Capitalismo de Vigilância e a Coleta de Dados Online

Antony Seabra, Claudio Fraga, Sergio Lifschitz

[1]Departamento de Informatica - PUC-Rio

{amedeiros, cfraga, sergio}@inf.puc-rio.br

Resumo. *Este artigo investiga o fenômeno do capitalismo de vigilância e suas implicações para a coleta e uso de dados pessoais na web. Através da análise empírica do tráfego de rede, demonstramos como as técnicas de rastreamento e direcionamento de dados, inicialmente desenvolvidas para fins comerciais, permeiam a experiência online dos usuários, revelando um ecossistema complexo onde informações pessoais são continuamente extraídas, compartilhadas e utilizadas para influenciar o comportamento dos usuários. Além disso, discutimos a preocupante possibilidade da aplicação dessas técnicas para a disseminação de notícias falsas e desinformação, representando uma ameaça direta à integridade democrática. Ao fornecer evidências empíricas, podemos contribuir para o desenvolvimento de políticas e ferramentas eficazes para proteger a integridade da informação e fortalecer a democracia na era digital.*

1. Introdução

Observa-se, no cenário digital atual, uma alteração substancial: a ascensão do capitalismo de vigilância, onde os dados pessoais não são mais meros subprodutos da atividade online, mas um ativo primário, alimentando uma vasta e intrincada maquinaria econômica. Neste ecossistema orientado por dados, a capacidade de extrair, analisar e monetizar informações do usuário tornou-se a base de inúmeros modelos de negócios lucrativos. Gigantes da era digital, como Google e Meta, construíram impérios basedos na manipulação sofisticada de dados pessoais para publicidade direcionada, uma prática que, como destaca [Zuboff 2023], representou impressionantes 89% da receita da Alphabet por meio dos programas de anúncios direcionados do Google já em 2016. A magnitude absoluta dessa extração de dados é exemplificada pelo domínio da internet do Google, processando uma média de mais de 40.000 consultas de pesquisa por segundo, traduzindo para mais de 3,5 bilhões de pesquisas diárias e 1,2 trilhão de pesquisas anuais em todo o mundo em 2017.

Além da busca e das mídias sociais, a influência do capitalismo de vigilância permeia diversos setores. Gigantes do comércio eletrônico, como a Amazon, alavancam algoritmos complexos para personalizar as experiências do usuário, adaptando as recomendações de produtos com base em pontos de dados granulares, como compras anteriores, padrões de navegação e históricos de pesquisa. Essa hiperpersonalização não apenas promove um maior envolvimento do usuário, mas também impulsiona um crescimento substancial da receita. Dentro do crescente setor de tecnologia de fitness e saúde, empresas como Fitbit e Strava coletam dados extensos sobre atividades do usuário e métricas fisiológicas. Esses dados, embora potencialmente benéficos para insights de saúde personalizados, também levantam preocupações sobre seu uso potencial por terceiros, como seguradoras ou instituições de pesquisa. Esses exemplos mostram a onipresença de dados pessoais como um recurso estratégico nos modelos de negócios

contemporâneos, onde a profundidade e a amplitude das informações do usuário se correlacionam diretamente com o valor percebido.

Este artigo procura fornecer validação empírica das práticas generalizadas de compartilhamento de dados que sustentam o capitalismo de vigilância. Especificamente, pretendemos documentar instâncias de transmissão de dados de dispositivos pessoais para serviços baseados na Internet, iluminando assim os mecanismos muitas vezes opacos pelos quais os dados do usuário são coletados e disseminados. Além disso, exploraremos as implicações potenciais dessas práticas para a privacidade e autonomia do usuário. Para garantir a transparência e facilitar novas pesquisas, todos os dados capturados durante nossas investigações empíricas estão publicamente acessíveis em [Seabra 2025].

2. Capitalismo de Vigilância

O surgimento do capitalismo de vigilância, termo cunhado por Shoshana Zuboff em sua obra inovadora, *A Era do Capitalismo de Vigilância* [Zuboff 2019], significa uma transformação fundamental na dinâmica da criação de valor econômico. Essa nova forma de capitalismo depende da extração e análise sistemáticas de dados pessoais, coletados dos vastos ecossistemas digitais de plataformas e serviços baseados na Internet. Esses fluxos de dados são então meticulosamente processados para construir perfis comportamentais complexos, capturando as nuances das preferências e ações de usuários individuais. Os insights resultantes não são meramente retidos em uma única entidade, mas sim disseminados por uma rede de empresas interconectadas, alimentando um mercado preditivo onde as expectativas sobre o comportamento do usuário são negociadas como *commodities*.

Em seus estágios iniciais, a Internet se caracterizava pela descentralização e pela defesa da troca livre de informações e do anonimato. No entanto, a trajetória desta fronteira digital sofreu uma mudança dramática com a ascensão de gigantes tecnológicos como Google e Facebook (hoje Meta). Essas entidades pioneiras reconheceram o potencial econômico latente nos vastos repositórios de dados gerados pelo usuário, transformando consultas de pesquisa e interações sociais em matérias-primas de um modelo de negócios revolucionário. Ao longo do tempo, essas corporações refinaram suas metodologias de aquisição de dados, implantando técnicas como *cookies* e rastreamento de *pixels* para monitorar meticulosamente o comportamento do usuário em todo o cenário digital.

O crescimento exponencial nas capacidades de armazenamento e processamento de dados facilitou o acúmulo de volumes sem precedentes de dados pessoais. Esses dados alimentaram não apenas a publicidade direcionada, mas também a criação de perfis de usuário sofisticados, permitindo a previsão e a manipulação do comportamento do usuário. O modelo transcendeu a publicidade tradicional, à medida que as empresas começaram a monetizar insights preditivos, oferecendo-os a empresas terceirizadas. Embora essa personalização possa melhorar a experiência do usuário em determinados domínios, e muitos associam essa experiência a conveniência, ela levanta preocupações sobre a privacidade de dados e a autonomia individual. Além disso, [Zuboff 2022] argumenta que o capitalismo de vigilância representa uma ameaça inerente às instituições democráticas, levando à instabilidade social e à erosão dos direitos civis. Segundo a autora, isso exige o desenvolvimento de novas instituições públicas, estruturas de direitos e salvaguardas legais adaptadas aos imperativos democráticos da era digital, protegendo

assim os cidadãos do potencial exploratório de seus dados pessoais.

Os mecanismos de rastreamento generalizados inerentes ao capitalismo de vigilância também apresentam riscos significativos para a esfera política. Essas técnicas permitem o direcionamento preciso de indivíduos e grupos com mensagens políticas personalizadas, incluindo a disseminação de desinformação e notícias falsas. Ao alavancar perfis de usuário detalhados, os atores políticos podem explorar vulnerabilidades e manipular a opinião pública, minando a integridade dos processos democráticos. Essa capacidade de influência granular representa uma ameaça direta à integridade eleitoral, ao discurso cívico e à estabilidade das instituições democráticas. A capacidade de rastrear e influenciar o comportamento do usuário se estende além dos interesses comerciais, potencialmente se transformando em uma ferramenta poderosa de manipulação política e controle social.

Embora o discurso acadêmico tenha abordado extensivamente as preocupações com a privacidade de dados e as propostas regulatórias, como evidenciado por trabalhos como [Andrew and Baker 2021, Wu et al. 2023], persiste uma lacuna de pesquisa significativa. Especificamente, há uma escassez de investigações empíricas que fornecem evidências concretas dos mecanismos e impactos do capitalismo de vigilância. Este artigo visa contribuir para preencher essa lacuna, fornecendo uma análise detalhada das práticas de coleta e compartilhamento de dados que sustentam o capitalismo de vigilância.

3. Metodologia

Este estudo adota uma abordagem de análise de tráfego de rede para investigar empiricamente as práticas de coleta e compartilhamento de dados que sustentam o capitalismo de vigilância. Nosso objetivo principal é capturar e analisar os pacotes de dados transmitidos de dispositivos pessoais, como *smartphones* e computadores, quando os usuários interagem com sites e serviços online. Este método nos permite obter informações sobre os mecanismos muitas vezes opacos pelos quais os dados do usuário são coletados, transmitidos e potencialmente compartilhados com entidades terceirizadas.

A metodologia utiliza o conceito de um servidor proxy *Man in the Middle* (MITM), uma técnica comumente usada em pesquisa de segurança de rede para interceptar e analisar o tráfego de rede [mit 2024]. Nesta configuração, um servidor proxy dedicado é posicionado entre o dispositivo do usuário e a internet. Toda a comunicação entre o dispositivo e servidores externos é roteada através deste proxy, permitindo a interceptação e inspeção de pacotes de dados em trânsito. Para facilitar a descriptografia e análise do tráfego HTTPS, o certificado do servidor proxy é instalado no dispositivo do usuário. Isso permite que o proxy descriptografe conexões HTTPS criptografadas, fornecendo acesso aos dados transmitidos em formato de texto simples.

O procedimento detalhado para configurar o servidor proxy MITM e configurar os dispositivos do usuário está documentado em nosso repositório publicamente acessível em [Github, 2025]. A ferramenta fornece uma interface web amigável para visualizar dados interceptados, filtrar o tráfego com base em critérios específicos e conduzir uma análise aprofundada de pacotes capturados. Ao monitorar os pacotes trafegados nesse acesso, procuraremos:

Identificar o destino dos dados: Determinar os serviços web específicos e domínios

de terceiros que recebem dados do usuário ao acessar um website ou realizar uma busca na Web.

Analisar o conteúdo dos dados: Examinar o conteúdo dos pacotes de dados transmitidos para entender os tipos de informações que estão sendo coletadas e compartilhadas, incluindo identificadores pessoais, histórico de navegação e dados comportamentais.

Descobrir mecanismos de rastreamento: Investigar as tecnologias de rastreamento específicas empregadas, como *cookies*, rastreadores de *pixels* e técnicas de *fingerprinting* de navegador.

Quantificar a transmissão de dados: Medir o volume e a frequência da transmissão de dados para diferentes entidades.

Ao empregar esta metodologia para analisar o tráfego dos dados a partir do acesso a sites, pretendemos identificar completamente os fluxos de dados complexos que caracterizam o capitalismo de vigilância, fornecendo evidências concretas da extensão e natureza das práticas de coleta e compartilhamento de dados na Web.

4. Estudos de Caso

4.1. Navegação na Web

Este estudo de caso examina a extensão e os mecanismos de coleta de dados durante sessões típicas de navegação na web em um contexto brasileiro. Ao navegar por vários sites, o estudo revela mecanismos de rastreamento digital. Começamos com a navegação direta para o site *magalu.com.br*. Essa abordagem é escolhida para monitorar e analisar a pegada digital deixada por tal navegação, especialmente a comunicação subsequente com serviços externos. Após a conclusão das interações dentro do *magalu.com.br*, nossa análise se concentra nos serviços externos contatados como resultado dessa visita inicial. Especificamente, queremos identificar e listar as entidades fora do *magalu.com.br* que foram acessadas, conforme indicado pela inclusão de *magalu.com.br* no *payload* dos pacotes de rede.

Ao visitar *magalu.com.br*, nossa investigação revelou diversos acessos externos a uma variedade de serviços, incluindo, entre outros, facebook.com, twitter.com, tiktok.com e pinterest.com. Cada um desses serviços desempenha um papel distinto no ecossistema de publicidade digital, contribuindo para uma abordagem multifacetada para rastreamento, criação de perfil e publicidade direcionada de usuários online.

Por exemplo, o Criteo é especializado em retargeting, exibindo anúncios para usuários que visitaram sites específicos, sugerindo que visitar *magalu.com.br* pode levar a anúncios direcionados do Magazine Luiza em outros sites. Da mesma forma, os pixels de rastreamento do Facebook (por meio de connect.facebook.net e www.facebook.com) permitem a coleta de interações detalhadas do usuário no site do Magazine Luiza, permitindo anúncios altamente personalizados nas plataformas do Facebook. Os serviços de análise do X (antigo Twitter) estendem essa capacidade para o domínio do engajamento de mídia social, potencialmente influenciando os anúncios e o conteúdo que os usuários veem no X com base em seu histórico de navegação. Os serviços de análise do TikTok são projetados para rastrear e analisar as interações do usuário relacionadas ao conteúdo do TikTok incorporado ou compartilhado nesse site.

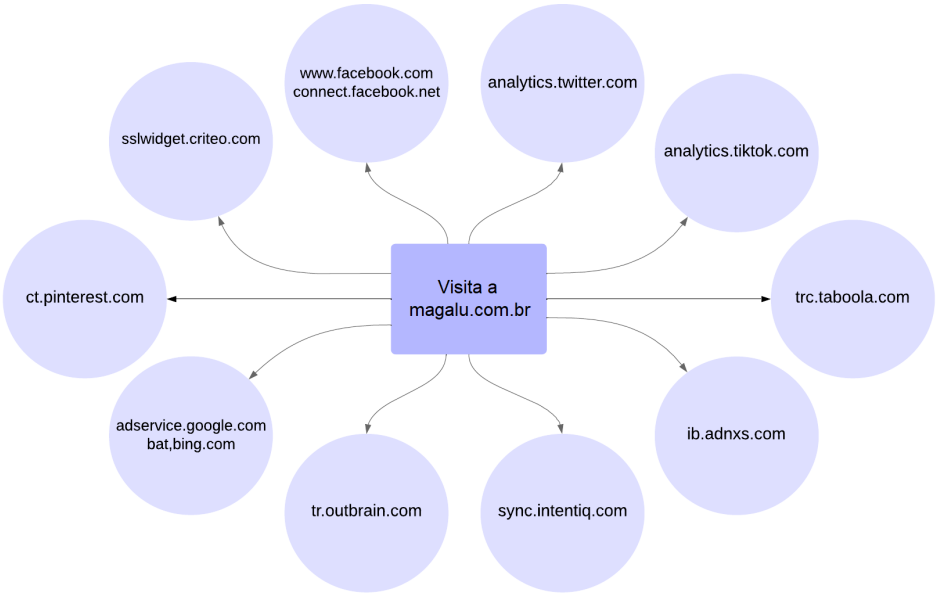

Figure 1. Acessos externos após visitar magalu.com.br. Fonte: Autores.

O serviço de rastreamento do Bing (bat.bing.com) e os serviços de anúncios do Google oferecem insights sobre o comportamento e as preferências de pesquisa do usuário, refinando ainda mais os recursos de segmentação de anúncios na Web. Taboola e Outbrain são especializados em recomendação de conteúdo, indicando que os usuários podem ver conteúdo sugerido relacionado a produtos do Magazine Luiza ou interesses relacionados com base em sua visita a *magalu.com.br*.

Adnxs.com (AppNexus) representa uma plataforma de publicidade programática que facilita lances em tempo real para espaço publicitário, sugerindo que os dados dos usuários podem ser usados para leiloar espaço publicitário em tempo real para o licitante mais alto, com base no valor percebido da oportunidade de publicidade. Intentiq e serviços semelhantes (sync.intentiq.com) se concentram na resolução de identidade, ajudando os anunciantes a vincular a atividade em vários dispositivos a um único usuário, aprimorando as estratégias de segmentação entre dispositivos.

Após a captura e análise detalhadas do tráfego de rede iniciado por uma visita a *magalu.com.br*, uma navegação subsequente a um site de notícias, como *g1.globo.com*, eventualmente revela uma observação significativa: uma série de anúncios originados de *magalu.com.br* em destaque no site. Essa ocorrência serve como uma evidência direta dos mecanismos sofisticados das redes de publicidade online e sua capacidade de fornecer anúncios altamente direcionados com base na atividade recente do usuário. A presença de anúncios do Magazine Luiza em *www.globo.com*, logo após a visita ao site oficial do Magazine Luiza, destaca a eficácia das tecnologias de rastreamento e cookies no rastreamento de interesses e comportamentos do usuário na web.

Além dos rastreadores de grandes corporações globais, o cenário brasileiro de publicidade digital também inclui atores locais que desempenham um papel ativo na coleta de dados.

Figure 2. Publicidade. Fonte: Autores.

Empresas brasileiras especializadas em marketing digital e publicidade programática frequentemente implantam seus próprios rastreadores, complementando ou, em alguns casos, competindo com as soluções internacionais. Esses trackers, muitas vezes integrados a plataformas de Demand-Side (DSPs) e Data Management Platforms (DMPs) locais, permitem a coleta de dados de navegação, comportamento online e informações demográficas de usuários brasileiros. O objetivo principal é a criação de perfis detalhados para segmentação de anúncios, otimização de campanhas publicitárias e análise de desempenho.

4.2. Busca na Web

Este estudo de caso investiga os mecanismos de vigilância ativados por mecanismos de busca em resposta a consultas de usuários. Ao analisar o tráfego de rede gerado a partir da pesquisa de termos específicos, é possível verificar como os mecanismos de busca coletam, armazenam e possivelmente compartilham dados de pesquisa. Adicionalmente, explora o potencial para a construção de perfis com base no histórico de pesquisa e examina as preocupações de privacidade relacionadas a resultados de pesquisa personalizados e publicidade direcionada.

Para ilustrar este caso, realizamos buscas por termos variados utilizando Google. A análise do tráfego de rede revelou uma série de interações com serviços externos, além da simples exibição dos resultados da pesquisa. Observamos, por exemplo, a comunicação com plataformas de publicidade como o Google Ads e o Facebook Ads, evidenciando a coleta de dados da pesquisa para personalização de anúncios. Notavelmente, a análise inclui acessos a domínios utilizados para transmitir dados analíticos e de desempenho, permitindo que as empresas monitorem e otimizem seus serviços. Além disso, a análise revela interações com *endpoints* que contribuem para aprimorar a experiência do usuário em vários serviços.

As preocupações com a privacidade decorrentes dessas práticas são multifacetadas. À medida que os mecanismos de busca rastreiam e armazenam grandes quantidades de dados do usuário, incluindo histórico de pesquisa, localização e preferências pessoais, eles constroem perfis detalhados que podem ser usados para veicular anúncios altamente direcionados. Essa coleta e criação de perfil de dados levantam questões significativas de privacidade, pois os usuários geralmente não sabem a extensão das informações coletadas e como elas são utilizadas. Há uma falta de transparência e, com algumas exceções, como a linha do tempo do Google, nenhum consentimento do usuário nesses processos.

4.3. Notícias falsas e desinformação

As mesmas técnicas de rastreamento e direcionamento que alimentam o motor econômico do capitalismo de vigilância podem ser facilmente reaproveitadas para fins políticos nefastos. A capacidade de coletar dados granulares sobre o comportamento online, preferências e vulnerabilidades dos indivíduos cria um terreno fértil para a disseminação de notícias falsas, desinformação e propaganda enganosa. Ao aproveitar esses *insights*, agentes mal-intencionados podem elaborar campanhas altamente direcionadas, projetadas para manipular a opinião pública, semear a discórdia e minar os processos democráticos.

Na Web, isso geralmente se manifesta através da disseminação de artigos de notícias fabricados, vídeos manipulados e postagens em mídias sociais, projetadas para desencadear respostas emocionais e contornar o pensamento crítico. Essas mensagens enganosas são então micro-direcionadas a indivíduos com base em sua suscetibilidade a narrativas específicas, amplificando seu impacto e promovendo câmaras de eco de desinformação. A natureza pervasiva do rastreamento online permite que essas campanhas transcendam as fronteiras das plataformas, seguindo os usuários na web e nas mídias sociais, criando uma enxurrada constante de desinformação que erode a confiança em fontes e instituições legítimas. Essa instrumentalização da informação representa uma ameaça significativa à integridade das eleições, ao discurso público e ao próprio tecido das sociedades democráticas.

Sejam três perfis de usuários com base em características ideológicas multidimensionais, evitando o tradicional eixo binário entre esquerda e direita, naveguem por seus endereços preferidos, e façam buscas na Web, conforme mostrado anteriormente. Cada perfil combina posições distintas em temas econômicos, sociais, institucionais e de consumo de mídia, representando configurações realistas do cenário político brasileiro. O Perfil A reflete um usuário com valores econômicos liberais e conservadorismo social, favorável à privatização e à ordem pública, com consumo de mídia alinhado à Jovem Pan e influenciadores conservadores. O Perfil B representa um ativista progressista, com pautas ligadas à justiça social, direitos civis e proteção ambiental, que consome mídias independentes como Brasil de Fato e Nexo Jornal. Já o Perfil C simula um eleitor ambivalente e desconfiado, propenso a narrativas populistas e anti-establishment, que alterna entre fontes tradicionais e sensacionalistas, incluindo vídeos no YouTube e encaminhamentos no WhatsApp.

Com base nesses perfis, diferentes tipos de desinformação poderiam ser direcionados para explorar suas vulnerabilidades específicas. Para o Perfil A, uma notícia falsa poderia afirmar que “o governo quer proibir pais de educarem seus filhos em casa”, apelando ao conservadorismo moral. Para o Perfil B, um conteúdo desinformativo poderia alegar que “o novo plano econômico retira verbas das universidades e prejudica minorias”, ativando pautas de justiça social com distorções. Já para o Perfil C, mais suscetível a teorias conspiratórias, mensagens como “o STF quer fechar o Congresso” ou “a vacina foi criada para implantar chips” seriam típicas de campanhas baseadas em medo e desconfiança. Esses exemplos ilustram como a personalização de conteúdo, quando combinada com perfis comportamentais, pode ser explorada para direcionar narrativas falsas com alta eficácia. Quando a personalização algorítmica se transforma em manipulação comportamental, a democracia deixa de ser um espaço de escolha livre e se torna um experimento invisível de controle social.

5. Conclusões e Trabalhos Futuros

Neste estudo, exploramos as nuances do capitalismo de vigilância, destacando como as técnicas de rastreamento e direcionamento de dados, inicialmente desenvolvidas para fins comerciais, permeiam a experiência online dos usuários. Através da análise empírica do tráfego de rede, evidenciamos a extensão da coleta de dados por terceiros, revelando um ecossistema complexo onde informações pessoais são continuamente extraídas, compartilhadas e utilizadas para influenciar o comportamento dos usuários. Além disso, discutimos a preocupante possibilidade da aplicação dessas técnicas para a disseminação de notícias falsas e desinformação, representando uma ameaça direta à integridade democrática.

Trabalhos futuros devem se concentrar em aprofundar a investigação empírica do uso de técnicas de capitalismo de vigilância na disseminação de desinformação. Propomos utilizar técnicas de captura de tráfego de rede para mapear o fluxo de dados em campanhas de desinformação, identificando os atores envolvidos e as técnicas de rastreamento utilizadas, e analisar como os algoritmos de plataformas de mídia social amplificam a disseminação de notícias falsas, investigando o papel dos dados de perfil do usuário na personalização da desinformação. Desenvolver metodologias para rastrear a origem de notícias falsas, identificando os sites e serviços que as disseminam inicialmente e como elas se propagam pela web, e investigar a infraestrutura de rastreamento utilizada por agentes mal-intencionados.

6. References

References

(2024). How to inspect network traffic using mitmproxy. `https://lucaslegname.github.io/mitmproxy/2020/04/10/mitmproxy.html`. Accessed: 2024-06-22.

Andrew, J. and Baker, M. (2021). The general data protection regulation in the age of surveillance capitalism. *Journal of Business Ethics*, 168:565–578.

Seabra, A. (2025). Github repository. `https://github.com/antonyseabramedeiros/surveillance_capitalism`. Accessed: 2025-05-05.

Wu, Y., Bice, S., Edwards, W. K., and Das, S. (2023). The slow violence of surveillance capitalism: How online behavioral advertising harms people. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency*, pages 1826–1837.

Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier.

Zuboff, S. (2022). Surveillance capitalism or democracy? the death match of institutional orders and the politics of knowledge in our information civilization. *Organization Theory*, 3(3):26317877221129290.

Zuboff, S. (2023). The age of surveillance capitalism. In *Social theory re-wired*, pages 203–213. Routledge.