos sites visitados podem ser obtidos pelo servidor das propagandas e utilizados para traçar um perfil dos gostos, necessidades e traços demográficos do usuário [39]. Baseado nesse perfil, as redes de anúncios podem exibir propagandas que chamem mais a atenção do usuário. As propagandas que utilizam o perfil do usuário para possuírem mais relevância, recebem o nome de “propagandas direcionais”.

Na área da publicidade e das propagandas direcionadas, a seleção de propagandas através da utilização de informações centradas diretamente na atividade do usuário online, recebe o nome de *Behavioral Targeting* (segmentação comportamental). Propagandas exibidas através dessa estratégia podem receber quase sete vezes mais atenção e cliques do que não possuem relações com o usuário [40]. O que por um lado aumenta a efetividade e relevância das propagandas, por outro lado, reduz a privacidade dos usuários, que podem ser rastreados pelas propagandas exibidas.

Outro uso dos dados do usuário na área da publicidade pode ser visto na prática de *remarketing.* Nessa prática, as propagandas de produtos visitados recentemente pelo usuário o seguem entre as diversas páginas acessadas, como um lembrete de que ele ainda pode retornar àquela loja e comprar o produto. Essa prática é ilustrada na Figura 13.

Figura 13 – Remarketing: Após busca por "monitor" aparecem anúncios de lojas.



Fonte: Elaboração própria

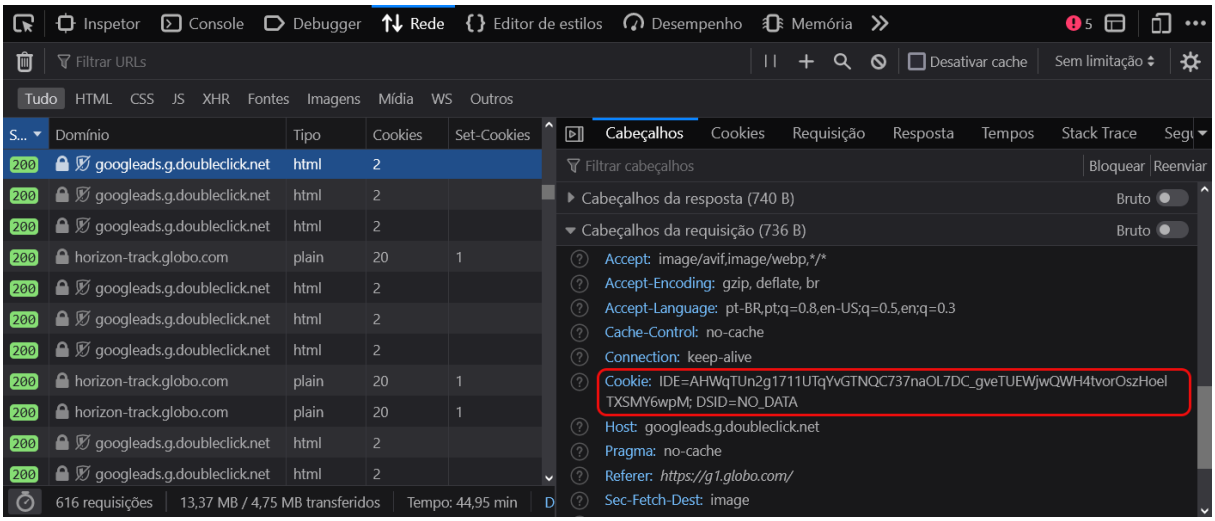
No momento em que o usuário acessa, por exemplo, o site de uma loja que contém propagandas, o serviço responsável pela exibição das propagandas deposita um Cookie no navegador do usuário contendo informações sobre esse site. Essas informações podem ser quais produtos são vendidos no site, qual é o tema do site ou quais produtos o usuário teve interesse e acessou. Em outro momento, quando o usuário entrar em outro site, contendo anúncios da mesma rede de propaganda, os Cookies depositados podem ser utilizados para identificá-lo e exibir anúncios mais relevantes, como os produtos que o usuário viu durante seu acesso à loja anterior.

5.3 IDENTIFICANDO COOKIES DE TERCEIROS

Como demonstrado no capítulo anterior, a janela de *DevTools* do navegador permite que o usuário tenha acesso às mensagens de requisições e de respostas entre o servidor e o navegador de cada objeto que compõe a página Web. Qualquer Cookie gerado por objeto que não pertençam ao domínio do site, serão Cookies de Terceiros.

Como demonstrado na Figura 14, o site do G1 [41] foi acessado e todos os objetos que compõe a página foram listados na aba de Redes das *DevTools*. Todos os objetos que não pertencem ao domínio globo.com são considerados objetos de terceiros e, consequentemente, todos Cookies criados e utilizados por esses objetos são Cookies de Terceiros. Portanto, os Cookies vindos do domínio doubleclick.net, são de terceiros.

Figura 14 – Identificação de Cookies de Terceiros através das *DevTools*.



Fonte: Elaboração Própria

Além de ser possível distinguir os Cookies, existem meios de identificar a funcionalidade dos Cookies de Terceiros mais utilizados por ferramentas disponíveis online. O site CookieYes [42] oferece uma ferramenta capaz de verificar qualquer site desejado e gerar uma análise dos Cookies utilizados por ele. Essa análise contém diversas informações de cada Cookie, como o nome, domínio, descrição sobre a sua funcionalidade e duração, além de separar os Cookies por tipo. Outras ferramentas úteis para encontrar informações de Cookies de Terceiros são o site do Cookiedatabase [43] e o repositório do github Open-Cookie-Database [44], ambos possuindo um vasto banco de dados contendo informações de Cookies utilizados ao redor da Internet.

Figura 15 – Tipos de Cookies presentes na página principal do G1.

Fonte: CookieYes [45].

Através das informações desses sites, é possível identificar os propósitos dos Cookies exibidos na Figura 14: o IDE e o DSID. O Cookie IDE é um Cookie de propaganda/marketing e pertence ao serviço de anúncios do Google Doubleclick. É um Cookie responsável por armazenar informações de como o usuário utiliza os sites visitados para gerar propagandas mais relevantes [45]. O Cookie DSID possui praticamente o mesmo propósito, é um Cookie responsável por direcionar, analisar e otimizar campanhas de propagandas [46].

5.4 PRIVACIDADE E COOKIES DE TERCEIROS

Os navegadores devem lidar com os Cookies de Terceiros da mesma forma que lidam com os Cookies Primários, permitindo que o usuário tenha controle sobre a sua criação e seja capaz de bloqueá-los. Além de compartilhar dos mesmos problemas de segurança que os Cookies Primários, os Cookies de Terceiros ainda têm um problema a mais: a violação da privacidade.

Mesmo que a palavra privacidade seja comum ao vocabulário, não há uma definição considerada a mais correta [18]. Uma das muitas definições forjadas define privacidade como “o direito de ser deixado em paz” [19], o que pode ser interpretado como o direito de uma pessoa de manter segredos longe dos outros ou de não revelar tudo sobre sua vida. Na legislação brasileira, é dito que a intimidade, a vida pessoal e a honra do indivíduo são invioláveis [20]. Então, a privacidade pode ser considerada, de maneira geral, como o direito dado a cada indivíduo de omitir e divulgar informações próprias apenas conforme os seus interesses. Portanto, essas informações não podem e nem devem ser violadas por terceiros.

A partir do momento que as redes de anúncio digital sabem quais são os sites acessados pelos usuários, elas podem criar um perfil mais detalhado do comportamento de cada indivíduo, podendo até mesmo adivinhar informações pessoais como o gênero e faixa etária do usuário. Mesmo que essas informações não sejam totalmente precisas [47], um serviço de terceiros armazenando informações pessoais é, por si só, um grande risco a privacidade dos usuários. Além disso, por ser armazenado por terceiros, nem sempre o usuário poderá ter controle sobre as suas próprias informações pessoais e a sua privacidade.

Redes de anúncios já foram autuadas por comercializar dados dos usuários sem autorização. Como nos anos 2000, onde a rede de propagandas *DoubleClick* foi processada após tentar vender mais de 100 mil perfis de usuários sem o conhecimento deles [48]. O compartilhamento de informações privadas dos usuários se tornou mais controlada após a criação de leis de proteção de dados como o GDPR e a Diretiva de Privacidade Eletrônica (EPD), que impôs a necessidade do consentimento dos usuários antes da capitação de dados [2].

5.5 O AUMENTO DA REJEIÇÃO AOS COOKIES DE TERCEIROS

Sabendo dos problemas envolvendo a privacidade dos usuários, as companhias de tecnologia começaram a tomar atitudes, reduzindo o suporte aos Cookies de Terceiros. Em 2019, a Mozilla Foundation, responsável pelo desenvolvimento do navegador Firefox, bloqueou por padrão qualquer Cookie de terceiro conhecido por ser capaz de rastrear o usuário [49]. No Safari, em 2020, a Apple introduziu diversas melhorias visando a privacidade do usuário, incluindo o bloqueio de absolutamente todos os Cookies de Terceiros no seu navegador [6].

O Google também anunciou o interesse de tornar os Cookies de Terceiros obsoletos em seus produtos, como o Chrome, até 2024 através de uma iniciativa chamada *Privacy Sandbox* [50]. Porém, para o Google abandonar por completo o uso dos Cookies de Terceiros sem um grande prejuízo para a empresa, deve existir algum meio de substituí-los. Isso ocorre devido ao fato de a empresa utilizar os Cookies de Terceiros como modelo de negócio, através de seus serviços de monetização de propagandas [51].

6 PROPAGANDAS SEM OS COOKIES DE TERCEIROS

Este capítulo discute a necessidade, principalmente pelo Google e pelas plataformas de propaganda, de desenvolver tecnologias que substituam os Cookies de Terceiros. Além disso, são descritas algumas das tecnologias que estão sendo desenvolvidas como substitutas: FLoC, Topics e FLEDGE.

6.1 A NECESSIDADE DE UM SUBSTITUTO PARA OS COOKIES DE TERCEIROS

Como discutido no Capítulo 5, questões de privacidade estão levando ao fim dos Cookies de Terceiros. O fim ao suporte aos Cookies de Terceiros afetará, pelo menos, três grandes partes envolvidas com as propagandas negativamente: os sites que cedem o espaço publicitário, as marcas anunciantes e as redes de anúncio digital. O usuário final também poderá ser afetado também, mas de maneira mais branda. Da perspectiva do usuário, as propagandas podem apenas deixarem de ser tão interessantes quanto são atualmente.

Muitos dos sites espalhados pela Internet oferecem seu conteúdo de graça, por isso, a forma que eles encontram para gerar receita para financiar suas operações é através dos espaços de propaganda disponíveis em suas páginas. O valor recebido pelos sites varia conforme o número de exibições das propagandas e cliques do usuário [52]. Com o fim dos Cookies de Terceiros, as propagandas se tornam menos relevantes ao usuário, o que pode fazer com que ele clique menos nelas, diminuindo a renda dos sites.

Para as marcas que desejam exibir seus anúncios, o fim dos Cookies de Terceiros pode prejudicar o seu alcance. Através das propagandas direcionadas, os usuários que mais se encaixam no perfil desejado pela marca que recebem as propagandas. Sem os Cookies de Terceiros, o direcionamento das propagandas para um público específico se torna mais difícil, diminuindo a eficiência das propagandas exibidas e aumentando o custo para exibir o anúncio para todos.

As redes de anúncio digital também são fortemente impactadas em uma Internet sem Cookies de Terceiros. Incapacitadas de rastrear os interesses do usuário, as redes de anúncios perdem um dos seus diferenciais: a exibição de propagandas

personalizadas para cada usuário. Isso pode levar a perda de clientes e, posteriormente, a uma queda na receita.

Em 2021, o Google gerou cerca de 209 bilhões de dólares apenas através da exibição de propagandas nos seus serviços, representando cerca de 80% de toda a sua receita [24]. Com as propagandas compondo uma fatia tão grande de sua renda anual, o Google iniciou o desenvolvimento de novas tecnologias que consigam direcionar propagandas sem a necessidade de rastrear os usuários individualmente. Essas mudanças chamam a atenção de todo o mercado de propagandas, pois o navegador do Google é o mais popular do mercado, sendo utilizado por mais de 63% dos usuários da Internet [23], e seus sites são uns dos mais utilizados toda a Internet [22].

6.2 FLOC

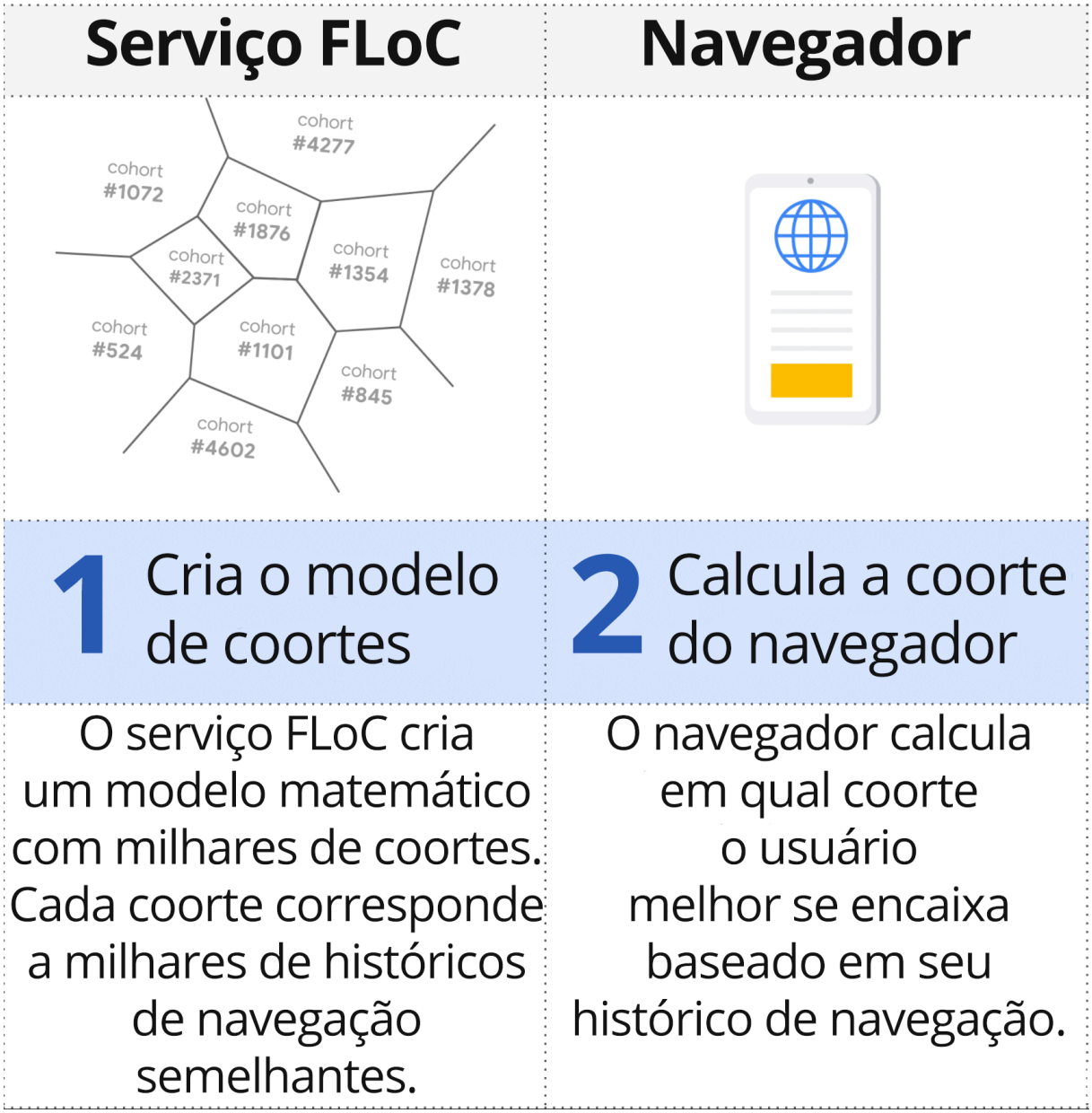
O *Federated Learning of Cohorts* (Aprendizado Federado de Coortes - FLoC), apresentado no início de 2021, foi uma das primeiras soluções criadas pelo Google para substituir os Cookies de Terceiros [53]. A função do FLoC é identificar os interesses de um grande grupo de pessoas ao invés dos interesses individuais de cada usuário. Para isso, cada usuário é associado a uma coorte, um grande grupo de pessoas com interesses em comum [54], e as redes de anúncio exibem propagandas baseadas no interesse geral dessa coorte, sem saber quem são os integrantes individuais.

O funcionamento do FLoC inicia com a criação das coortes. Através de Inteligência Artificial e modelos de históricos de navegação, o serviço FLoC cria diversas coortes baseadas em possíveis interesses dos usuários. A criação das coortes não utiliza dados de usuários reais, apenas modelos hipotéticos de navegação [53]. Cada coorte possui seu próprio número de identificação, para facilitar o seu uso.

O navegador, baseado no histórico de navegação do usuário e após receber os dados das coortes criadas, seleciona qual coorte representa melhor os interesses daquele usuário. A escolha de coorte é feita no próprio navegador, sem que nenhum dado do histórico seja enviado ao serviço FLoC ou a serviços de terceiros. Os históricos de navegação de usuários podem ser diferentes, mas caso seus interesses sejam semelhantes o suficiente, o FLoC considera que devem estar na mesma coorte.

Cada usuário faz parte de apenas uma coorte por vez, podendo migrar de coorte de tempos em tempos, baseado na sua mudança de interesses [53].

Figura 16 – Etapas de criação das coortes do FLoC

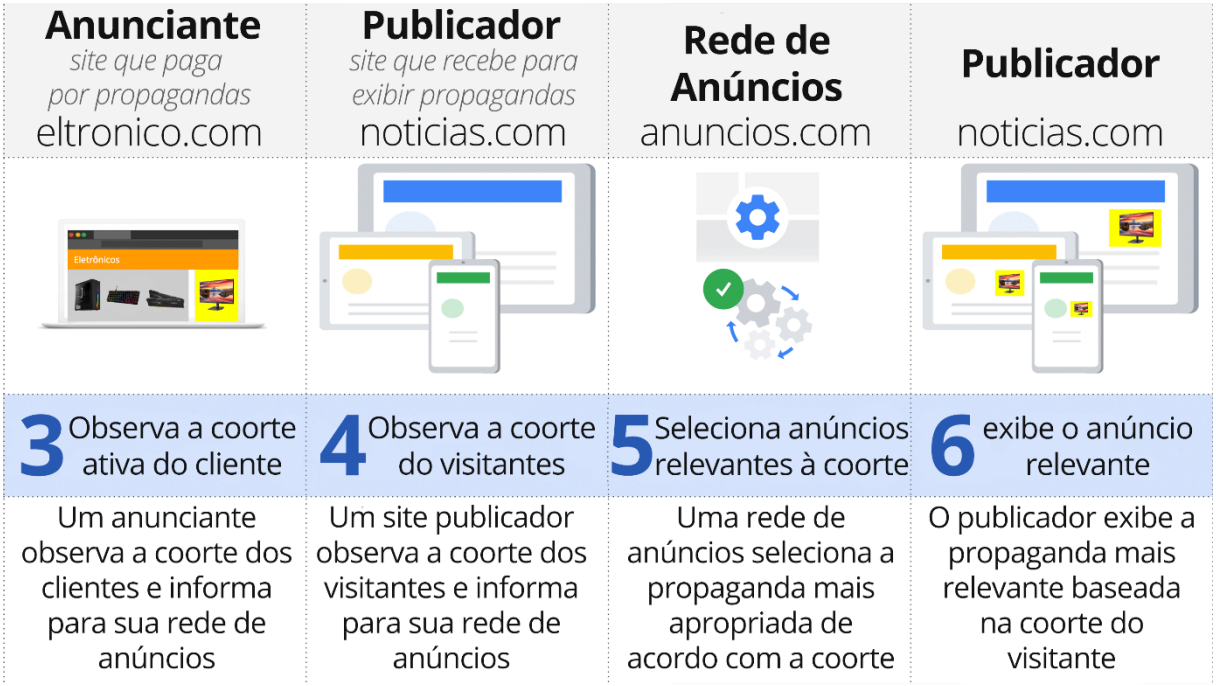


Fonte: Adaptado de Google [53].

Através do número de identificação das coortes, as lojas podem rastrear quais coortes possuem usuários interessados em seus produtos e compartilhar esses dados com sua rede de anúncios. Assim, sempre que um usuário acessar um site que contenha propagandas, utilizando o número de sua coorte, a rede de propagandas pode exibir um conteúdo relevante [53].

Por exemplo, dois usuários chamados João e Maria são colocados na coorte 1101 devido à semelhança de seus históricos de pesquisa. Quando João acessar uma loja de eletrônicos, a loja pode requisitar o número de identificação de sua coorte. A partir desse momento, a loja sabe que a coorte 1101 pode ter interesse em seus produtos e pede para que sua rede de anúncios inclua aquela a coorte na sua lista de exibição de propagandas. Caso Maria acesse um site de notícias que possua um espaço de propagandas da mesma rede de anúncios, essa rede identifica qual é a sua coorte. Quando a rede de anúncios identifica a coorte de Maria, ela exibe propagandas que sejam relevantes para a coorte 1101, como as propagandas da loja de eletrônicos. Mesmo que Maria nunca tenha entrado na loja de eletrônicos, ela pode receber propagandas de eletrônicos devido os interesses de sua coorte.

Figura 17 – Etapas de exibição de propagandas através do FLoC



Fonte: Adaptado de Google [53].

6.2.1 Desenvolvimento Interrompido

Especialista e desenvolvedores não receberam a ideia do FLoC com muito agrado [55]. Uma das primeiras críticas exercidas foi em relação ao seu nome [56]. A

ferramenta não utilizava o *Federated Learning* (aprendizado federado), que é uma técnica do aprendizado de máquina, que permite um aprendizado coletivo de maneira descentralizada [57]. Questionado sobre isso, o Google informou que inicialmente havia o interesse da utilização do aprendizado federado, mas após iniciado o desenvolvimento, essa técnica se mostrou desnecessária [58].

Outro problema identificado no FLoC é a possibilidade de coortes com conteúdo sensível serem criadas [55]. Como o algoritmo responsável por criar as coortes se baseia em modelos de histórico, ele poderia ser capaz de gerar coortes e separar os usuários baseados em informações sensíveis, como gênero, idade ou classe social. O navegador deveria ser capaz de identificar o que é conteúdo sensível e removê-lo da sua coleta de dados, mas isso não significa que todo conteúdo sensível estaria seguro [59]. Como solução, caso uma coorte com conteúdo sensível fosse criada, ela seria bloqueada e o algoritmo deveria ser parcialmente retrabalhado para impedir novas ocorrências.

Além das críticas, o FLoC também apresenta diversos problemas de privacidade [60]. Um dos seus principais problemas é o *fingerprinting*. Quando o cliente faz requisições a sites, informações sobre o seu navegador e dispositivo podem ser identificadas pelo servidor. Em um universo com centenas de milhões de usuários, essas informações são muito mais difíceis de serem relacionadas a uma única pessoa. Porém, em uma coorte, aonde existem muito menos usuários, é mais fácil de ligar essas informações a um usuário específico e o identificar sempre que fizer uma requisição [61]. No início de 2022, o Google anunciou estar encerrando o desenvolvimento do FLoC [62] e migrando seus esforços para uma nova tecnologia que seria mais segura e permitiria mais controle por meio dos usuários [63].

6.3 GOOGLE TOPICS

O Google Topics segue uma proposta semelhante ao do seu antecessor, o FLoC. Porém, foi desenvolvido para corrigir os problemas que eram mais evidentes no FLoC, baseando-se nos feedbacks recebidos através da comunidade [63]. A intenção do Google com o Topics continua sendo separar os usuários em grupos de interesse, contudo, de maneira mais geral e permitindo maior controle do usuário sobre sua privacidade [63].

A base do Topics é uma lista com diversos tópicos de possíveis interesses dos usuários, chamada de taxonomia [61]. A taxonomia atual, criada para o período de testes, possui 350 tópicos escolhidos por humanos. Esses tópicos cobrem diversas áreas de interesse geral do público, contendo tópicos como “maquiagem”, “filmes de ação” e “educação à distância”. O Google tem interesse de aumentar essa lista em mais algumas centenas de tópicos, mas de forma que continue sendo filtrada por humanos para não permitir que tópicos sensíveis sejam adicionados, como ocorreu como FLoC [64]. Porém, quem decide quais dados que são sensíveis ou não é o próprio Google. Dessa forma, uma informação que o usuário considere confidencial pode não ser vista do mesmo jeito pelo Google.

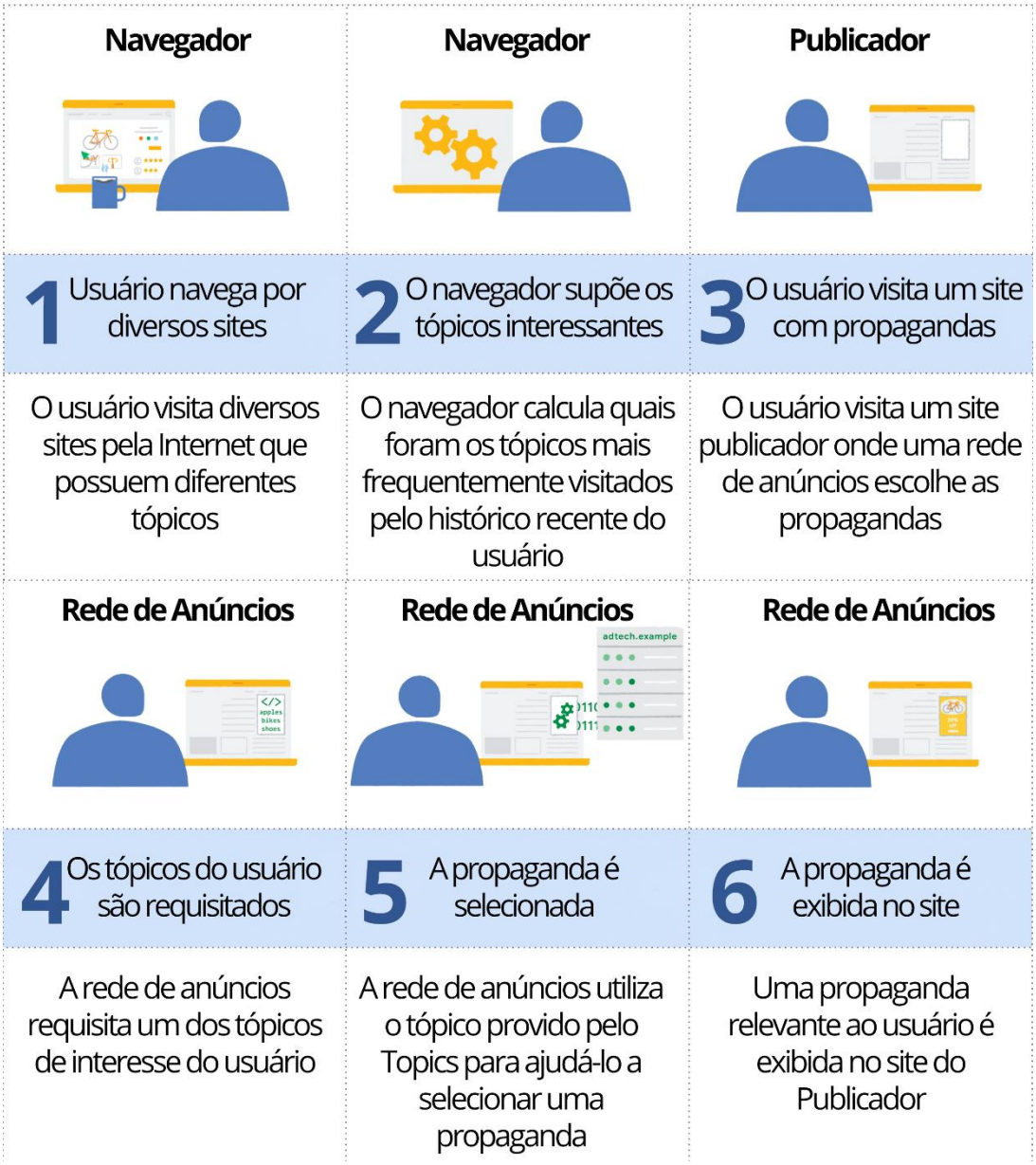
O navegador fica responsável por analisar os sites acessados pelo usuário e através de metadados ou aprendizado de máquina é ser capaz de identificar os tópicos que mais se encaixam com o site acessado [64]. Sabendo quais os tópicos dos sites acessados recentemente, o navegador gera uma lista com os tópicos 5 mais interessantes para aquele usuário naquela semana [61]. Caso o usuário não possua dados de navegação suficientes, o navegador escolhe aleatoriamente os tópicos faltantes. O usuário é capaz de ver e gerenciar os tópicos relacionados ao seu comportamento, podendo excluir os que não lhe agradam [63].

Caso o usuário acesse um site que contenha propagandas, a rede de anúncios requisita ao navegador por um tópico de interesse. O navegador seleciona um dos cinco tópicos de interesse do usuário e envia ao servidor da rede de anúncios, para a escolha da propaganda. No entanto, em 5% das vezes, o navegador pode enviar um tópico aleatório para a rede de anúncios, com o intuito de assegurar que todos os tópicos possuam membros [61]. Durante toda a próxima semana que o usuário acessar aquele site, o navegador continuará fornecendo o mesmo tópico à rede de anúncios. O navegador só poderá fornecer um novo tópico após um período de uma semana. Mesmo que os tópicos de interesse do usuário mudem durante esse período, os sites não deixarão de receber o mesmo tópico até a semana passar [61].

Os serviços que requisitarem por um tema, só podem receber os tópicos das páginas que observam por mais de três semanas [64]. Isto é, se uma rede de anúncios está presente em sites que possuam apenas os tópicos “gato”, “saúde” e “jogos digitais”; ela é capaz de receber somente esses tópicos do navegador. Isso é uma forma de impedir que as redes de anúncios tenham acesso a mais informações do

que tem com os Cookies de Terceiros, onde uma rede de anúncios só cria propagandas de sites em que ela está presente [64].

Figura 18 – Etapas do funcionamento do Topics.



Fonte: Adaptado de Google [53].

Através desse modelo de tópicos, as propagandas passam a se concentrar em um grupo de pessoas com interesses mais gerais, sem possuir se concentrar no gosto de cada indivíduo. O que impede o uso técnicas como as propagandas de *remarketing*, que dependem do rastreio de cada usuário pelos sites acessados. Contudo, essa é a moeda de troca pela privacidade, quanto menos informações do

usuário são utilizadas, menos direcionadas a um indivíduo específico as propagandas se tornam.

O Google planeja disponibilizar o Topics para todos os usuários até a segunda metade de 2023, após um período de testes públicos [65]. Até lá, todas as informações descritas sobre o Google Topics estão sujeitas a mudanças, já que o projeto está em constante desenvolvimento e a partir de sugestões da comunidade poderá sofrer alterações ou até mesmo sofrer uma interrupção no seu desenvolvimento.

6.4 FLEDGE

O fim dos Cookies de Terceiros dificulta o uso de algumas estratégias de publicidade, incluindo o *remarketing*. Sabendo disso, o Google iniciou o desenvolvimento de tecnologia *First "Locally-Executed Decision over Groups" Experiment* (Primeiro Experimento de Decisão Executada Localmente sobre Grupos - FLEDGE), que visa permitir a exibição de propagandas de *remarketing* de uma maneira que respeite a privacidade do usuário [66].

No modelo FLEDGE, os anunciantes são responsáveis por criar grupos de interesses que acreditam ser relevantes para os seus clientes. Quando um usuário acessa o site de um anunciante, como uma loja, o navegador recebe uma solicitação para fazer parte do grupo de interesses da loja ou dos produtos acessados. Caso a solicitação seja aceita, o navegador fica responsável por salvar as informações do grupo de interesses localmente. Entre as informações salvas, estão o nome do proprietário do grupo, quais propagandas podem ser exibidas para esse grupo, as funções, códigos e metadados utilizados para escolher e exibir as propagandas e em qual servidor o navegador deve buscar por atualizações do grupo [67].

Figura 19 – Criação dos grupos de interesse no FLEDGE.

Fonte: Adaptado de Agência Mestre e Google [66], [68].

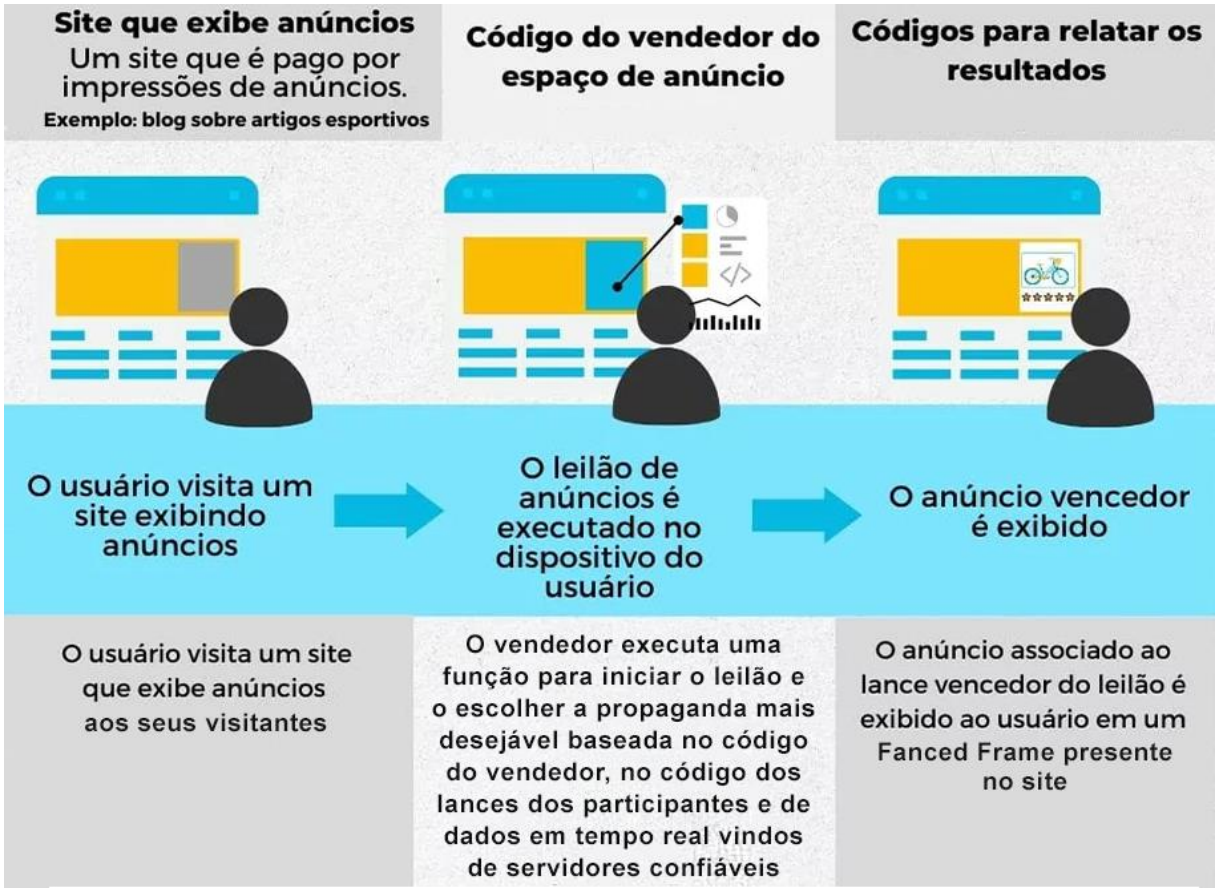
No momento que o usuário acessar um site que possua um espaço publicitário que utilize o FLEDGE, o navegador executa um leilão para decidir qual propaganda deverá ser exibida. O leilão utiliza os parâmetros de venda do espaço publicitário definidos pelo site acessado e as informações dos grupos de interesse, que contêm a lógica do leilão e os lances de cada anúncio. Toda a fase de leilão é executada localmente no navegador do usuário [67].

Com o leilão sendo executado diretamente no navegador, é possível que existam problemas de desempenho no dispositivo do usuário [67]. Por isso, o navegador pode utilizar dados em tempo real vindos de um "servidor confiável" para auxiliar no leilão. Os parâmetros de servidor confiável ainda não foram completamente definidos pelo Google, mas apenas o navegador deve ser capaz de se comunicar com ele de forma segura [67].

Após ser definido um vencedor, o leilão é encerrado e as informações sobre seu resultado são enviadas a todos os participantes. A propaganda vencedora, então,

é exibida em um *Fenced Frame* presente no site. O *Fanced Frame* é um mecanismo HTML que está sendo proposto e desenvolvido pelo Google, seu propósito é bloquear a comunicação entre o que está contido nele e o restante da página. Assim, o site não é capaz de ver os interesses do usuário baseado na propaganda que foi exibida, nem a propaganda consegue enxergar o conteúdo do site em que está [67].

Figura 20 – Leilão de anúncios através do FLEDGE.



Fonte: Adaptado de Agência Mestre e Google [66], [68].

Os grupos de interesses devem possuir um número mínimo de integrantes para preservar o anonimato de seus usuários. Cada grupo é mantido no navegador por um tempo limitado. A duração é especificada quando o anunciante solicita ao navegador a sua participação. A duração máxima definida é de 30 dias e após esse período o grupo é removido. Caso o anunciante tenha o interesse em estender essa duração, ele deve criar uma nova solicitação de grupo de interesse com o mesmo nome e proprietário, para substituir o grupo antigo [67].

7 COMPARAÇÃO ENTRE AS TECNOLOGIAS DE PUBLICIDADE DIGITAL

Este capítulo exibe uma comparação entre as características das tecnologias FLoC, Topics e FLEDGE, desenvolvidas como substitutas, e dos Cookies de Terceiros. O capítulo também presenta as melhorias sofridas pelo Topics em relação ao FLoC, seu antecessor. Outro assunto discutido neste capitulo é o interesse do Google de utilizar as tecnologias Topics e FLEDGE simultaneamente. Por último são apresentadas a análise dessa comparação.

7.1 COMPARAÇÃO COM OS COOKIES DE TERCEIROS

As tecnologias de exibição de propaganda possuem características distintas. Essas características podem ser mais adequadas em certas situações e não tão adequadas em outras. A partir das informações apresentadas nos capítulos anteriores, tornou-se possível elaborar um quadro comparativo (Quadro 1) entre os Cookies de Terceiros e as tecnologias FLoC, Topics e FLEDGE. As características comparadas nesse quadro são: os tipos de propagandas exibidas em cada tecnologia; de que forma a tecnologia separa os usuários baseados em seus gostos e interesses; em que momento os dados de navegação dos usuários são captados; em qual local esses dados são armazenados e processados; qual o grau de relevância das propagandas exibidas; e quais os problemas encontrados em cada tecnologia em envolvendo a privacidade e segurança do usuário.

Quadro 1 – Comparação entre as tecnologias de exibição de propagandas

	Cookies de Terceiros	**FLoC**	**Topics**	**FLEDGE**
Tipos de propagandas exibidas	Propaganda direcional e de *remarketing*	Propaganda direcional	Propaganda direcional	Propaganda de *Remarketing*

Separação dos usuários conforme os interesses	Perfil de interesses individuais de cada usuário gerado pelo servidor responsável pelo Cookie de terceiro	Coortes de interesses criadas e gerenciadas por inteligência artificial. Cada coorte contém alguns milhares de usuários	Lista de tópicos criadas por pessoas e gerenciada por inteligência artificial. Contém centenas de milhares de usuários em cada tópico	Grupo de interesses criados pelo serviço de propagandas, com usuários suficientes para preservar o anonimato
Obtenção de dados de navegação	Em tempo real durante a navegação entre sites	Através do histórico de navegação	Através do histórico de navegação semanal	Durante da navegação do usuário em tempo real do usuário
Local de processamento dos dados de navegação	No servidor responsável pelos Cookies	No navegador do usuário	No navegador do usuário	No navegador do usuário com o auxílio de um "servidor confiável"
Relevância das propagandas	Exibe propagandas relevantes segundo o perfil de cada usuário	Exibe propagandas que tentam ser relevantes a todo um grupo, podendo não ser interessantes a todos	Propagandas exibidas são relacionadas a uma área genérica e predefinida por uma lista, que o usuário tenha tido interesse	Exibe propagandas com remarketing semelhantes aos Cookies de Terceiros
Problemas relacionados à segurança e privacidade	Perfil detalhado dos usuários; armazenamento de dados sensíveis; vendas de dados sem autorização, roubo de Cookies, falta de controle sobre os dados compartilhados	Coortes contendo dados sensíveis; Possibilidade de *Fingerprinting*	O Google que define que dados do usuário são sensíveis	Envio de dados a um “servidor confiável” que ainda não teve suas características oficialmente definidas

Fonte: Elaboração Própria.

7.2 TOPICS: MELHORANDO O FLOC

O FLoC foi desenvolvido como um substituto aos Cookies de Terceiros e prometia melhorias na área da privacidade. Porém, ao mesmo tempo que o FLoC oferecia soluções, acabou gerando novos problemas [58]. Desta forma, o seu desenvolvimento foi interrompido e acabou sendo substituído pelo Google Topics, ou

simplesmente Topics. O Topics ficou responsável por corrigir os problemas do FLoC e apresentar as melhorias sugeridas por profissionais da área [58]. Porém, mesmo corrigindo os problemas do FLoC, o Topics não se tornou “perfeito” e possui ainda áreas cinzas que podem levantar dúvidas sobre a sua segurança [69].

Um dos problemas encontrados no FLoC que foi corrigido pelo Topics era a possibilidade de identificar os usuários por meio de *fingerprinting*. Isso ocorria devido ao baixo número de usuários alocados em cada coorte. Para corrigir isso, cada item da taxonomia do Topics é composto por um número muito maior de usuários. Além disso, cada site só recebe um item entre os cinco itens de interesse disponíveis, o que impede que o site aprenda os gostos do usuário. Assim, dificultando a reidentificação do usuário por meio de *fingerprinting* [64].

Outro problema que o FLoC enfrentava eram as coortes com conteúdo sensível devido à geração por meio de inteligência artificial. Isso também foi abordado no desenvolvimento do Topics. Nesse modelo, os tópicos são escolhidos e filtrados por pessoas ao invés de serem gerados por uma inteligência artificial. Porém, quem decide se um conteúdo é sensível ou não, é o próprio Google. Isso pode gerar divergências com usuário, já que apenas ele pode definir quais dos seus dados são considerados sensíveis. Mesmo assim, o Topics apresentou mais uma melhoria em relação ao FLoC, uma que com o Topics, o usuário é capaz de visualizar e gerenciar de quais tópicos faz parte e remover os que desejar, algo que não era possível com as coortes do FLoC.

7.3 USO SIMULTÂNEO DAS TECNOLOGIAS

A partir das características comparadas, pode-se afirmar que os Cookies de Terceiros possuem uma versatilidade muito maior do que os seus sucessores em relação à diversidade dos tipos de propagandas exibidas. Por isso, é possível deduzir que desenvolver uma única ferramenta capaz de ocupar todas as funcionalidades dos Cookies de Terceiros e continuar respeitando a privacidade do usuário é uma tarefa complicada. Com isso, cada tecnologia se dedica a uma das funções exercidas pelos Cookies de Terceiros. O FLoC e o Topics, por exemplo, têm como objetivo exibir propagandas direcionadas aos interesses gerais do usuário, mas são incapazes de

gerar propagandas de *remarketing*. Já com FLEDGE o oposto ocorre. O seu objetivo é exibir apenas propagandas de *remarketing*, sem se preocupar em gerar propagandas direcionadas.

Cada uma das tecnologias fica responsável por substituir uma das características dos Cookies de Terceiros. No entanto, o uso exclusivo de uma ou de outra tecnologia não é interessante para todos os tipos de negócio. O Google tem o interesse de testar o uso dessas duas tecnologias ao mesmo tempo [70]. A estratégia consiste em fazer o Topics exibir propagandas direcionadas e o FLEDGE exibir propagandas de remarketing. Deste modo, num mesmo site poderão existir algumas propagandas geradas pelo Topics e algumas propagandas geradas pelo FLEDGE. Desta forma, dois dos principais modelos de propagandas utilizados da Internet, remarketing e propaganda direcionada, são cobertos.

7.4 CONCLUSÕES

Após analisar e comparar os Cookies de Terceiros e as novas tecnologias, é possível chegar à conclusão de que o Topics e o FLEDGE serão, provavelmente, tecnologias de grande relevância para o Google. Porém, essa relevância não está relacionada ao aumento da privacidade do usuário. A real aposta do Google com essas tecnologias é de se manter no topo, possuindo os principais serviços de exibição de propagandas. E o Google tem o poder para isso, pois mais da metade dos usuários da Internet utilizam o seu navegador [23]. Portanto, é possível deduzir que a real preocupação do Google não é direcionada à privacidade do usuário, mas direcionada a não perder usuários e eventualmente capital, por conta de problemas com a privacidade.

De acordo com a análise feita por este capítulo pelo capítulo anterior, as novas tecnologias realmente oferecem algum tipo de melhoria em relação a privacidade dos usuários, mesmo que não sejam tão efetivos quanto os Cookies de Terceiros. Porém, essas tecnologias ainda possuem problemas que continuam chamando a atenção dos especialistas. O que pode fazer com que a reputação do Topics e do FLEDGE não sejam tão melhores do que a reputação dos Cookies de Terceiros.

8 CONCLUSÃO

Este trabalho apresenta um estudo sobre os Cookies Primários, com ênfase nos Cookies de Terceiros e as consequências do seu uso. O trabalho discorre sobre a origem, o funcionamento e as diferenças entre os Cookies Primários e os Cookies de Terceiros. No trabalho são levantadas questões envolvendo a privacidade e a segurança do usuário. Além disso, é abordado como a influência dos Cookies de Terceiros na área da publicidade digital, os tornam tão difíceis de serem abandonados mesmo com seus problemas. Também são descritas as tecnologias de FLoC, Topics e FLEDGE, que estão sendo desenvolvidas com o objetivo de substituírem os Cookies de Terceiros. Por último é feita uma comparação entre os Cookies de Terceiros e as novas tecnologias, analisando quais melhorias foram alcançadas e quais problemas ainda existem. Portanto, considerando todo o conteúdo apresentado por este trabalho, é possível concluir que os objetivos propostos foram alcançados.

Em trabalhos futuros é possível estender o conteúdo sobre as novas tecnologias que estão sendo desenvolvidas para substituir os Cookies de Terceiros, recorrendo ao conteúdo que já foi agregado por este trabalho. Tendo como o objetivo, analisar os impactos dessas novas tecnologias na privacidade do usuário e como o mercado da publicidade digital será afetado por elas. Isso se dá, devido ao fato dessas tecnologias ainda estarem em fase de desenvolvimento. O que não permite ter uma noção completa de como elas serão recebidas pelos anunciantes e pelas redes de anúncios, nem é possível saber quais alterações poderão ocorrer até o fim de seus desenvolvimentos ou que outras tecnologias poderão surgir.

REFERÊNCIAS

[1] Brasil, *LEI Nº 13.709, DE 14 DE AGOSTO DE 2018*. Brasília : Presidência da República, 2018. Acessado: ago. 19, 2022. [Online]. Available: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709compilado.htm

[2] R. Koch, “Cookies, the GDPR, and the ePrivacy Directive - GDPR.eu”. https://gdpr.eu/cookies/ (acessado set. 18, 2022).

[3] L. C. de Souza, “Carole Crema fala de cookies e dá receita para a sua privacidade online”, jul. 13, 2022. https://blog.avast.com/pt-br/video-receita-cookies-carole-crema-exclusiva-avast-privacidade-online (acessado ago. 19, 2022).

[4] D. M. Kristol, “HTTP Cookies: Standards, Privacy, and Politics”, *ACM Trans. Internet Technol.*, vol. 1, nº 2, p. 151–198, nov. 2001, doi: 10.1145/502152.502153.

[5] “Privacidade - Google Trends”, 2022. https://trends.google.com.br/trends/explore?date=2017-01-01%202022-01-08&geo=BR&q=%2Fm%2F06804 (acessado ago. 19, 2022).

[6] J. Wilander, “Full Third-Party Cookie Blocking and More | WebKit”, mar. 04, 2020. https://webkit.org/blog/10218/full-third-party-cookie-blocking-and-more/ (acessado ago. 19, 2022).

[7] J. F. Kurose e K. W. Ross, *COMPUTER NETWORKING A Top-Down Approach*, 6º ed. Pearson, 2013.

[8] H. T. Alvestrand, “A Mission Statement for the IETF”, nº 3935. RFC Editor, out. 2004. doi: 10.17487/RFC3935.

[9] A. S. Tanenbaum e D. Wetherall, *REDE DE COMPUTADORES*, 5º ed. São Paulo: Pearson Prentice Hall, 2011.

[10] J. Postel, “Internet Protocol”, set. 1981, doi: 10.17487/RFC0791.

[11] W. M. Eddy, “Transmission Control Protocol (TCP)”, ago. 2022, doi: 10.17487/RFC9293.

[12] A. Uzelac e Y. Lee, “Voice over IP (VoIP) SIP Peering Use Cases”, nov. 2011, doi: 10.17487/RFC6405.

[13] J. Klensin, “Simple Mail Transfer Protocol”, out. 2008, doi: 10.17487/RFC5321.

[14] M. Belshe, R. Peon, e M. Thomson, “Hypertext Transfer Protocol Version 2 (HTTP/2)”, maio 2015, doi: 10.17487/RFC7540.

[15] H. Nielsen *et al.*, “Hypertext Transfer Protocol – HTTP/1.1”, nº 2616. RFC Editor, jun. 1999. doi: 10.17487/RFC2616.

[16] R. T. Fielding, M. Nottingham, e J. Reschke, “RFC 9110: HTTP Semantics”, 2022, Acessado: nov. 15, 2022. [Online]. Available: https://www.rfc-editor.org/info/rfc9110

[17] “Scopus preview - Scopus - Welcome to Scopus”. https://www.scopus.com/ (acessado ago. 19, 2022).

[18] J. DeCew, “Privacy”, em *The Stanford Encyclopedia of Philosophy*, Spring 2018., E. N. Zalta, Org. Metaphysics Research Lab, Stanford University, 2018. Acessado: set. 22, 2022. [Online]. Available: https://plato.stanford.edu/archives/spr2018/entries/privacy/

[19] S. D. Warren e L. D. Brandeis, “The Right to Privacy”, *Harv Law Rev*, vol. 4, nº 5, p. 193, dez. 1890, doi: 10.2307/1321160.

[20] Brasil, *Constituição da República Federativa do Brasil*. Brasilia, 1988. Acessado: set. 22, 2022. [Online]. Available: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

[21] “GitHub: Let’s build from here · GitHub”. https://github.com/ (acessado nov. 09, 2022).

[22] “Sites Mais Visitados - Ranking dos Principais Sites | Similarweb”. https://www.similarweb.com/pt/top-websites/ (acessado set. 23, 2022).

[23] “Browser Market Share Worldwide | Statcounter Global Stats”. https://gs.statcounter.com/browser-market-share (acessado set. 23, 2022).

[24] Alphabet Inc., “UNITED STATES SECURITIES AND EXCHANGE COMMISSION FORM 10-K”, Washington, D.C., 2021. Acessado: set. 24, 2022. [Online]. Available: https://abc.xyz/investor/static/pdf/20220202_alphabet_10K.pdf

[25] “Form 10-K | Investor.gov”. https://www.investor.gov/introduction-investing/investing-basics/glossary/form-10-k (acessado nov. 14, 2022).

[26] E. S. Raymond, *The New Hacker’s Dictionary, third edition*. 1996.

[27] D. Kristol e L. Montulli, “HTTP State Management Mechanism”, fev. 1997, doi: 10.17487/RFC2109.

[28] A. Barth, “HTTP State Management Mechanism”, nº 6265. RFC Editor, abr. 2011. doi: 10.17487/RFC6265.

[29] MDN contributors, “Set-Cookie - HTTP | MDN”, 2020. https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Set-Cookie (acessado set. 07, 2022).

[30] KirstenS, “Cross Site Scripting (XSS) | OWASP Foundation”. https://owasp.org/www-community/attacks/xss/ (acessado dez. 01, 2022).

[31] C. Sidoti, “How HttpOnly cookies help mitigate XSS attacks”. https://clerk.dev/blog/how-httponly-cookies-help-mitigate-xss-attacks (acessado dez. 01, 2022).

[32] E. Rescorla, “RFC 2818: HTTP Over TLS”, maio 2000, doi: 10.17487/RFC2818.

[33] MDN contributors, “Using HTTP cookies - HTTP | MDN”. https://developer.mozilla.org/en-US/docs/Web/HTTP/Cookies (acessado dez. 03, 2022).

[34] HITCON, “HITCON CMT 2019 - The cookie monster in your browsers - YouTube”, jul. 23, 2020. https://youtu.be/njQcVWPB1is (acessado dez. 03, 2022).

[35] S. Langkemper, “Overwriting HttpOnly cookies using cookie jar overflow”. https://www.sjoerdlangkemper.nl/2020/05/27/overwriting-httponly-cookies-from-javascript-using-cookie-jar-overflow/ (acessado dez. 03, 2022).

[36] “What is CSRF (Cross-site request forgery)? Tutorial & Examples | Web Security Academy”. https://portswigger.net/web-security/csrf (acessado dez. 01, 2022).

[37] “Cross-Site Request Forgery Prevention - OWASP Cheat Sheet Series”. https://cheatsheetseries.owasp.org/cheatsheets/Cross-Site_Request_Forgery_Prevention_Cheat_Sheet.html (acessado dez. 02, 2022).

[38] “CSRF tokens | Web Security Academy”. https://portswigger.net/web-security/csrf/tokens (acessado dez. 02, 2022).

[39] A. Epasto *et al.*, “Clustering for Private Interest-based Advertising”, *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 2802–2810, ago. 2021, doi: 10.1145/3447548.3467180.

[40] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, e Z. Chen, “How much can behavioral targeting help online advertising?”, em *Proceedings of the 18th international conference on world wide web*, 2009, p. 261–270.

[41] “G1 - O portal de notícias da Globo”. https://g1.globo.com/ (acessado nov. 05, 2022).

[42] “CookieYes Cookie Consent Solution”. https://www.cookieyes.com/ (acessado out. 15, 2022).

[43] “Home - Cookiedatabase.org”. https://cookiedatabase.org/ (acessado out. 15, 2022).

[44] J. Kwakman, “Open-Cookie-Database/open-cookie-database.csv at master · jkwakman/Open-Cookie-Database · GitHub”. https://github.com/jkwakman/Open-Cookie-Database/blob/master/open-cookie-database.csv (acessado out. 15, 2022).

[45] “Cookie Checker - CookieYes”. https://www.cookieyes.com/cookie-checker/?ref=cybasicscanner&website=https://g1.globo.com/ (acessado out. 15, 2022).

[46] J. Kwakman, “Open-Cookie-Database/open-cookie-database.csv at master · jkwakman/Open-Cookie-Database · GitHub”, out. 2019. https://github.com/jkwakman/Open-Cookie-Database/blob/master/open-cookie-database.csv (acessado out. 15, 2022).

[47] N. Neumann, C. E. Tucker, e T. Whitfield, “Frontiers: How effective is third-party consumer profiling? evidence from field studies”, *Marketing Science*, vol. 38, nº 6, p. 918–926, 2019, doi: 10.1287/MKSC.2019.1188.

[48] D. F. Gray, “CNN - DoubleClick sued for privacy violations - January 28, 2000”, 2000. http://edition.cnn.com/2000/TECH/computing/01/28/double.click.lawsuit.idg/index.html (acessado set. 17, 2022).

[49] M. Wood, “Today’s Firefox Blocks Third-Party Tracking Cookies and Cryptomining by Default”, set. 03, 2019. https://blog.mozilla.org/en/products/firefox/todays-firefox-blocks-third-party-tracking-cookies-and-cryptomining-by-default/ (acessado set. 18, 2022).

[50] A. Chavez, “Expanding testing for the Privacy Sandbox for the Web”, jul. 27, 2022. https://blog.google/products/chrome/update-testing-privacy-sandbox-web/ (acessado set. 23, 2022).

[51] “Como o Google AdSense usa cookies - Ajuda do Google AdSense”. https://support.google.com/adsense/answer/7549925?hl=pt-BR (acessado set. 24, 2022).

[52] “Como funciona o Google AdSense - Ajuda do Google AdSense”. https://support.google.com/adsense/answer/6242051?hl=pt-BR#zippy=%2Cqual-%C3%A9-a-diferen%C3%A7a-entre-o-google-adsense-e-as-outras-redes-de-an%C3%BAncios (acessado set. 23, 2022).

[53] S. Dutton, “What is FLoC?”, mar. 30, 2021. https://web.dev/i18n/en/floc/ (acessado set. 24, 2022).

[54] “Coorte - Dicionário Online Priberam de Português”. https://dicionario.priberam.org/coorte (acessado nov. 14, 2022).

[55] B. Cyphers, “Google’s FLoC Is a Terrible Idea | Electronic Frontier Foundation”, mar. 03, 2021. https://www.eff.org/deeplinks/2021/03/googles-floc-terrible-idea (acessado set. 24, 2022).

[56] E. Rescorla e M. Thomson, “Technical Comments on FLoC Privacy”, jun. 2021, Acessado: set. 24, 2022. [Online]. Available: https://mozilla.github.io/ppa-docs/floc_report.pdf

[57] B. McMahan e D. Ramage, “Google AI Blog: Federated Learning: Collaborative Machine Learning without Centralized Training Data”, abr. 06, 2017. https://ai.googleblog.com/2017/04/federated-learning-collaborative.html (acessado set. 24, 2022).

[58] J. Karlin e M. Kleber, “GitHub - The Topics API - Evolution from FLoC”, jan. 21, 2022. https://github.com/patcg-individual-drafts/topics#evolution-from-floc (acessado set. 24, 2022).

[59] Y. Xiao *et al.*, “GitHub - WICG/floc: FLoC”, ago. 22, 2019. https://github.com/WICG/floc (acessado set. 30, 2022).

[60] F. Turati, “Analysing and exploiting Google’s FLoC advertising proposal”, mar. 2022, doi: 10.3929/ETHZ-B-000539945.

[61] J. Karlin, M. Kleber, e S. Dutton, “GitHub - patcg-individual-drafts/topics: The Topics API”, jan. 21, 2022. https://github.com/patcg-individual-drafts/topics (acessado set. 30, 2022).

[62] “The Privacy Sandbox: Technology for a More Private Web”. https://privacysandbox.com/intl/en_us/proposals/floc (acessado set. 24, 2022).

[63] V. Goel, “Get to know the new Topics API for Privacy Sandbox”, jan. 25, 2022. https://blog.google/products/chrome/get-know-new-topics-api-privacy-sandbox/ (acessado set. 25, 2022).

[64] S. Dutton, “The Topics API - Chrome Developers”, jan. 25, 2022. https://developer.chrome.com/en/docs/privacy-sandbox/topics/ (acessado set. 30, 2022).

[65] “How We’re Protecting Your Online Privacy - The Privacy Sandbox”. https://privacysandbox.com/intl/en_us/open-web/#the-privacy-sandbox-timeline (acessado set. 30, 2022).

[66] S. Dutton, “FLEDGE API - Chrome Developers”, jan. 27, 2022. https://developer.chrome.com/docs/privacy-sandbox/fledge/ (acessado out. 07, 2022).

[67] M. Kleber, “turtledove/FLEDGE.md at main · WICG/turtledove · GitHub”, jan. 22, 2021. https://github.com/WICG/turtledove/blob/main/FLEDGE.md (acessado out. 07, 2022).

[68] R. Pereira, “O Futuro dos Anúncios: Conheça a API FLEDGE”, ago. 03, 2022. https://www.agenciamestre.com/trafego-pago/api-fledge-google-ads/ (acessado out. 15, 2022).

[69] P. Snyder, “Google’s Topics API: Rebranding FLoC Without Addressing Key Privacy Issues | Brave Browser”. https://brave.com/web-standards-at-brave/7-googles-topics-api/ (acessado nov. 11, 2022).

[70] “turtledove/Proposed_First_FLEDGE_OT_Details.md at main · WICG/turtledove · GitHub”. https://github.com/WICG/turtledove/blob/main/Proposed_First_FLEDGE_OT_Details.md (acessado nov. 18, 2022).