# Atingidos Pelas Redes Sociais


ORGANIZADORES

R. Marie Santini
Débora Salles
Bruno Mattos
Nicole Sanchotene
Luciane Belin


Os impactos da indústria da desinformação nos consumidores brasileiros


Editora Sulina

MINISTÉRIO DA JUSTIÇA E SEGURANÇA PÚBLICA

BRASIL

FDD
Fundo de Defesa de Direitos Difusos

PRESIDENTE DA REPÚBLICA:
Luiz Inácio Lula da Silva

VICE-PRESIDENTE DA REPÚBLICA:
Geraldo Alckmin

MINISTRO DA JUSTIÇA E SEGURANÇA PÚBLICA:
Ricardo Lewandowski

SECRETÁRIO NACIONAL DO CONSUMIDOR
Wadih Damous

DIRETOR DO FUNDO DE DIREITOS DIFUSOS
Vitor Guimarães

UFRJ
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

REITOR
Roberto de Andrade Medronho

VICE-REITORA
Cassia Curan Turci

net lab
Laboratório de Estudos de Internet e Redes Sociais

DIREÇÃO
R. Marie Santini

APOIO EXECUTIVO
Alda Rosana Almeida
Flávia Cherullo
Cecília Fortes

APOIO TÉCNICO
Amanda Borges da Fonseca - Bernardo Yoneshigue - Bianca Melo - Felipe Maia - Julia Dias - Marina Loureiro - Rafael Tadeu dos Santos - Rafaela Campos - Renata Seade Bastos - Stéphanie Gomes - Thamyres Magalhães - Vitor do Carmo

# ATINGIDOS PELAS REDES SOCIAIS

## Os impactos da indústria da desinformação nos consumidores brasileiros


AUTORES

Alékis Moreira
Arthur Mendes
Bernardo Dias
Bruno Mattos
Carlos Eduardo Barros
Danielle Pinho
Daphane Silva
Débora Salles
Felipe Grael
Fernando Ferreira
João Gabriel Haddad
Lucas Murakami
Luciane Belin
Marcela Canavarro
Marcio Borges
Matheus Gomes
Nicole Sanchotene
Priscila Medeiros
R. Marie Santini
Thiago Ciodaro
Vinicius B. Scortegagna



*Editora Sulina*

Organizadores
R. Marie Santini - Débora Salles - Bruno Mattos
Nicole Sanchotene - Luciane Belin

Autores
Alékis Moreira - Arthur Mendes - Bernardo Dias
Bruno Mattos - Carlos Eduardo Barros - Danielle Pinho
Daphane Silva - Débora Salles - Felipe Grael
Fernando Ferreira - João Gabriel Haddad - Lucas Murakami
Luciane Belin - Marcela Canavarro - Marcio Borges
Matheus Gomes - Nicole Sanchotene - Priscila Medeiros
R. Marie Santini - Thiago Ciodaro - Vinícius Branco Scortegagna

Capa: Felipe Loureiro
Projeto gráfico e editoração: Cristiano Marques
Revisão: Erick Dau e Adriano Belisário
Editor: Luis Antonio Paim Gomes

Dados Internacionais de Catalogação na Publicação (CIP)
Bibliotecária Responsável: Denise Mari de Andrade Souza – CRB 10/960

A872

Atingidos pelas redes sociais: os impactos da indústria da desinformação nos consumidores brasileiros / organizado por R. Marie Santini [et al.]. -- Porto Alegre: Sulina, 2025.

ISBN: 978-65-5759-219-9 [Livro digital]

DOI: 10.29327/5566995

1. Cultura Digital. 2. Redes Sociais – Desinformação. 3. Fake News – Impacto – Economia. 4. Economia – Fake News. 5. Estatística – Economia – Fake News. 6 Jornalismo – Fake News – Economia.

CDU: 070
316.422
CDD: 070
306
318



www.editorasulina.com.br
e-mail: sulina@editorasulina.com.br

Maio/2025

# Sumário

# Introdução

Com a adoção de tecnologias de vigilância digital e microssegmentação, as plataformas de redes sociais se consolidaram como os principais espaços para anúncios publicitários no ecossistema contemporâneo de mídia, permitindo o direcionamento preciso de mensagens a públicos específicos. Esse modelo de negócios, amplamente explorado por *big tech* como Google e Meta, baseia-se na coleta massiva de dados dos usuários e no uso de ferramentas de inteligência artificial para prever comportamentos de consumo (Zuboff, 2019). Embora argumentos favoráveis destaquem os benefícios desse tipo de estratégia de financiamento de mídia – como a viabilidade econômica e logística da publicidade nas redes sociais para pequenos negócios ou empreendedores – esse ecossistema tem sido instrumentalizado para operações de influência política, manipulação da opinião pública e fraudes financeiras (Andrejevic; O'Neill; Mahoney, 2025; Zeng *et al*., 2020). Essas práticas se beneficiam da falta de transparência, da ausência de regulamentação específica e da inexistência de fiscalização (Crain; Nadler, 2019). Por conta desses problemas regulatórios e de governança, há descumprimento recorrentemente das normas brasileiras e das políticas de uso criadas pelas plataformas. Ou seja, essas empresas têm dificuldade de cumprir as próprias regras. Esse cenário cria obstáculos significativos para o monitoramento e a responsabilização dos diversos atores envolvidos em práticas de publicidade irregulares nos ambientes online.

O modelo de negócios das plataformas é estruturado para maximizar a atenção dos usuários, consolidando-se como um modelo lucrativo devido à transformação dos dados em *commodities* (Aaltonen; Alaimo; Kallinikos, 2021). Trata-se de um conjunto de processos de gerenciamento, análise e interpretação de dados massivos sobre a audiência, que articula indústrias e organizações de mídia e de tecnologia. Esse modelo se baseia no uso de estratégias de microssegmentação para direcionar conteúdos a públicos altamente específicos, na priorização de conteúdos que geram maior engajamento e do uso de sistemas de recomendação algorítmica construídos a partir de dados comportamentais e sociopsicológicos dos usuários (Gehl; Lawson, 2022). Embora isso seja promovido como um modo de otimizar as campanhas publicitárias online, na prática, esse sistema tem facilitado a disseminação de conteúdos enganosos impulsionados por anúncios pagos que exploram as vulnerabilidades dos usuários, sem que haja efetiva fiscalização e responsabilização pelas irregularidades praticadas.

Diferentemente dos anúncios exibidos em mídias tradicionais, que são públicos, os anúncios digitais são distribuídos de forma opaca e indiscriminada. Ou seja, a falta de transparência dificulta o monitoramento por parte de pesquisadores, reguladores, até dos próprios anunciantes e concorrentes (Jamison *et al.*, 2020). Além disso, a falta de verificação mínima de legitimidade dos anunciantes e marcas causa enormes estragos econômicos e sociais. Essa estrutura permite que anúncios fraudulentos e enganosos sejam disseminados livremente, sem que as plataformas assumam qualquer tipo de responsabilidade. Como resultado, práticas como golpes financeiros, promoção de desinformação e incentivo ao comércio ilegal tornaram-se comuns no ambiente digital, gerando impactos negativos significativos para os consumidores e a sociedade como um todo (Andrejevic; O'Neill; Mahoney, 2025).

A publicidade nativa, um dos formatos mais populares na publicidade digital contemporânea, agrava esse cenário ao integrar anúncios diretamente ao conteúdo editorial das plataformas, tornando-os visualmente indistinguíveis de notícias e postagens orgânicas (Wojdynski;

Golan, 2016; Campbell; Grimm, 2018). Embora essa abordagem seja defendida como uma estratégia eficaz de persuasão e engajamento (IAB, 2019), o fato de os anúncios adotarem o mesmo formato visual do conteúdo editorial e gerado pelos usuários dificulta que os usuários façam a distinção entre informação e publicidade, levando-os a interagir inadvertidamente com conteúdos potencialmente enganosos (Zeng *et al.*, 2020). Pesquisas demonstram que anúncios nativos frequentemente utilizam técnicas como *clickbait* e imagens sensacionalistas para capturar a atenção dos usuários e induzi-los ao clique, muitas vezes direcionando-os a páginas de baixa qualidade, repletas de desinformação ou até mesmo esquemas fraudulentos (Santos Junior, 2024).

Além disso, a combinação da veiculação programática com a publicidade nativa criou um ambiente extremamente favorável para a proliferação de campanhas de desinformação e fraudes financeiras (Ali *et al.*, 2023; Sadeghpour; Vlajic, 2021). Como os anunciantes pagam apenas pelos cliques gerados, há um incentivo para a promoção de estratégias enganosas que maximizem esse engajamento, independentemente da veracidade ou do impacto social do conteúdo promovido (Turow, 2012; McStay, 2017). Essas práticas colocam em risco a segurança dos consumidores e prejudicam anunciantes legítimos, que acabam competindo em um mercado distorcido e repleto de conteúdos abusivos (O'Neil, 2020).

Esse novo cenário, em que as plataformas estabelecem suas próprias políticas e operam com pouca ou nenhuma fiscalização externa, facilita a disseminação de anúncios fraudulentos e conteúdos desinformativos, que passaram a ser usados como ferramentas eficazes de manipulação e exploração econômica. Cada vez mais, crimes tradicionalmente cometidos no ambiente offline estão migrando para o digital, como fraudes financeiras e estelionato. Isso ocorre, em grande parte, porque o que é considerado cumplicidade ou corresponsabilidade em crimes fora da internet ainda não tem uma tradução clara para o ambiente online – seja pela ausência de normas específicas (Meireles; Pasit-

to, 2024), seja pela dificuldade das instituições em interpretar e aplicar as normas existentes a esses novos contextos.

De acordo com uma pesquisa de 2024 do DataFolha, fraudes envolvendo Pix e boletos são as principais fontes de receita de crimes digitais no Brasil, gerando um prejuízo anual estimado em R$ 25,5 bilhões aos consumidores – um valor superior ao causado por fraudes com cartões de crédito e roubos de celulares (Kruse, 2024). Uma característica central desse cenário é que grande parte dessas fraudes têm origem nas redes sociais: segundo o estudo Golpes com Pix (Silverguard & SOS Golpes, 2024), 79% das transações financeiras baseadas em fraudes e golpes denunciados entre janeiro e junho de 2024 começaram nas plataformas da Meta – sendo 39% no WhatsApp. Além disso, a importância das plataformas de redes sociais como meio para aplicação de fraudes é reforçada por pesquisa da Global Anti-Scam Alliance, ScamAdviser e Whoscall, que identificou que quase 50% dos 1.322 brasileiros entrevistados relataram ter sido abordados por golpistas através de anúncios digitais (Rogers, 2024). Esses dados evidenciam como o ambiente de publicidade nas redes sociais se tornou terreno fértil para a atuação de criminosos, que encontraram um ambiente propício para captar vítimas ideais e aplicar golpes sem grandes obstáculos. A dificuldade em localizar os criminosos é agravada pela subnotificação, já que muitas vítimas não registram boletins de ocorrência – um requisito para a recuperação de valores por parte das instituições financeiras (Kruse, 2024).

Além disso, dados da *Federal Trade Commission* (FTC), dos Estados Unidos, mostram que, em 2022, os consumidores norte-americanos relataram perdas superiores a 1,2 bilhão de dólares devido a fraudes iniciadas em mídias sociais, consolidando as plataformas de redes sociais como o principal canal de contato entre estelionatários e suas vítimas (Federal Trade Commission, 2023b). Já no Reino Unido, uma pesquisa do órgão comercial *UK Finance* revelou que, em 2022, quase 3 milhões de cidadãos perderam 1,2 bilhões de libras em golpes financeiros, sendo que 80% dessas fraudes tiveram origem em plataformas online (Barret, 2023). Esses golpes ocorrem principalmente por meio de mídias sociais,

*marketplaces* digitais, aplicativos de relacionamento e serviços de mensagens, reforçando a necessidade de que as plataformas assumam maior responsabilidade pela publicidade e pelas interações que circulam em seus ambientes, possibilitadas e facilitadas por sua própria arquitetura.

Entretanto, é importante destacar que as próprias plataformas de redes sociais possuem os meios para identificar irregularidades e práticas comerciais suspeitas como os casos de golpes e fraudes financeiras, já que elas coletam informações sobre os anunciantes e são responsáveis pelo processamento dos pagamentos relacionados a veiculação de anúncios (Ali *et al.*, 2019). Assim, considerando a constante utilização de mecanismos de compra de impulsionamento de conteúdo dentro das plataformas para promoção de golpes, seria viável aumentar a auditabilidade da publicidade em redes sociais por meio do cruzamento de informações de pagamento, dados dos anunciantes e registros de campanhas publicitárias – caso houvesse comprometimento das plataformas com a aplicação efetiva das leis, de políticas de transparência e da proteção dos usuários.

A popularização de ferramentas de Inteligência Artificial tem ampliado estes problemas, tornando as estratégias de engenharia social cada vez mais sofisticadas e difíceis de detectar. Entre as táticas mais preocupantes está a falsificação de identidade (*impersonation*), que consiste na simulação de figuras públicas, empresas ou instituições para enganar usuários e conferir legitimidade a golpes financeiros (Algarni *et al.*, 2014; Baltezarević; Baltezarević, 2024; Goel, 2021). O avanço dos *deepfakes* e de outras tecnologias de IA generativa permite a criação de vídeos e áudios hiper-realistas que misturam conteúdo sintético com material autêntico, tornando os golpes mais convincentes e reduzindo a percepção de risco entre as vítimas (de Rancourt-Raymond; Smaili, 2023).

À medida que modelos de IA generativa se tornam mais acessíveis e potentes, criminosos conseguem automatizar a criação e veiculação de anúncios fraudulentos em larga escala, diante da falta de controle efetivo sobre esse tipo de conteúdo nos sistemas de monitoramento e mo-

deração das plataformas (Dhaliwal, 2025; Volkova, 2025). No Brasil, investigações recentes revelam que fraudadores têm utilizado *deepfakes* para se passar por políticos eleitos (NetLab UFRJ, 2024) e personalidades da mídia (Rudnitzki, 2024), promovendo esquemas financeiros enganosos por meio de anúncios pagos. Essas campanhas resultam em danos materiais e causam impactos emocionais e financeiros às vítimas dos golpes, além de prejuízos reputacionais para pessoas, instituições e empresas que têm suas imagens usadas indevidamente, corroendo a confiança na informação digital e ampliando o ambiente de desinformação. Diante desse cenário, a ausência de regulamentação específica para a veiculação de anúncios com IA e a inexistência, limitação ou baixa prioridade dada ao desenvolvimento de ferramentas eficazes para identificar conteúdos gerados por essas tecnologias agrava a vulnerabilidade dos consumidores e reforça a necessidade de políticas públicas que exijam maior transparência, auditoria e responsabilização das plataformas (Causin, 2025; Fukuzawa, 2024; Sensity, 2024).

Para evitar que um ecossistema digital nocivo continue se expandindo e consolidando incentivos financeiros para golpes e manipulação online, é essencial repensar a governança das plataformas de redes sociais. Além de aprimoramentos regulatórios, é necessário questionar práticas de mercado e modelos de negócio que priorizam o lucro em detrimento da ética, da proteção dos consumidores e da transparência. Diferentemente do que ocorre na publicidade veiculada nos meios de comunicação tradicionais – que é exibida de modo uniforme a toda audiência e que, portanto, pode ser fiscalizável –, a publicidade nas plataformas de redes sociais é distribuída por algoritmos que operam de maneira opaca, ou seja, não há transparência sobre o conteúdo dos anúncios tampouco sobre seus critérios de distribuição (Carah *et al.*, 2024; Jamison *et al.*, 2020). A dificuldade de fiscalização externa não se limita à natureza personalizada e volátil desses anúncios: ela se agrava pelo fato de que essas plataformas operam sob um regime de desresponsabilização, estruturado como uma caixa-preta. Orientadas por seus próprios termos de uso, elas também negligenciam as normas e leis já

existentes nos diferentes países onde atuam, ignorando indícios claros de irregularidades e desconsiderando evidências de crimes e riscos sistêmicos que poderiam apontar caminhos para mitigar danos individuais e coletivos.

Nesse cenário, a pesquisa acadêmica desempenha um papel central na produção de conhecimento sobre esses fenômenos e na orientação de políticas públicas baseadas em evidências. O acesso a dados digitais permite identificar e descrever padrões de desinformação, fraude e manipulação online, oferecendo subsídios para a fiscalização, a regulação e a atuação responsável das plataformas de redes sociais. Ou seja, a compreensão desse mercado e o combate às práticas comerciais predatórias no ambiente digital exigem investigações detalhadas sobre os mecanismos utilizados por fraudadores e pelas plataformas, os públicos mais impactados, as ferramentas tecnológicas exploradas – particularmente aquelas baseadas em IA. Em relação à publicidade fraudulenta, o problema é menos a falta de leis e mais as falhas na aplicação das normas existentes que garantem impunidade daqueles que produzem, compram, vendem e veiculam anúncios abusivos e enganosos nas redes sociais.

Porém, só é possível fazer pesquisa sobre padrões de desinformação, fraude e manipulação online se os pesquisadores tiverem acesso a dados. Nesse sentido, a transparência das plataformas torna-se um dos eixos fundamentais da governança digital. Como apontam Urman e Makhortykh (2023), a transparência consiste na disponibilização de informações internas de interesse público, permitindo que governos, pesquisadores e sociedade civil avaliem como as plataformas operam e quais políticas aplicam. Indicadores de transparência são essenciais para mapear e avaliar as ferramentas atualmente oferecidas, identificando dados inconsistentes, incompletos ou de baixa qualidade, que dificultam a análise sistemática de riscos e a responsabilização de plataformas por práticas abusivas. Diversas organizações internacionais, como a OCDE (2024) e a UNESCO (2023), já apontaram a necessidade global de ampliar a transparência sobre publicidade digital, e os resultados de

pesquisa apresentados neste livro buscam inserir o Brasil nesse debate, considerando as especificidades do mercado nacional e suas consequências para a população.

Apesar da crescente demanda por transparência, pesquisadores frequentemente enfrentam obstáculos para acessar informações relevantes, consistentes e representativas das atividades que ocorrem dentro das plataformas de redes sociais. Dados de alta qualidade são essenciais para assegurar a reprodutibilidade e a confiabilidade dos estudos, além de possibilitar generalizações robustas sobre os objetos analisados (Srivastava; Mishra, 2021). Segundo a Electronic Code Management Association (ECCMA), dados de baixa qualidade aumentam os custos de compliance e são uma das principais causas de falhas na transparência (ECCMA, [*S.d.*]). No contexto da pesquisa social online, a qualidade dos dados não apenas aprimora a análise acadêmica, mas também desempenha um papel fundamental na formulação e no aprimoramento de medidas de transparência das plataformas. Além disso, permite avaliar se as iniciativas de transparência anunciadas por essas empresas estão sendo efetivamente implementadas, garantindo maior controle sobre suas práticas.

Em todo o mundo, diferentes países vêm adotando iniciativas para aumentar a transparência e a responsabilização das plataformas, reconhecendo que as *big tech* têm permitido que práticas fraudulentas se proliferem sem controle adequado em suas plataformas. Apesar das crescentes evidências sobre os riscos da publicidade predatória, as propostas regulatórias enfrentam forte resistência das grandes empresas de tecnologia (Le Pochat *et al.*, 2022; Napoli; Caplan, 2017). No Brasil, essa oposição se manifestou de maneira explícita quando as próprias plataformas veicularam anúncios contra o PL 2630, – projeto de lei que visava a regulamentação – interferindo diretamente no debate público e no processo legislativo ao burlar seus próprios termos de uso (NetLab UFRJ, 2023). Além disso, há indícios de que as plataformas restringem o acesso a dados críticos que poderiam expor sua participação na amplificação de desinformação e em crises de imagem (Ben-David, 2020).

Esse "teatro da transparência", no qual as empresas alegam compromisso com a abertura de dados, mas impõem barreiras à fiscalização independente, reforça a necessidade de indicadores objetivos para avaliar a real transparência dessas plataformas e detectar possíveis violações de leis locais e das próprias políticas empresariais.

Diversos países têm identificado populações mais vulneráveis à exposição a fraudes e publicidade abusiva, como crianças, idosos, aposentados e pessoas endividadas. No caso das crianças, estudos demonstram que menores de oito anos frequentemente interpretam anúncios como informações neutras sobre produtos, em vez de reconhecê-los como estratégias de venda – o que os torna especialmente suscetíveis à publicidade predatória (Global Action Plan, 2020). Nos Estados Unidos, a FTC tem desempenhado um papel fundamental no processo de responsabilizar mais as plataformas, impondo multas e exigindo maior prestação de contas das *big tech*. Um dos casos mais emblemáticos envolveu a Google e o YouTube, que foram multados em 170 milhões de dólares pelo uso indevido de dados pessoais de crianças sem o consentimento dos pais. Apesar da penalização, investigações posteriores revelaram que a empresa continuava exibindo anúncios mesmo em conteúdos marcados como "*made for kids*", levantando questionamentos sobre a eficácia das medidas de fiscalização (Adalytics, [*S.d.*]; Federal Trade Commission, 2019). Esse caso evidencia duas falhas graves: por um lado, há um dano direto aos consumidores, com a coleta indevida de dados para segmentação publicitária; por outro, há um prejuízo para os anunciantes, que são enganados quanto à entrega real de seus anúncios e sua eficácia dentro das diretrizes prometidas pelas plataformas (Khan; Bedoya; Slaughter, 2023).

Embora o Brasil tenha um dos debates públicos mais avançados sobre transparência e responsabilização das *big tech* fora da Europa, o país ainda carece de um marco regulatório específico para lidar com esses crimes nos ambientes online. O Marco Civil da Internet (Lei nº 12.965/2014) estabelece diretrizes importantes para a governança digital, mas não prevê regulamentações claras para combater fraudes di-

gitais e publicidade enganosa online. Além disso, o artigo 19 da lei confere imunidade às plataformas em relação ao conteúdo gerado por terceiros, criando uma brecha argumentativa que tem sido usada para imunidade dessas empresas pelos danos causados por conteúdos pagos, ou seja, publicidade. À medida que iniciativas internacionais avançam no sentido de implementar abordagens mais incisivas para responsabilizar as plataformas de redes sociais, cresce a urgência de que o Brasil acompanhe esse movimento e fortaleça suas respostas institucionais para mitigar os impactos negativos da publicidade digital e das fraudes online. É preciso criar um modelo de governança mais transparente e responsável em que as plataformas assumam a responsabilidade pelos conteúdos que promovem e que geram lucro aos donos dessas empresas.

Em outros países, está cada vez mais claro que o argumento de que o que vale no mundo *offline* também deve valer no mundo online, garantindo que consumidores tenham o mesmo nível de proteção em qualquer ambiente (European Commission, [*S.d.*]). Ainda assim, a maioria dos países enfrenta grandes dificuldades para implementar de forma efetiva suas próprias leis no ambiente digital, o que evidencia que o problema está menos na criação de normas e mais nos desafios políticos e econômicos envolvidos na sua aplicação. Por exemplo, casos recentes mostram como as *big tech* seguem violando legislações recentes, mesmo sob o *Digital Services Act* (DSA) e o *Digital Markets Ac*t (DMA), que regulam especificamente as práticas online (Satariano, 2025, Jahangir, 2025). Isso evidencia a dificuldade de fazer com que essas empresas cumpram efetivamente as regras estabelecidas, mesmo quando existem instrumentos legais específicos para o ambiente digital. Assim, o avanço em direção a um ecossistema digital mais seguro e confiável depende, sobretudo, da capacidade de garantir a aplicação efetiva das normas já existentes e do fortalecimento dos mecanismos de fiscalização, com o envolvimento ativo de pesquisadores, governos e da sociedade civil.

A epidemia de fraudes que assola as redes sociais no Brasil é um problema de grandes proporções, mas sua real dimensão permanece

desconhecida devido à falta de transparência das plataformas de redes sociais. Essa opacidade não apenas impede uma avaliação precisa do cenário, mas também expõe milhões de brasileiros a riscos e prejuízos financeiros. Sem mecanismos eficazes de fiscalização e de garantia do cumprimento de leis e políticas de proteção aos usuários, as grandes empresas de tecnologia continuam lucrando com anúncios fraudulentos, enquanto os usuários permanecem vulneráveis a esquemas cada vez mais sofisticados.

Este livro é um dos resultados das pesquisas desenvolvidas no âmbito do *Observatório da Indústria da Desinformação e seu impacto nas relações de consumo no Brasil*, projeto em parceria entre o NetLab UFRJ e a Secretaria Nacional do Consumidor do Ministério da Justiça e Secretaria Pública (Senacon/MJSP). O principal objetivo do Observatório é prover insumos que sirvam de base para políticas públicas que protejam os consumidores das operações de desinformação e influência – caracterizadas pela disseminação de informações falsas, distorcidas ou enganosas, com o objetivo de manipular percepções, gerar confusão e direcionar opiniões e comportamentos em benefício de interesses políticos ou comerciais (Santini *et al.*, 2022; Briant; Bakir, 2024). Na medida em que essas operações vêm se desenvolvendo e ganhando espaço nas plataformas de redes sociais, o Observatório visa produzir análises sobre sua infraestrutura, a economia política e as estratégias de manipulação das relações de consumo e da opinião pública.

Aqui estão reunidos estudos que analisam criticamente a atuação das plataformas, os impactos da falta de mecanismos de fiscalização da publicidade digital e as barreiras impostas ao acesso a dados de interesse público. Ao longo do projeto, esses estudos foram publicados como relatórios e amplamente divulgados na imprensa com o objetivo de incidir no debate público. Com base em metodologias empíricas e análises aprofundadas, os capítulos traçam um panorama detalhado sobre as formas pelas quais as plataformas digitais são organizadas, reguladas e responsabilizadas no Brasil — o que chamamos aqui de governança das plataformas. Ao examinar as limitações dos atuais mecanismos de

transparência, discutimos caminhos para torná-los mais eficazes e compatíveis com valores democráticos.

O capítulo 2, intitulado Índice de Transparência de Dados das Plataformas de Redes Sociais parte da seguinte pergunta norteadora: quais são as condições para que a academia possa estudar e o poder público fiscalizar o que acontece nessas plataformas? A partir desse questionamento, apresentamos uma avaliação detalhada sobre o grau de qualidade e acessibilidade dos dados fornecidos por plataformas como YouTube, Facebook, Instagram, X/Twitter, Telegram, TikTok, Kwai e WhatsApp. A metodologia do índice apresentado no capítulo avalia critérios como disponibilidade de APIs oficiais, clareza na documentação, granularidade dos dados e acesso a relatórios de transparência. O estudo revela que nenhuma dessas plataformas atinge um nível satisfatório de transparência, dificultando a análise independente de fluxos informacionais e da governança algorítmica. Isso significa que pesquisadores, a sociedade civil e o poder público têm limitações críticas no acesso a informações sobre como as plataformas moderam e recomendam conteúdo, como realizam a curadoria de informações e de que modo decidem o que é promovido ou removido nas plataformas. Além disso, o capítulo demonstra que algumas empresas aplicam políticas de transparência desiguais entre as diferentes regiões do mundo, adotando práticas de maior transparência em países do Norte Global do que em países como o Brasil.

A principal conclusão do estudo é que a falta de padronização e a ausência de meios para acesso a dados resultam em um cenário de "apagão de dados", que prejudica pesquisadores, jornalistas e formuladores de políticas públicas. A pesquisa reforça que a transparência das plataformas não é uma questão meramente técnica, mas uma decisão política das empresas, que têm o poder de limitar ou expandir o acesso a informações essenciais para a fiscalização de suas práticas e recortar o que pode ou não ser visto por observadores externos a essas empresas. O capítulo também propõe recomendações para melhorar a governança da transparência de dados, incluindo a adoção de padrões mínimos de

acesso a dados de interesse público e a garantia de mecanismos de fiscalização e escrutínio independente sobre as práticas das plataformas.

O capítulo 3 apresenta o Índice de Transparência da Publicidade nas Plataformas de Redes Sociais, um estudo que avalia os níveis de transparência das plataformas no que diz respeito a dados sobre veiculação de anúncios, sejam eles políticos ou comerciais. No entanto, há uma notável falta de transparência sobre como anúncios são distribuídos, direcionados e moderados. A pesquisa revela que, apesar das promessas de maior transparência, empresas como Meta e Google fornecem informações limitadas sobre anúncios impulsionados, dificultando a identificação de práticas abusivas.

O estudo também analisa como as plataformas classificam e lidam com anúncios pagos, distinguindo-os em categorias como "políticos" ou "comerciais". A preocupação com conteúdos políticos nas redes sociais ganhou destaque internacional após o escândalo da Cambridge Analytica, quando dados de milhões de usuários do Facebook foram usados, sem consentimento, para direcionar campanhas e manipular o comportamento eleitoral (Confessore, 2018). Diante da pressão pública, as plataformas passaram a adotar iniciativas de autorregulação, anunciando medidas para aumentar a transparência e permitir algum grau de fiscalização. Em 2018, a Meta foi uma das primeiras *big tech* a criar um repositório online de anúncios, de modo a evitar acusações de manipulação eleitoral em um contexto de uso indevido de dados de usuários em campanhas de microsegmentação nocivas. No entanto, essa iniciativa acumulou críticas de pesquisadores devido às limitações de acesso aos dados (Bossetta, 2020), sendo frequentemente interpretada como um "teatro" (Bouko; van Ostaeyen; Voué, 2021), mais preocupado em sinalizar compromisso público do que em garantir abertura real à fiscalização.

Este capítulo reforça essas críticas ao demonstrar como as plataformas adotam critérios arbitrários e inconsistentes para classificar anúncios como políticos ou comerciais, o que impacta diretamente na fiscalização e no controle social da publicidade digital. Casos recentes

em diferentes países – incluindo o escândalo da Cambridge Analytica e outros episódios envolvendo o uso das plataformas e de anúncios para manipulação política (Albert, 2024; Ellis-Petersen, 2024) – evidenciam como conteúdos microssegmentados podem ser utilizados para interferir em processos democráticos. Além disso, a separação rígida entre anúncios políticos e anúncios comerciais, combinada com a adoção de critérios de transparência distintos para cada categoria, compromete a fiscalização efetiva da publicidade digital, resultando em um sistema no qual conteúdos de teor político e de interesse público muitas vezes escapam de monitoramento mais rigoroso. A pesquisa aponta que essa falta de transparência beneficia anunciantes mal-intencionados, que exploram um ambiente pouco fiscalizável para disseminar desinformação ou aplicar golpes financeiros. O capítulo conclui que a ausência de um padrão claro para a transparência publicitária compromete a segurança dos consumidores e a integridade do ambiente informacional, sugerindo recomendações para ampliar a transparência das plataformas.

O capítulo 4, *Anúncios falsos, seus mecanismos e potenciais danos à sociedade,* investiga como golpistas utilizam publicidade paga para enganar o público e manipular a opinião pública, analisando casos concretos em que anúncios fraudulentos exploraram a credibilidade de figuras públicas e instituições. Entre os exemplos analisados, destacam-se o uso de *deepfakes* para simular falas de políticos e líderes religiosos, além da utilização indevida de logotipos e imagens de entidades governamentais para promover falsas ofertas financeiras. A pesquisa evidencia que as plataformas de redes sociais, seja por ausência ou ineficácia dos mecanismos de verificação, permitem a circulação de conteúdos publicitários irregulares e a atuação de anunciantes ilegítimos, enquanto restringem o acesso a informações sobre essas campanhas. Mesmo diante de denúncias públicas e ações de fiscalização, muitas dessas peças publicitárias permanecem ativas, demonstrando a ineficiência dos mecanismos de controle das empresas de tecnologia. O capítulo também discute como a distribuição de anúncios por meio de microssegmentação torna os consumidores mais vulneráveis a fraudes, expondo populações espe-

cíficas a práticas predatórias. Por fim, o estudo destaca a necessidade de ações para mitigar esses impactos, como a ampliação da transparência na veiculação de anúncios e a imposição de responsabilidades mais rígidas às plataformas que lucram com publicidade digital.

Já o capítulo 5, *Ecossistema de desinformação e fraudes com marcas de programas governamentais*, investiga como criminosos têm se apropriado de iniciativas públicas, explorando a credibilidade de programas governamentais como Desenrola Brasil e Voa Brasil, para enganar cidadãos vulneráveis e aplicar fraudes disfarçadas de ofertas legítimas. A pesquisa analisa a estrutura desses golpes, identificando como perfis falsos, páginas fraudulentas e anúncios impulsionados simulam ser representantes oficiais do governo para induzir usuários a fornecerem seus dados pessoais ou realizarem pagamentos indevidos.

O estudo demonstra que as plataformas de redes sociais desempenham um papel central na disseminação dessas fraudes, permitindo que anúncios enganosos continuem circulando mesmo após medidas cautelares para sua remoção. Além disso, o capítulo analisa a falha dos sistemas de autenticação de anunciantes, que não exigem comprovação de identidade para a veiculação de publicidade paga. A persistência dessas práticas evidencia a necessidade de maior transparência e responsabilização das plataformas, bem como de políticas públicas para combater golpes financeiros no ambiente digital. A pesquisa discute os potenciais impactos financeiros e reputacionais sobre os usuários e instituições públicas e privadas, propondo recomendações para o fortalecimento da fiscalização da publicidade digital, visando proteger consumidores e garantir um ecossistema mais seguro e confiável.

O Capítulo 6, *Big tech, direito do consumidor e interesse público*, discute o papel das plataformas de redes sociais na mediação do espaço informacional e as implicações de suas políticas para os direitos dos consumidores e o interesse público. Os estudos apresentados nos capítulos anteriores demonstram que as estratégias de governança adotadas pelas *big tech* têm sido ineficazes no combate a conteúdos irregulares, como campanhas de desinformação e práticas comerciais enganosas.

Neste cenário, a predominância de um modelo autorregulatório e ausência de meios para fiscalização externa efetiva permitem que essas empresas atuem priorizando seus interesses comerciais em detrimento da integridade da informação, evidenciando a necessidade de um maior envolvimento do Estado na governança digital.

Partindo do conceito de interesse público, este capítulo amplia o debate sobre a governança das plataformas de redes sociais para além das abordagens tradicionalmente focadas nos meios de comunicação convencionais. Argumentamos que as *big tech* não atuam meramente como intermediárias, mas exercem influência direta sobre fluxos de informação, comportamentos, práticas de consumo e opiniões políticas, moldando ativamente o ambiente digital de maneira opaca e desalinhada com o bem-estar coletivo. Ao priorizarem seus interesses comerciais em detrimento da segurança dos usuários e da transparência, essas empresas frequentemente violam normas locais e desafiam a soberania dos Estados. Diante desse cenário, o capítulo examina o papel das políticas públicas na governança das plataformas, apresenta esforços regulatórios recentes em outros países e discute a importância de conciliar inovação tecnológica, proteção aos direitos dos consumidores e fortalecimento do interesse público.

## Referências

AALTONEN, A.; ALAIMO, C.; KALLINIKOS, J. The Making of Data Commodities: Data Analytics as an Embedded Process. Journal of Management Information Systems, [S.l.], v. 38, n. 2, p. 401–429, 3 abr. 2021. Disponível em: <https://doi.org/10.1080/07421222.2021.1912928>. Acesso em: 1 ago. 2024.

ADALYTICS. Are YouTube Advertisers Inadvertently Harvesting Data From Millions of Children?. *Adalytics*, [*S.d.*]. Disponível em: https://adalytics.io/blog/are-youtube-ads-coppa-compliant. Acesso em 31 out. 2024.

ALBERT, John. TikTok and the Romanian elections: A stress test for DSA enforcement. *DSA Observatory*, 20 dez. 2024. Disponível em: ht-

tps://dsa-observatory.eu/2024/12/20/tiktok-and-the-romanian-elections/. Acesso em: 26 abr. 2025.

ALGARNI, Abdullah.; XU, Yue; CHAN, Tayzan; TIAN, Yu-Chu. Social Engineering In Social Networking Sites: How Good Becomes Evil. *PACIS 2014 Proceedings*, [*S.l.*], 2014. Disponível em: https://aisel.aisnet.org/pacis2014/271. Acesso em: 8 abr. 2025.

ALI, Muhammad; SAPIEZYNSKI, Piotr; BOGEN, Miranda; KOROLOVA, Aleksandra; MISLOVE, Alan; RIEKE, Aaron. Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes. ACM ON HUMAN-COMPUTER INTERACTION, [*S.l.*], v. 3, n. CSCW, p. 199:1-199:30, 7 nov. 2019. *Anais* [...]. Nova Iorque: Association for Computing Machinery, 2019. Disponível em: https://doi.org/10.1145/3359301. Acesso em: 1 ago. 2024.

ALI, Muhammad; GOETZEN, Angelica; MISLOVE, Alan; REDMILES, Elissa M.; SAPIEZYNSKI, Piotr. Problematic Advertising and its Disparate Exposure on Facebook. In: 32ND USENIX SECURITY SYMPOSIUM (USENIX SECURITY 23), 2023. *Anais* [...], [*S.l.*], 2023. p. 5665–5682. Disponível em: https://www.usenix.org/conference/usenixsecurity23/presentation/ali. Acesso em: 9 abr. 2025.

ANDREJEVIC, Mark; O'NEILL, Christopher; MAHONEY, Isabella. "The scandal that shocked the world": conspirituality and online scam ads. *Journal of Information Technology & Politics*, [*S.l.*], 2025. Disponível em: https://www.tandfonline.com/doi/full/10.1080/19331681.2025.2453920. Acesso em: 9 abr. 2025.

BALTEZAREVIĆ, Radoslav; BALTEZAREVIĆ, Ivana. Students' Attitudes on The Role of Artificial Intelligence (Ai) In Personalized Learning. *International Journal of Cognitive Research in Science, Engineering and Education*, [*S.l.*], v. 10, n. 2, p. 123–145, 2024. Disponível em: https://ijcrsee.com/index.php/ijcrsee/article/view/3006. Acesso em: 8 abr. 2025.

BARRET, C. It's time social media platforms unfriended fraudsters. *Financial Times*, 2023. Disponível em: https://www.ft.com/content/884cd0c4-c8bc-438c-b9c2-ae74448e33ae. Acesso em 20 maio 2023.

BEN-DAVID, Anat. Counter-archiving Facebook. *European Journal of Communication*, [*S.l.*], v. 35, n. 3, p. 249–264, 1 jun. 2020. Disponível em: https://doi.org/10.1177/0267323120922069. Acesso em: 1 ago. 2024.

BOSSETTA, Michael. Scandalous design: how social media platforms' responses to scandal impacts campaigns and elections. *Social Media + Society*, v. 6, n. 2, p. 205630512092477, abr. 2020. Disponível em: https://doi.org/10.1177/2056305120924777. Acesso em: 26 abr. 2025.

BOUKO, C.; VAN OSTAEYEN, P.; VOUÉ, P. Facebook's policies against extremism: Ten years of struggle for more transparency. First Monday, [S.l.], v. 26, n. 9, p. 1–22, 2021. Disponível em: https://firstmonday.org/ojs/index.php/fm/article/view/11705. Acesso em: 1 ago. 2024.

BRIANT, E.; BAKIR, V. *Routledge Handbook of the Influence Industry*. Routledge: Abington; Nova York, 2024.

CAMPBELL, Colin; GRIMM, Pamela E. The Challenges Native Advertising Poses: Exploring Potential Federal Trade Commission Responses and Identifying Research Needs. *Journal of Public Policy & Marketing*, [*S.l.*], v. 38, n. 1, p. 110–123, 1 jan. 2019. Disponível em: https://doi.org/10.1177/0743915618818576. Acesso em: 1 ago. 2024.

CARAH, N.; HAYDEN, L.; BROWN, M.-G.; ANGUS, D.; BROWNBILL, A.; HAWKER, K.; TAN, X. Y.; DOBSON, A.; ROBARDS, B. Observing "tuned" advertising on digital platforms. *Internet Policy Review*, [S.l.], v. 13, n. 2, p. 1–16, 26 jun. 2024. Disponível em: https://policyreview.info/articles/analysis/observing-tuned-advertising-digital-platforms. Acesso em: 1 ago. 2024.

CAUSIN, Juliana. Golpes com Pix vão gerar prejuízo anual de R$ 11 bilhões para bancos e consumidores até 2028. *O Globo*, [*S.l.*], 2025. Disponível em: https://oglobo.globo.com/economia/noticia/2025/01/21/golpes-com-pix-vao-gerar-prejuizo-anual-de-r-11-bilhoes-para-bancos-e-consumidores-ate-2028.ghtml. Acesso em: 8 abr. 2025.

CONFESSORE, Nicholas. Cambridge Analytica and Facebook: The Scandal and the Fallout So Far. *The New York Times*, 04 abr. 2018. Dis-

ponível em: https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html. Acesso em: 26 abr. 2025.

CRAIN, Matthew; NADLER, Anthony. Political Manipulation and Internet Advertising Infrastructure. *Journal of Information Policy*, [*S.l.*], v. 9, p. 370–410, 2019. Disponível em: https://doi.org/10.5325/jinfopoli.9.2019.0370. Acesso em: 9 abr. 2025.

DHALIWAL, Jasdev. State of the Scamiverse - How AI is Revolutionizing Online Fraud. *In*: *MCAFEE BLOG*, 6 jan. 2025. Disponível em: https://www.mcafee.com/blogs/internet-security/state-of-the-scamiverse/. Acesso em: 9 abr. 2025.

DE-RANCOURT-RAYMOND, Audrey; SMAILI, Nadia. The unethical use of deepfakes. *Journal of Financial Crime*, [*S.l.*], v. 30, n. 4, p. 1066–1077, 2022. Disponível em: https://econpapers.repec.org/article/emejfcpps/jfc-04-2022-0090.htm. Acesso em: 8 abr. 2025.

ECCMA. Code Management Association. The Code Management Association. What is ISO 8000?. *ECCMA*, [*S.d.*]. Disponível em: https://eccma.org/what-is-iso-8000/. Acesso em: 1 ago. 2024.

ELLIS-PETERSEN, Hannah. Revealed: Meta approved political ads in India that incited violence. *The Guardian*, 20 maio 2024. Disponível em: https://www.theguardian.com/world/article/2024/may/20/revealed--meta-approved-political-ads-in-india-that-incited-violence. Acesso em: 26 abr. 2025.

EUROPEAN COMMISSION. About the Digital Markets Act. *European Commission*, [*S.d.*]. Disponível em: https://digital-markets-act.ec.europa.eu/about-dma_en. Acesso em: 1 ago. 2024.

FEDERAL TRADE COMMISSION. Google and YouTube Will Pay Record $170 Million for Alleged Violations of Children's Privacy Law. *Federal Trade Commission*, 4 set. 2019. Disponível em: https://www.ftc.gov/news-events/news/press-releases/2019/09/google-youtube-will-pay--record-170-million-alleged-violations-childrens-privacy-law. Acesso em: 31 out. 2024.

FEDERAL TRADE COMMISSION. FTC Issues Orders to Social Media and Video Streaming Platforms Regarding Efforts to Address Surge in Advertising for Fraudulent Products and Scams. Federal Trade Commission, 2023b. Diponível em: https://www.ftc.gov/news-events/news/press-releases/2023/03/ftc-issues-orders-social-media-video-streaming--platforms-regarding-efforts-address-surge-advertising. Acesso em: 10 maio 2023.

FUKUZAWA, Joel. The Growing Threat of AI-Generated Fraud in Online Advertising. *MEDIUM*, 2 jun. 2024. Disponível em: https://joelhu.medium.com/the-growing-threat-of-ai-generated-fraud-in-online-advertising-a633b0f33448. Acesso em: 8 abr. 2025.

GEHL, Robert W.; LAWSON, Sean T. Social Engineering: How Crowdmasters, Phreaks, Hackers, and Trolls Created a New Form of Manipulative Communication. [*S.l.*]: *The MIT Press*, 2022. Disponível em: https://direct.mit.edu/books/oa-monograph/5281/Social-EngineeringHow-Crowdmasters-Phreaks-Hackers. Acesso em: 8 abr. 2025.

GLOBAL ACTION PLAN. Kids for Sale: Online Advertising and the manipulation of children. Londres: *Global Action Plan*, 2020. Disponível em: https://www.globalactionplan.org.uk/files/kids_for_sale.pdf. Acesso em: 1 ago. 2024.

GOEL, Rajeev K. Masquerading the Government: Drivers of Government Impersonation Fraud. *Public Finance Review*, *[S.l.]*, v. 49, n. 4, p. 548-572, 2021. Disponível em: https://doi.org/10.1177/10911421211029305. Acesso em: 9 abr. 2025.

IAB. "Native Advertising Playbook 2.0". *IAB*, 2019. Disponível em: https://www.iab.com/wp-content/uploads/2019/05/IAB-Native-Advertising-Playbook-2_0_Final.pdf. Acesso em: 24 abr. 2025.

JAHANGIR, Ramsha. Understanding the EU's Digital Services Act Enforcement Against X. *Tech Policy Press*, 05 abr. 2025. Disponível em: https://www.techpolicy.press/understanding-the-eus-digital-services-act--enforcement-against-x/. Acesso em: 26 abr. 2025.

JAMISON, Amelia M.; BRONIATOWSKI, David A.; DREDZE, Mark; WOOD-DOUGHTY, Zach; KHAN, DureAden; QUINN, Sandra C. Vaccine-related advertising in the Facebook Ad Archive. *Vaccine*, [*S.l.*], v. 38, n. 3, p. 512–520, 16 jan. 2020. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S0264410X1931446X. Acesso em: 1 ago. 2024.

KHAN, Lina M.; BEDOYA, Alvaro; SLAUGHTER. Rebecca K. FTC Request For Investigation. Destinatário: *Federal Trade Commission*, 23. ago. 2023. Disponível em: https://fairplayforkids.org/wp-content/uploads/2023/08/FTCRequestForInvestigationAug23.pdf. Acesso em: 31 out. 2024.

KRUSE, Tulio. Fraude digital e roubo de celular dão prejuízo de R$ 71 bi em 1 ano, aponta Datafolha. *Folha de São Paulo*, São Paulo, 12 ago. 2024. Disponível em: https://www1.folha.uol.com.br/cotidiano/2024/08/fraude-digital-e-roubo-de-celular-dao-prejuizo-de-r-71-bi-em-1-ano-aponta-datafolha.shtml. Acesso em: 23 jan. 2025.

LE POCHAT, Victor; EDELSON, Laura; GOETHEM, Tom. V.; JOOSEN, Wouter; MCCOY, Damon; LAUINGER, Tobias. An audit of Facebook's political ad policy enforcement. *In*: USENIX Security Symposium, 31, ago. 2022, Boston: *Anais* [...] USENIX Association, *[S.l.]*, 2022. Disponível em: https://www.usenix.org/system/files/sec22-lepochat.pdf. Acesso em: 14 jan. 2025.

MCSTAY, Andrew. *Digital Advertising*. [*S.l.*]: Bloomsbury Publishing, 2017.

MEIRELES, Isys G.; PASITTO, Fernando T. Estelionato e suas Implicações: O Constante Crescimento dos Golpes Virtuais. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, [*S.l.*], v. 10, n. 11, p. 6303–6316, 2024. Disponível em: https://periodicorease.pro.br/rease/article/view/17063. Acesso em: 9 abr. 2025.

NAPOLI, Philip M; CAPLAN, Robyn. Why media companies insist they're not media companies, why they're wrong, and why it matters. *First Monday*, *[S.l.]*, v. 22, n. 5, 2017. Disponível em: https://firstmonday.org/ojs/index.php/fm/article/view/7051. Acesso em: 1 ago. 2024.

NETLAB UFRJ. A guerra das plataformas contra o PL 2630. *NetLab UFRJ*, maio 2023. Disponível em: https://www.netlab.eco.br/post/a-guerra-das-plataformas-contra-o-pl-2630. Acesso em: 8 abr. 2025.

NETLAB UFRJ. Anúncios com IA usam imagem de políticos brasileiros para aplicar golpes. *NetLab UFRJ*, 17 jun. 2024. Disponível em: https://www.netlab.eco.br/post/an%C3%BAncios-com-ia-usam-imagem-de-pol%C3%ADticos-brasileiros-para-aplicar-golpes. Acesso em: 8 abr. 2025.

O'NEIL, Cathy. *Algoritmos de destruição em massa*. Santo André: Editora Rua do Sabão, 2020.

OCDE. Organização para a Cooperação e Desenvolvimento Econômico. Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity. Paris: *OECD Publishing*, 2024. Disponível em: https://doi.org/10.1787/d909ff7a-en. Acesso em: 1 ago. 2024.

ROGERS, Sam. 1-in-3 Brazilians Targeted by Scammers in the Last 12 Months as Estimated Losses Reach US$54 Billion. *Global Anti-Scam Alliance*, 30 out. 2024. Disponível em: https://www.gasa.org/post/1-in-3-brazilians-targeted-by-scammers-in-2024-state-of-scam-report. Acesso em: 23 jan. 2025.

RUDNITZKI, Ethel. Centenas de anúncios na Meta promovem falso bolão da Mega da Virada. *Aos Fatos*, [*S.l.*], 2024. Disponível em: https://www.aosfatos.org/noticias/anuncios-meta-falso-bolao-mega-virada/. Acesso em: 8 abr. 2025.

SADEGHPOUR, Shadi; VLAJIC, Natalija. Ads and Fraud: A Comprehensive Survey of Fraud in Online Advertising. *Journal of Cybersecurity and Privacy*, [*S.l.*], v. 1, n. 4, p. 804–832, 2021. Disponível em: https://www.mdpi.com/2624-800X/1/4/39. Acesso em: 9 abr. 2025.

SANTINI, R. Marie; SALLES, D.; REGATTIERI, L. L.; BARROS, C. E. Computational propaganda effects. In: CERON, Andrea (org.). *Encyclopedia of technology and politics*. Cheltenham: Edward Elgar, 2022. v. 1, p. 273–276.

SANTOS JUNIOR, Marcelo Alves dos. Estudo exploratório do financiamento da desinformação na web: fraudes, apostas, trading e clickbaits. *Contracampo*, v. 43, n. 1, 2024. Disponível em: https://periodicos.uff.br/contracampo/article/view/56987. Acesso em: 9 abr. 2025.

SATARIANO, Adam. Apple and Meta Are First to Be Hit by E.U. Digital Competition Law. *The New York Times*, 23 abr. 2025. Disponível em: https://www.nytimes.com/2025/04/23/technology/apple-meta-eu-fines--competition-law.html. Acesso em: 26 abr. 2025.

SENSITY. The State of Deepfakes 2024. *Sensity*, 2024. Disponível em: https://sensity.ai/reports/. Acesso em: 8 abr. 2025.

SILVERGUARD; SOS GOLPE. *Estudo Golpes com Pix 2024*. 2024. Disponível em: https://static1.squarespace.com/static/672922c4b034ed7793cab948/t/673368ac9f58876b76a71175/1731422402579/Estudo+Golpes+com+Pix+2024_Silverguard_SOSGolpe.pdf. Acesso em: 14 fev. 2025.

SRIVASTAVA, Amit. K.; MISHRA, Rajhans. Analyzing Social Media Research: A Data Quality and Research Reproducibility Perspective. *IIM Kozhikode Society & Management Review*, Calicute, v. 12, n. 1, p. 39–49, 26 maio 2021. Disponível em: https://doi.org/10.1177/22779752211011810. Acesso em: 1 ago. 2024.

TUROW, Joseph. *The Daily You*: How the New Advertising Industry Is Defining Your Identity and Your Worth. New Haven: Yale University Press, 2012. Disponível em: https://yalebooks.yale.edu/book/9780300188011/the-daily-you/. Acesso em: 01 nov. 2024.

UNESCO. *Guidelines for the governance of digital platforms*: Safeguarding freedom of expression and access to information through a multi-stakeholder approach. Paris: Unesco, 2023. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000387339. Acesso em: 1 ago. 2024.

URMAN, Aleksandra; MAKHORTYKH, Mykola How transparent are transparency reports? Comparative analysis of transparency reporting across online platforms. *Telecommunications Policy*, *[S.l.]*, v. 47, n. 3, abr.

2023. Disponível em: https://doi.org/10.1016/j.telpol.2022.102477. Acesso em: 19 dez. 2024.

VOLKOVA, Svetlana. The Dark Side of Deepfakes: Fraud and Cybercrime. *In*: Deepfakes and their Impact on Business. [*S.l.*]: *IGI Global Scientific Publishing*, 2025. p. 221–242. Disponível em: https://www.igi-global.com/chapter/the-dark-side-of-deepfakes/www.igi-global.com/chapter/the-dark-side-of-deepfakes/364354. Acesso em: 9 abr. 2025.

WOJDYNSKI, Bartosz W.; GOLAN, Guy J. Native Advertising and the Future of Mass Communication. *American Behavioral Scientist*, [*S.l.*], v. 60, n. 12, p. 1403–1407, 2016. Disponível em: https://doi.org/10.1177/0002764216660134. Acesso em: 8 abr. 2025.

ZENG, Eric; KOHNO, Tadayoshi; ROESNER, Franziska; ALLEN, Paul G. Bad News: Clickbait and Deceptive Ads on News and Misinformation Websites. *In*: Bad News, 2020. *Anais* [...], 2020. Disponível em: https://www.semanticscholar.org/paper/Bad-News%3A-Clickbait--and-Deceptive-Ads-on-News-and-Zeng-Kohno/a7d7a8e1eca1dee63e871751dad4e0481079acb4. Acesso em: 8 abr. 2025.

ZUBOFF, Shoshana. "We Make Them Dance": Surveillance Capitalism, the Rise of Instrumentarian Power, and the Threat to Human Rights. *In*: JØRGENSEN, Rikke Frank. *Human Rights in the Age of Platforms*. Cambridge: The MIT Press, 2019. Disponível em: https://direct.mit.edu/books/oa-edited-volume/4531/chapter/202528/We-Make-Them--Dance-Surveillance-Capitalism-the. Acesso em: 18 fev. 2025.

# Capítulo 1

# Índice de Transparência de Dados das Plataformas de Redes Sociais

O crescimento vertiginoso do número de usuários de plataformas de redes sociais ao fim da década de 2000 representou um importante ponto de virada nas Ciências Sociais Aplicadas. O arcabouço teórico, epistemológico e sobretudo metodológico destas áreas mudou significativamente com a então recém-adquirida capacidade de produzir bases de dados abundantes sobre os fluxos de informação e as relações sociais construídas nestes espaços, estimulando a criação de institutos de pesquisa e novos periódicos científicos dedicados à investigação do desenrolar de fenômenos culturais, políticos e econômicos em rede (Bruns, 2019; Edelmann *et al*., 2020).

O principal ganho para a pesquisa foi de escala: qualquer pesquisador com conhecimentos básicos em programação conseguiria coletar volumes de dados sem precedentes sobre temáticas de seu interesse com relativa facilidade, em uma proporção que métodos tradicionais como *surveys*, entrevistas e etnografias não seriam capazes de alcançar. Para conseguir analisar este volume inédito de dados, as Ciências Sociais passaram a ser influenciadas pelas Ciências da Computação, levando à formação das Ciências Sociais Computacionais (Edelmann *et al*., 2020) e campos correlatos como as Ciências Computacionais da Comunicação (van Atteveldt; Peng, 2018). Assim, pesquisadores visavam não a mera transposição de teorias sociais e comportamentais já existentes para o mundo digital, mas a criação de novas abordagens e quadros

conceituais com base nos fenômenos que se propunham a analisar neste ambiente (Edelmann *et al*., 2020).

A "era de ouro" da pesquisa em plataformas de redes sociais foi, então, marcada por uma maior transparência – aqui entendida como a possibilidade de acessar informações de interesse público sobre as funcionalidades, governança e modos de uso das plataformas para avaliação da sociedade, governos e outros *stakeholders* (Urman; Makhortykh, 2023) –, ainda que o estímulo à pesquisa não fosse o principal motivo dessa abertura de dados (Tromble, 2021). Não tardou muito, porém, para que esta era chegasse a um fim, entre 2016 e 2017, sobretudo graças à utilização indevida de dados de usuários do Facebook no que ficou conhecido como o *Escândalo de dados Facebook–Cambridge Analytica* (Bruns, 2019; Tromble, 2021)

Por exemplo, foi graças à possibilidade de coletar dados massivos do Twitter que pesquisadores de todo o mundo começaram a desenvolver classificadores e métodos de detecção personalizados para a identificação de contas inautênticas e automatizadas, apontando para operações de influência sobre diversas eleições ao redor do globo (Woolley; Howard, 2018). Escândalos de privacidade envolvendo o uso intenso e indevido de dados pessoais de usuários para manipulação eleitoral nos Estados Unidos e na Europa também levaram as plataformas a se fecharem cada vez mais, colocando um fim a diversas ferramentas de coleta e análise de dados utilizadas por pesquisadores (Bruns, 2019; Tromble, 2021). Para as *big tech*, é conveniente ignorar as questões expostas por estes pesquisadores e acabar com os meios utilizados por eles para divulgá-las, ajudando a subdimensionar problemas reais. Neste movimento, as plataformas frequentemente introduzem novas ferramentas pouco eficientes voltadas à coleta de dados e introduzem outras funcionalidades que supostamente tornam as experiências de seus usuários mais "seguras", como forma de evitar crises reputacionais e pedidos de maior regulação governamental (Bossetta, 2020). Por conta disso, pesquisadores estão à mercê da própria capacidade de gerar soluções para coletar dados e investigar estas plataformas, muitas vezes baseadas em técnicas extraoficiais desautorizadas por elas, sujeitando-os a maior insegurança jurídica (Freelon, 2018).

Diante de um cenário de escassez geral de informação qualificada para pesquisas sociais aplicadas baseadas em dados oriundos de plataformas de redes sociais, apresentamos o Índice de Transparência de Dados das Plataformas de Redes Sociais (ITD) no Brasil. Com ele, buscamos indicar caminhos para melhorias no acesso gratuito e universal a dados de plataformas de redes sociais e, assim, contribuir para a realização de pesquisas acadêmicas de interesse público.

Consideramos como plataformas de redes sociais os espaços em que usuários produzem conteúdo por conta própria, consumindo e interagindo com publicações de outros usuários com os quais se conectam neste processo (Ellison; Boyd, 2013), ideia que também abarca os chamados aplicativos de mensageria, como Telegram e WhatsApp. Estes aplicativos rompem a barreira entre a comunicação interpessoal e massiva, desempenhando papel relevante na disseminação e no compartilhamento de conteúdos produzidos por usuários em grupos e canais públicos. Permitem, ainda, a criação de redes de interação com base em afinidades (Júnior *et al*., 2021; Rogers, 2020), nas quais os usuários tornam-se simultaneamente produtores de conteúdo empoderados e alvos para a exploração comercial de empresas.

No ITD, avaliamos e respondemos:

- Quais são as medidas de transparência e acesso a dados das principais plataformas de redes sociais no Brasil?
- Qual é o nível da qualidade dos dados disponibilizados por essas plataformas para atividades de pesquisa?
- Neste sentido, nossos objetivos incluem:
- **Padronizar:** Definir parâmetros de avaliação sobre o nível de acesso e qualidade dos dados de interesse público provenientes de plataformas de redes sociais.
- **Avaliar:** Identificar, de forma sistemática e transparente, os pontos fortes e fracos no acesso e na qualidade dos dados provenientes das plataformas de redes sociais.

- **Comparar:** Aferir comparativamente a performance de cada plataforma a partir de critérios comuns e metodologia padronizada.
- **Aprimorar:** Indicar, pública e objetivamente, o que precisa ser melhorado na disponibilização de dados para pesquisa por parte das plataformas de redes sociais.

Para tanto, o ITD segue um roteiro de avaliação estruturado, sistematizado e reprodutível, baseado em critérios de qualidade de dados, para avaliar os mecanismos existentes de acesso a dados de interesse público resultantes de publicações orgânicas e públicas nas principais plataformas de redes sociais que atuam no Brasil, vistas na Figura 1:

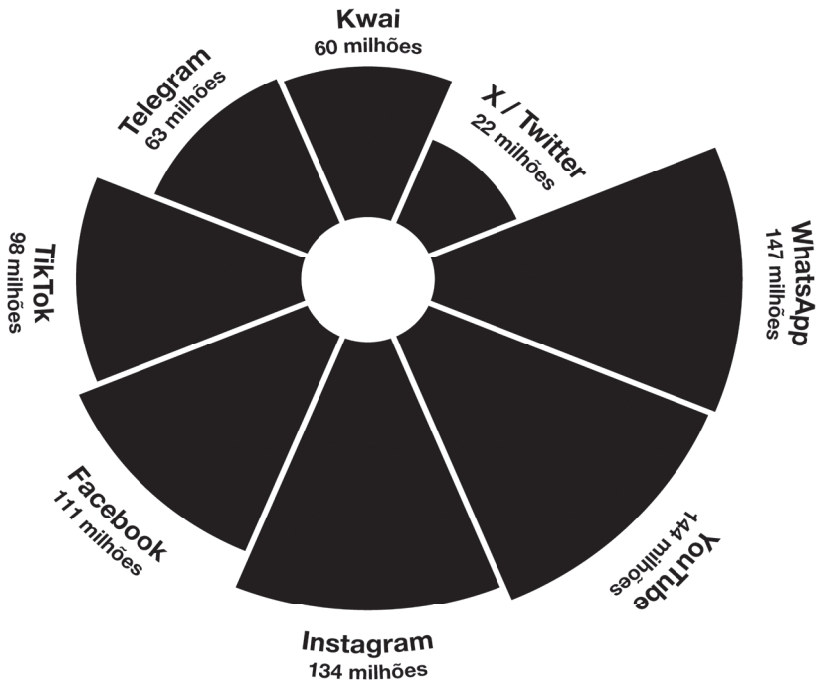



**Figura 1:** Número de usuários de cada plataforma analisada no Brasil (Bianchi, 2024; Data Reportal, 2024; Global AD, 2024; Opinion Box, 2024; Singh, 2024)[1]

[1] Ao contrário do que ocorre na União Europeia (European Commission, 2023), as plataformas digitais não possuem qualquer obrigação legal de divulgar números oficiais de suas bases de usuários no Brasil. Diante da falta dessas informações, baseamo-nos primordialmente em estimativas não oficiais.

As plataformas foram selecionadas conforme princípio do *Digital Services Act* (DSA), regulação vigente na União Europeia que estabelece medidas de responsabilidade e transparência das plataformas digitais que são utilizadas por mais de 10% da população do bloco. Por serem as mais relevantes e as capazes de gerar maior impacto social no contexto brasileiro, consideramos que estas plataformas devem garantir maiores investimentos em infraestrutura robusta de transparência e seguir as melhores práticas do mercado por terem os recursos necessários para tanto. Embora, segundo estimativas não oficiais, plataformas como LinkedIn e Pinterest atinjam 10% da população brasileira, não as analisamos, por elas terem pouco presença em investigações acadêmicas e científicas no Brasil e no resto do mundo (Kapoor *et al*., 2018; Zuckerman, 2021).

Nossa premissa é a de que a padronização de critérios para avaliar sistematicamente a transparência e a qualidade dos dados das plataformas de redes sociais tende a impactar positivamente na transparência desses espaços. A interpretação da transparência dos dados a partir de preceitos de qualidade é especialmente importante, posto que a qualidade foi, por muito tempo, deixada de lado por pesquisadores de plataformas de redes sociais, mais interessados no grande volume de dados que conseguiam coletar do que em seus problemas de representatividade e outros vieses (Tromble, 2021).

O roteiro de avaliação do ITD, composto por 40 parâmetros de avaliação que embasam o cálculo e a aferição de uma nota para cada plataforma, toma por base seis dimensões de qualidade de dados preconizadas pela literatura científica, independentemente das particularidades das plataformas avaliadas. Entre as dimensões endógenas aos dados, são avaliadas a completude, a atualidade, a consistência e a acessibilidade (Barbieri, 2019; Batini; Scannapieco, 2006; McGilvray, 2008; Loshin, 2008). Outras dimensões avaliadas, como conformidade e relevância (Barbieri, 2019), dependem de fatores exógenos e, por isso, podem variar de acordo com as normas legais em vigor em cada país ou com os objetivos específicos da pesquisa.

Neste capítulo, primeiramente, nós apresentamos o surgimento das Ciências Sociais Computacionais e áreas correlatas, e as oportunidades e desafios por elas enfrentados, especialmente no que diz respeito à disponibilização de dados para pesquisa por parte das plataformas. Em seguida, explicamos a abordagem metodológica do ITD, apresentando seu roteiro de avaliação e os cálculos para a composição das notas das plataformas analisadas. Por fim, detalhamos um panorama da avaliação individual de cada plataforma, salientando boas e más práticas percebidas em suas medidas e ações de transparência e disponibilização de dados para pesquisa. Em nossa avaliação, nenhuma plataforma atingiu uma pontuação ideal e apenas uma foi avaliada satisfatoriamente, indicando um cenário preocupante para pesquisadores brasileiros quando comparados a suas contrapartes no Norte Global.

## A importância da pesquisa com dados digitais

Cidadãos comuns estão sujeitos ao uso de seus dados para uma infinidade de ações comerciais e operações de influência que, hoje, dificilmente são auditáveis, embora gerem grande impacto na vida individual e coletiva. As *big tech*, como o Google e a Meta, acumulam dados sobre todos os tipos de ações e interações realizadas por seus usuários, bem como sobre seus comportamentos e atitudes (Tromble, 2021), mas o usuário pouco sabe sobre o que é feito com as informações que ele fornece à plataforma, saindo prejudicado nesta relação assimétrica (Dobber *et al.*, 2023). Quando o usuário aceita um contrato de adesão para utilizar a plataforma, concordando com seus termos de uso, cria-se um vínculo comercial e uma relação de consumo que, como todas as outras relações desse tipo, deve oferecer meios para sua proteção. A pesquisa acadêmica cumpre exatamente esse papel, ao pressionar as plataformas a adotarem medidas que tragam algum grau de equilíbrio – ainda que moderado e limitado – a essa relação.

Diferentes fatores concedem aos dados digitais o potencial de não apenas identificar, descrever e compreender fenômenos e problemas sociais relevantes, mas também de ajudar a lidar com eles ou até resolvê-

-los. A relevância dos dados de redes sociais para pesquisa fica evidente na crescente influência das plataformas na vida pública – hoje, as principais arenas de debate público (Staab; Thiel, 2022; Yasseri, 2023), envolvendo tanto temáticas políticas e sociais que circulam organicamente quanto campanhas planejadas para influenciar gostos, hábitos de consumo, modos de vida, opiniões e comportamentos (Woolley; Howard, 2018). Por conta de sua enorme influência sobre debates públicos, o acesso à informação e a liberdade de expressão, suas atribuições são de interesse público e por isso elas devem ser publicamente escrutinadas, com prestação de contas à população, e não apenas a seus acionistas (Bromell, 2022).

Logo, por estas plataformas estarem cada vez mais implicadas em processos e dinâmicas sociais e políticas de grande relevância, pesquisadores dependem de seus dados para conduzir pesquisas empíricas pertinentes e de alta qualidade (Bruns, 2019). Com uma capacidade de prover detalhes riquíssimos sobre as relações sociais de grandes populações à medida que elas se desenrolam (Edelmann *et al.*, 2020), pesquisas realizadas com estes dados são capazes de informar políticas públicas e de governança que terão impacto no desenvolvimento das próprias tecnologias e no tipo de espaços de sociabilidade que oferecem.

O ganho mais abrangente da pesquisa com dados digitais é justamente a possibilidade de mensurar o impacto das próprias tecnologias na vida social, já que boa parte do recente desenvolvimento dos métodos computacionais "é voltado para analisar a estrutura e dinâmica da comunicação humana" (van Atteveldt; Peng, 2018, p. 81). Dados nativamente digitais fornecem informações extraídas ou inferidas em amostras coletadas no ambiente onde a comunicação de fato se desdobra, sendo produzidos sem a interferência dos pesquisadores sobre o objeto de estudo (Edwards *et al.*, 2013; Lee *et al.*, 2008; Marres, 2017; Rogers, 2009). A pesquisa com dados digitais, portanto, é capaz de oferecer, com agilidade, importantes *insights* sobre fenômenos sociais, culturais, econômicos e políticos capazes de impactar milhões de pessoas (van Atteveldt; Peng, 2018), além de colaborar com a responsabilização de

plataformas por problemas que elas, indireta ou diretamente, ajudam a perpetuar (Tromble, 2021).

Para analisar esses dados, pesquisadores devem conhecer a fundo o funcionamento do ambiente digital de onde foram extraídos, já que os dados "tendem a ser fortemente marcados pelos efeitos da plataforma, como a busca por termos sugeridos em funções de autocompletar ou a utilização de *hashtags* que estão em alta na plataforma em questão" (Shaw, 2015, p. 2). Essa é uma das razões pelas quais os dados digitais trouxeram, junto aos novos tipos de informações a serem analisadas, novos tipos de problemas sociais e desafios de pesquisa. Também emergiram novos métodos para explorá-los e para compreender diferentes características das plataformas digitais e, em particular, das plataformas de redes sociais, que moldam as dinâmicas sociais nesses ambientes (Rogers, 2009), devendo ser aplicados com foco no interesse público e nos problemas sociais contemporâneos.

Por serem voltados para o processamento e a descoberta de informações relevantes em meio a quantidades massivas de dados, os métodos digitais também têm sido amplamente utilizados em diversas áreas do conhecimento. Esse é o aspecto que coloca as Ciências Sociais Computacionais como uma disciplina científica instrumental para outras áreas (Cioffi-Revilla, 2018). Estudos epidemiológicos e de farmacovigilância, por exemplo, utilizam métodos digitais e computacionais há anos (El-Sayed *et al*., 2012; Pappa; Stergioulas, 2019), abordagens que foram úteis durante a pandemia de Covid-19, no início dos anos 2020, em aspectos como controle de aglomerações (Cecilia *et al*., 2020), avaliação de sofrimentos psicológicos (Dhelim *et al*., 2023), predição da evolução da doença e definição de prognósticos (Senthilraja, 2021). Neste contexto, um dos objetivos sociais e de interesse público mais urgentes é colaborar com o acúmulo de evidências, a produção de diagnósticos e a indicação de soluções para fenômenos como a desinformação, o uso de tecnologias por crianças e adolescentes, os crimes cibernéticos e o impacto das redes sociais na saúde mental dos cidadãos, entre outros.

A ideia de que os dados para pesquisa em plataformas de redes sociais são de interesse público e, portanto, devem ser publicamente acessíveis é central ao ITD, e não está em contradição com a integridade da privacidade dos usuários das plataformas de redes sociais. Embora o conceito de "interesse público" seja abrangente, o entendimento da Autoridade Nacional de Proteção de Dados (ANPD) para aplicar condições especiais no tratamento de dados pessoais, de acordo com a Lei Geral de Proteção de Dados (LGPD), está ligado à natureza do órgão que realiza a pesquisa – compreensão que também adotamos para definir pesquisas de interesse público. A LGPD também estipula que dados tornados manifestamente públicos pelo seu titular não requerem consentimento para uso (Brasil, 2018), definição que viabiliza a pesquisa com dados originados em espaços de opinião pública com potencial de impactar inúmeros cidadãos. Pode ser entendido, portanto, que deve ter acesso a dados gerados a partir de conteúdos públicos, abertos e acessíveis por qualquer usuário de plataformas de redes sociais o "órgão ou entidade da administração pública direta ou indireta ou pessoa jurídica de direito privado sem fins lucrativos legalmente constituída sob as leis brasileiras, com sede e foro no País, que inclua em sua missão institucional ou em seu objeto social ou estatutário a pesquisa básica ou aplicada de caráter histórico, científico, tecnológico ou estatístico" (ANPD, [*S.d.*], p. 26).

No ITD, também levamos em consideração a disponibilização de dados que se originam em grupos e canais públicos no WhatsApp e no Telegram, quando divulgados abertamente na internet com o intuito de captar novos membros e nos quais qualquer usuário pode entrar (Evangelista; Bruno, 2019; Garimella; Tyson, 2018; Resende *et al.*, 2019). Visto que qualquer pessoa com o link de acesso ao grupo pode integrá-lo, pesquisadores estabeleceram critérios éticos sobre o que pode ser considerado público nesses aplicativos. Listas online que elencam os grupos públicos e ferramentas de busca que possibilitam encontrá-los têm sido úteis para a identificação desse tipo de grupo por pesquisadores (Garimella, Tyson, 2018; Melo, 2022; Resende *et al.*, 2019). O próprio Telegram, inclusive, apresenta esta definição de grupos e canais

públicos em sua documentação (Telegram, [*S.d.*]a; [*S.d.*]b). Por serem importantes vias para a mobilização social e a propagação de informação política, confiável ou não, esses aplicativos são cada vez mais relevantes para a pesquisa social (Calvo-Gutiérrez; Marín-Lladó, 2023; Evangelista; Bruno, 2019; Melo *et al*., 2019; Ozawa *et al*., 2023; Smith *et al*., 2023; Wendratama; Yusuf, 2023).

Assim sendo, dentro do escopo proposto, consideramos que os dados digitais públicos de plataformas de redes sociais se originam a partir de publicações com visibilidade definida como pública por seus autores e conteúdos temporários também públicos, mas apenas enquanto estiverem no ar, bem como de mensagens enviadas em grupos e canais públicos de aplicativos de mensageria, como WhatsApp e Telegram. Já os dados digitais não públicos são gerados a partir de publicações definidas como privadas por seus autores e que só podem ser visualizadas por eles e outras pessoas selecionadas, bem como de conteúdos feitos por perfis privados e de mensagens enviadas em quaisquer conversas privadas em aplicativos como Telegram, WhatsApp, Instagram, Messenger e X/Twitter.

## Apagão de dados digitais

A transparência e o conhecimento público sobre o funcionamento de plataformas de redes sociais não são proporcionais ao seu potencial de impacto na vida social e à sua relevância como espaços de formação de opinião pública. Além da dificuldade de acesso aos dados (Bruns, 2019), as plataformas de redes sociais mudam constantemente, nem sempre de forma transparente, as regras de funcionamento embutidas em seus algoritmos, gerando importantes preocupações éticas e políticas (Selinger; Hartzog, 2016; Guess *et al*., 2023) e reforçando a opacidade sobre os algoritmos (Lu, 2021). Essas restrições não ameaçam só o trabalho de pesquisadores, mas ainda a capacidade da sociedade de entender a dimensão das implicações sociais do funcionamento das plataformas digitais (Bruns, 2019).

O domínio das plataformas digitais na produção e distribuição de conteúdo permitiu que elas influenciassem ativamente na forma como são vistas pela opinião pública. Ao mesmo tempo que as plataformas de redes sociais buscam ser percebidas como facilitadoras da distribuição de conteúdos gerados por usuários, elas se esforçam para se afastar das responsabilidades sobre o que esses usuários publicam e da percepção de que realizam uma curadoria algorítmica de conteúdo (Gillespie, 2010). Frequentemente, diminuir o acesso aos dados é parte fundamental desses esforços.

Embora dados digitais sejam vitais para compreender o modo como as plataformas online impactam dinâmicas políticas e sociais, o acesso a dados completos, atualizados, consistentes, disponíveis no tempo necessário, em conformidade legal e relevantes para pesquisa é cada vez mais limitado, dificultado ou mesmo extinto pelas *big tech* (Bruns, 2019; Greene; Martens; Schmueli, 2022; Tromble, 2021). Entre os exemplos mais emblemáticos, está o encerramento do CrowdTangle pela Meta (Soares, 2024), em agosto de 2024, ferramenta que garantia acesso a dados do Facebook e Instagram tanto por uma interface de programação de aplicações (API)[2] quanto por uma interface de usuário[3]. Esse movimento acompanha a mudança do X/Twitter que, desde

[2] Do original *application programming interface*, uma API é um dos principais meios programáticos utilizados por pesquisadores para a coleta de dados de plataformas de redes sociais e outros serviços digitais. A funcionalidade básica de uma API é estabelecer a comunicação entre dois componentes de *software* (Goodwin, 2024) – um servidor, como um banco de dados, e um "cliente", que pode ser o dispositivo de um pesquisador, por exemplo –, mediante pedidos de dados e informações que devem cumprir definições e protocolos próprios. Dessa forma, as APIs possibilitam o compartilhamento de dados entre diferentes aplicações, sistemas, dispositivos e plataformas, facilitando a interoperabilidade entre sistemas diversos (Postman, [*S.d.*]). Para acessar uma API, é preciso gerar *tokens*, que são pequenos códigos entregues a cada cliente cadastrado e que funcionam como senhas para autenticar e validar requisições. Já para fazer pedidos a uma API, o cliente deve realizar "chamadas" aos *endpoints*, de acordo com os parâmetros indicadas em sua documentação, que retornam dados e informações de diferentes naturezas, a depender do uso pretendido e das instruções fornecidas em linhas de código de programação (CloudFlare, [*S.d.*]; Postman, [*S.d.*]).

[3] Uma interface gráfica de usuário (*graphical user interface*) é qualquer ambiente digital que permita a interação entre um usuário e um banco de dados por meio de elementos gráficos, como ícones, janelas e menus, sem a necessidade de digitar instruções em códigos de programação. Além de suas interfaces de utilização padrão, algumas plataformas mantêm um outro tipo de interface dedicada à coleta de dados, avaliada nos parâmetros do ITD. Por meio destas interfaces de coleta,

março de 2023, cobra um alto valor pelo acesso à sua API, que até então era gratuita e amplamente explorada para pesquisa (Mozelli, 2023). A tendência, entretanto, não é restrita aos últimos anos: logo após o escândalo da Cambridge Analytica em 2018, o Facebook também modificou e restringiu funcionalidades de sua API sem aviso prévio, prejudicando o desenvolvimento de pesquisas acadêmicas (Bruns, 2019).

Ao analisar o cenário do Facebook, Instagram e X/Twitter, Bruns (2019) considera que as *big tech* procuram limitar o desenvolvimento de pesquisas críticas fundamentadas em dados sobre suas plataformas, ao mesmo tempo que lançam ferramentas de acesso a dados incompletos e enviesados para gerar publicidade positiva com base em uma falsa capacidade de resolução de problemas. Por muito tempo, as plataformas prezavam pela imagem positiva oriunda de pesquisas que davam dimensão de sua relevância, diante de críticas de que elas seriam excessivamente "banais", o que também levou à contratação de pesquisadores e à cessão de dados para institutos de pesquisa (Bruns, 2019). Entretanto, em face de grandes escândalos, as *big tech* entendem que elas não se beneficiam mais dessa relação (Bruns, 2019), substituindo ferramentas de acesso a dados existentes por outras pouco efetivas e que inutilizam todo um arcabouço operacional e técnico construído por pesquisadores ao longo do tempo (Freelon, 2018). Afinal, para as plataformas digitais, a transparência é menos prioritária do que a imagem pública que elas construíram de si mesmas, então a inefetividade das atuais medidas de transparência que promovem é cuidadosamente pensada para que esta imagem não seja destruída (Bossetta, 2020; Bromell, 2022).

Diante desse cenário, pesquisadores comumente tecem críticas às novas restrições impostas ao acesso às ferramentas e aos mecanismos de

---

é possível explorar e extrair os dados de publicações geradas por usuários. Idealmente, estas interfaces permitem a busca de publicações de acordo com termos de busca ou páginas de interesse, gerar sumarizações e visualizações de dados e exportá-los em formato estruturado. Diferentemente de uma API, estas interfaces, porém, não permitem a coleta e exploração programática e automatizada de dados e, por consequência, não garantem um monitoramento constante e sistemático de publicações de temáticas de interesse. Por isso, não devem ser encaradas como uma solução por si só, mas como um complemento a estratégias de coleta garantidas por uma API, facilitando o acesso a dados por usuários com poucos conhecimentos técnicos em programação.

transparência das plataformas (Coalition for Independent Technology Research, 2023; Mozilla Foundation, 2024). Em contrapartida, outros pesquisadores (ver Tromble, 2021) argumentam que a relação entre plataformas digitais e a academia sempre foi turbulenta e extremamente mediada e controlada, de maneira que o uso de dados massivos para pesquisa nunca teria sido um fim esperado por elas, mas uma consequência da abertura de suas APIs para que desenvolvedores pudessem criar aplicativos capazes de atrair novos usuários e, logo, aumentar o alcance de suas redes de publicidade. A motivação, portanto, nunca teria sido a de apoiar pesquisas que expusessem as implicações sociais de seus sistemas, mas de sempre expandir o desenvolvimento de inovação por terceiros e sem custos, visando a maximização de seus lucros (Tromble, 2021). De todo modo, pesquisadores têm explorado caminhos alternativos para continuarem coletando dados das plataformas, especialmente por meio de técnicas de raspagem (*data scraping* ou *web scraping*)[4], em sua maioria desautorizadas pelas plataformas, o que pode acarretar desde banimentos dos usuários a problemas jurídicos (Conger, 2016; Freelon, 2018; Krotov; Johnson; Silva, 2020; Roth, 2023).

Hoje, as *big tech* não têm qualquer obrigação legal sobre como seus dados devem ser disponibilizados a partes interessadas para pesquisa no Brasil. Além disso, quando promovem medidas autorregulatórias de transparência, como a publicação de relatórios de transparência sobre suas ações de moderação e governança, o fazem sem qualquer tipo de critério ou padronização, apresentando dados com granularidade insatisfatória, por exemplo (Kosta; Brewczyńska, 2019; Suzor *et al*., 2019). Documentos tradicionalmente voluntários, esses relatórios reúnem estatísticas e informações sobre as ações motivadas pelo descumprimento

[4] A raspagem de dados envolve processos de combinação, estruturação e extração de informações de páginas e outros conteúdos disponibilizados online, antecedendo o desenvolvimento das primeiras APIs de coleta de dados (Bar-Ilan, 2001; Freelon, 2018; Mooney; Westreich; El-Sayed, 2015). Por conta do progressivo fechamento de APIs e outras ferramentas para coleta de dados de plataformas de redes sociais, pesquisadores têm se valido cada vez mais de técnicas de raspagem para desenvolver seus trabalhos. Mesmo no caso de plataformas que oferecem APIs, não é incomum que pesquisadores recorram a técnicas de raspagem para complementar suas extrações de dados com informações que não são retornadas oficialmente ou, ao menos, não em volume ideal.

de seus termos de uso e por pedidos de autoridades e entes estatais, como a remoção de determinados conteúdos e o fornecimento de outras informações não disponíveis publicamente. Em muitos casos, servem somente a uma "visibilidade moderada" (Wagner *et al.*, 2020), já que, sem meios para uma auditoria independente, não é possível saber se há a liberação apenas de certas informações e enquadramentos que promovam uma perspectiva favorável às plataformas. O DSA passou a obrigar que grandes plataformas online e ferramentas de busca divulguem esse tipo de relatório sobre suas atividades em países da União Europeia. Os primeiros relatórios após a nova regra foram divulgados em novembro de 2023, mas têm recebido críticas pela falta de padronização entre eles, que impossibilita comparações analíticas entre plataformas, e pelas discrepâncias de qualidade e granularidade dos dados disponibilizados (Miller, 2023).

À luz de estudos anteriores que apontam que fornecer dados de interesse público é insuficiente se não houver a indicação de critérios objetivos de qualidade (Transparência Brasil, 2018; Braga; Cunha, 2022; Santos, 2023), consideramos que é importante vincular dados públicos de plataformas de redes sociais a parâmetros satisfatórios mínimos de acesso e qualidade. Iniciativas legislativas recentes buscam mitigar os efeitos sociais negativos do alto poder de influência das plataformas de redes sociais na vida social (Zittrain, 2019) por meio da ampliação de sua responsabilidade como intermediárias – como o PL 2.630/2020, que tramitou no Brasil por quatro anos e foi retirado de pauta em abril de 2024 – e da regulação de seus serviços – como o DSA, em vigor da União Europeia (European Commission, [*S.d.*]). No entanto, embora algumas dessas iniciativas a nível internacional versem sobre o acesso a dados de redes sociais para pesquisa, nenhuma delas estabelece parâmetros de qualidade para tal. Afinal, por muito tempo, dados digitais foram considerados "permanentes e invariantes" e as próprias APIs e outras ferramentas de coleta de dados, "neutras" (Tromble, 2021), o que torna ainda mais urgente uma avaliação das práticas de disponibilização de dados estabelecidas pelas plataformas.

Assim, em paralelo ao esforço para garantir e ampliar o acesso a dados em plataformas de redes sociais para pesquisa acadêmica, a definição de indicadores de qualidade é fundamental para mapear e avaliar as medidas de transparência oferecidas por elas. Dessa forma, é possível fornecer evidências sobre dados incompletos, inconsistentes e de baixa qualidade que impedem ou dificultam a análise sistemática, o diagnóstico de riscos e a responsabilização por eventuais danos. A padronização de critérios de qualidade e transparência também estabelece parâmetros comuns que possibilitam comparações e acompanhamento sistemáticos de diferentes empresas.

## Abordagem metodológica

Para mensurar a transparência das principais plataformas de redes sociais disponíveis no Brasil em relação a dados públicos de conteúdos gerados por usuários, elaboramos um roteiro de avaliação em um processo iterativo e deliberativo, estabelecendo os parâmetros e critérios de avaliação e suas definições conceituais. Os parâmetros foram avaliados e justificados por oito pesquisadores do NetLab UFRJ, divididos em duplas que incluíam um especialista em coleta, infraestrutura e processamento de dados e outro com experiência em análise de dados e desenho de pesquisa em Ciências Sociais Computacionais. As duplas também ficaram responsáveis pela revisão de respostas feitas por outros pares, conforme a divisão apresentada na **Tabela 1**. A distribuição de plataformas entre os pesquisadores levou em consideração o conhecimento prévio e a participação em pesquisas envolvendo dados da plataforma avaliada. Ao longo do processo de elaboração do índice, a adequação dos parâmetros e a pertinência de suas justificativas foram continuamente deliberadas em conjunto pelos avaliadores e outros pesquisadores envolvidos no estudo. A avaliação foi realizada e revisada ao longo do primeiro semestre de 2024.

**Tabela 1:** Divisão das respostas dos parâmetros por duplas de especialista ($E_n$)

| Plataforma | Pesquisadores responsáveis pela resposta | Pesquisadores responsáveis pela revisão |
|---|---|---|
| YouTube | E1 e E2 | E5 e E7 |
| Facebook | E3 e E4 | E2 e E6 |
| Instagram | E3 e E4 | E2 e E6 |
| X/Twitter | E3 e E5 | E6 e E8 |
| Telegram | E2 e E6 | E3 e E4 |
| Kwai | E2 e E3 | A revisão foi realizada em conjunto |
| TikTok | E5 e E7 | E1 |
| WhatsApp | E6 e E8 | E3 e E4 |

O roteiro é composto por 40 parâmetros que analisam seis dimensões de qualidade de dados: acessibilidade, conformidade, completude, consistência, relevância e atualidade. As avaliações foram realizadas e justificadas com base em cinco diferentes fontes de informação: i) a experiência acumulada do NetLab UFRJ; ii) a realização de testes de acesso e coleta de dados; iii) a documentação oficial da plataforma e de sua API, quando disponível; iv) os relatórios de transparência de moderação de conteúdo das plataformas de redes sociais, quando disponíveis; e v) a literatura acadêmica sobre o tema.

Sobretudo, levamos em consideração os obstáculos enfrentados e as soluções desenvolvidas pelo NetLab UFRJ na construção de sua infraestrutura própria e customizada para garantir o monitoramento constante de diferentes plataformas ao longo dos anos. Desde 2020, o laboratório desenvolve e mantém uma coleta contínua e ininterrupta para o monitoramento constante de diversas plataformas de redes sociais. A construção dessa infraestrutura de coleta depende do manuseio de sistemas variados, assim como do entendimento pleno dos tipos de dados disponibilizados pelas plataformas e de como coletá-los. Conduzimos, ainda, experimentos controlados, simulando situações reais de uso e extração de dados, para testar e verificar fatores como a consistência das respostas da API, a persistência de conteúdos removidos,

a viabilidade da raspagem via navegador e a ocorrência de bloqueios à raspagem de dados.

Além disso, consultamos e referenciamos os termos de usos e políticas das plataformas e, quando disponíveis, as documentações oficiais de suas APIs. A documentação de uma API relata, detalha e explica o seu funcionamento, indicando aos usuários como utilizá-la. Plataformas que disponibilizam APIs comumente incluem documentações para que desenvolvedores possam entendê-las durante a elaboração de requisições. Além da documentação, consultamos as políticas e os termos de uso da plataforma e das APIs para responder aos critérios. Em casos excepcionais, contatamos diretamente o suporte da plataforma para maiores esclarecimentos.

Além disso, consideramos a disponibilização de relatórios de transparência por parte das plataformas analisadas e, nos casos em que são disponibilizados, a periodicidade de sua publicação e o detalhamento de suas informações. Por fim, também levamos em consideração a produção acadêmica nacional e internacional publicada em periódicos de impacto, com metodologias desenvolvidas, testadas e aprovadas por pares. A literatura acadêmica foi usada, principalmente, para identificar e conferir o uso de métodos não oficiais de coleta de dados.

Os parâmetros poderiam ser respondidos de duas maneiras diferentes, segundo avaliações positivas ou negativas, que, ao fim, embasaram a realização de cálculos de notas para cada uma das plataformas analisadas. Casos em que um parâmetro não foi aplicável à avaliação de uma plataforma também foram adequadamente indicados e desconsiderados dos cálculos finais de sua pontuação.

## Critérios de avaliação

Os 40 parâmetros avaliados para cada plataforma se dividem da seguinte maneira:

*Acessibilidade (16 parâmetros)*

Dimensão mais importante analisada, já que os parâmetros de avaliação referentes às outras dimensões dependem do acesso aos dados. A acessibilidade se refere à disponibilidade de dados e facilidade de localizá-los, acessá-los, obtê-los e explorá-los para um determinado fim (Mahanti, 2018). Portanto, não basta torná-los acessíveis. É preciso, também, haver condições para que sejam facilmente compreendidos e analisados por pesquisadores com variados graus de conhecimento técnico, principalmente no que diz respeito à necessidade de programação. Nesta dimensão, foram analisados fatores como a disponibilização gratuita ou paga de API e se ela permite a extração total ou parcial de dados de interesse público.

Parâmetros que compõem a dimensão de *Acessibilidade*

P1: A plataforma disponibiliza API oficial para acesso aos dados públicos publicados por usuários?

Aqui, observamos se a plataforma oferece uma API com ao menos um *endpoint* para acesso aos dados de conteúdos gerados por usuários.

P2: O universo a ser monitorado é recuperável pela API da plataforma?

Neste campo, busca-se verificar se a plataforma permite a descoberta e a coleta de dados, de forma programática, oriundos de quaisquer publicações publicamente visíveis.

P3: O acesso à API da plataforma é gratuito?

Verificamos se a plataforma oferece acesso geral à sua API de forma gratuita.

P4: A plataforma oferece a pesquisadores acesso gratuito e específico à API?

Neste campo, é verificado se a API da plataforma disponibiliza algum tipo de acesso a pesquisadores, seja na forma de um *token* específico ou de *endpoints* exclusivos.

P5: A plataforma oferece uma interface para coletar os dados por meio de busca customizável?

Neste campo, é verificado se a plataforma oferece uma interface de usuário para observação e coleta de dados, direcionada a pessoas sem conhecimento técnico de programação.

P6: É possível extrair os dados requisitados diretamente da resposta da API da plataforma?

Neste campo, verificamos se a API retorna dados estruturados como resposta ao pedido, ao invés de entregar um link que redireciona para os dados. Dados de mídias audiovisuais, como arquivos de imagens, vídeos e áudio, não foram considerados para a avaliação deste parâmetro.

P7: A API da plataforma provê uma forma de autenticação que permite a renovação automática, sem bloqueios à aquisição de dados?

Neste campo, é avaliado se os *tokens* disponibilizados para o uso da API expiram e se a renovação dos mesmos pode ser feita de forma automática.

P8: A criação de *tokens* de acesso à API da plataforma pode ser feita de forma gratuita?

Neste campo, foi avaliado se a plataforma permite a criação de novos *tokens* de forma gratuita, sem a necessidade de cadastramento de mais de uma conta para a utilização da API, possivelmente burlando seus termos de uso.

P9: É possível criar novos *tokens* de acesso à API da plataforma sem limitações de quantidade?

Aqui, é verificado se a plataforma limita a quantidade de *tokens* de acesso à API que podem ser criados por um mesmo usuário.

P10: O processo para que pesquisadores tenham acesso à API da plataforma é claro, descomplicado e com prazo bem definido?

Aqui, se verifica se o processo para solicitação de acesso à API exclusivo a pesquisadores é bem descrito pela plataforma. Avaliamos se a plataforma deixa clara a documentação necessária para comprovar vínculo institucional e o tempo para que o acesso seja concedido, em caso de aprovação da solicitação do pesquisador.

P11: A API da plataforma disponibiliza um *endpoint* para recuperar dados de uma publicação específica?

Neste campo, verifica-se a possibilidade de recuperar dados de uma publicação pública específica na plataforma em questão, por meio de um identificador único e não necessariamente por termos de busca ou outros parâmetros e filtros.

P12: A API da plataforma disponibiliza um *endpoint* para recuperar dados de um autor específico?

Neste campo, é verificado se a API permite a recuperação de dados de publicações públicas feitas na plataforma em questão por um autor específico, por meio de seu nome de usuário ou identificador único.

P13: A API da plataforma disponibiliza um *endpoint* para recuperar dados por meio de termos de busca?

Este campo verifica se é possível recuperar dados de publicações públicas da plataforma em questão por meio de termos de busca, ou seja, montar uma base de dados com publicações mencionando esses termos.

P14: É possível realizar aquisição de dados por raspagem, sem necessidade de autenticação, por meio da interface de usuário da plataforma?

Neste campo, é avaliado se é possível fazer a coleta dos dados por meio de técnicas de raspagem sem necessidade de criação de contas e logins.

P15: É possível realizar aquisição de dados por raspagem sem necessidade de outros dispositivos?

Esse campo verifica se há a necessidade de outros dispositivos eletrônicos (como celulares) para a realização da coleta por meios alternativos, não sendo possível realizá-la por meio da interface de usuário da plataforma em um computador.

P16: É possível recuperar os dados por meio de raspagem, sem necessidade de contornar ferramentas e técnicas que visam impedir o acesso programático aos dados?

Aqui, é verificado se a plataforma utiliza ferramentas como Cloudflare para impedir o acesso aos dados por meio de *paywall*, taxa de tráfego limitada e/ou bloqueio por detecção de comportamento automatizado.

*Conformidade (11 parâmetros)*

Avalia se a documentação oficial e os dados recuperados são de fácil compreensão e adequados a formatos comumente utilizados. Trata-se de uma dimensão exógena, ou seja, relacionada mais ao "'entorno' dos dados do que com eles próprios" e, portanto, mais atrelada "à sua governança e gerência do que ao seu próprio conteúdo" (Barbieri, 2019). Nesta dimensão, é avaliado, por exemplo, se datas e URLs são entregues em acordo com padrões internacionais, além de aspectos relacionados à documentação das APIs e aos relatórios de transparência sobre ações de moderação.

Parâmetros que compõem a dimensão de *Conformidade*

P17: A estrutura dos dados disponibilizados pela API da plataforma é estável?

Neste campo, é verificada a estabilidade da estrutura das respostas retornadas pela API. Ela é considerada estável se não mudar constantemente e sem aviso prévio de, pelo menos, 30 dias, com ampla divulgação. Também é avaliado se as mudanças na API impactam o funcionamento de aplicações integradas a elas.

P18: Os dados retornados pela API da plataforma estão em formato padronizado?

Aqui, é avaliado se os tipos de dados disponibilizados pela API correspondem ao consenso e/ou aos padrões da área como, por exemplo, o formato de datas segundo a norma ISO 8601, já que a estruturação padronizada facilita o armazenamento e a utilização dos dados.

P19: A documentação da API da plataforma é publicada em acesso aberto?

Aqui, é verificado se a plataforma publica na internet a documentação para o uso de sua API, de forma aberta, sem necessidade de cadastro e login.

P20: A documentação da API da plataforma está escrita de forma clara e exemplificada?

Aqui, é avaliado se a documentação para uso da API da plataforma está escrita de forma clara, é completa e tem exemplos de implementação.

P21: A documentação da API da plataforma descreve seus termos de uso?

Neste campo, é verificado se a documentação da API apresenta, de forma clara e sem ambiguidades, os termos para sua utilização e seus aspectos legais.

P22: A documentação descreve o formato da resposta dos *endpoints* da API da plataforma?

Neste campo, é verificado se a documentação da API descreve o formato de cada resposta, incluindo exemplos e os possíveis erros.

P23: A documentação da API da plataforma é disponibilizada nativamente em português?

Aqui, é verificado se a plataforma disponibiliza a documentação de sua API em língua portuguesa, para melhor compreensão de usuários brasileiros.

P24: A plataforma permite raspagem e outros tipos de acesso automático nos seus termos de uso?

Aqui, é verificado se a plataforma expressamente proíbe técnicas de raspagem para obtenção de dados em seus termos de uso.

P25: A plataforma produz relatórios de transparência periódicos sobre a moderação de conteúdos no Brasil e os disponibiliza publicamente, sem necessidade de requisição?

Neste campo, é verificado se a plataforma produz e disponibiliza, publicamente e sem necessidade de requisição pelas partes interessadas,

relatórios de transparência com periodicidade mínima semestral. Nestes relatórios, devem ser detalhadas informações sobre a aplicação de suas políticas de governança e as ações de moderação no Brasil, como a quantidade de publicações removidas ou restritas e/ou de usuários suspensos no país.

P26: Em seus relatórios de transparência, a plataforma indica o volume de cada tipo de violação identificada no Brasil de acordo com as políticas de moderação vigentes?

Aqui, é verificado se os relatórios de transparência da plataforma apresentam informações sobre o volume de violações identificadas, separadas por tipo de violação, durante a aplicação de suas políticas de governança e ações de moderação no Brasil. Os tipos de violação podem incluir, por exemplo, a disseminação de conteúdo ilegal, discurso de ódio e informações falsas.

P27: Os relatórios de transparência especificam informações sobre a quantidade e o tipo de requisições feitas por entes do Estado brasileiro à plataforma, além da quantidade e do tipo de solicitações acatadas?

Este campo verifica se os relatórios de transparência produzidos pela plataforma elencam os pedidos de moderação e de entrega de dados realizados por entes do Estado brasileiro, detalhando a natureza do pedido, o total de requisições e o volume de solicitações deferidas e indeferidas.

*Completude (6 parâmetros)*

Indica se os dados recuperados apresentam os atributos indispensáveis para sua compreensão e se é possível realizar um monitoramento íntegro de cada uma das plataformas analisadas ao coletá-los (Mahanti, 2018), considerando como completos os dados que podem ser utilizados e aplicados em situações diversas de pesquisa. Assim, é avaliado, por exemplo, se os dados disponibilizados pelas APIs oficiais das plataformas refletem os dados exibidos na interface de usuário e se a frequência permitida para coleta de dados por vias oficiais possibilita o monitoramento consistente.

Parâmetros que compõem a dimensão de *Completude*

P28: É possível recuperar dados dos comentários de uma publicação por meio da API da plataforma?

Aqui, verifica-se a possibilidade da recuperação de dados de comentários, incluindo o conteúdo destes, quando disponíveis na plataforma, seja junto aos dados da publicação, seja por meio de um *endpoint* específico.

P29: É possível recuperar dados de conteúdos temporários por meio da API da plataforma?

Neste campo, verificamos se a API da plataforma fornece ao menos um *endpoint* para a recuperação de dados de publicações temporárias, como stories e mensagens temporárias.

P30: É possível recuperar dados históricos por meio da API da plataforma?

Aqui, é avaliado se a API oferece *endpoints* que permitam indicar um período de tempo específico que abranja, ao menos, os últimos 365 dias para a coleta de dados, contados a partir do momento da requisição.

P31: A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 1 milhão de publicações em 24 horas?

Aqui, avalia-se a possibilidade de recuperar dados, sem interrupções e perdas, de requisições que acumulem mais de 1 milhão de publicações em 24 horas, por meio da API da plataforma, de forma a garantir condições para uma coleta de dados robusta em tempo real, sem necessidade de ser suplementada.

P32: A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 100 mil publicações em 24 horas?

Aqui, avalia-se a possibilidade de recuperar dados, sem interrupções e perdas, de requisições que acumulem mais de 100 mil publicações em 24 horas, por meio da API da plataforma, de forma a garantir

condições para uma coleta de dados satisfatória, que, a depender das necessidades do pesquisador, ainda poderia ser retomada e suplementada posteriormente.

P33: A quantidade de requisições permitidas pela API da plataforma é suficiente para monitorar mais de 10 mil publicações em 24 horas?

Aqui, avalia-se a possibilidade de recuperar dados, sem interrupções e perdas, de requisições que acumulem mais de 10 mil publicações em 24 horas, por meio da API da plataforma, de forma a garantir mínimas condições à coleta de dados, porém com mais chances de esta precisar se estender, acarretando maior demora na formação da base de dados de interesse.

*Consistência (4 parâmetros)*

Avalia se o formato e a apresentação dos dados são consistentes e idênticos em todas as bases extraídas e, em especial, em requisições idênticas entre si (Mahanti, 2018). Também verifica se os termos de busca e filtros usados recuperam dados coerentes e sem contradições, duplicações e discrepâncias. A consistência é imprescindível para produzir relatórios precisos e ágeis, pois evita a necessidade de conferência e/ou correção constante dos dados e permite maior auditabilidade.

Parâmetros que compõem a dimensão de *Consistência*

P34: Os dados retornados pela API da plataforma são persistentes?

Este campo verifica se os dados recuperados pela API da plataforma estão imunes à expiração, ainda que sejam links. Espera-se que alguns metadados das publicações removidas da plataforma não sejam excluídas da resposta da API, mas que haja uma sinalização da remoção do conteúdo.

P35: Os dados recuperados pela API da plataforma refletem o que é exibido em sua interface de usuário?

Neste campo, é verificado se os dados retornados pela API correspondem às informações exibidas na interface de usuário das plata-

formas. Deve ser possível identificar na resposta da API, por exemplo, informações como autoria, conteúdo completo e principais interações.

P36: A resposta retornada pela API da plataforma é sempre a esperada?

Neste campo, é avaliado se os dados recuperados por meio da API são sempre os mesmos, de acordo com os parâmetros e filtros utilizados, ou ao menos coerentes com outras coletas feitas de forma idêntica, excetuando-se publicações apagadas ou realizadas entre elas.

P37: A resposta retornada pela API da plataforma é coerente com os parâmetros e filtros utilizados na requisição?

Neste campo, é verificado se os dados recuperados pela API da plataforma refletem as escolhas de parâmetros e filtros determinadas no momento da requisição.

*Relevância (2 parâmetros)*

Avalia se os dados são pertinentes para a finalidade à qual se destinam (Mahanti, 2018), ou seja, se estão de acordo com os objetivos da pesquisa e da requisição, devendo ser suficientes para embasar uma análise robusta.

Parâmetros que compõem a dimensão de *Relevância*

P38: As entidades retornadas pela API da plataforma são suficientes para compreender os dados em todos os seus níveis de detalhes?

Aqui, é avaliado se a resposta da API entrega todos os dados necessários para compreender o ciclo de vida completo de uma publicação específica, incluindo comentários, compartilhamentos, respostas e outros possíveis relacionamentos, assim como seus autores e conteúdos referenciados, como nos casos de compartilhamentos e menções.

P39: A API da plataforma permite a utilização de filtros para refinar a requisição de dados?

Este campo verifica se a API da plataforma permite a utilização de filtros de busca, como a localização do publicador, idioma ou período específico, entre outros.

*Atualidade (1 parâmetro)*

Qualifica o impacto da passagem de tempo na disponibilização dos dados, de modo a avaliar a agilidade com que o processo de atualização dos dados ocorre (Mahanti, 2018).

Parâmetros que compõem a dimensão de *Atualidade*

P40: É possível recuperar dados recém-publicados, quase em tempo real à publicação, por meio da API da plataforma?

Aqui, é avaliado se a API da plataforma permite a recuperação de dados de um conjunto de publicações específicas em até uma hora após sua publicação.

## Composição das notas

Entre os 40 parâmetros avaliados, cinco parâmetros relativos à dimensão de *acessibilidade* foram desdobrados em dois critérios especiais de avaliação que compõem 50% da nota, de forma que cada critério especial corresponde a 25% da nota final. Os outros 35 parâmetros correspondem aos 50% restantes da pontuação total e valem cerca de 1,43 ponto cada.

Para os critérios especiais, consideramos os parâmetros de avaliação imprescindíveis para a realização de coletas sistemáticas e não enviesadas de dados das plataformas analisadas, de forma que um bom desempenho no restante dos parâmetros esteja condicionado a eles. Dessa forma, a distribuição das pontuações das plataformas que compõem o índice se organiza da seguinte maneira:

25 pontos correspondentes ao Critério Especial #1 ("É possível acessar pela API o universo de dados de forma gratuita para fins de pesquisa?"), alcançados por plataformas que permitam uma recuperação de dados completa, sistemática e gratuita para pesquisadores.

Para pontuar neste critério especial, é preciso atender aos parâmetros *P1* ("*A plataforma disponibiliza API oficial para acesso aos dados públicos publicados por usuários?*") e *P2* ("*O universo a ser monitorado é recuperável pela API da plataforma?*"), além de um entre o *P3* ("*O acesso*

*à API da plataforma é gratuito?*") e o *P4* ("*A plataforma oferece a pesquisadores acesso gratuito e específico à API?*"). O universo a ser monitorado é composto por todas as postagens públicas imediatamente localizáveis, acessíveis e recuperáveis por qualquer usuário. Assim, entendemos que, caso o pesquisador assim deseje, as plataformas devem disponibilizar todo o conjunto de dados das publicações públicas correspondentes às requisições feitas, ao invés de apenas disponibilizar conjuntos recortados e definidos arbitrariamente por elas, de forma a permitir a escalabilidade do monitoramento e a reprodutibilidade de pesquisas, evitando possíveis enviesamentos. Além disso, a gratuidade de acesso a estes dados é outro fator importante, uma vez que coloca em pé de igualdade todos os pesquisadores que decidirem coletá-los e analisá-los, sem quaisquer barreiras econômicas que possam reforçar o colonialismo acadêmico que privilegia o Norte Global.

25 pontos correspondentes ao Critério Especial #2 ("*A plataforma oferece uma interface para coletar os dados por meio de busca customizável?*"), alcançados por plataformas que oferecem uma interface de usuário que permita que pesquisadores com pouca ou nenhuma habilidade de programação coletem os mesmos dados fornecidos pela API oficial.

Para pontuar neste critério especial, é preciso atender ao parâmetro *P5* ("*A plataforma oferece uma interface para coletar os dados por meio de busca customizável?*"). Esse tipo de ferramenta colabora com a democratização da pesquisa e a transparência dos dados, mas sem dispensar a necessidade de APIs, já que estas garantem maior customização aos processos de coleta de dados e permitem que estes processos sejam automatizados e, então, ganhem escala.

50 pontos correspondentes ao desempenho da plataforma nos 35 parâmetros restantes, dependentes da soma de pontos obtidos a partir de avaliações positivas em relação ao total de parâmetros aplicáveis.

Assim, a pontuação final do índice é formalmente representada por:

Em que ***CEi*** é a resposta ao Critério Especial #1, valendo 25 pontos; ***CEii*** é a resposta ao Critério Especial #2, valendo 25 pontos; ***Positivas*** é o número de avaliações positivas de cada plataforma; e ***Pa-***

***râmetros*** é o número de parâmetros de avaliação restantes e aplicáveis[5] no questionário.

## Níveis de transparência de dados

Com base nas pontuações finais e para facilitar a interpretação das mesmas, classificamos e dividimos as plataformas segundo cinco níveis de transparência de dados:

- **Transparência irrelevante ou nula (0 a 20 pontos):** Plataformas que não investem em quaisquer medidas de transparência e acesso a dados. Recebem poucos pontos graças às possibilidades de raspagem de dados, que, em geral, não são oficialmente permitidas. Não costumam publicar relatórios de transparência periódicos sobre suas ações de moderação.
- **Transparência precária (21 a 40 pontos):** Plataformas que impõem barreiras técnicas, operacionais e/ou financeiras significativas às suas medidas de acesso a dados, inviabilizando o monitoramento para a maioria dos pesquisadores. Também não publicam relatórios de transparência periódicos sobre suas ações de moderação de conteúdo.
- **Transparência regular (41 a 60 pontos):** Plataformas que apresentam algumas medidas de transparência e acesso a dados, mas com diversas limitações relacionadas ao tipo de conteúdo que pode ser acessado e à amostra do universo passível de coleta. Em geral, divulgam relatórios de transparência sobre ações de moderação, mas sem o detalhamento esperado.
- **Transparência satisfatória (61 a 80 pontos):** Plataformas que disponibilizam dados sem restrições financeiras, porém com limitações quanto ao volume de dados que pode ser requisitado e/ou apresentando problemas de qualidade, especialmente

[5] Nas avaliações de WhatsApp e YouTube, desconsideramos um dos parâmetros de avaliação e readequamos o restante dos cálculos em torno desta decisão.

de consistência. Divulgam relatórios de transparência sobre suas ações de moderação no Brasil com alguma periodicidade.

- **Transparência ideal (81 a 100 pontos):** Plataformas com soluções oficiais eficientes para coleta de dados, que incluem APIs e interface de coleta de dados, com documentações bem exemplificadas e sem impedimento à raspagem. Costumam divulgar relatórios de transparência periódicos, com detalhes sobre violações e remoções a pedido do Estado.

## Resultados

Os resultados do ITD apontam que nenhuma plataforma avaliada alcança pontuação ideal no que diz respeito às medidas de transparência e de acesso a dados de conteúdos públicos gerados por usuários e à qualidade dos dados retornados. A plataforma mais bem avaliada é o YouTube, com 63,2 pontos, garantindo-lhe um nível de transparência considerado satisfatório.

A seguir, apresentamos uma visão geral do que foi observado em cada plataforma analisada. O detalhamento e as justificativas completas das avaliações de cada plataforma estão disponíveis no site do NetLab UFRJ[6].

## YouTube

*Transparência de dados satisfatória*

Dentre as plataformas pesquisadas, o YouTube foi a que obteve a melhor pontuação, de 63,2 pontos, sendo sua transparência de dados considerada satisfatória[7]. A plataforma disponibiliza uma API oficial (YouTube, [*S.d.*]) de acesso gratuito para qualquer usuário (*P1* e *P3*),

[6] Disponível em https://netlab.eco.ufrj.br/itd.

[7] Para o cálculo final da nota do YouTube, desconsideramos o parâmetro *P29* ("É possível recuperar dados de conteúdos temporários por meio da API da plataforma?"), referente à dimensão de completude, uma vez que a plataforma não permite que usuários publiquem conteúdos temporários. Portanto, além dos dois critérios especiais, consideramos 34 dos 35 parâmetros aplicáveis e readequamos todos os demais cálculos.

permitindo a busca em todo o universo de vídeos considerados públicos da plataforma (*P2*), pontuando no Critério Especial #1. A plataforma também disponibiliza uma API específica para pesquisadores (*P4*), cujo acesso pode ser solicitado facilmente (*P10*). Positivamente, o YouTube também permite a recuperação ágil de dados recém-publicados (*P40*) e de dados históricos (*P30*). Por outro lado, a plataforma não pontua no Critério Especial #2 em virtude da ausência de uma interface de coleta de dados (*P5*).

Parâmetros relativos às dimensões de consistência e relevância prejudicam o desempenho da plataforma, uma vez que encontramos diversos problemas relacionados a respostas inconsistentes (*P36*) e com dados incoerentes em relação aos parâmetros e filtros aplicados nos momentos da coleta (*P37*). Além disso, os dados recuperados pela API da plataforma não são persistentes (*P34*): ela não disponibiliza metadados de publicações removidas, ao contrário da interface da plataforma, que exibe uma mensagem sobre a remoção e seus motivos, caso o usuário tente assistir a um vídeo que não está mais no ar.

Em seu "*Relatório sobre cumprimento das diretrizes da comunidade do YouTube*" (Google, [*S.d.*]), a plataforma disponibiliza trimestralmente o volume de vídeos removidos que foram publicados por usuários no Brasil e que violaram suas diretrizes (*P25*). Entretanto, não são disponibilizadas informações como o número de canais removidos, o número de remoções com base em denúncias e agregações por tipo de violações no país (*P26*), apesar desse tipo de informação ser providenciada em outros países sobre os quais o YouTube publica relatórios específicos. A plataforma também aparece em um relatório semestral junto a outros serviços associados ao Google, no qual constam o volume de solicitações governamentais de remoção de conteúdo por tipo de violação no Brasil (*P27*).

## Facebook

*Transparência de dados regular*

A pontuação da transparência de dados do Facebook de 53,6 pontos é considerada regular. Quase metade da nota da plataforma dependia dos 25 pontos recebidos graças ao CrowdTangle ([*S.d.*]), interface (*P5*) e API (*P1*) para coleta de dados que a Meta disponibilizava[8] para pesquisadores e jornalistas coletarem amostras de dados públicos do Instagram e do Facebook.

No entanto, a API do CrowdTangle apresentava várias limitações. A mais importante delas era a impossibilidade de recuperação de todo o universo de dados públicos da plataforma (*P2*), uma vez que só podiam ser recuperados dados de publicações feitas: (i) por páginas com mais de 25 mil seguidores ou curtidas, (ii) por perfis verificados, e (iii) em grupos públicos com mais de 95 mil membros (CrowdTangle, [*S.d.*]). A ferramenta não permitia acesso a comentários vinculados às publicações originais (*P28*) ou a conteúdos temporários (*P29*), como *stories*, o que prejudicava a completude dos dados retornados. Além disso, também não indicava quando conteúdos eram removidos da plataforma (*P34*): publicações deletadas eram tratadas pelo CrowdTangle como se nunca tivessem existido.

Na Central de Transparência da Meta (Meta, [*S.d.*]a), a empresa disponibiliza semestralmente informações sobre publicações, perfis e comentários moderados em suas plataformas conforme as leis locais de cada país em que atua, incluindo o Brasil (*P25*). A Meta também divulga dados semestrais de requisições feitas por entes do Estado brasileiro para a moderação de conteúdo em suas plataformas (*P27*) (Meta, [*S.d.*] b). Porém, dados sobre o total de ações de moderação, incluindo remo-

[8] Durante a elaboração deste índice, a Meta anunciou que o serviço seria oficialmente descontinuado a partir de agosto de 2024 (CrowdTangle, 2024). Com o objetivo de suprir a lacuna deixada pela extinção do CrowdTangle, a Meta lançou sua nova Biblioteca de Conteúdo (Clegg, 2023). No entanto, até o encerramento deste trabalho, não havíamos conseguido aprovação para utilização da ferramenta, ainda subutilizada em todo o mundo (Iyer, 2024) e com relatos de problemas (Clark, 2024; Iyer, 2024) que, se confirmados, tendem a fazer com que a pontuação do Facebook seja ainda menor em avaliações futuras.

ções por violação dos termos de uso da Meta, são disponibilizados apenas a nível global ou para alguns países com relatórios específicos — o que não é o caso do Brasil (*P26*).

## Instagram

*Transparência de dados regular*

Tendo atingido 52,1 pontos, a pontuação da transparência de dados do Instagram é considerada regular. Assim como no Facebook, utilizamos o serviço CrowdTangle para responder aos parâmetros de avaliação da plataforma. A ferramenta oferecia[9] uma API oficial (CrowdTangle, [*S.d.*]) para extração de dados (*P1*), além de uma interface de usuário dedicada à coleta de dados (*P5*).

Como também ocorria com o Facebook, porém, o CrowdTangle não permitia a recuperação de todo o universo de dados públicos da plataforma (*P2*), uma vez que apenas dados de publicações de perfis com mais de 50 mil seguidores e/ou verificados podiam ser recuperadas, e pelo fato de dados gerados a partir de *reels* não poderem ser requisitados. O Instagram também foi penalizado pela falta de completude dos dados retornados pelo CrowdTangle, que não permitia a coleta de comentários e stories (*P28* e *P29*), por exemplo.

Como ocorria com o Facebook, o CrowdTangle não indicava quando uma publicação era removida pelo Instagram (*P34*). Ainda como faz para o Facebook, a Meta disponibiliza semestralmente informações, discriminadas a nível de país, sobre publicações, perfis e comentários moderados no Instagram (*P25*) e um breve relatório sobre requisições governamentais de moderação e/ou acesso a dados da plataforma feitas por entes do Estado Brasileiro (*P27*) (Meta, [*S.d.*]b).

A principal diferença entre as avaliações do Facebook e do Instagram reside na dimensão de consistência. No caso do Instagram, o CrowdTangle retornava links de redirecionamento para acessar as ima-

[9] O Instagram está, no geral, sujeito aos mesmos problemas de transparência e acesso a dados que o Facebook, graças ao encerramento do CrowdTangle e à introdução da Biblioteca de Conteúdo da Meta. Para mais, ver nota anterior.

gens presentes nas publicações identificadas e recuperadas, mas estes expiravam quase instantaneamente e não podiam ser analisados de forma sistemática, fazendo com que os dados coletados não refletissem o que era exibido na interface da plataforma (*P35*). Isso é especialmente problemático no caso do Instagram, já que os conteúdos publicados são majoritariamente visuais, em formatos de foto e/ou vídeo.

## X/Twitter

*Transparência de dados precária*

Marcando apenas 30 pontos, a transparência de dados da plataforma é considerada precária. No passado, o X/Twitter havia se consolidado como uma das plataformas mais acessíveis para coleta de dados públicos (Zuckerman, 2021). No entanto, em 2023, a API do X/Twitter passou a condicionar o acesso aos dados à adesão a planos pagos (X/Twitter, [*S.d.*]a), com taxas proibitivas para a recuperação de volumes substanciais de dados. São eles o plano *Basic*, que permite a recuperação de dados de 10.000 publicações ao custo mensal de US$100 (cerca de R$599,00, segundo cotação de novembro de 2024); o *Pro*, que permite a recuperação de dados de 1 milhão de publicações ao custo mensal de US$5.000 (cerca de R$29,8 mil, segundo cotação de novembro de 2024); e o plano *Enterprise*, que custa a partir de US$42.000 (cerca de R$251,1 mil, segundo cotação de novembro de 2024). Para a avaliação da plataforma, consideramos as funcionalidades dos planos *Basic* e *Pro*, já que o plano *Enterprise* é apenas comercializado para empresas aprovadas mediante solicitação (X/Twitter, [*S.d.*]b).

Com a implementação destes planos pagos, o X/Twitter deixou de garantir uma forma gratuita de acesso aos dados para qualquer usuário (*P3*), e não garante mais qualquer forma de acesso específica para pesquisadores brasileiros (*P4*). Embora alegue disponibilizar dados para pesquisadores que atuam na União Europeia por conta da vigência do DSA, a plataforma é acusada de nem sequer cumprir devidamente esta determinação (European Commission, 2023).

Devido à alta qualidade dos dados retornados por meio da API, ainda que paga, a plataforma garante pontos nas dimensões de completude e consistência. Entre eles, o X/Twitter é a única plataforma que sinaliza adequadamente publicações removidas nas respostas de sua API (*P34*). Porém, a plataforma não disponibiliza nenhum relatório com as informações totais sobre conteúdos removidos no Brasil ou submetidos a outras formas de moderação, como faz para outros países (*P25*).

## Telegram

*Transparência de dados precária*

Também somando 30 pontos, a transparência de dados do Telegram é considerada precária. A plataforma disponibiliza uma API oficial e gratuita (*P1* e *P3*), possibilitando o acesso programático a dados de maneira confiável, consistente e, em grande parte, em conformidade com boas práticas. No entanto, a plataforma só permite buscar por mensagens sobre um determinado tema em grupos ou canais conhecidos previamente. Como não há maneira de buscar na API por grupos e canais públicos, o acesso ao universo de dados é limitado (*P2*), o que contribui para sua baixa pontuação. A plataforma também não oferece uma interface de usuário dedicada à coleta de dados (*P5*).

Por um lado, o Telegram é a única plataforma analisada que atende a todos os critérios de completude por permitir a coleta de dados de comentários (*P28*), dados de conteúdos temporários (*P29*) e dados históricos (*P30*), além de garantir a recuperação de um grande volume de dados sem dificuldades (*P31*, *P32* e *P33*). Por outro, a plataforma é penalizada na dimensão de conformidade. O Telegram não publica relatórios de transparência públicos e periódicos, o que impossibilita a identificação de conteúdos moderados no Brasil (*P25*, *P26* e *P27*). A documentação de sua API também é deficiente por não definir diretamente como chamar *endpoints* (*P22*), não descrever os formatos dos dados retornados (*P18*) e por não ser disponibilizada em português (*P23*).

## TikTok

*Transparência de dados irrelevante*

O TikTok não oferece acesso a uma API gratuita de coleta de dados oficial para interessados no Brasil, apenas para pesquisadores de instituições localizadas nos Estados Unidos e na Europa (TikTok, [*S.d.*] a). Esta restrição privilegia países do Norte Global, que têm acesso a meios mais efetivos para avaliar os impactos da plataforma, aprofundando assimetrias e desigualdades entre diferentes regiões.

Em grande parte por conta disso, a transparência de dados da plataforma alcança uma pontuação considerada irrelevante, de apenas 7,1 pontos. Além da indisponibilidade de API no país (*P1*) e da consequente impossibilidade de se acessar programaticamente todo o universo de dados públicos de interesse (*P2*), a plataforma também não oferece qualquer interface de usuário dedicada à coleta de dados (*P5*).

Os poucos pontos alcançados pela plataforma na dimensão de acessibilidade decorrem, em parte, das possibilidades de raspagem de dados a partir da interface web da plataforma (*P14* e *P15*). No entanto, a necessidade de contornar processos de verificação por *Captcha*, empregados pela plataforma para se evitar o uso automatizado, limita substancialmente o processo de coleta de dados (*P16*). O TikTok, além disso, é a única plataforma a pontuar nos três parâmetros referentes a relatórios públicos de transparência (*P25*, *P26* e *P27*) por disponibilizar informações sobre o volume e tipos de violações moderadas pela plataforma (TikTok, 2024) e por apontar, por meio de relatórios semestrais, os pedidos de moderação e requisições de dados feitas pelo Estado brasileiro desde 2019 (TikTok, 2023).

## Kwai

*Transparência de dados irrelevante*

O Kwai não disponibiliza API para coleta de dados no Brasil ou em qualquer outro lugar do mundo (*P1*), inviabilizando a auditabilidade do universo de conteúdos públicos que são publicados na plataforma (*P2*). A plataforma tampouco disponibiliza uma interface de usuário

dedicada à coleta de dados (*P5*). Outro ponto negativo é a ausência de relatórios de transparência dedicados a publicações removidas e usuários suspensos no Brasil (*P25*, *P26* e *P27*), uma vez que o Kwai disponibiliza informações de transparência apenas no nível do continente (Kwai, [*S.d.*]). Estes fatores contribuem amplamente para os 4,3 pontos alcançados na avaliação de sua transparência de dados, considerada irrelevante.

Os poucos pontos obtidos pelo Kwai na avaliação do ITD decorrem de três parâmetros da dimensão de acessibilidade, devido à possibilidade de raspagem de dados (*P14*, *P15* e *P16*) sem necessidade de autenticação em interface web da plataforma. No entanto, os dados extraídos são limitados e enviesados pelo baixo volume de publicações retornadas a cada busca em sua interface padrão de utilização.

## WhatsApp

### *Transparência de dados irrelevante*

O WhatsApp aparece em último lugar na nossa avaliação, marcando 1,5 ponto apenas na dimensão de acessibilidade e com transparência de dados considerada irrelevante[10]. A plataforma não disponibiliza API oficial (*P1*), diferentemente do Telegram, seu concorrente direto, e tampouco interface dedicada à coleta de dados (*P5*).

O WhatsApp teve resposta positiva em apenas um dos parâmetros de avaliação, relativo ao processo de raspagem de dados (*P16*). De todo modo, ainda que seja possível obter dados pela raspagem, é impossível ter acesso total ao universo de mensagens de interesse no WhatsApp, visto que a coleta se limita a uma amostra de grupos previamente selecionados pelos pesquisadores. A falta de transparência, influenciada

[10] Para o cálculo final da nota do WhatsApp, desconsideramos o parâmetro *P28* ("É possível recuperar dados dos comentários de uma publicação por meio da API da plataforma?"), referente à dimensão de completude, uma vez que a plataforma não apresenta suporte para comentários em suas mensagens. Portanto, além dos dois critérios especiais, consideramos 34 dos 35 parâmetros aplicáveis restantes e readequamos todos os demais cálculos.

também pela ausência de relatórios públicos que apresentem as ações de moderação empreendidas pela plataforma (*P25*), acende um preocupante sinal de alerta, uma vez que o WhatsApp é visto como uma das principais plataformas para a disseminação de desinformação no Sul Global, especialmente no Brasil (Kalogeropoulos; Rossini, 2023).

## Boas e más práticas na disponibilização de dados de conteúdos gerados por usuários

A partir das evidências observadas, apresentamos um panorama de medidas que devem ser amplamente adotadas ou evitadas pelas plataformas de redes sociais a fim de se garantir um nível ideal de transparência e disponibilização de dados para pesquisas de interesse público no Brasil. A maioria das plataformas analisadas não oferece uma API gratuita e nem uma interface de usuário para coleta de dados no país. Apenas o Facebook e o Instagram permitiam a coleta de dados por ambas API e interface de usuário, mas a ferramenta CrowdTangle, que as reunia, foi descontinuada em agosto de 2024. Ainda assim, só era possível coletar amostras enviesadas e muito limitadas dos dados de ambas as plataformas.

Dentre as plataformas que ainda oferecem uma API de coleta de dados funcional no Brasil, apenas o YouTube permite a recuperação gratuita de seu universo de dados públicos referentes a conteúdos gerados por usuários. Apesar de o X/Twitter não necessariamente limitar quais dados de postagens públicas podem ser descobertos a partir de sua API, sua utilização é paga em dólares americanos desde 2023, afetando de forma desproporcional países do Sul Global, cujos pesquisadores não só desfrutam de menos recursos financeiros, mas também enfrentam taxas de câmbio adversas. O TikTok disponibiliza API oficial para coleta de dados públicos de suas publicações nos Estados Unidos e na Europa, porém não no Brasil, o que constitui medida discriminatória facilmente evitável.

Mesmo que os pesquisadores aceitem aderir aos planos pagos da API do X/Twitter, há níveis de utilização com diferentes limitações de

volume de dados. Por exemplo, o plano *Basic*, que custa US$100 mensais, permite a recuperação de dados de 10 mil publicações todo mês, enquanto o *Pro*, que custa US$5.000 mensais, permite a recuperação de dados de até 1 milhão de publicações. No entanto, a limitação a um baixo volume de dados coletáveis não é um problema apenas do X/Twitter ou de quaisquer outras plataformas que venham a cobrar pelo acesso de sua API. Mesmo as plataformas que fornecem meios gratuitos para coletar dados impõem restrições que dificultam a otimização de extrações simultâneas ou muito volumosas, como limitações à criação de novos *tokens* de acesso à API.

A API do CrowdTangle, por exemplo, não permitia a recuperação de mais de 10 mil publicações do Facebook e do Instagram por vez e abertamente desencorajava seu uso para extrações de grandes volumes de dados. A API do Telegram não limita, explicitamente, o volume de dados que podem ser requisitados, mas a equipe do NetLab UFRJ foi banida de utilizá-la repetidas vezes sem quaisquer justificativas por parte da plataforma, mesmo quando questionada através de seus canais de contato oficiais.

Além de disponibilizar uma API e uma interface de usuário para coleta de dados, é fundamental que estas permitam a aplicação de filtros de pesquisa personalizáveis, tais como a seleção de páginas e perfis de interesse, buscas por palavras-chave e seleção de idiomas. Filtros de datas que permitam a recuperação de dados históricos, como disponibilizados por Facebook, Instagram e YouTube, também são essenciais para a realização de análises longitudinais. No entanto, os nossos testes mostraram que o YouTube frequentemente entrega dados inconsistentes com os filtros de busca aplicados nas requisições, prejudicando negativamente a relevância, pertinência e replicabilidade de pesquisas realizadas sobre a plataforma.

A fim de garantir maior completude dos dados, é essencial que estas APIs e interfaces de usuário forneçam acesso a dados de publicações de diferentes naturezas, incluindo conteúdos temporários, como *stories*, como seria essencial no caso de Facebook e Instagram, mensagens tem-

porárias, como permite a API do Telegram, e comentários, tal qual faz a API do YouTube. Independentemente do tipo de dado, é imprescindível que URLs retornadas não expirem rapidamente, sobretudo se forem referentes a publicações que ainda continuam no ar, como acontecia com o Instagram, e que publicações removidas sejam devidamente identificadas, mesmo que o conteúdo delas permaneça restrito, como no caso do X/Twitter. Positivamente, YouTube, Telegram, X/Twitter, Facebook e Instagram, as plataformas que, no momento da análise, ofereciam APIs oficiais para coleta de dados no Brasil, disponibilizaram dados em tempo real para coleta e análise em nossos testes.

O WhatsApp e o Kwai foram as plataformas de pior desempenho em nossa análise, deixando de pontuar em cinco das seis dimensões avaliadas. Os únicos parâmetros em que pontuam, entretanto, decorrem da falta de supervisão de seus próprios termos de uso, já que não impedem a raspagem de dados por completo, mesmo que expressamente desautorizem a prática. Consideramos que a autorização da raspagem de dados – que, atualmente, só não é explicitamente proibida pelo Telegram – precisa ser concedida a pesquisadores segundo condições específicas, conferindo-lhes maior segurança jurídica.

As documentações das APIs de coleta de dados das plataformas precisam ser disponibilizadas publicamente, sem necessidade de requisição individual, e em português brasileiro, algo a que apenas o YouTube adere por completo. Idealmente, a documentação oficial da API deve descrever possíveis erros ao usar cada *endpoint* disponível e oferecer exemplos claros e informações detalhadas sobre como executar solicitações de dados. Além de não ser traduzida para o português brasileiro, a documentação da API do Telegram se destaca negativamente pela ausência de exemplos claros para sua utilização e pela má redação.

Por fim, o TikTok é a única plataforma que disponibiliza relatórios de transparência de suas ações de moderação e governança detalhados e periódicos, pontuando em todos os parâmetros de avaliação adequados. Embora Facebook e Instagram também publiquem relatórios de transparência, eles não apresentam o volume de cada tipo de violação

identificada no Brasil, sendo mais robustos em países com legislações que estabelecem critérios mínimos para sua divulgação. Já X/Twitter, Telegram, Kwai e WhatsApp não disponibilizam quaisquer relatórios de transparência sobre o contexto brasileiro, o que é especialmente preocupante por conta de seu papel em campanhas de desinformação, majoritariamente focadas no comportamento do eleitor.

## Perspectivas e considerações finais

Os resultados de nossas avaliações apontam que a pesquisa qualificada e de interesse público com base em dados sobre conteúdos gerados por usuários nas plataformas de redes sociais está sob forte ameaça, reforçando o preocupante cenário de apagão relatado por pesquisadores do mundo todo.

No Brasil, apenas o YouTube garante acesso satisfatório a dados para a elaboração de pesquisas de interesse público, apesar de longe do ideal. Facebook e Instagram foram avaliadas como plataformas regulares, mas suas notas dependem do CrowdTangle, ferramenta de transparência descontinuada em agosto de 2024, ainda disponível à época da elaboração do ITD. O X/Twitter, antes uma das plataformas mais abertas para pesquisa e uma das mais estudadas globalmente, alcançou uma pontuação considerada precária, graças a seu progressivo fechamento desde o fim de 2022, bem como o Telegram. Já TikTok, Kwai e WhatsApp, plataformas cada vez mais utilizadas pelos brasileiros, não promovem medidas de transparência amplas e relevantes. Mesmo que o acesso a uma API gratuita seja primordial para sistematizar e automatizar processos de coleta de dados em larga escala, entendemos que a experiência autorregulatória de plataformas que já promoveram interfaces de usuário para coleta de dados evidencia que esta é uma prática que poderia e deveria ser adotada de modo mais amplo.

É necessário, ainda, que o acesso a quaisquer ferramentas de coleta de dados seja gratuito, ao menos para pesquisadores, para não criar barreiras econômicas ao desenvolvimento de pesquisas de interesse público, especialmente em países do Sul Global, cujos pesquisadores

têm menos recursos para a pesquisa científica. Estima-se que o acesso a apenas 0,3% das publicações feitas mensalmente no X/Twitter, que passou a cobrar pelo uso de sua API no início de 2023, custaria cerca de US$500 mil, ou R$3 milhões, na cotação de novembro de 2024, o que pesquisadores classificam como "dinheiro demais para dados de menos" (Stokel-Walker, 2023). Ao dificultar o acesso sistemático e gratuito a dados, as plataformas impõem barreiras técnicas e econômicas à pesquisa de interesse público, comprometendo a representatividade das evidências e prejudicando a reprodutibilidade das análises.

Mesmo quando as plataformas oferecem meios oficiais de coleta de dados, persistem restrições que impedem que os processos de extração sejam otimizados e ganhem escala, com inconsistências frequentes, resultados insuficientes e limitações para a criação de novos *tokens* de acesso à API, prejudicando a completude e a relevância dos dados. Em nossa avaliação, entendemos que pouco adianta disponibilizar ferramentas de coleta de dados se elas limitarem substancialmente o volume a ser coletado, considerando que determinadas discussões podem atingir quantidades massivas de publicações a serem monitoradas. Todas as APIs disponíveis e analisadas apresentam problemas em diversas dimensões de qualidade. Por exemplo, o YouTube entrega dados inconsistentes e incoerentes com os parâmetros de busca e filtros aplicados, o que ocasiona ruídos e prejudica significativamente sua relevância para pesquisa

Para facilitar e democratizar o acesso aos recursos existentes, é fundamental que suas documentações sejam publicadas em acesso aberto e em língua portuguesa – ou, no caso de o *framework* analítico do índice ser aplicado a outros contextos locais, em quaisquer idiomas aplicáveis. Muitas das documentações analisadas na formulação deste índice não apresentam o detalhamento esperado ou são escritas de maneira pouco clara. A ausência de exemplos numerosos de diferentes situações de uso e erros também é evidente. Há uma necessidade clara de aprimorar as documentações atualmente existentes, diante do entendimento deste processo como uma prática primária de transparência.

Sem uma padronização mínima e a adoção de critérios vinculativos básicos, nota-se que a disponibilização de relatórios de transparência sobre a moderação das plataformas no Brasil é muito precária. Como esta é outra medida autorregulatória das plataformas para promover uma transparência controlada, os relatórios analisados frequentemente apresentam dados com granularidade insatisfatória e variada, de forma que eles não podem ser comparados entre si. O ideal é que os relatórios que tratam de ações de moderação sejam disponibilizados a nível do país. Contudo, aqueles disponibilizados por algumas das plataformas analisadas só as abordam a nível continental, o que ainda é insuficiente para análises aprofundadas e contextualizadas de sua governança.

A título de exemplo, o TikTok se destaca positivamente na transparência de suas ações de moderação, sejam elas proativas ou a pedidos de entes governamentais. Já o X/Twitter mantém uma página destinada à publicação de relatórios de transparência no Brasil em sua Central de Transparência, mas esta não é atualizada desde 2021, o que serve, no fim, para alimentar as alegações de que as plataformas direcionam esforços para manter um "teatro da transparência", em que dificultam o acesso e fornecem dados incompletos (Bouko; van Ostaeyen; Voué, 2021). Assim, recomendamos enfaticamente a publicação periódica destes relatórios, com detalhamento do volume de publicações, comentários e perfis removidos e suspensos destas plataformas e dos motivos que levaram à moderação. Consideramos que a transparência das políticas e práticas de moderação de conteúdo das plataformas de redes sociais é o principal caminho para a garantia dos direitos dos usuários em rede, de modo a assegurar a justiça na aplicação das diretrizes de uso das plataformas e a permitir que violações ao direito à liberdade de expressão e vieses em sistemas algorítmicos de remoção de conteúdo sejam mais facilmente detectados e tratados.

Ainda neste ponto, também é importante que as plataformas sinalizem, nas respostas de suas APIs, quando publicações são removidas e usuários são suspensos, oferecendo acesso a metadados de moderação, ainda que o conteúdo das publicações fique restrito. Esse tipo de dado

é essencial para a pesquisa de temas relacionados à desinformação e à circulação de conteúdos ilegais ou abusivos nas plataformas de redes sociais, bem como para compreender suas práticas de governança.

Como também será discutido no Índice de Transparência da Publicidade nas Plataformas de Redes Sociais, no próximo capítulo, as principais propostas de regulação de plataformas de redes sociais no mundo não se fundamentam em uma perspectiva de transparência aliada à qualidade dos dados. Isso é especialmente relevante quando consideramos que algumas plataformas, incluindo o X/Twitter e o TikTok, oferecem bases de dados com diferentes níveis de transparência, qualidade e completude entre países do Norte e do Sul Global. Diferenças regionais significativas na transparência de dados de interesse público sugerem que é preciso estabelecer critérios mínimos satisfatórios de acesso e qualidade como obrigação das plataformas, de forma justa e igualitária em todos os territórios onde elas atuam.

## Referências


ANPD. Autoridade Nacional de Proteção de Dados. Glossário de proteção de dados pessoais e privacidade: 2.0. *Governo Federal*, [*S.d.*]. Disponível em: https://www.gov.br/anpd/pt-br/documentos-e-publicacoes/glossario-anpd-protecao-de-dados-pessoais-e-privacidade.pdf. Acesso em: 3 abr. 2024.

BAR-ILAN, Judit. Data collection methods on the Web for infometric purposes — A review and analysis. *Scientometrics*, [*S.l.*], v. 96, n. 3, p. 7-32, 2001. Disponível em: https://link.springer.com/article/10.1023/A:1005682102768. Acesso em: 9 abr. 2024.

BARBIERI, Carlos. *Governança de Dados*: Práticas, conceitos e novos caminhos. Rio de Janeiro: Alta Books, 2019.

BATINI, Carlo; SCANNAPIECO, Monica. *Data Quality Concepts, Methodologies and Techniques*. Nova Iorque: Springer Berlin Heidelberg, 2006.

BIANCHI, Tiago. Topic: WhatsApp in Brazil. *Statista*, [*S.l.*], 10 jan. 2024. Disponível em: https://www.statista.com/topics/7731/whatsapp-in--brazil/. Acesso em: 3 abr. 2024.

BOSSETTA, Michael. Scandalous Design: How Social Media Platforms' Responses to Scandal Impacts Campaigns and Elections. *Social Media + Society*, [*S.l.*], v. 6, n. 2, p. 1-4, jun. 2020. Disponível em: https://journals.sagepub.com/doi/10.1177/2056305120924777. Acesso em: 9 abr. 2024.

BOUKO, Catherine; VAN OSTAEYEN, Pieter; VOUÉ, Pieter. Facebook's policies against extremism: Ten years of struggle for more transparency. *First Monday*, [*S.l.*], v. 26, n. 9, p. 1–22, 2021. Disponível em: https://firstmonday.org/ojs/index.php/fm/article/view/11705. Acesso em: 01 ago. 2024.

BRAGA, Luciana. CUNHA, Brenda. Violações dos direitos de povos tradicionais e as barreiras de acesso à informação ambiental: Uma análise em transparência ativa. *In*: Fórum de Direito de Acesso a Informações Públicas (Org.). *A LAI É 10*: O Brasil após uma década da Lei de Acesso à Informação. São Paulo: Abraji, 2022. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/a_lai_e_10_ebook.pdf. Acesso em: 9 abr. 2024.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 15 ago. 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/Lei/L13709.htm. Acesso em: 10 abr. 2024.

BROMELL, David. *Regulating Free Speech in a Digital Age*: Hate, Harm and the Limits of Censorship. Cham: Springer International Publishing, 2022.

BRUNS, Axel. After the 'APIcalypse': social media platforms and their fight against critical scholarly research. *Information, Communication & Society,* [*S.l.*], v. 22, n. 11, p. 1544–1566, 19 set. 2019. Disponível em: https://www.tandfonline.com/doi/full/10.1080/1369118X.2019.1637447. Acesso em: 9 abr. 2024

CALVO-GUTIÉRREZ, Elvira.; MARÍN-LLADÓ, Carles. Combatting Fake News: A Global Priority Post COVID-19. *Societies*, [*S.l.*], v. 13, n. 7, p. 1–13, 2023. Disponível em: https://www.mdpi.com/2075-4698/13/7/160. Acesso em: 15 abr. 2024.

CECILIA, José M.; CANO, Juan Carlos; HERNÁNDEZ-ORALLO, Enrique.; CALAFATE, Carlos T.; MANZONI, Pietro. Mobile crowdsensing approaches to address the COVID-19 pandemic in Spain. *IET Smart Cities*, [*S.l.*], v. 2, n. 2, p. 58–63, 2020. Disponível em: https://ietresearch.onlinelibrary.wiley.com/doi/full/10.1049/iet-smc.2020.0037. Acesso em: 9 abr. 2024.

CIOFFI-REVILLA, Claudio. *Introduction to Computational Social Science*: Principles and Applications. Cham: Springer International Publishing, 2018.

CLARK, L. Researchers find Meta's withdrawal of misinformation tool hard to swallow. *The Register*, [*S.l.*], 18 jun. 2024. Disponível em: https://www.theregister.com/2024/06/18/metas_decision_to_withdraw_misinformation/. Acesso em: 10 fev. 2025.

CLEGG, Nick. New Tools to Support Independent Research. *Meta*, [*S.l.*], 21 nov. 2023. Disponível em: https://about.fb.com/news/2023/11/new-tools-to-support-independent-research/. Acesso em: 10 fev. 2025.

CLOUDFLARE. O que é um endpoint de API?. *CloudFlare*, [*S.d.*]. Disponível em: https:// www.cloudflare.com/pt-br/learning/security/api/what-is-api-endpoint/. Acesso em: 3 abr. 2024.

COALITION FOR INDEPENDENT TECHNOLOGY RESEARCH. Letter: Imposing Fees to Access the Twitter API Threatens Public-Interest Research. *Coalition for Independent Technology Research*, 2023. Disponível em: https://independenttechresearch.org/letter-twitter-api-access-threatens-public-interest-research/. Acesso em: 4 abr. 2024.

CONGER, Kate. LinkedIn sues anonymous data scrapers. *TechCrunch*, [*S.l.*], 15 ago. 2016. Disponível em: https://techcrunch.com/2016/08/15/linkedin-sues-scrapers/. Acesso em: 15 abr. 2024.

CROWDTANGLE. What data is CrowdTangle tracking? *CrowdTangle*, [*S.d.*]. Disponível em: https://help.crowdtangle.com/en/articles/1140930-what-data-is-crowdtangle-tracking. Acesso em: 4 abr. 2024.

CROWDTANGLE. Important Update to CrowdTangle | March 2024. *CrowdTangle*, 2024. Disponível em: https://help.crowdtangle.com/en/arti-

cles/9014544-important-update-to-crowdtangle-march-2024. Acesso em: 3 abr. 2024.

DATA REPORTAL. Digital 2024: Brazil. *Data Reportal*, 23 fev. 2024. Disponível em: https://datareportal.com/reports/digital-2024-brazil. Acesso em: 3 abr. 2024.

DHELIM, Sahraoui; CHEN, Liming; DAS, Sajal K.; NING, Huansheng; NUGENT, Chris; LEAVEY, Gerard; PESCH, Dirk; BANTRY-WHITE, Eleanor; BURNS, Devin. Detecting Mental Distresses Using Social Behavior Analysis in the Context of COVID-19: A Survey. *ACM Computing Surveys,* [*S.l.*], v. 55, n. 14, p. 1-30, 17 jul. 2023. Disponível em: https://dl.acm.org/doi/10.1145/3589784. Acesso em: 9 abr. 2024.

DOBBER, Tom; KRUIKEMEIER, Sanne; HELBERGER, Nanali; GOODMAN, Ellen. Shielding citizens? Understanding the impact of political advertisement transparency information. *New Media & Society*, [*S.l.*], v. 26, n. 11 p. 6715–6735, 2023. Disponível em: https://doi.org/10.1177/14614448231157640. Acesso em: 1 ago. 2024.

EDELMANN, Achim; WOLFF, Tom; MONTAGNE, Danielle; BAIL, Christopher A. Computational Social Science and Sociology. *Annual review of sociology*, [*S.l.*], v. 46, n. 1, p. 61-81, 2020. Disponível em: https://www.annualreviews.org/content/journals/10.1146/annurev-soc-121919-054621. Acesso em: 7 fev. 2025

EDWARDS, A.; HOUSLEY, W.; WILLIANS, M.; SLOAN, L.; WILLIANS, M. Digital social research, social media and the sociological imagination: Surrogacy, augmentation and re-orientation. *International Journal of Social Research Methodology*, [*S.l.*], v. 16, n. 3, p. 245–260, maio 2013. Disponível em: https://www.tandfonline.com/doi/abs/10.1080/13645579.2013.774185. Acesso em: 9 abr. 2024.

EL-SAYED, Abdulrahman M.; SCARBOROUGH, Peter; SEEMANN, Lars; GALEA, Sandro. Social network analysis and agent-based modeling in social epidemiology. *Epidemiologic Perspectives & Innovations*, [*S.l.*], v. 9, n. 1, p. 1–9, fev. 2012. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC3395878/. Acesso em: 9 abr. 2024.

ELLISON, Nicole; BOYD, Danah M. Sociality through social network sites. *In*: DUTTON, W. H. (Ed.). The Oxford Handbook of Internet Studies. Oxford: *Oxford University Press*, 2013. Disponível em: https://academic.oup.com/edited-volume/34364. Acesso em: 7 fev. 2025.

EUROPEAN COMMISSION. Commission opens formal proceedings against X under the DSA. *European Commission*, Bruxelas, 18 dez. 2023. Disponível em: https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6709. Acesso em: 3 abr. 2024

EUROPEAN COMMISSION. The EU's Digital Services Act. *European Commission*, [*S.l.*], [*S.d.*]. Disponível em: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en. Acesso em: 3 abr. 2024.

EVANGELISTA, Rafael; BRUNO, Fernanda. WhatsApp and political instability in Brazil: Targeted messages and political radicalisation. *Internet Policy Review*, [*S.l.*], v. 8, n. 4, p. 1–23, 31 dez. 2019. Disponível em: https://policyreview.info/articles/analysis/whatsapp-and-political-instability-brazil-targeted-messages-and-political. Acesso em: 9 abr. 2024.

FREELON, Deen. Computational research in the post-API age. *Political Communication*, [*S.l.*], v. 35, n. 4, p. 665-668, 2018. Disponível em: https://www.tandfonline.com/doi/full/10.1080/10584609.2018.1477506. Acesso em: 7 fev. 2025.

GARIMELLA, Kiran; TYSON, Gareth. WhatsApp, doc? A first look at WhatsApp public group data. *In*: INTERNATIONAL AAAI CONFERENCE ON WEB AND SOCIAL MEDIA (ICWSM 2018), 12., jun. 2018, Califórnia. *Anais* [...]. [*S.l.*]: Association for the Advancement of Artificial Intelligence, 2018. Disponível em: https://ojs.aaai.org/index.php/ICWSM/article/view/14989. Acesso em: 9 abr. 2024.

GILLESPIE, Tarleton. The politics of 'platforms'. *New Media & Society*, [*S.l.*], v. 12, n. 3, p. 347–364, maio 2010. Disponível em: https://journals.sagepub.com/doi/10.1177/1461444809342738. Acesso em: 9 abr. 2024.

GLOBAL AD. Kwai em 2024: Um panorama detalhado da plataforma. Global AD, [*S.l.*], 22 fev. 2024. Disponível em: https://globalad.com.br/

blog/explorando-o-kwai-em-2024-um-pano-rama-detalhado/. Acesso em: 3 abr. 2024.

GOODWIN, Michael. What Is an API (Application Programming Interface)?. *IBM*, [*S.l.*], 9 abr. 2024. Disponível em: https://www.ibm.com/topics/api. Acesso em: 26 abr. 2024.

GOOGLE. Cumprimento das diretrizes da comunidade do YouTube – Google Relatório de Transparência. Google, [*S.d.*]. Disponível em: https://transparencyreport.google.com/youtube-policy/removals. Acesso em: 16 abr. 2024.

GREENE, Travis; MARTENS, David; SHMUELI, Galit. Barriers to academic data science research in the new realm of algorithmic behaviour modification by digital platforms. *Nature Machine Intelligence*, [*S.l.*], v. 4, n. 4, p. 323–330, abr. 2022. Disponível em: https://www.nature.com/articles/s42256-022-00475-7. Acesso em: 9 abr. 2024.

GUESS, Andrew M.; MALHOTRA, Neil; PAN, Jennifer; BARBERÁ, Pablo; ALLCOTT, Hunt; BROWN, Taylor; CRESPO-TENORIO, Adriana; DIMMERY, Drew; FREELON, Deen; GENTZKOW, Matthew; GONZÁLEZ-BAILÓN, Sandra; KENNEDY, Edward; KIM, Young M.; LAZER, David; MOEHLER, Devra; NYHAN, Brendan; RIVERA, Carlos V.; SETTLE, Jaime; THOMAS, Daniel R.; THORSON, Emily; TROMBLE, Rebekah; WILKINS, Arjun; WOJCIESZAK, Magdalena; XIONG, Beixian; DE JONGE, Chad K.; FRANCO, Annie; MASON, Winter; STROUD, Natalie J.; TUCKER, Joshua. A. How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*, [*S.l.*], v. 381, n. 6656, p. 398– 404, 28 jul. 2023. Disponível em: https://www.science.org/doi/10.1126/science.abp9364. Acesso em: 9 abr. 2024.

IYER, Prithvi. Researchers Consider the Impact of Meta's CrowdTangle Shutdown. *Tech Policy Press*, [*S.l.*], 4 ago. 2024. Disponível em: https://www.techpolicy.press/researchers-consider-the-impact-of-metas-crowdtangle-shutdown/. Acesso em: 10 fev. 2025.

JÚNIOR, MANOEL; MELO, PHILIPE; SILVA, ANA P. C. DA; BENEVENUTO, FABRÍCIO; ALMEIDA, JUSSARA. Towards understan-

ding the use of Telegram by political groups in brazil. *In*: BRAZILIAN SYMPOSIUM ON MULTIMEDIA AND THE WEB, 27., p. 237-244, 5-12 nov. 2021, Belo Horizonte. *Anais* [...]. Nova Iorque: Association for Computing Machinery, 2021. Disponível em: https://dl.acm.org/doi/proceedings/10.1145/3470482. Acesso em: 7 fev. 2025.

KALOGEROPOULOS, Antonis; ROSSINI, Patrícia. Unraveling WhatsApp group dynamics to understand the threat of misinformation in messaging apps. *New Media & Society*, [*S.l.*], v. 0, n. 0, n.p., 2023. Disponível em: https://doi.org/10.1177/14614448231199247. Acesso em: 16 abr. 2024.

KAPOOR, Kawaljeet K.; TAMILMANI, Kuttimani; RANA, Nripendra P.; PATIL, Pushp; DWIVEDI, Yogesh, K.; NERUR, Sridhar. Advances in social media research: Past, present and future. *Information Systems Frontiers*, [*S.l.*], v. 20, p. 531-558, 2018. Disponível em: https://link.springer.com/article/10.1007/s10796-017-9810-y. Acesso em 7 fev. 2025.

KOSTA, Eleni; BREWCZYŃSKA, Magdalena. Government Access to User Data: Towards More Meaningful Transparency Reports. *In*: BALLARDINI, R.; KUOPPAMÄKI, P.; PITKÄNEN, O. (Eds.). *Regulating industrial internet through IPR, data protection and competition law*. Países Baixos: Kluwer Law International, 2019. Disponível em: https://papers.ssrn.com/abstract=3601661. Acesso em: 15 abr. 2024.

KROTOV, Vlad; JOHNSON, Leigh; SILVA, Leiser. Tutorial: Legality and Ethics of Web Scraping. *Communications of the Association for Information Systems*, [*S.l.*], v. 47, n.p., 2020. Disponível em: https://doi.org/10.17705/1CAIS.04724. Acesso em: 7 fev. 2025.

KWAI. Safety Center. *Kwai*, [*S.d.*]. Disponível em: https://www.kwai.com/safety. Acesso em: 16 abr. 2024.

LEE, R. M.; FIELDING, N.; BLANK, G. The Internet as a research medium: An editorial introduction to the sage handbook of online research methods. *In*: FIELDING, N.; LEE, R. M.; BLANK, G. (Eds.). *The Sage handbook of online research methods*. Oxford: SAGE Research Methods, 2008. Disponível em: https://uk.sagepub.com/en-gb/eur/the-sage-handbook-of-online-research-methods/book245027. Acesso em 7 fev. 2025.

LOSHIN, David. *Master Data Management*. Burlington: Morgan Kaufmann, 2008.

LU, Sylvia. Algorithmic Opacity, Private Accountability, and Corporate Social Disclosure in the Age of Artificial Intelligence. *Vanderbilt Journal of Entertainment & Technology Law 99*, [*S.l.*], v. 23, n. 1, p. 99–159, 20 out. 2021. Disponível em: https://papers.ssrn.com/abstract=3582222. Acesso em: 3 abr. 2024.

MAHANTI, Rupa. Data quality: dimensions, measurement, strategy, management, and governance. Quality Press, 2019.

MARRES, N. *Digital Sociology*: The Reinvention of Social Research. Cambridge: Wiley, 2017.

MCGILVRAY, Danette. *Executing Data Quality Projects*: Ten Steps to Quality Data and Trusted Information. São Francisco: Morgan Kaufmann Publishers, 2008.

MELO, Philipe de F. *Activism and Misinformation on WhatsApp*. 218 p. Tese [Doutorado em Ciência da Informação] – Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, Minas Gerais, 2022. Disponível em: https://repositorio.ufmg.br/bitstream/1843/49506/4/Tese_Philipe-Melo_ActicismAndMisinformationOnWhatsApp.pdf. Acesso em: 15 abr. 2024.

META. Meta Transparency Center. *Meta*, [*S.d.*]a. Disponível em: https://transparency.meta.com/pt-br/. Acesso em: 29 abr. 2024.

META. Meta Transparency Center. Government Requests for User Data. Reports. Brazil. *Meta*, [*S.d.*]b. Disponível em: https://transparency.meta.com/reports/government-data-requests/country/BR/. Acesso em: 10 fev. 2025.

MILLER, Gabby. First Transparency Reports Under Digital Services Act Are Difficult to Compare. *Tech Policy Press*, [*S.l.*], 22 nov. 2023. Disponível em: https://techpolicy.press/first-transparency-reports-under-digital-services-act-are-difficult-to-compare. Acesso em: 15 abr. 2024.

MOONEY, Stephen J.; WESTREICH, Daniel J.; EL-SAYED, Abdulrahman M. Commentary: Epidemiology in the Era of Big Data. *Epide-*

*miology*, [*S.l.*], v. 26, n. 3, p. 390394, maio 2015. Disponível em: https://journals.lww.com/epidem/fulltext/2015/05000/. Acesso em: 4 abr. 2024.

MOZELLI, Rodrigo. Twitter: API cara impede uso para pesquisas acadêmicas. *Olhar Digital*, [*S.l.*], 01 jun. 2023. Disponível em: https://olhardigital.com.br/2023/06/01/internet-e-redes-sociais/twitter-api-cara-impede-uso-para-pesquisas-academicas/. Acesso em: 3 abr. 2024.

MOZILLA FOUNDATION. Open Letter To Meta: Support CrowdTangle Through 2024 and Maintain CrowdTangle Approach. *Mozilla Foundation*, 2024. Disponível em: https://foundation.mozilla.org/en/campaigns/open-letter-to-meta-support-crowdtangle-through-2024-and-maintain-crowdtangle-approach/. Acesso em: 4 abr. 2024.

OPINION BOX. Relatório Kwai no Brasil 2024. *Opinion Box*, 2024. Disponível em: https://content.app-us1.com/JY8yY/2024/05/24/f32d--1193-1919-4170-8727-db73a82d84a1.pdf. Acesso em: 20 set. 2024

OZAWA, João V. S.; WOOLEY, Samuel C.; STRAUBHAAR, Jacob; RIEDL, Martin; JOSEFF, Katie; GURSKY, Jacob. How Disinformation on WhatsApp Went From Campaign Weapon to Governmental Propaganda in Brazil. *Social Media + Society*, [*S.l.*], v. 9, n. 1, n.p., 2023. Disponível em: https://journals.sagepub.com/doi/full/10.1177/20563051231160632. Acesso em: 15 abr. 2024.

PAPPA, Dimitra; STERGIOULAS, Lampros K. Harnessing social media data for pharmacovigilance: A review of current state of the art, challenges and future directions. *International Journal of Data Science and Analytics*, [*S.l.*], v. 8, n. 2, p. 113–135, set. 2019. Disponível em: https://link.springer.com/article/10.1007/s41060-019-00175-3. Acesso em: 9 abr. 2024.

POSTMAN. What is an API? A Beginner's Guide to APIs. *Postman*, [*S.d.*]. Disponível em: https://www.postman.com/what-is-an-api/. Acesso em: 3 abr. 2024.

RESENDE, Gustavo; MELO, Philipe; SOUSA, Hugo; MESSIAS, Johnnatan; VASCONCELOS, Marisa; ALMEIDA, Jussara; BENEVENUTO, Fabricio. (Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures. *In*: THE WORLD WIDE WEB CON-

FERENCE, 28., maio 2019, Califórnia. *Anais* [...]. Nova Iorque: Association for Computing Machinery (ACM), 13 maio 2019. Disponível em: https://homepages.dcc.ufmg.br/~fabricio/download/resende-www2019.pdf. Acesso em: 3 abr. 2024.

ROGERS, Richard. *The end of the virtual*: Digital methods. Amsterdã: Amsterdam University Press, 2009.

ROGERS, Richard. Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication*, [*S.l.*], v. 35, n. 3, p. 213-229, 2020. Disponível em: https://journals.sagepub.com/doi/10.1177/0267323120922066. Acesso em: 7 fev. 2025.

ROTH, Emma. Elon Musk's X sues anti-hate researchers for allegedly scraping data from Twitter. *The Verge*, [*S.l.*], 2023. Disponível em: https://www.theverge.com/2023/8/1/23815515/twitter-ccdh-anti-hate-research--group-lawsuit. Acesso em: 15 abr. 2024.

SANTOS, Edilaine. Após 11 anos de LAI, governos locais ainda dificultam pedidos de acesso à informação. *Open Knowledge Brasil*, [*S.l.*], 18 maio 2023. Disponível em: https://ok.org.br/noticia/apos-11-anos-de-lai-governos-locais-ainda-dificultam-pedidos-de-acesso-a-informacao/. Acesso em: 3 abr. 2024

SELINGER, Evan; HARTZOG, Woodrow. Facebook's emotional contagion study and the ethical problem of co-opted identity in mediated environments where users lack control. *Research Ethics*, [*S.l.*], v. 12, n. 1, p. 35–43, jan. 2016. Disponível em: https://journals.sagepub.com/doi/10.1177/1747016115579531. Acesso em: 9 abr. 2024.

SENTHILRAJA, M. Application of Artificial Intelligence to Address Issues Related to the COVID-19 Virus. *SLAS*, [*S.l.*], v. 26, n. 2, p. 123–126, abr. 2021. Disponível em: https://journals.sagepub.com/doi/full/10.1177/2472630320983813. Acesso em: 9 abr. 2024.

SHAW, Ryan. Big Data and reality. *Big Data & Society*, [*S.l.*], v. 2, n. 2, p. 1-4, dez. 2015. Disponível em: https://journals.sagepub.com/doi/10.1177/2053951715608877 Acesso em 9 abr. 2024.

SINGH, Shubham. Telegram Users Statistics (2025) –New Global Data. *DemandSage*, [*S.l.*], 1 jan. 2025. Disponível em: https://www.demandsage.com/telegram-statistics/. Acesso em: 7 fev 2025.

SMITH, Rory; CHEN, Kung; WINNER, Daisy; FRIEDHOFF, Stefanie; WARDLE, Claire. A Systematic Review Of COVID-19 Misinformation Interventions: Lessons Learned. *Health Affairs*, [*S.l.*], v. 42, n. 12, p. 1738–1746, 2023. Disponível em: https://pubmed.ncbi.nlm.nih.gov/37967291/. Acesso em: 15 abr. 2024.

SOARES, Matheus. Meta encerra CrowdTangle e impõe alternativa limitada e burocrática. *Desinformante*, [*S.l.*], 14 ago. 2024. Disponível em: https://desinformante.com.br/meta-crowtangle-alternativa. Acesso em: 12 set. 2024.

STAAB, Philipp.; THIEL, Thorsten. Social Media and the Digital Structural Transformation of the Public Sphere. *Theory, Culture & Society*, [*S.l.*], v. 39, n. 4, p. 129–143, 5 set. 2022. Disponível em: https://journals.sagepub.com/doi/10.1177/02632764221103527. Acesso em: 26 abr. 2024.

STOKEL-WALKER, Chris. X nega a pesquisadores que estudam desinformação acesso a sua API. *Fast Company Brasil*, [*S.l.*], 1 mar. 2024. Disponível em: https://fastcompanybrasil.com/tech/x-nega-a-pesquisadores-que-estudam-desinformacao-acesso-a-sua-api/. Acesso em: 4 abr. 2024.

SUZOR, Nicolas. P; ESTE, Sarah M.; QUODLING, Andrew; YORK, Jilian. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. *International Journal of Communication*, [*S.l.*], v. 13, n.0, p. 1526–1543, 2019. Disponível em: https://ijoc.org/index.php/ijoc/article/view/9736. Acesso em: 15 abr. 2024

TELEGRAM. Canais do Telegram. *Telegram*, [*S.d.*]a. Disponível em: https://telegram.org/tour/channels/pt-br. Acesso em: 4 abr. 2024.

TELEGRAM. Grupos do Telegram. Telegram, [*S.d.*]b. Disponível em: https://telegram.org/tour/groups/pt-br. Acesso em: 4 abr. 2024.

TIKTOK. Research API. *TikTok*, [*S.d.*]. Disponível em: https://developers.tiktok.com/products/research-api/. Acesso em: 3 abr. 2024.

TIKTOK. Government Removal Requests Report. *TikTok*, 2023. Disponível em: https://www.tiktok.com/transparency/en/government-removal-requests-2023-1. Acesso em: 15 abr. 2024.

TIKTOK. Community Guidelines Enforcement Report. *TikTok*, 2024. Disponível em: https://www.tiktok.com/transparency/en/community-guidelines-enforcement-2023-4. Acesso em: 16 abr. 2024.

TRANSPARÊNCIA BRASIL. O que a população quer saber do poder público? Uma análise de respostas a pedidos de acesso à informação de órgãos de todos os poderes e níveis federativos. *Transparência Brasil*, 2018. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/RelatorioLAI_TransparenciaBrasil_2018_vf.pdf. Acesso em: 3 abr. 2024.

TROMBLE, Rebekah. Where have all the data gone? A critical reflection on academic digital research in the post-API age. *Social Media + Society*, [*S.l.*], v. 7, n. 1, n.p., 2021. Disponível em: https://journals.sagepub.com/doi/full/10.1177/2056305121988929. Acesso em: 7 fev. 2025.

URMAN, Aleksandra; MAKHORTYKH, Mykola. How transparent are transparency reports? Comparative analysis of transparency reporting across online platforms. *Telecommunications policy*, [*S.d.*] v. 47, n. 3, n.p., 102477, 2023. Disponível em: https://www.sciencedirect.com/science/article/pii/S0308596122001793. Acesso em: 7 fev. 2025.

WAGNER, Ben; ROZGONYI, Krisztina; SEKWENZ, Marie-Therese; COBBE, Jennifer; SINGH, Jatinder. Regulating Transparency? Facebook, Twitter and the German Network Enforcement Act. *In*: Conference on Fairness, Accountability, and Transparency, 20., 2020, Barcelona. *Anais* [...]. [*S.l.*]: Association for Computing Machinery (ACM), 2020, p. 261-271. Disponível em: https://dl.acm.org/doi/10.1145/3351095.3372856. Acesso em 15 abr. 2024.

WOOLLEY, Samuel C.; HOWARD, Philip N. (Eds.). Computational propaganda: Political parties, politicians, and political manipulation on social media. Oxford: *Oxford University Press*, 2018.

VAN ATTEVELDT, Wouter; PENG, Tai-Quan. When communication meets computation: Opportunities, challenges, and pitfalls in computatio-

nal communication science. *Communication Methods and Measures*, [*S.l.*] v. 12, n. 2-3, p. 81-92, 2018. Disponível em: https://doi.org/10.1080/19312458.2018.1458084. Acesso em: 7 fev. 2025.

WENDRATAMA, Engelbertus; YUSUF, Iwan. A. COVID-19 Falsehoods on WhatsApp: Challenges and Opportunities in Indonesia. *In*: SOON, C. (Org.). *Mobile Communication and Online Falsehoods in Asia*: Trends, Impact and Practice. Países Baixos: Springer, 2023. Disponível em: https://link.springer.com/chapter/10.1007/978-94-024-2225-2_2. Acesso em: 15 abr. 2024.

X/TWITTER. X Developer Platform. X API. *X/Twitter*, [*S.d.*]a. Disponível em: https://developer.x.com/en/docs/x-api. Acesso em: 10 fev. 2025.

X/TWITTER. X Developer Platform. Enterprise data customers. *X/Twitter*, [*S.d.*]b. Disponível em: https://developer.x.com/en/products/x-api/enterprise/directory. Acesso em: 10 fev. 2025.

YASSERI, Taha. From Print to Pixels: The Changing Landscape of the Public Sphere in the Digital Age. *SSRN Scholarly Paper*, 2019. [Manuscrito em pré-publicação] Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4543907. Acesso em 26 abr. 2024

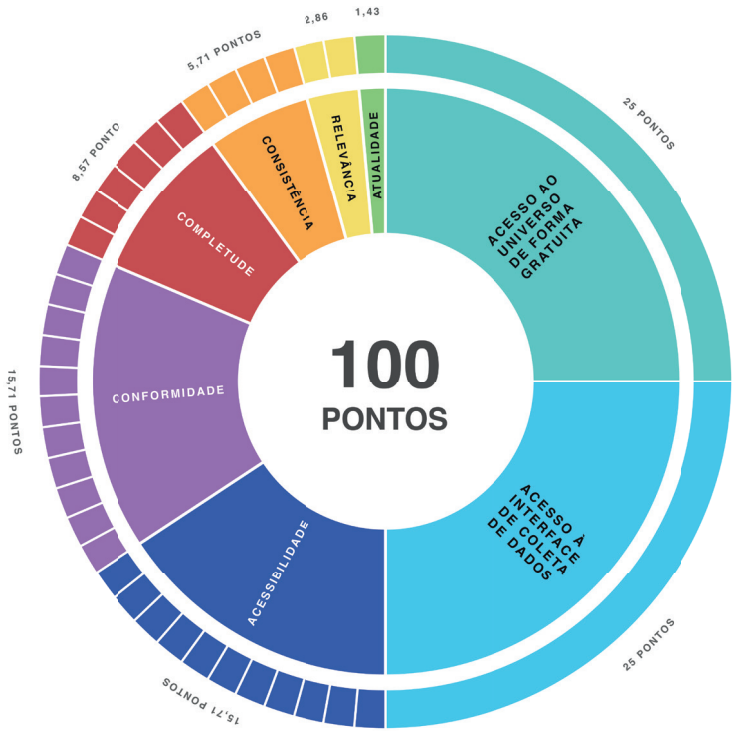
YOUTUBE. YouTube Research - How It Works. *YouTube*, [*S.d.*]. Disponível em: https://research.youtube/how-it-works/. Acesso em: 4 abr. 2024.

ZITTRAIN, Jonathan. Three eras of digital governance. *SSRN Scholarly Paper*, 2019. [Manuscrito em pré-publicação] Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3458435. Acesso em: 30 jan. 2025.

ZUCKERMAN, Ethan. Why study media ecosystems? *Information, Communication & Society*, [*S.l.*], v 24, n. 10, p. 1495-1513, 2021. Disponível em: https://www.tandfonline.com/doi/abs/10.1080/1369118X.2021.1942513. Acesso em: 7 fev. 2024.

# Avaliação das Plataformas de Redes Sociais quanto a Transparência de Dados

**Figura 1:** Representação visual da pontuação passível de ser obtida pelas plataformas analisadas, considerando a aplicabilidade de todos os parâmetros de avaliação propostos




**ACESSO AO UNIVERSO DE FORMA GRATUITA (CRITÉRIO ESPECIAL 1)**
**PESO:** 25 PTS
**Nº DE PARÂMETROS:** 4

**ACESSO À INTERFACE DE COLETA DE DADOS (CRITÉRIO ESPECIAL 2)**
**PESO:** 25 PTS
**Nº DE PARÂMETROS:** 1

**ACESSIBILIDADE**
**PESO:** 15,71 PTS
**Nº DE PARÂMETROS:** 11

**CONFORMIDADE**
**PESO:** 15,71 PTS
**Nº DE PARÂMETROS:** 11

**COMPLETUDE**
**PESO:** 8,57 PTS
**Nº DE PARÂMETROS:** 6

**CONSISTÊNCIA**
**PESO:** 5,71 PTS
**Nº DE PARÂMETROS:** 4

**RELEVÂNCIA**
**PESO:** 2,86 PTS
**Nº DE PARÂMETROS:** 2

**ATUALIDADE**
**PESO:** 1,43 PT
**Nº DE PARÂMETROS:** 1

**Figura 2:** Visão geral das avaliações das plataformas, da maior à menor nota obtida

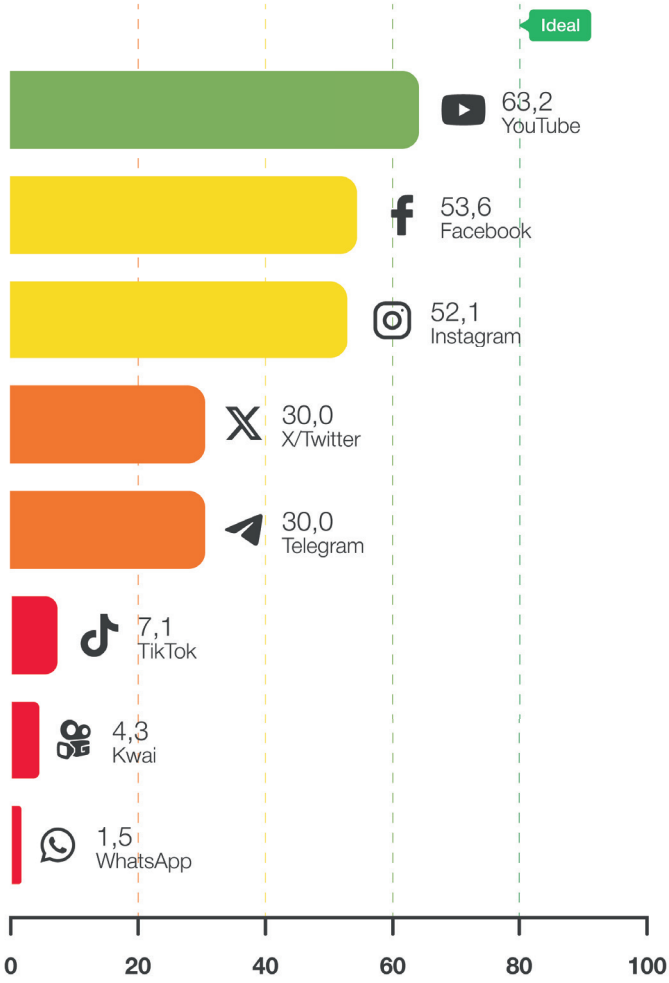


**Figura 3:** Representação visual da nota obtida pelo YouTube em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

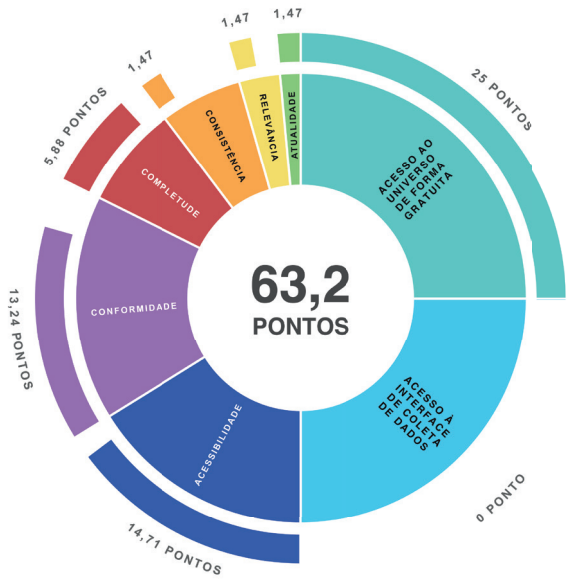



**Figura 4:** Representação visual da nota obtida pelo Facebook em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

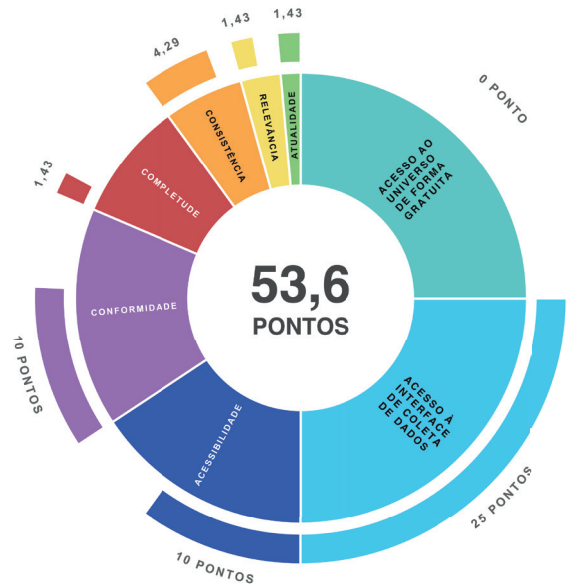


**Figura 5:** Representação visual da nota obtida pelo Instagram em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

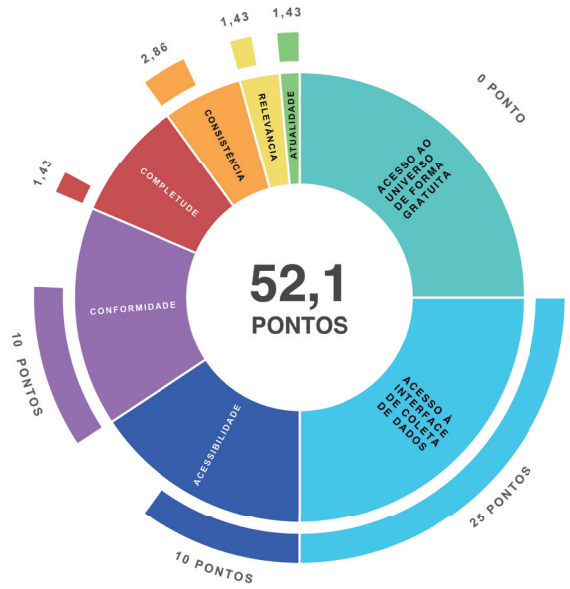



**Figura 6:** Representação visual da nota obtida pelo X/Twitter em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

**Figura 7:** Representação visual da nota obtida pelo Telegram em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

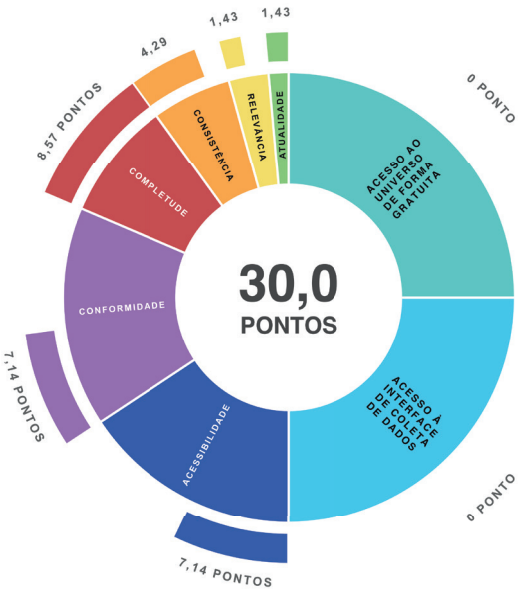



**Figura 8:** Representação visual da nota obtida pelo TikTok em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

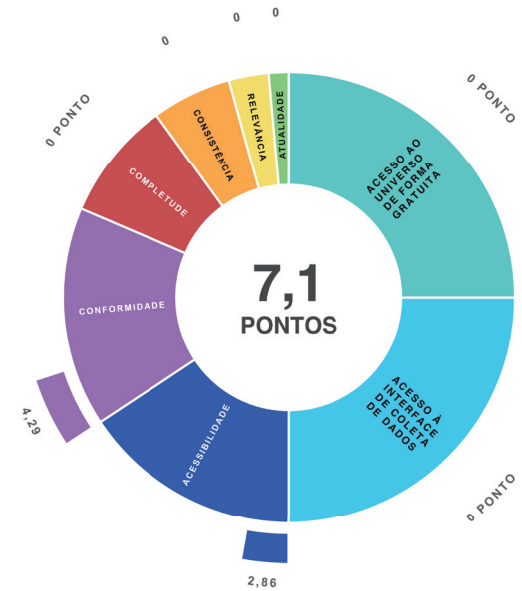


**Figura 9:** Representação visual da nota obtida pelo Kwai em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

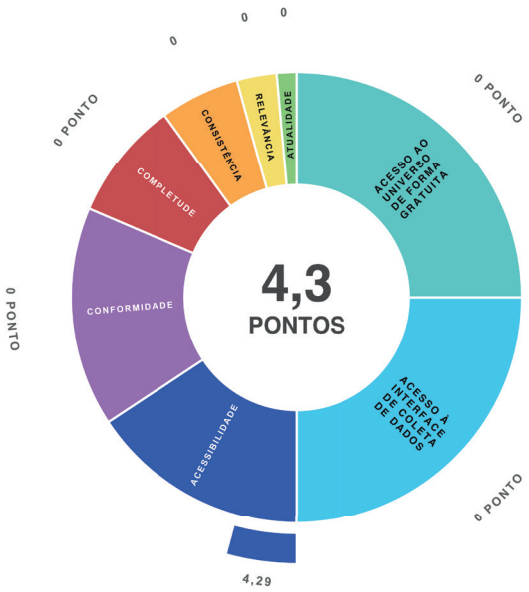
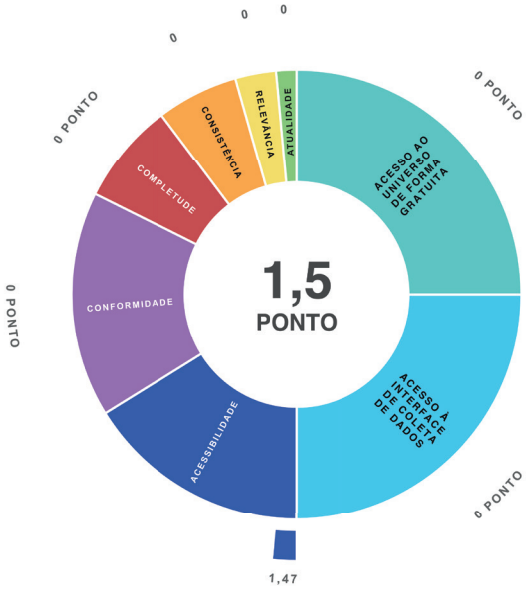



**Figura 10:** Representação visual da nota obtida pelo WhatsApp em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas



# Capítulo 2

# Índice de Transparência da Publicidade nas Plataformas de Redes Sociais

A fim de que pudessem lucrar com seus serviços ao mesmo tempo que mantinham a gratuidade de sua utilização, as plataformas digitais desenvolveram e aperfeiçoaram uma nova forma de distribuição de conteúdo publicitário. Baseada no processamento de volumes massivos de dados comportamentais de seus bilhões de usuários em todo o planeta, essa nova publicidade microssegmentada permite que cada consumidor seja impactado por anúncios cuidadosamente elaborados para apelar a seus maiores interesses, anseios e desejos e assim capturar sua atenção. É uma lógica de mercado que se impõe no ambiente digital: as plataformas e aplicativos mais acessados no mundo são inundados por um volume cada vez maior de anúncios indiscretos, mesmo que os usuários assinem planos pagos (Hermann, 2023).

Essa nova forma de publicidade personalizada possibilitou um ganho sem precedentes de escala do mercado publicitário e interrompeu disputas por espaços físicos para a veiculação de anúncios, com a possibilidade de um mesmo conteúdo ser replicado indefinidamente para novos públicos e em momentos diferentes, caso anunciantes assim desejem. Como define o Conselho Administrativo de Defesa Econômica (CADE), a publicidade online se tornou "um serviço sem substituto" (CADE, 2023, p. 102) por suas características singulares de segmentação da audiência. As plataformas digitais e redes sociais emergiram

como agentes poderosos no mercado publicitário, capazes de reter a atenção dos usuários, um recurso cada vez mais escasso diante do excesso de informação online, e de distribuir conteúdo de forma altamente personalizada.

A promessa da publicidade online microssegmentada é de que todos os envolvidos na relação comercial com as plataformas obtêm vantagens. Os anunciantes pagariam menos, mas teriam um retorno maior por se comunicarem diretamente com pessoas mais propensas a se interessarem pelo que têm a oferecer. Já os usuários seriam impactados por anúncios de produtos e serviços que lhe interessam mais, além de poderem continuar utilizando as plataformas de forma gratuita e despreocupada. E por fim, as plataformas aumentariam seus lucros e aprimorariam seus sistemas de distribuição e recomendação de anúncios, tornando-os mais "irresistíveis" e oniscientes, capturando a atenção voluntária e involuntária dos usuários. Os resultados são cifras impressionantes: a publicidade hoje constitui quase 98% das receitas anuais da Meta (Meta, 2025), empresa dona de plataformas como Facebook, Instagram e WhatsApp, e cerca de 76% do Google (Statista, 2025a; 2025b).

Por mais que esta relação possa parecer benéfica à primeira vista, os consumidores finais, na prática, são seu elo fraco, estando vulneráveis e desprotegidos nos ambientes online. Afinal, as plataformas de redes sociais que utilizam são pouquíssimo transparentes em relação a como seus dados são instrumentalizados e monetizados para manter seu modelo de negócios em pleno funcionamento (Bromell, 2022). Enquanto elas sabem muito mais sobre seus usuários do que quaisquer Estados jamais imaginaram saber sobre seus cidadãos, os usuários pouco sabem sobre os dados que estão em posse das plataformas, para quê e como eles são usados. As plataformas têm como matéria-prima a captura de dados pessoais de seus usuários e seu modelo de negócios é transformá-los em mercadoria e assim gerar lucro, em uma relação assimétrica baseada na vigilância constante, acarretando diretamente a perda de privacidade (Zuboff, 2025; Dobber *et al.*, 2023). Nos termos de Zuboff (2015), este

é um modelo que prospera em virtude de um desconhecimento generalizado dos usuários sobre o que acontece e também sobre seus próprios direitos à privacidade. Este desconhecimento é fomentado pelas plataformas por meio da opacidade e ausência de responsabilidade legal diante dessa nova forma de exploração em massa.

A formação destas relações assimétricas é apenas uma das muitas consequências da falta de transparência da publicidade em plataformas de redes sociais. Auditar a publicidade online já seria, naturalmente, uma tarefa difícil, dado que, ao contrário da publicidade veiculada em outros meios, que pode ser vista igualmente por todo o público, os anúncios em redes sociais só são vistos por aqueles diretamente impactados por eles, graças a ferramentas de microssegmentação (Jamison *et al*., 2020). Além disso, as atuais ferramentas de transparência de anúncios disponibilizadas voluntariamente por algumas plataformas de redes sociais ofuscam questões fundamentais sobre a publicidade online, fazendo com que não seja possível ir além de análises e investigações superficiais (ver Carah *et al*., 2024; Pershan; Lesplingart, 2024).

Como cada plataforma decide, segundo seus próprios termos e interesses, o que deve ser disponibilizado para escrutínio público nestas ferramentas, isto acaba resultando em uma "transparência" que mais confunde do que colabora à auditabilidade de suas operações comerciais (ver Ananny; Crawford, 2016). Esta opacidade é um dentre muitos elementos que ajudam a viabilizar e a alavancar as atividades de atores mal-intencionados, que atuam de modo coordenado para afetar as percepções, opiniões e comportamentos dos usuários, com diferentes propósitos políticos e/ou econômicos em mente (ver Briant; Bakir, 2024; Klein, 2024; Tufekci, 2017). Notoriamente, a microssegmentação tem sido uma importante peça do ecossistema de manipulação da opinião pública em períodos eleitorais e de maior efervescência política (ver Armitage *et al*., 2023; Medert; Otto; Perczel, 2024). Além disso, o anonimato dos anunciantes nas redes sociais, combinado à capacidade de segmentar usuários vulneráveis com base em seus interesses, fez da

publicidade microssegmentada uma ferramenta valiosa para a indústria do crime online (Santini *et al.*, 2023; 2024a).

Perante a esta notória impossibilidade de se investigar profundamente os sistemas de distribuição de anúncios das plataformas de redes sociais, que suscita intensas e constantes demandas por maior transparência por parte de pesquisadores de todo o mundo (ver Edelson *et al.*, 2021; Mozilla Foundation, [*S.d.*]; Santini *et al.*, 2024b), apresentamos o Índice de Transparência da Publicidade nas Plataformas de Redes Sociais (ITP) no Brasil. Segundo uma abordagem guiada pela qualidade de dados e um entendimento de que as operações comerciais das plataformas digitais não podem estar acima do interesse público, este capítulo se propõe a indicar critérios para uma transparência efetiva da publicidade veiculada nas principais plataformas de redes sociais que operam no país.

De acordo com Gillespie (2018), as plataformas de redes sociais são serviços digitais que, apesar de não produzirem seus próprios conteúdos, hospedam, organizam e fazem circular conteúdos e interações de terceiros, ao mesmo tempo que processam seus dados para a venda de anúncios personalizados para gerar lucro. A definição do que constitui um anúncio nas plataformas de redes sociais pode variar radicalmente conforme seus elementos visuais e persuasivos e sua forma de distribuição. Tradicionalmente, um anúncio divulga informações que marcas, empresas e instituições querem transmitir a potenciais consumidores, ajudando a vender seus produtos e serviços ou a aumentar o alcance de ideias e campanhas de cunho social, cívico, político ou eleitoral, por exemplo. Neste capítulo, nós chamamos de "anúncios" todos os conteúdos que têm visibilidade elevada e privilegiada para determinados grupos e segmentos escolhidos pelo anunciante em plataformas de redes sociais mediante pagamentos – sejam eles realizados por intermediários, como agências de publicidade, ou não.

Dessa forma, há diferenças classificatórias entre anúncios online que costumam ser mais exploradas pelo mercado publicitário, mas que podemos destrinchar. Comumente, o chamado "post patrocinado" ou

"post impulsionado" é qualquer publicação que fica disponível organicamente em uma página ou perfil em plataformas de redes sociais e que também é impulsionada algoritmicamente e de forma microssegmentada, de acordo com critérios definidos pelo anunciante, para atingir novos públicos. Já o chamado "dark post" é um tipo de conteúdo impulsionado que não é exibido na página ou no perfil do anunciante, ficando visível apenas para usuários alcançados pelas ferramentas de microssegmentação das plataformas (Mirago, 2024). No fim, ambos os tipos de conteúdo recebem um rótulo de "conteúdo pago" (ou similar) pelas plataformas e serão considerados anúncios para os fins deste capítulo.

Aliás, esta comunicação clara e efetiva sobre o caráter publicitário dos conteúdos é uma forma de dar mais transparência à publicidade online, principalmente aos usuários impactados por ela, como defendem diferentes pesquisadores (ver Campbell; Grimm, 2019; Dobber *et al.*, 2023; Reijmersdal; Rozendaal, 2020) e órgãos internacionais que definem normas e condutas para a publicidade (ver Conar, 2021; FTC, 2013; 2015a; 2015b; 2023; ICC, 2018). Contudo, embora informar sobre o caráter comercial dos conteúdos pagos seja uma medida de transparência primordial, ela não promove, por conta própria, a auditabilidade sistemática e independente dos anúncios veiculados em plataformas de redes sociais, se não for acompanhada do acesso a dados qualificados sobre esses mesmos anúncios.

As plataformas de redes sociais também têm sido inundadas pelos chamados "publiposts" – uma abreviação do termo "post publicitário" –, publicações feitas por influenciadores digitais, celebridades e criadores de conteúdo de médio e grande alcance que recebem para promover marcas, empresas ou serviços, envolvendo a participação criativa de quem publica (Schnaider, 2022). Geralmente, são publicações orgânicas, feitas como quaisquer outras, mas com cunho comercial, e que, por isso, devem vir acompanhadas de hashtags como #publi ou #parceriapaga para melhor sinalização de seu caráter comercial, segundo regras instituídas pelo Conselho Nacional de Autorregulamentação Publici-

tária (Conar) (Ferreira, 2022). A natureza publicitária dos publiposts é inegável, mas seria necessário desenvolver outro quadro metodológico específico para analisá-los.

Mesmo com todos os desafios sociais e econômicos impostos por suas práticas comerciais baseadas no uso de dados pessoais e comportamentais, a pesquisa sobre o ecossistema de publicidade online no Brasil ainda é incipiente. Hoje, há poucos estudos empíricos sobre a dinâmica deste mercado, dando conta de todos os relacionamentos sociotécnicos que ela implica, que incluem anunciantes, patrocinadores, estratégias de segmentação, competição de mercado, critérios de precificação, formatos de mídia, dentre outros. Tendo em vista seus impactos sociais e econômicos em potencial, o mercado publicitário das plataformas de redes sociais precisa ser passível de análises cuidadosas e detalhadas. Para garantir a segurança do ambiente online, é necessário reforçar os mecanismos de transparência da publicidade nas plataformas de redes sociais e garantir a possibilidade de monitoramento sistemático e de auditabilidade dos anúncios a partir do estabelecimento de critérios rigorosos.

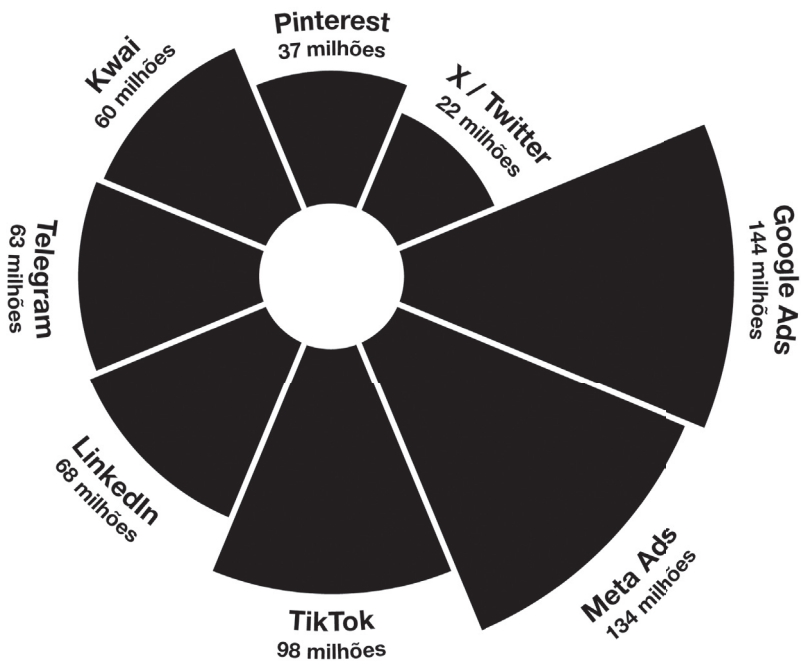
No ITP, avaliamos e respondemos:

- Quais são as medidas de transparência e acesso a dados sobre anúncios veiculados nas principais plataformas de redes sociais no Brasil?
- Qual é o nível da qualidade dos dados sobre anúncios disponibilizados por essas plataformas?
- Nesse sentido, nossos objetivos incluem:
- **Padronizar:** Definir parâmetros de avaliação sobre o nível de acesso e qualidade dos dados sobre anúncios provenientes de plataformas de redes sociais.
- **Avaliar:** Identificar, de forma sistemática e transparente, os pontos fortes e fracos do acesso e da qualidade dos dados sobre anúncios.

- **Comparar:** Aferir comparativamente a performance de cada plataforma de rede social a partir de critérios comuns e metodologia padronizada.
- **Aprimorar:** Indicar, pública e objetivamente, o que precisa ser melhorado na disponibilização de dados sobre anúncios.

Para tanto, o ITP segue um roteiro de avaliação estruturado, sistematizado e reprodutível, baseado em critérios de qualidade de dados, para avaliar os mecanismos existentes de acesso a dados resultantes do impulsionamento de anúncios nas principais plataformas de redes sociais que operam no Brasil, vistas na Figura 1[11]:

**Figura 1:** Número de usuários de cada plataforma analisada no Brasil (Kemp, 2024; Global AD, 2024; Singh, 2024)




[11] Como Google e Meta não veiculam anúncios apenas em plataformas de redes sociais, consideramos os usuários de suas plataformas mais populares no Brasil: o YouTube e o Instagram, respectivamente. Por controlarem outros serviços digitais com anúncios, o número de cidadãos brasileiros impactados por suas redes de publicidade é, possivelmente, ainda maior. A rede de publicidade do Google inclui anúncios veiculados em sites e aplicativos de terceiros, buscadores, vídeos no YouTube, Google Discover, Play Store, Google Maps, Google Shopping e Gmail.

As plataformas foram selecionadas conforme princípio do *Digital Services Act* (DSA), regulação vigente na União Europeia que estabelece medidas de responsabilidade e transparência das plataformas digitais, incluindo as de redes sociais, que são utilizadas por mais de 10% da população do bloco, adaptando-o ao contexto brasileiro. Isto posto, foram avaliadas as plataformas de redes sociais de maior impacto social que ofertam serviços de publicidade no país. Por serem as mais relevantes, tanto em número de usuários como em receita, consideramos que estas empresas possuem os recursos necessários para garantir investimento em infraestrutura robusta de transparência e seguir as melhores práticas do mercado. Devido ao foco em plataformas de redes sociais, outros *players* relevantes no mercado da publicidade online não são abarcados pelo escopo do ITP, a exemplo de redes de publicidade programática[12]. Apesar de Google e Meta explorarem outros espaços que não apenas as suas plataformas de redes sociais, essas empresas são precursoras neste mercado e têm maior alcance em comparação a concorrentes.

Em diálogo com autores que abordam a qualidade de dados em áreas técnicas (Barbieri, 2019; Mahanti, 2018) e nas Ciências Sociais Aplicadas (Dommett; Power, 2024; Michener; Bersch, 2013), nossa premissa é a de que a padronização de critérios para avaliar sistematicamente o acesso e a qualidade dos dados sobre anúncios nas plataformas sociais tende a impactar positivamente na transparência destes espaços e na responsabilização destes por seus serviços enquanto importantes agentes econômicos. Nosso *framework* se baseia na noção de que a governança de dados das grandes plataformas de redes sociais é um tema de interesse público e que, logo, ultrapassa a esfera corporativa e privada destas empresas (ver Finger, 2019).

[12] Conteúdos publicitários que são distribuídos e alocados de forma automática por meio de algoritmos pertencentes a plataformas intermediárias (Van Looy, 2016), prática adotada por sites, plataformas de *streaming*, plataformas de redes sociais e, até mesmo, a mídia *out-of-home* (Borges, 2023; Fulgêncio, 2023). A distribuição se dá por meio de leilões, que ocorrem instantaneamente quando um usuário acessa um espaço digital atendido por uma rede de publicidade programática ou este é atualizado (Marotta; Abhishek; Acquisti, 2019).

O roteiro de avaliação do ITP, composto por 60 parâmetros de avaliação que embasam o cálculo e aferição de uma nota para cada plataforma, toma por base seis dimensões de qualidade de dados preconizadas pela literatura científica, independentemente das particularidades das plataformas. Entre as dimensões endógenas aos dados, são avaliadas a completude, a acessibilidade, a consistência e a acurácia (Barbieri, 2019; Batini; Scannapieco, 2006; McGilvray, 2008; Loshin, 2008). Outras dimensões avaliadas, como conformidade e relevância (Barbieri, 2019), dependem de fatores exógenos e, por isso, podem variar de acordo com as normas legais em vigor em cada país ou com os objetivos específicos da pesquisa.

Neste capítulo, primeiramente, nós apresentamos a base do modelo de negócios comum às grandes plataformas digitais, incluindo as de redes sociais, explicando como este gira em torno da veiculação massiva e personalizada de publicidade. Em seguida, explicamos como funcionam as técnicas de microssegmentação de audiência propiciada pelas plataformas digitais e vendidas como a forma mais eficiente de distribuição de publicidade na internet, apresentando benefícios e riscos associados a elas. Finalizando esta parte inicial, detalhamos alguns dos desafios colocados sobre a transparência da publicidade online e como as plataformas se beneficiam pela manutenção da opacidade de suas operações comerciais, em detrimento do interesse público.

Feita essa introdução, explicamos a abordagem metodológica do ITP, apresentando seu roteiro de avaliação e os cálculos para a composição das notas das plataformas analisadas. Por fim, apresentamos um panorama da avaliação individual de cada uma delas, ressaltando boas e más práticas nas medidas que promovem (ou não) para aumentar a transparência de seus ecossistemas de publicidade. A Meta – cujo ecossistema abrange Facebook, Instagram, Messenger e Audience Network – obteve a melhor nota em nossa avaliação, com 49,8 pontos, o que é considerado apenas regular, enquanto quatro plataformas nem sequer pontuaram em nenhum dos parâmetros de avaliação propostos. Ou seja, mesmo que a Meta apresente uma transparência insatisfatória,

outras empresas e plataformas são ainda mais opacas, acendendo um preocupante sinal de alerta para a pesquisa e a auditabilidade da publicidade online. Experiências no Norte Global, porém, já indicam possíveis caminhos a serem seguidos no Brasil – e no Sul Global, de modo geral – para a melhoria desse cenário.

## A publicidade como o modelo de negócios das plataformas de redes sociais

As grandes plataformas digitais existem hoje apenas por um motivo: a venda de publicidade microssegmentada com base nos dados que seus usuários deixam para trás (Bromell, 2022). Seu modelo de negócios está calcado no que Zuboff (2015; 2019; 2021) chama de *Capitalismo de Vigilância*, um sistema econômico baseado na extração, modelagem e posterior monetização de volumes massivos de dados sobre os comportamentos, anseios, interesses e crenças de usuários de plataformas e outros serviços digitais. Nesse cenário, a monetização da atenção dos usuários por meio da venda de espaços publicitários torna-se a principal contrapartida oferecida pelas plataformas digitais, incluindo as redes sociais, permitindo que seus usuários continuem a utilizá-las "gratuitamente". Em troca, cedem seus dados comportamentais para o aprimoramento de algoritmos de recomendação e distribuição de conteúdo patrocinado (Bromell, 2022).

Estima-se que os gastos globais com publicidade online tenham atingido 694 bilhões de dólares americanos em 2024 (Statista, 2024a). As receitas globais totais, bem como as provenientes da veiculação de publicidade, das plataformas analisadas neste estudo são apresentadas na Tabela 1.

**Tabela 1:** Receitas globais totais e de publicidade em 2024 das plataformas analisadas (Curry, 2025; Goel, 2024; Iqbal, 2025; Kuaishou, 2025; Lunden, 2025; Meta, 2025; Pinterest, 2025; Statista, 2024b; 2025a; 2025b)

| Plataforma | Receita global (em bilhões de dólares americanos) | Receita global de publicidade (em bilhões de dólares americanos) |
|---|---|---|
| Google | 348,1 | 264,6 |
| Meta | 164,5 | 160,6 |
| TikTok | 23 | 18,4 |
| Kwai | 17,4 | 9,9 |
| LinkedIn | 16,2 | 5,4[13] |
| Pinterest | 3,6 | -[14] |
| X/Twitter | 2,5 | 1,7 |
| Telegram | 1 | -[15] |

Desde a ascensão da internet comercial, nos anos 1990, as *big tech* buscaram caminhos para tornar o ecossistema digital rentável em meio ao crescente senso comum de que "na internet, o conteúdo é grátis" (Macnamara, 2010). Daí em diante, empresas como Google e Meta ajudaram a catapultar novas formas de publicidade, inexistentes nas mídias offline, como os links patrocinados, os posts impulsionados (ou posts patrocinados) e a mídia programática, dentre outros novos formatos.

O Google, parte da *holding* Alphabet, foi pioneiro ao encontrar um modelo de publicidade online economicamente sustentável ao lançar os links patrocinados em 2000. Com a contratação deste serviço, a plataforma garante aos anunciantes uma posição no topo da página de resultados de seu buscador, com anúncios que mimetizam a estética de resultados de busca orgânicos e com apenas uma discreta sinalização de que se trata de um conteúdo patrocinado (Van Looy, 2016). Ainda em

[13] Segundo Statista (2024b), cerca de 33% das receitas do LinkedIn entre junho de 2023 e junho de 2024 foram oriundas da veiculação de publicidade. Em 2024, a plataforma registrou receitas totais de 16,2 bilhões de dólares, mas não detalhou quanto desse valor veio de anúncios (Lunden, 2025). Dessa forma, utilizamos o percentual de 33% para estimar, de maneira não oficial, suas receitas de publicidade.

[14] No balanço referente a 2024, a plataforma não detalhou a parcela da receita proveniente de anúncios. Para mais, ver Pinterest (2025).

[15] Em 2024, a plataforma registrou lucro pela primeira vez em sua história, mas não especificou quanto de sua receita veio de anúncios. Para mais, ver Goel (2024).

2007, o Google deu um passo importante para expandir esse modelo ao permitir a monetização de vídeos no YouTube por meio de anúncios, compartilhando a receita com os criadores de conteúdo (Burns, 2021).

Os links patrocinados marcaram o rompimento da lógica propagandeada pelo Google de que suas práticas seriam guiadas somente pela eficiência e neutralidade técnica de seus sistemas de recomendação, agora subordinados a seus interesses comerciais (Gonzalez, 2012). Assim, revelam um dos traços mais marcantes da veiculação de anúncios em plataformas de redes sociais: a confusão dos limites entre conteúdos comerciais e não comerciais (Campbell; Grimm, 2019; Reijmersdal; Rozendaal, 2020). Ao incorporar a publicações impulsionadas uma estética semelhante à das mensagens que circulam organicamente nos mesmos espaços digitais, elas podem ser facilmente visualizadas por usuários sem que eles saibam que se trata de uma forma de publicidade (FTC, 2015a).

Já a Meta, quando ainda se chamava (e se resumia ao) Facebook, começou suas operações comerciais de publicidade em 2004 cobrando pela segmentação de público na distribuição de banners online – peças gráficas unindo ilustrações ou vídeos e chamadas em texto, muito semelhantes à estética de anúncios em revistas e jornais impressos e ainda massivamente utilizados em espaços digitais (CADE, 2023; Van Looy, 2016). Nos anos seguintes, o modelo de publicidade do Facebook incorporou novos serviços (Fuchs, 2021) até o lançamento, em 2014, do *Lookalike Audiences*, que inaugurou os posts impulsionados na plataforma (Meta, [*S.d.*]b). Desde então, as receitas advindas do impulsionamento de publicações se tornaram tão fundamentais ao modelo de negócios da Meta que ela vem anunciando seguidas reduções no alcance de publicações orgânicas para dar mais visibilidade às impulsionadas (Peterson, 2016; Samsing, 2018; Wang, 2017).

As experiências pioneiras do Google e da Meta com o impulsionamento e o patrocínio de conteúdo pavimentaram o caminho para que outras plataformas de redes sociais entrassem no mercado da publicidade online, incluindo TikTok, Twitter, LinkedIn, Pinterest e Kwai. A

relevância de ambas empresas é tamanha que muitos autores têm apontado a uma concentração do mercado de publicidade online em um possível duopólio controlado por elas (Fuchs, 2018; van Dijck; Nieborg; Poell, 2019). Não só elas praticamente inauguraram um novo modelo de publicidade como entenderam o que é necessário para a manutenção de sua relevância neste mercado: a retroalimentação de seus sistemas algorítmicos de distribuição de conteúdo com volumes torrenciais de novos dados sobre seus usuários, facilitando a previsão de comportamentos e preferências individuais (Arogyaswamy, 2020; Hermann, 2023). Como as *big tech* desfrutam de maiores recursos financeiros e tecnológicos, elas coletam muitos mais dados sobre seus usuários que quaisquer outras empresas a nível global (Crain, 2021), e logo têm maior capacidade de desenvolver algoritmos sofisticados de distribuição de publicidade, impondo novas barreiras econômicas à entrada de outros competidores neste mercado (CADE, 2023).

Por acumularem tantos dados sobre seus usuários de forma exclusiva, as plataformas vêm ampliando sua relevância como intermediárias na circulação de publicidade online, papel antes reservado a veículos de comunicação tradicionais e empresas de mídia. Embora atuem na curadoria e na distribuição de conteúdo orgânico e de publicidade, as corporações por trás das plataformas insistem que não são empresas de mídia, mas prestadoras de serviços de tecnologia (Napoli; Caplan, 2017). Tratar as plataformas de redes sociais como meras empresas de tecnologia subdimensiona as dimensões culturais, políticas e econômicas de suas operações e desconsidera seu enorme papel no atual ecossistema de mídia e sua capacidade de influenciar no comportamento de audiências diversas. Pesquisadores têm criticado esse enquadramento por ofuscar o elevado grau de determinação que elas exercem sobre a distribuição de conteúdo, especialmente sobre aquele que é pago, e que é embutido em suas ferramentas de microssegmentação e algoritmos de recomendação (Napoli; Caplan, 2017).

Diante de pressões políticas, econômicas e regulatórias, as plataformas de redes sociais investem em esforços de comunicação e *lobby*

para que sejam consideradas como intermediárias, sem responsabilidade tanto pelo conteúdo publicado organicamente por seus usuários quanto pelo conteúdo impulsionado de forma paga e microssegmentada (Ali *et al*., 2019; Gillespie, 2018; Popiel, 2018). A ideia é evitar novas regulações que possam acarretar a perda de vantagem competitiva no mercado de conteúdo digital (Napoli; Caplan, 2017). Independentemente de serem consideradas empresas de mídia ou não, as plataformas de redes sociais, na prática, atuam como agentes econômicos do setor de publicidade, mas não estão sujeitas às mesmas normas aplicadas a outros atores que intermedeiam a oferta de espaços publicitários para anunciantes, incluindo radiodifusores, veículos tradicionais de imprensa e empresas de publicidade *out-of-home,* o que indica um grave problema de concorrência baseado na assimetria regulatória.

Reside aí uma diferença crítica entre a publicidade offline e aquela veiculada nas plataformas de redes sociais: enquanto a primeira é passível de escrutínio público por ser exibida igualmente a toda a audiência, a publicidade nas plataformas de redes sociais é distribuída por algoritmos que operam sem qualquer transparência (Carah *et al*., 2024; Jamison *et al*., 2020). Isso significa que conteúdos diferentes são exibidos para cada usuário e não é possível saber, precisamente, o que é veiculado num dado momento para diferentes segmentos de audiência (Jamison *et al*., 2020). Com a microssegmentação, a publicidade nas plataformas de redes sociais conseguiu dominar o mercado online, apoiando-se na opacidade de sua arquitetura e na complacência do poder público com as suas políticas de governança.

## Microssegmentação da audiência e opacidade

A partir da modelagem de dados pessoais dos usuários, incluindo seus interesses, histórico de navegação e de busca, atividades e interações como curtidas, tempo de visualização, localização, rede de seguidores e conexões, a microssegmentação (do original, *microtargeting*) possibilita o direcionamento detalhado de anúncios, personalizado de acordo com critérios definidos pelos anunciantes.

Com isso, a ideia é encontrar os "usuários certos", que apresentam maior probabilidade de clicar, engajar e comprar aquilo que é anunciado (Carah *et al*., 2024; Papakyriakopoulos *et al*., 2018; Ribeiro *et al*., 2019; Turow, 2011). A publicidade online é, portanto, um grande laboratório em constante evolução, voltado a análises preditivas de padrões de consumo em larga escala (Napoli, 2010; O'Neil, 2021), no qual as plataformas de redes sociais trabalham em busca da formulação de perfis comportamentais cada vez mais detalhados de seus usuários.

Para que anúncios possam circular de forma massiva e personalizada para públicos específicos, a publicidade nas plataformas de redes sociais depende da automatização de sua distribuição (Silveira; Morisso, 2018), que funciona com base em um sistema de leilões (Van Looy, 2016). A definição dos parâmetros do leilão se baseia nas preferências dos anunciantes e nos dados comportamentais dos usuários que se pretende atingir (Nekipelov; Wang, 2017). O Google, por exemplo, diz avaliar e precificar anúncios conforme a relação entre cliques feitos por usuários e a quantidade de impressões (Varian, 2010). Logo, o preço do lance de um anúncio na rede de publicidade da empresa é determinado pela forma como a própria empresa avalia sua "relevância". Para chegar a esta métrica de relevância, a empresa diz levar em consideração a concorrência com outros anunciantes e a análise comportamental dos usuários na plataforma (Zuboff, 2021). Já a Meta afirma considerar também o valor investido pelo anunciante, a probabilidade de transformar impressões em cliques e a "qualidade" do anúncio (Meta, [*S.d.*]d). A única certeza é que, quanto mais segmentado e personalizado para as audiências que se pretende atingir, mais alto será o custo para a veiculação do anúncio (Bromell, 2022). Apesar de as plataformas alegarem que estes são os critérios que elas adotam, não é possível fazer uma verificação externa dos parâmetros que, de fato, influenciam estes leilões, uma vez que este processo é opaco e inauditável, o que limita a compreensão de seu real funcionamento.

Os critérios de precificação e distribuição de anúncios são influenciados pelo algoritmo de distribuição sem o conhecimento pleno do anunciante e do público (Ali *et al*., 2021). Esse modelo opaco permite distorções de mercado e enviesamentos, como o favorecimento de determinados anunciantes em detrimento de outros. Na Índia, por exemplo, jornalistas identificaram que a Meta cobrou valores menores para anúncios veiculados pelo partido que governava o país (Bharatiya Janata Party; BJP) em comparação com os anúncios da oposição, e atribuíram a diferença aos algoritmos da plataforma (Sambhav; Ranganathan, 2022).

A despeito de a microssegmentação ser vendida pelas plataformas como um método mais eficaz e efetivo que outros tipos de distribuição de publicidade no que diz respeito à persuasão das audiências, Armitage *et al*. (2023) apontam que não há evidências contundentes produzidas por organizações independentes das *big tech* que comprovem que essa técnica realmente oferece uma vantagem significativa em relação a outros modos de distribuição. É a própria falta de transparência deste modelo que, inclusive, dificulta a comparação com abordagens alternativas que são menos dependentes de dados pessoais. Como a eficácia da microssegmentação se tornou um consenso no mercado, os anunciantes se sentem dependentes dos serviços de publicidade das plataformas para alcançar bons resultados online. A consequência disso é que aqueles que vendem espaços publicitários online sem as promessas da microssegmentação tendem a enfrentar maiores dificuldades na conquista de clientes e na obtenção de receitas mais altas, pois a *percepção* dos anunciantes sobre a eficácia dos anúncios impacta diretamente nos valores pagos (Armitage *et al*., 2023). Desse modo, a crença na eficácia da microssegmentação é baseada menos em evidências e mais em sua adoção generalizada pelo mercado publicitário.

Os problemas e riscos da publicidade digital baseada em microssegmentação, porém, não se resumem à inconsistência de informações sobre sua verdadeira eficácia. Muitas críticas são direcionadas à intrusividade e à onipresença associadas a esta "vigilância digital constante",

sobre a qual os consumidores não têm controle, mesmo que as plataformas disponibilizem ferramentas de controle de dados que apenas dão um poder de decisão ilusório aos usuários acerca do que eles irão receber (Amitage *et al*., 2023; Ur *et al*., 2012). Para contornar essa situação, usuários enfrentam interfaces pouco intuitivas e configuram preferências em cada plataforma separadamente, o que é pouco prático (Armitage *et al*., 2023).

Ainda que, historicamente, os consumidores tenham aceitado a presença da publicidade para que pudessem ter acesso a conteúdo midiático, a perda de controle sobre seus dados pessoais e, em essência, de sua privacidade, não fazia parte deste "acordo" (Helberger *et al*., 2020). De fato, o controle dos indivíduos sobre seus próprios dados é parte dos direitos fundamentais à proteção de dados e da privacidade, inclusive em ambientes digitais, de forma a proteger a liberdade individual, além da autonomia e dignidade dos cidadãos (ANPD, 2022; Armitage *et al*., 2023). Sem garantias de segurança, pesquisadores apontam para o risco de vazamento de dados pessoais, que podem ser comercializados e usados contra os próprios indivíduos. Por conseguinte, o funcionamento da publicidade digital, como se tem hoje, oferece riscos à segurança dos consumidores e impede que eles exerçam plenamente seus direitos (Armitage *et al*., 2023): trata-se de um modelo em que o lucro se sobrepõe ao direito à privacidade (Crain, 2021).

Um exemplo dessa situação pode ser ilustrada com um dos caso de violações de direitos que foi exposto pela *Federal Trade Commission* nos Estados Unidos, que multou o YouTube e o Google em 170 milhões de dólares pelo uso ilegal de dados pessoais de crianças para a distribuição de publicidade, sem o consentimento dos pais; mais tarde, outro relatório mostrou que a empresa veiculava anúncios mesmo em conteúdos categorizados como "*made for kids*" – isto é, produzidos para o público infantil (Adalytics, [*S.d.*]; FTC, 2019). Há ao menos dois problemas, sendo um para consumidores e outro para anunciantes: primeiro, porque se há a veiculação de anúncios em conteúdos "*made for kids*", há dano aos usuários pela coleta indevida de dados para segmentação

de anúncios; em segundo lugar, se a empresa nega coletar dados de crianças e veicula anúncios mesmo assim, os anunciantes são enganados quanto à eficácia dos serviços de microssegmentação e seus verdadeiros critérios de distribuição (Khan; Bedoya; Slaughter, 2023).

A análise do comportamento dos usuários pode ser feita de forma oportunista, visando explorar suas vulnerabilidades, influenciar suas opiniões e hábitos e induzi-los ao erro (Tufekci, 2014). Isso pode impactar tanto a vida em sociedade quanto a tomada de decisões individuais: os usuários podem ser influenciados a adotar práticas prejudiciais à saúde, cair em fraudes financeiras e rejeitar recomendações de políticas públicas que visam seu próprio bem-estar, entre outras atitudes nocivas (Andreou *et al*., 2019; Cotter *et al*., 2021; Kruikemeier *et al*., 2022; OMS, 2022). A título de exemplo, um vazamento de informações do Facebook revelou que a plataforma podia prever o estado emocional de adolescentes para que anunciantes direcionassem peças àqueles que demonstrassem maior fragilidade, partindo da premissa de que consumidores emocionalmente vulneráveis são mais facilmente persuadidos (Crain, 2021).

O modelo de funcionamento da microssegmentação abre margem para diversas práticas discriminatórias, com base na sexualidade, religião, posicionamento político e estado de saúde dos usuários. No Brasil, estes dados são considerados sensíveis pela Lei Geral de Proteção de Dados (LGPD) e, portanto, exigem maior cuidado em sua utilização, justamente por conta de seu alto potencial discriminatório (Laboratório Nacional de Computação Científica, 2025; Serpro [*S.d.*]). Plataformas da Meta já foram criticadas, nesse sentido, pelo oferecimento de opções de segmentação baseada em critérios como raça, gênero e "afinidade étnica" (Cotter *et al*., 2021; Armitage *et al*., 2023). Estas categorias permitem a reprodução de discriminações sociais, como em um caso em que homens receberam mais anúncios para empregos bem remunerados do que mulheres ao acessarem plataformas do Google (Datta; Tschantz; Datta, 2015).

A microssegmentação também pode ser mal utilizada em campanhas eleitorais, tendo como possíveis fins a manipulação dos votos dos cidadãos, inclusive a partir da veiculação de desinformação (Armitage *et al*., 2023). Como não são visíveis para todos os usuários, anúncios personalizados atuam silenciosamente e longe dos holofotes para limitar o debate público, dificultar a contestação de informações falsas ou enganosas e reduzir a confiança de uma população na democracia e no processo eleitoral (Medert; Otto; Perczel, 2024). A Meta foi uma das empresas a se comprometer publicamente com o combate da publicidade política nociva, mas uma investigação sobre as eleições nos Estados Unidos de 2024 revelou que redes de anunciantes usaram o Facebook e o Instagram para veicular mais de 160 mil anúncios problemáticos sobre temas eleitorais e sociais, exibidos cerca de 900 milhões de vezes para os usuários destas plataformas (Silverman; Bengani, 2024).

A desinformação nas plataformas de redes sociais é intimamente ligada à veiculação de publicidade, uma vez que anúncios, além de conterem desinformação, também podem aparecer ao lado de conteúdos orgânicos nocivos e, dependendo da plataforma em questão, as receitas de publicidade podem ser divididas com os autores destes conteúdos. Assim, a reputação de marcas e anunciantes legítimos pode ser comprometida pela associação a conteúdos ilegais, tóxicos e/ou inapropriados (Hsu, 2022). No Brasil, o NetLab UFRJ tem se dedicado a investigar a indústria online de publicidade fraudulenta, mostrando como anúncios desinformativos, normalmente impulsionados por páginas falsas, que se passam por autoridades e figuras públicas de renome, buscam lesar moral e financeiramente consumidores com a promoção de serviços e ofertas falsas (ver NetLab UFRJ 2023c; 2024a; 2024b; Santini *et al*., 2023; 2024a; 2025). Mesmo que as plataformas afirmem atuar para que conteúdos desinformativos ou fraudulentos não apareçam próximos de anúncios ou sejam monetizados, diferentes pesquisadores têm mostrado que essas medidas não têm sido eficazes (Armitage *et al*., 2023).

Um dos motivos que leva a publicidade microssegmentada nas plataformas de redes sociais a ser tão visada por anunciantes nocivos é

a falta de rigor no seu controle: não é necessário passar por processos de verificação rígidos ou submeter documentos para começar a impulsionar anúncios em plataformas como Facebook, Instagram, Telegram, Google, TikTok e X/Twitter no Brasil (Santini *et al*., 2024b). Para se tornar um anunciante na Meta, por exemplo, basta ter uma conta na plataforma e um método de pagamento, como cartão de crédito (Andreou *et al*., 2019). Portanto, este é um mercado duplamente vantajoso para anunciantes ilegítimos e mal-intencionados, que podem utilizar ferramentas de microssegmentação para atingir pessoas vulneráveis de maneira otimizada e aproveitar o anonimato garantido pelas plataformas para promover práticas prejudiciais aos usuários sem serem identificados e responsabilizados. No fim, a opacidade das redes de publicidade das grandes plataformas é benéfica somente a estes anunciantes, enquanto os riscos à segurança e ao pleno exercício dos direitos fundamentais dos consumidores, a marcas e anunciantes legítimos e à integridade democrática são continuamente aprofundados pela falta de transparência.

## Quem tem medo da transparência da publicidade online?

Todo mercado que movimenta cifras bilionárias todos os anos, como a publicidade em plataformas de redes sociais, necessita de confiabilidade e auditabilidade, e garantir um acesso público aos dados sobre anúncios que circulam nelas é primordial para a proteção dos direitos do consumidor. A expansão da transparência da publicidade veiculada em plataformas e serviços digitais tem sido apontada como uma necessidade global por instituições como a Unesco (2023) e a OCDE (2024). Porém, as *big tech* têm assumido apenas o compromisso de maximização do lucro frente a seus acionistas, deixando de lado a responsabilidade para com seus usuários, a sociedade civil, governos e a comunidade acadêmica (Bromell, 2022).

Diante de escândalos que emergiram de falhas decorrentes de seus sistemas de distribuição microssegmentada de anúncios (ver Ali; Hollgren, 2022; Bossetta, 2020; Ghosh; Scott, 2018; Kreiss; Mcgre-

gor, 2019; Milano; Mittelstadt; Wachter, 2021; Tuttle, 2018), as plataformas têm comunicado alguns esforços pontuais para aumentar sua transparência de anúncios como estratégias de relações públicas. Na prática, poucos avanços significativos ocorreram de fato. Apesar de uma abordagem que parece amigável e conciliadora, as plataformas têm, na verdade, dificultado o acesso a dados fundamentais para o desenvolvimento de pesquisas e a auditabilidade de seus serviços, com soluções ineficientes para proteger consumidores e anunciantes legítimos da atividade ilícita (Armitage *et al*., 2023; Ben-David, 2020; Hoffman, 2022; Leerssen *et al*., 2019).

Os esforços pontuais de transparência das plataformas já foram caracterizados como parte de um "teatro" (Bouko; van Ostaeyen; Voué, 2021) por direcionarem a observação pública para informações mais superficiais, como o conteúdo dos anúncios, ao invés de viabilizarem análises completas e sistemáticas sobre os aspectos técnicos e institucionais dos sistemas de publicidade algorítmica (Carah *et al*., 2024). Contudo, cada vez mais pesquisadores têm direcionado sua atenção às (im)possibilidades de auditoria da publicidade em plataformas de redes sociais, buscando entender como se dão possíveis estratégias discriminatórias de microssegmentação, em que medidas anúncios descumprem legislações locais e servem à disseminação de conteúdo danoso e qual é o papel que deve ser delegado a plataformas digitais, que, por padrão, se isentam da responsabilidade pelo conteúdo veiculado (Conger, 2023; de Vreese; Tromble, 2023).

Estes esforços pontuais por parte das plataformas estão majoritariamente concentrados na criação de repositórios de transparência de anúncios (também chamados apenas de "repositórios de anúncios" ou de "bibliotecas de anúncios"), embora informações sobre a moderação e remoção de anúncios também possam aparecer em relatórios de transparência[16]. Popularizados nos últimos anos como parte de medidas de

[16] Tradicionalmente, um relatório de transparência é um documento voluntário publicado por plataformas digitais para dar maior visibilidade pública a suas ações de moderação e remoção de conteúdo gerado por usuário, sejam elas proativas ou a pedido de entes estatais, também podendo

transparência voluntárias das plataformas de redes sociais e, agora, exigidos por lei em determinadas regiões, como a União Europeia (Comissão Europeia, 2024), os repositórios de transparência de anúncios são bases de dados acessíveis publicamente e atualizadas automaticamente que armazenam informações sobre as peças publicitárias veiculadas nestas plataformas (Leerssen *et al.*, 2019). Os primeiros repositórios de transparência de anúncios foram disponibilizados de forma proativa pelas próprias *big tech* como uma resposta a pressões sociais e para que elas pudessem se blindar de quaisquer possibilidades de regulações rígidas (Leerssen *et al.*, 2019). Entre outros dados encontrados nestes repositórios, figuram o conteúdo das peças; os dias em que elas circularam; quantas vezes elas foram vistas; e quanto elas custaram a seus anunciantes. Idealmente, esses repositórios podem ser acessados tanto por meio de uma interface de programação de aplicações (API)[17] quanto por uma interface de usuário.

No entanto, diante da falta de critérios vinculativos mínimos na maior parte do mundo, estas ferramentas já foram criticadas por serem pouco confiáveis, superficiais e adotarem parâmetros de transparência muito discrepantes entre si (ver Bossetta, 2020; Leerssen *et al.*, 2019; Pershan; Lesplingart, 2024; Rosenberg, 2019; Santini *et al.*, 2024b; Sosnovik; Goga, 2021). Além disso, a disponibilização dos repositórios por parte das plataformas transferiu, em grande medida, a responsabilidade de identificação de anúncios danosos para acadêmicos e para a sociedade civil (Carah *et al.*, 2024), agora incumbidos de utilizar ferramentas que limitavam estrategicamente a busca e a análise sistemática e contextual de dados (Bossetta, 2020; Carah *et al.*, 2024).

O exemplo mais notório é a Biblioteca de Anúncios da Meta, lançada em 2018, logo após as crises de reputação da empresa decor-

incluir detalhes sobre a moderação de conteúdo publicitário. Para mais detalhes, ver o Índice de Transparência de Dados das Plataformas de Redes Sociais, no segundo capítulo deste livro.

[17] Uma API é um dos principais meios programáticos utilizados por pesquisadores para a coleta de dados de plataformas de redes sociais e, no caso deste trabalho, para a extração de dados de repositórios de transparência de anúncios. Para mais detalhes, ver o Índice de Transparência de Dados das Plataformas de Redes Sociais, no segundo capítulo deste livro.

rentes do uso indevido de dados pessoais de usuários do Facebook e do Instagram durante as eleições presidenciais nos Estados Unidos em 2016 e o plebiscito do Brexit (Leerssen *et al*., 2021). Em seguida a seu lançamento, pesquisadores passaram a apontar problemas na utilização da ferramenta, como a remoção indevida de anúncios de sua interface de usuário (Rosenberg, 2019), limitações relacionadas à identificação de anúncios políticos (Pochat *et al*., 2022; Sosnovik; Goga, 2021) e as diferenças nos protocolos de transparência adotados entre diferentes países (Santini *et al*., 2024b). Até hoje, a Biblioteca de Anúncios da Meta oferece mais transparência para anúncios sobre moradia, emprego e crédito nos Estados Unidos e no Reino Unido do que no Brasil (Santini *et al.,* 2023).

Seguindo a iniciativa da Meta, o antigo Twitter lançou, em 2018, o *Ads Transparency Center* (Central de Transparência de Anúncios), um repositório que consistia em uma interface de usuário para acessar o conteúdo dos anúncios veiculados nas plataformas nos sete dias anteriores à data de consulta; anúncios publicados por políticos em campanha nos Estados Unidos também eram acompanhado de dados de segmentação e investimento (Falck, 2018). O X/Twitter, entretanto, declarou proibir a veiculação de anúncios políticos no ano seguinte e, em 2021, suspendeu a ferramenta de transparência (BBC, 2021).

Já o Google lançou sua primeira iniciativa de transparência de anúncios, a Central de Transparência de Anúncios, em maio de 2018, com informações sobre anúncios eleitorais dos Estados Unidos (Walker, 2018). No Brasil, lançou um repositório de anúncios políticos em 2022, após ter firmado uma parceria com o Tribunal Superior Eleitoral (TSE) com o objetivo de diminuir a desinformação eleitoral (Poder 360, 2022). A princípio, eram disponibilizadas apenas peças referentes a pleitos nacionais, como campanhas à Câmara dos Deputados, ao Senado Federal e à Presidência da República, mas a pressão de pesquisadores e da sociedade civil fez com que o repositório fosse ampliado e passasse a abranger, também, candidaturas de nível estadual e distrital (Abraji, 2022).

As diferenças entre estes repositórios escancaram alguns dos maiores problemas das medidas proativas de transparência das plataformas de redes sociais: cada empresa decide dar transparência ao que bem entende, de acordo com as diferentes motivações políticas e interesses econômicos que estiverem em jogo (Bechmann, 2020; Bossetta, 2020). Entre os repositórios de anúncios disponibilizados pelas plataformas, isso é notório com a classificação e separação arbitrária entre a publicidade classificada como "política" e a publicidade comercial. Na maioria dos casos, as plataformas apenas dão transparência a anúncios que as mesmas consideram "políticos", porém a interpretação do que é "político" varia significativamente entre as empresas (Leerssen *et al*., 2019). As disparidades dos graus de transparência entre repositórios também incluem se e como as plataformas verificam o conteúdo e a identidade de seus anunciantes (Santini *et al*., 2024a; 2024b).

A classificação de anúncios como "políticos" é uma decisão estratégica tomada pelas plataformas, sem critério técnico ou regulatório claro e objetivo, para promover alguma transparência e abertura de dados para apenas uma amostra arbitrária de anúncios. Sob esse pretexto, as plataformas se eximem da responsabilidade de disponibilizar dados relevantes e pertinentes sobre todas as peças publicitárias veiculadas, ofuscando questões mais substantivas e fundamentais sobre seus modos de governança e seus serviços de publicidade (Zalnieriute, 2021).

A literatura acadêmica, por sua vez, propõe diversas possibilidades para diferenciar anúncios políticos daqueles que são puramente comerciais, mas a falta de consenso entre pesquisadores favorece a permanência de ambiguidades classificatórias. Dommett e Zhu (2023) mostram que, entre os sentidos que possivelmente podem ajudar nesta distinção, estão a promoção ou os ataques a candidatos e políticos e a bandeiras políticas de indivíduos, partidos, grupos, governos e outras organizações, em especial durante eleições. Os autores também apontam que é possível se apoiar em um sentido ainda mais amplo, que define atividade política como qualquer padrão das relações humanas que

envolva poder, autoridade ou governo, ou ainda tentativas de redefinir parâmetros econômicos, sociais e políticos (Dommett; Zhu, 2023).

De fato, Sosnovik e Goga (2021) demonstram como a categorização de anúncios políticos baseada em definições vagas e pouco delimitadas pelas plataformas é complexa e propensa a erros por parte de anunciantes, moderadores e classificadores baseados em aprendizado de máquina. Analisando anúncios que circularam no ecossistema da Meta, a pesquisa das autoras ilustra a inconsistência da categoria de "anúncios políticos", mostrando que há um alto grau de divergência entre os participantes de um estudo na interpretação do conteúdo das peças analisadas, especialmente em anúncios referentes a questões humanitárias e sociais (Sosnovik; Goga, 2021).

Essa dificuldade de estabelecer contornos precisos para distinguir anúncios "políticos" daqueles "não políticos" se manifesta nas diferenças de classificação de conteúdo impulsionado pelas diferentes plataformas analisadas. Algumas plataformas como Facebook, Instagram e Telegram incorporam a ideia de "anúncios sensíveis" ("*issue ads*"), referentes a temas sociais que assumem um protagonismo em discussões políticas, tais como imigração, direitos humanos e pautas raciais, ao entendimento de anúncios políticos (Leerssen *et al*., 2019; Telegram, [*S.d.*]a). Porém, na prática, essa ideia de "questões sociais e políticas" é bastante ampla e subjetiva (Leerssen *et al*., 2019); Pochat *et al*. (2022) detectaram que 55% dos anúncios rotulados como políticos e/ou socialmente relevantes nas plataformas da Meta, na verdade, não o são. A mesma pesquisa indica ainda que 78% dos anúncios que atendem à definição de anúncios políticos e/ou socialmente relevantes circulam sem essa classificação.

Esta é uma consequência clara das diretrizes de rotulação de conteúdo pago nas plataformas de redes sociais, que delegam a responsabilidade da classificação do teor político e/ou social de um anúncio aos próprios anunciantes. Casos frequentes de peças irregulares ou fraudulentas mostram que anunciantes mal-intencionados se utilizam destas brechas para não declarar seus anúncios como políticos e/ou socialmen-

te relevantes, burlando os termos de uso das plataformas, infringindo leis locais e acarretando uma menor transparência do conteúdo impulsionado (FTC, 2022; Global Action Plan, 2020; Gong, 2019; Kim, 2024; Santini *et al.*, 2023; 2024c). Além disso, durante períodos e eventos importantes, como eleições, é comum encontrar anúncios que ferem regulamentações e resoluções locais, incluindo a insuficiência de dados sobre anunciantes políticos e o impulsionamento de conteúdo eleitoral por pessoas físicas e empresas, prática proibida no Brasil (Mello, 2023; NetLab UFRJ, 2022a; 2022b). Violações similares acontecem mesmo no caso de plataformas que declaram não permitir a veiculação de anúncios políticos, já que frequentemente permitem o impulsionamento de conteúdos que contrariam as suas próprias diretrizes e termos de uso (Dantas, 2023; Mello, 2023; NetLab UFRJ, 2023a).

Portanto, a mera declaração de proibição da veiculação de anúncios políticos não é suficiente para impedir que este tipo de conteúdo pago circule (Santini *et al.*, 2024d). Algumas plataformas alegam não permitir a veiculação de anúncios políticos, mas ao mesmo tempo se omitem de promover qualquer medida efetiva de transparência de publicidade, o que acaba impossibilitando o escrutínio público e a fiscalização de suas próprias diretrizes (Santini *et al.*, 2024d). Assim, por conta das falhas na classificação de anúncios e dado que não é possível transferir essa responsabilidade para as plataformas por completo, pesquisadores recomendam que as plataformas implementem políticas de transparência que abranjam todos os anúncios veiculados (Leerssen *et al.*, 2019; Sosnovik; Goga, 2021).

Em 2023, a entrada em vigor do DSA redefiniu o cenário da transparência de anúncios veiculados em plataformas digitais e de redes sociais na União Europeia por meio da regulamentação governamental. Fora do bloco, o projeto serve de inspiração para outras propostas regulatórias que visam a diminuição da opacidade do mercado de publicidade online em todo o mundo, inclusive no Brasil (Bueno; Canaan, 2024; Helberger; Samuelson, 2024). Embora não seja orientado por preceitos da qualidade de dados, o DSA estabelece critérios mínimos para que to-

das as plataformas e serviços digitais enquadrados disponibilizem dados sobre anúncios. A primeira exigência digna de nota é justamente o fato de o projeto não propor nenhuma diferenciação entre anúncios políticos e não políticos. Nesse cenário, todos os anúncios devem ser arquivados pelas plataformas, por ao menos um ano, em repositórios dedicados, com informações sobre seu alcance e os critérios de microssegmentação de público definidos por seus anunciantes (União Europeia, 2022).

Por exemplo, isso fez com que o Twitter (agora X) tivesse de lançar novamente uma ferramenta de transparência de anúncios na União Europeia, chamada de *X Ads Repository* (X/Twitter, [*S.d.*]b). No mesmo ano, a plataforma voltou a permitir o impulsionamento de anúncios políticos em diversos países (ABAP, 2023; Paul, 2023; X/Twitter, [*S.d.*]d), apesar de não ter expandido o acesso a seu repositório de anúncios por meio de API ou interface de usuário a outras regiões (X/Twitter, [*S.d.*]c). Na Europa, também em decorrência do DSA, o TikTok anunciou o lançamento de sua *Commercial Content Library* (Biblioteca de Conteúdo Comercial), em julho de 2023 (TikTok, 2023). Esta foi a primeira iniciativa de transparência de anúncios da empresa, permitindo o acesso a dados de todos os anúncios veiculados para usuários residentes em países-membros da União Europeia, Reino Unido e Suíça desde 1 de outubro de 2022 (TikTok, [*S.d.*]c). Além de novas ferramentas, o DSA redefiniu práticas adotadas por plataformas de redes sociais em ferramentas já existentes, como nos casos da Biblioteca de Anúncios da Meta e da Central de Transparência de Anúncios do Google, que tiveram de incorporar novas funcionalidades exigidas pela lei.

No entanto, a resistência das plataformas em expandir proativamente as medidas adotadas em decorrência do DSA para outras regiões acaba aprofundando disparidades regionais. Twitter/X e TikTok, para citar alguns casos, não sinalizaram, em nenhum momento, planos de oferecer suas novas ferramentas de transparência de publicidade para fora da União Europeia. Com isso, o recado é claro: apenas a saída regulatória será capaz de mudar o precário cenário atual de transparência.

No Brasil, o TSE é a única fonte de norma jurídica a tratar e regular publicações impulsionadas em plataformas de redes sociais como uma forma de publicidade online. Este entendimento foi instituído em 2017, como forma de estabelecer critérios mínimos para a veiculação de anúncios em campanhas eleitorais (Ferreira; Doederlein, 2018). Em 2024, a Resolução n.º 23.732 do TSE estabeleceu uma definição vinculativa de anúncios políticos no Brasil, determinando que as plataformas que permitem o seu impulsionamento disponibilizem um repositório público com todos eles (Brasil, 2024). As plataformas deveriam ter se adequado às novas regras da corte até o fim de abril seguinte, mas a resolução gerou uma onda de proibições de anúncios políticos, segundo as diretrizes das próprias plataformas. O Google alegou suspender o impulsionamento de anúncios políticos com o argumento de "incapacidade técnica de se adequar" aos termos previstos pelo TSE (Waltenberg, 2024), decisão replicada pelo Kwai (Trindade, 2024), mas anúncios deste tipo continuaram circulando em suas plataformas (Santini *et al.*, 2024d). O X/Twitter, sem anunciar sua decisão formalmente, retirou o Brasil da lista de países onde esse tipo de publicidade é, ao menos em parte, permitida (Iory, 2024), sem disponibilizar um repositório com informações mínimas para a detecção de anúncios irregulares no país.

## Abordagem metodológica

Para aferir o nível de transparência das principais plataformas de redes sociais no Brasil quanto a dados sobre anúncios, um roteiro de avaliação foi elaborado em um processo iterativo e deliberativo, que estabeleceu os parâmetros de análise, as definições conceituais e os critérios de avaliação.

Os parâmetros foram avaliados e justificados por nove pesquisadores do NetLab UFRJ, divididos em duplas que incluíam um especialista em coleta, infraestrutura e processamento de dados e outro com experiência em análises quantitativas e qualitativas de dados sobre anúncios. As duplas também ficaram responsáveis pela revisão de respostas feitas por outros pares, conforme a divisão apresentada na Tabela

2. A distribuição de plataformas entre os pesquisadores levou em consideração o conhecimento prévio e a participação em pesquisas envolvendo dados da plataforma avaliada.

Ao longo do processo de elaboração do índice, a adequação dos parâmetros e a pertinência de suas justificativas foram continuamente deliberadas em conjunto pelos avaliadores e outros pesquisadores envolvidos no estudo. A avaliação foi realizada e revisada ao longo do primeiro semestre de 2024.

**Tabela 2:** Divisão das respostas dos parâmetros por duplas de especialista ($E_n$)

| Plataforma | Pesquisadores responsáveis pela resposta | Pesquisadores responsáveis pela revisão |
|---|---|---|
| LinkedIn | E1 e E2 | A revisão foi realizada em conjunto |
| Meta | E3 e E4 | E5 e E6 |
| Google | E5 e E6 | E3 e E4 |
| Telegram | E7 e E8 | E1 e E2 |
| X/Twitter | E1 e E2 | E7 e E8 |
| TikTok | E4 e E9 | E7 e E8 |
| Kwai | E4 e E9 | E7 e E8 |
| Pinterest | As respostas foram debatidas em conjunto | A revisão foi realizada em conjunto |

O roteiro é composto por 60 parâmetros que analisam seis dimensões de qualidade de dados: completude, conformidade, acessibilidade, consistência, relevância e acurácia. As avaliações foram realizadas e justificadas com base em cinco diferentes fontes de informação: i) a experiência acumulada do NetLab UFRJ; ii) a realização de testes de acesso e coleta de dados sobre anúncios utilizando a API e interface do repositório de anúncios da plataforma de rede social, quando disponíveis; iii) a documentação oficial da API do repositório de anúncios da plataforma de rede social, quando disponível; iv) os relatórios de transparência de moderação de anúncios da plataforma de rede social, quando disponíveis; e v) a literatura acadêmica sobre o tema.

Os obstáculos enfrentados pelo NetLab UFRJ e as soluções desenvolvidas ao longo da construção de sua infraestrutura própria e customizada para o monitoramento constante de anúncios veiculados em diferentes plataformas de redes sociais, por meio de APIs e interfaces de repositório de anúncios, serviram de base para a maior parte das avaliações do ITP. Ao longo dos últimos anos, os pesquisadores do NetLab UFRJ publicaram uma série de estudos sobre diferentes ecossistemas de anúncios online (ver Medeiros *et al.*, 2024; NetLab UFRJ, 2022b; 2023c; 2024a; Santini *et al.*, 2024a; 2024d). Quando necessário, realizamos testes de coleta de dados e de usabilidade na interface do repositório de anúncios para melhor embasar nossas respostas e justificativas.

Também consideramos a documentação da API do repositório de anúncios. A documentação de uma API relata, detalha e explica o seu funcionamento, indicando aos usuários como utilizá-la. Plataformas que disponibilizam APIs comumente incluem documentações para que desenvolvedores possam entendê-las durante a elaboração de requisições. Além disso, consideramos a disponibilização e detalhamento de relatórios de transparência sobre ações de moderação de anúncios irregulares e ilegais por parte das plataformas analisadas. Ainda, nos baseamos na produção acadêmica nacional e internacional publicada em periódicos de impacto, com metodologias desenvolvidas, testadas e revisadas por pares.

Os parâmetros poderiam ser respondidos de três maneiras, segundo avaliações positivas, negativas ou parciais, que, ao fim, embasaram a realização de cálculos de notas para cada uma das plataformas analisadas. A avaliação parcial, que resulta em uma pontuação equivalente a 50% de uma positiva, corresponde a casos em que a plataforma de rede social atende ao mínimo esperado no parâmetro apenas em casos de anúncios sobre temas políticos, eleitorais e/ou de relevância social. Como plataformas online tradicionalmente promovem medidas de transparência mais efetivas para tais anúncios (Carah *et al.*, 2024; Sosnovik; Goga, 2021), a avaliação parcial é uma maneira de pontuar quando há diferenças marcantes entre medidas de transparência para

este tipo de publicidade e outras. Ao todo, a avaliação parcial é aplicável a 39 parâmetros de avaliação. Casos em que um parâmetro não foi aplicável à avaliação de uma plataforma também foram adequadamente indicados e desconsiderados dos cálculos finais de sua pontuação.

## Critérios de avaliação

Apresentamos abaixo a divisão dos 60 parâmetros avaliados para cada plataforma. Os parâmetros sinalizados com * podem receber uma avaliação parcial, caso a transparência se aplique apenas a anúncios políticos, eleitorais e/ou de relevância social, além de poderem ser classificados como positivos ou negativos:

*Completude (21 parâmetros)*

Considerando como completos os dados que podem ser utilizados em situações diversas de pesquisa, essa dimensão indica se os dados recuperados apresentam os atributos indispensáveis para sua compreensão e se é possível realizar um monitoramento sistemático ao coletá-los (Mahanti, 2018). É a dimensão mais importante de nossa avaliação, já que os parâmetros dizem respeito ao detalhamento dos dados sobre anúncios. Por serem peças publicitárias e não publicações orgânicas, entendemos que é essencial tornar públicas mais informações para escrutínio, especialmente sobre critérios de microsegmentação e dados do público alcançado

Parâmetros que compõem a dimensão de *Completude*

P1: A API fornece dados atualizados sobre o conteúdo do anúncio?*

Verifica se a API do repositório de anúncios da plataforma de rede social fornece dados atualizados relevantes sobre o conteúdo do anúncio, como textos e URLs para mídias, por pelo menos um ano após sua última exibição.

P2: A API retorna dados demográficos atualizados sobre o público para o qual o anúncio foi exibido?*

Verifica se a API do repositório de anúncios da plataforma de rede social retorna dados específicos e atualizados sobre a idade e o gênero do público atingido por anúncios, por pelo menos um ano após sua última exibição.

P3: A API disponibiliza dados geográficos atualizados sobre o público para o qual o anúncio foi exibido?*

Verifica se a API do repositório de anúncios da plataforma de rede social disponibiliza dados atualizados sobre a localização geográfica do público atingido por anúncios, por pelo menos um ano após sua última exibição. A unidade federativa do Brasil é a maior granularidade aceita.

P4: A API recupera todos os dados sobre a segmentação do público-alvo definida pelo anunciante?*

Avalia se a API do repositório de anúncios da plataforma de rede social recupera dados atualizados referentes a todos os critérios de segmentação de audiência definidos pelo anunciante no momento de criação e publicação dos anúncios, como a priorização ou a exclusão de segmentos demográficos e geográficos e informações sobre interesses, atitudes, comportamentos e palavras-chave, por pelo menos um ano após sua última exibição.

P5: A API retorna dados atualizados sobre anúncios inativos?*

Verifica se a API do repositório de anúncios da plataforma de rede social retorna, nas respostas às requisições, dados atualizados sobre anúncios inativos, por pelo menos um ano após sua última exibição.

P6: A API disponibiliza dados atualizados sobre os anunciantes que veicularam anúncios na plataforma de rede social?*

Examina se a API do repositório de anúncios da plataforma de rede social disponibiliza dados atualizados e pertinentes sobre os anunciantes que veicularam anúncios na plataforma, pelo menos, no último ano.

P7: A API fornece dados atualizados sobre os financiadores dos anúncios?*

Verifica se a API do repositório de anúncios da plataforma de rede social fornece dados atualizados e pertinentes sobre quem pagou pelo impulsionamento de anúncios na plataforma, pelo menos, no último ano.

P8: A API disponibiliza dados atualizados sobre o período de impulsionamento do anúncio?*

Verifica se a API do repositório de anúncios da plataforma de rede social disponibiliza dados atualizados e relevantes sobre os dias em que foram veiculados os anúncios, por pelo menos um ano após sua última exibição.

P9: A API recupera dados atualizados sobre o engajamento de usuários com o anúncio?*

Observa se, no caso de anúncios que permitem interações, a API do repositório de anúncios da plataforma de rede social recupera dados atualizados referentes ao total de interações realizadas por usuários, como curtidas, comentários, compartilhamentos e cliques, por pelo menos um ano após sua última exibição.

P10: A API permite a aplicação de filtros temporais para a recuperação de dados atualizados?*

Avalia se a API do repositório de anúncios da plataforma de rede social oferece meios para filtrar a recuperação de dados atuais de anúncios veiculados, pelo menos, no último ano, segundo seu período de veiculação.

P11: A API sinaliza, de forma clara e inequívoca, se os anúncios foram feitos por anunciantes verificados ou não verificados?*

Avalia se a API do repositório de anúncios da plataforma de rede social sinaliza claramente se anunciantes responsáveis por anúncios veiculados, pelo menos, no último ano, foram verificados ou não ao longo de seu processo de publicação.

P12: A interface do repositório exibe dados demográficos atualizados sobre o público para o qual o anúncio foi exibido?*

Verifica se a interface do repositório de anúncios da plataforma de rede social exibe dados atualizados sobre idade e gênero do público atingido por anúncios, por pelo menos um ano após sua última exibição.

P13: A interface do repositório exibe dados geográficos atualizados sobre o público para o qual o anúncio foi exibido?*

Verifica se a interface do repositório de anúncios da plataforma de rede social exibe dados atualizados sobre a localização geográfica do público atingido por anúncios, por pelo menos um ano após sua última exibição. A unidade federativa do Brasil é a maior granularidade aceita.

P14: A interface do repositório recupera todos os dados sobre a segmentação do público-alvo definida pelo anunciante?*

Avalia se a interface do repositório de anúncios da plataforma de rede social recupera dados referentes a todos os critérios de segmentação de audiência definidos pelo anunciante no momento de criação e publicação de anúncios, como a priorização ou a exclusão de segmentos demográficos e geográficos e informações sobre interesses, atitudes, comportamentos e palavras-chave, por pelo menos um ano após sua última exibição.

P15: A interface do repositório disponibiliza dados atualizados sobre anúncios inativos?*

Verifica se a interface do repositório de anúncios da plataforma de rede social permite encontrar e visualizar dados atualizados sobre anúncios inativos, por pelo menos um ano após sua última exibição.

P16: A interface do repositório retorna dados atualizados sobre os anunciantes que publicaram anúncios na plataforma de rede social?*

Examina se a interface do repositório de anúncios da plataforma de rede social retorna dados atualizados e pertinentes sobre os anunciantes que veicularam anúncios na plataforma, pelo menos, no último ano.

P17: A interface do repositório disponibiliza dados atualizados sobre os financiadores dos anúncios?*

Verifica se a interface do repositório de anúncios da plataforma de rede social disponibiliza dados atualizados e pertinentes sobre quem pagou pelo impulsionamento de anúncios na plataforma, pelo menos, no último ano.

P18: A interface do repositório fornece dados atualizados sobre o período de impulsionamento do anúncio?*

Observa se a interface do repositório de anúncios da plataforma de rede social fornece dados atualizados e relevantes sobre os dias em que foram veiculados os anúncios, por pelo menos um ano após sua última exibição.

P19: A interface do repositório recupera dados atualizados sobre o engajamento de usuários com o anúncio?*

Observa se, no caso de anúncios que permitem interações, a interface do repositório de anúncios da plataforma de rede social recupera dados atualizados referentes ao total de interações realizadas por usuários, como curtidas, comentários, compartilhamentos e cliques, por pelo menos um ano após sua última exibição.

P20: A interface do repositório permite a aplicação de filtros temporais para a recuperação de dados atualizados?*

Avalia se a interface do repositório de anúncios da plataforma de rede social oferece meios para filtrar a recuperação de dados atuais de anúncios veiculados, pelo menos, no último ano, segundo seu período de veiculação.

P21: A interface do repositório sinaliza, de forma clara e inequívoca, se os anúncios foram feitos por anunciantes verificados ou não verificados?*

Avalia se a interface do repositório de anúncios da plataforma de rede social sinaliza claramente se anunciantes responsáveis por anúncios veiculados, pelo menos, no último ano, foram verificados ou não ao longo de seu processo de publicação.

*Conformidade (12 parâmetros)*

Avalia se a documentação oficial e os dados recuperados estão adequados quanto aos formatos adotados e às normas legais vigentes no país (Mahanti, 2018). Trata-se de uma dimensão exógena, ou seja, relacionada mais ao "'entorno' dos dados do que a eles próprios" e, portanto, mais atrelada "à sua governança e gerência do que ao seu próprio conteúdo" (Barbieri, 2019). Nesta dimensão, avaliamos, por exemplo, a disponibilização e o detalhamento dos relatórios de transparência sobre ações de moderação de anúncios por parte das plataformas de redes sociais e a disponibilização de documentações de APIs facilmente acessíveis e compreensíveis.

Parâmetros que compõem a dimensão de *Conformidade*

P22: O processo de aquisição de dados e a estrutura na qual eles são disponibilizados pela API são estáveis?

Avalia se a estrutura das bases de dados disponibilizadas não muda com frequência e sem aviso prévio, de modo que as aplicações que se integram à API do repositório de anúncios da plataforma de rede social não sejam constantemente impactadas por esse tipo de mudança.

P23: A API sinaliza, de forma clara e inequívoca, conteúdos produzidos por Inteligência Artificial?*

Verifica se a API do repositório de anúncios da plataforma de rede social sinaliza os anúncios cujo conteúdo foi produzido com uso determinante de Inteligência Artificial.

P24: A API retorna dados em formato padronizado?

Verifica se os dados retornados pela API do repositório de anúncios da plataforma de rede social são estruturados de forma a facilitar seu armazenamento e utilização, sendo disponibilizados em formatos que correspondem ao consenso técnico e/ou aos padrões normatizados na área como, por exemplo, datas de acordo com a norma ISO 8601.

P25: A interface do repositório sinaliza, de forma clara e inequívoca, conteúdos produzidos por Inteligência Artificial?*

Verifica se a interface do repositório de anúncios da plataforma de rede social sinaliza os anúncios cujo conteúdo foi produzido com uso determinante de Inteligência Artificial.

P26: A documentação da API está publicada e disponível em acesso aberto?

Verifica se a plataforma de rede social publica na internet a documentação adequada e suficiente para o melhor uso de sua API, com acesso irrestrito e sem a necessidade de cadastro e login.

P27: A documentação da API disponibilizada está escrita de forma clara e exemplificada?

Verifica se a documentação da API do repositório de anúncios da plataforma de rede social está escrita de forma clara, completa e com exemplos específicos que simulam situações reais de utilização, de forma a facilitar a compreensão por usuários sem experiência prévia.

P28: A documentação descreve claramente quais são os termos de uso da API?

Verifica se a documentação da API do repositório de anúncios da plataforma de rede social apresenta seus termos de uso de forma clara e sem ambiguidades, tanto em relação às próprias normas quanto aos aspectos legais diretamente relacionados.

P29: A documentação da API é disponibilizada nativamente em português?

Verifica se a documentação da API do repositório de anúncios da plataforma de rede social é disponibilizada em língua portuguesa, em local fácil de encontrar e de acessar.

P30: A plataforma de rede social produz e disponibiliza publicamente e sem a necessidade de requisição relatórios de transparência detalhados, com dados sobre suas ações de moderação manual e/ou computacional proativa, para impedir o impulsionamento de publicidade ilegal, irregular ou abusiva?

Verifica se a plataforma de rede social disponibiliza publicamente e sem a necessidade de requisição relatórios de transparência, com perio-

dicidade mínima semestral, nos quais detalha informações de interesse público sobre sua atuação no Brasil no que tange à comercialização e à veiculação de anúncios, incluindo dados sobre ações de moderação manual e/ou computacional proativa (sem necessidade de ordem judicial ou requisição extrajudicial).

P31: Os dados dos relatórios de transparência sobre as ações de moderação de anúncios na plataforma de rede social são divididos de acordo com a localização geográfica?

Verifica se os dados dos relatórios de transparência sobre as ações de moderação de anúncios da plataforma de rede social estão agrupados pelas regiões em que residem os usuários impactados pelas peças. A unidade federativa do Brasil é a maior granularidade aceita.

P32: Os dados dos relatórios de transparência sobre as ações de moderação de anúncios da plataforma de rede social são divididos de acordo com o(s) tipo(s) de violação que motivaram a exclusão?

Verifica se os dados dos relatórios de transparência sobre as ações de moderação de anúncios da plataforma de rede social estão agrupados pelo tipo de violação identificada.

P33: Os relatórios de transparência sobre a moderação de anúncios especificam e apresentam informações sobre requisições feitas por entes do Estado à plataforma de rede social?

Verifica se os relatórios de transparência da plataforma de rede social elencam os pedidos realizados por entes do Estado brasileiro, detalhando a natureza dos pedidos, o total de requisições, o volume de solicitações deferidas e indeferidas, o ente estatal que fez a requisição e se o pedido foi feito por via judicial ou extrajudicial.

*Acessibilidade (11 parâmetros)*

Refere-se à facilidade de localizar, acessar, obter e visualizar dados para um determinado fim (Mahanti, 2018). Portanto, não basta que os dados estejam acessíveis, de modo que deve haver condições para sua fácil compreensão e análise por pesquisadores com variados graus de

conhecimento técnico. Nesta dimensão, foram analisados fatores como a disponibilização de API e de interface do repositório e se a plataforma permite a extração total de dados sobre anúncios.

Parâmetros que compõem a dimensão de *Acessibilidade*

P34: A plataforma de rede social disponibiliza API para acessar e coletar dados atualizados sobre todos os tipos de anúncios publicados?*

Verifica se a plataforma de rede social oferece, no Brasil, uma API com ao menos um *endpoint* para acesso e coleta de dados atualizados sobre publicações impulsionadas no último ano.

P35: O acesso à API é gratuito?

Verifica se é necessário algum pagamento para utilizar a API do repositório de anúncios da plataforma de rede social ou se há isenção, pelo menos, no caso de pesquisadores.

P36: A criação de *tokens* para acesso à API pode ser feita de forma gratuita?

Verifica a possibilidade de utilizar, de forma gratuita, mais de um *token* de API do repositório de anúncios da plataforma de rede social a partir de uma mesma conta de desenvolvedor.

P37: É possível criar novos *tokens* de acesso à API sem limitação de quantidade?

Verifica se a plataforma limita a quantidade de *tokens* por usuário para acesso à API do repositório de anúncios da plataforma de rede social.

P38: A API provê uma forma de autenticação que permita a renovação automática simplificada dos *tokens* de acesso, sem qualquer bloqueio à aquisição de dados?

Verifica se os *tokens* disponibilizados para o uso da API do repositório de anúncios da plataforma de rede social não expiram ou se a renovação pode ser feita de forma automática.

P39: É possível extrair os dados diretamente da resposta da API?

Verifica se os dados relativos ao conteúdo e à autoria são retornados diretamente na resposta da API do repositório de anúncios da plataforma de rede social, podendo ser extraídos sem necessidade de redirecionamento para outras janelas.

P40: A API disponibiliza meios para recuperar anúncios a partir de termos de busca?*

Identifica se é possível recuperar dados atualizados sobre anúncios a partir de termos de busca customizados pelo usuário por meio da API do repositório de anúncios da plataforma de rede social.

P41: A API disponibiliza meios para recuperar dados atualizados sobre um anúncio específico?*

Verifica se é possível recuperar dados atualizados sobre anúncios veiculados pelo menos no último ano, por meio da API do repositório de anúncios da plataforma de rede social, a partir de seus identificadores únicos.

P42: A plataforma de rede social disponibiliza interface do repositório de anúncios para acessar dados atualizados sobre todos os tipos de anúncios publicados?*

Verifica se a plataforma de rede social disponibiliza interface para acesso a dados atualizados sobre anúncios, facilitando o desenvolvimento de pesquisas com anúncios, sem necessidade de conhecimentos de programação.

P43: É possível extrair os dados exibidos na interface do repositório?*

Verifica a possibilidade de extrair, por meio de arquivos em formatos amplamente utilizados, dados atualizados exibidos na interface do repositório de anúncios da plataforma de rede social, para utilizá-los em outros aplicativos.

P44: É possível recuperar, na interface do repositório, anúncios atuais e dados atualizados sobre todos os anúncios por meio de termos de busca?*

Verifica a possibilidade de recuperar dados atualizados sobre anúncios, via interface do repositório de anúncios da plataforma de rede social, por meio de termos de busca determinados pelo usuário.

*Consistência (6 parâmetros)*

Avalia se o formato e a apresentação dos dados são consistentes e idênticos em toda a base de dados e para todas as instâncias (Mahanti, 2018). Esta dimensão analisa, por exemplo, se os termos de busca e filtros usados recuperam dados coerentes e não contraditórios e se os dados retornados são diferentes quando o acesso é feito em momentos distintos. A consistência é imprescindível para produzir relatórios precisos e ágeis, pois evita a necessidade de conferência e/ou correção constante dos dados, além de permitir maior auditabilidade.

Parâmetros que compõem a dimensão de *Consistência*

P45: A API indica quando um anúncio foi removido por violar os termos da plataforma de rede social?*

Verifica se a API do repositório de anúncios da plataforma de rede social disponibiliza dados atualizados sobre anúncios excluídos por violação das diretrizes da própria plataforma no Brasil e se eles são sinalizados como removidos.

P46: A API retorna dados persistentes?

Verifica se os dados retornados pela API do repositório de anúncios da plataforma de rede social expiram, sobretudo as URLs.

P47: A API retorna respostas consistentes?

Verifica se os dados retornados por meio da API do repositório de anúncios da plataforma de rede social são sempre os mesmos, ou quase os mesmos, quando mantidos os termos, parâmetros e filtros de busca de uma requisição.

P48: A API retorna respostas coerentes com os parâmetros e filtros utilizados na requisição?

Verifica se os dados retornados pela API do repositório de anúncios da plataforma de rede social de fato correspondem aos termos, pa-

râmetros e filtros utilizados na requisição, sendo possível avaliar essa correspondência a partir dos dados entregues.

P49: A interface do repositório sinaliza quando um anúncio foi removido por violar os termos da plataforma de rede social?*

Verifica se a interface do repositório de anúncios da plataforma de rede social disponibiliza dados atualizados sobre anúncios excluídos por violação das diretrizes da própria plataforma no Brasil e se eles são sinalizados como removidos.

P50: A API recupera os mesmos dados exibidos na interface do repositório?

Verifica se a API do repositório de anúncios da plataforma de rede social apresenta defasagens ou diferenças em relação ao que é exibido na interface do repositório de anúncios, de forma que todos os dados exibidos na interface devem também ser atualizados e passíveis de coleta via API.

*Relevância (6 parâmetros)*

Avalia se os dados são pertinentes para a finalidade à qual se destinam (Mahanti, 2018), ou seja, se estão de acordo com os objetivos da pesquisa e da requisição e se são suficientes para embasar uma análise robusta.

Parâmetros que compõem a dimensão de *Relevância*

P51: É possível filtrar os dados sobre anúncios na API por página ou perfil anunciante?*

Verifica se a API do repositório de anúncios da plataforma de rede social permite a utilização de filtros que busquem por dados atualizados sobre anúncios veiculados por anunciantes específicos a partir de seus identificadores únicos.

P52: A API permite filtrar os dados sobre anúncios de acordo com sua categoria?

Verifica a possibilidade de recuperar dados atualizados sobre anúncios pela API do repositório de anúncios da plataforma de rede

social, segundo as categorias oferecidas pela plataforma aos anunciantes no momento de criação do anúncio.

P53: A API permite filtrar os dados sobre anúncios por localização geográfica?*

Verifica a possibilidade de especificar na API do repositório de anúncios da plataforma de rede social uma localização geográfica, ou mais de uma, para filtrar a coleta de dados atualizados, sendo a unidade federativa do Brasil a maior granularidade aceita.

P54: É possível filtrar os dados sobre anúncios na interface do repositório por página ou perfil anunciante?*

Verifica se a interface do repositório de anúncios da plataforma de rede social permite a utilização de filtros para buscar dados atualizados a partir da definição e seleção de um anunciante específico.

P55: A interface do repositório permite filtrar os dados sobre anúncios de acordo com sua categoria?

Verifica a possibilidade de recuperar dados atualizados sobre anúncios pela interface do repositório de anúncios da plataforma de rede social, de acordo com as categorias oferecidas aos anunciantes no momento de criação e publicação do anúncio.

P56: A interface do repositório permite filtrar os dados sobre anúncios por localização geográfica?*

Verifica a possibilidade de especificar na interface do repositório de anúncios da plataforma de rede social uma localização geográfica, ou mais de uma, para filtrar a coleta de dados atualizados, sendo a unidade federativa do Brasil a maior granularidade aceita.

*Acurácia (4 parâmetros)*

Avalia o quanto os dados disponibilizados e armazenados refletem a realidade e o quão corretamente eles descrevem o objeto, entidade, situação ou fenômeno do mundo real analisado (Mahanti, 2018). Aqui, verificamos se os dados de impressões recebidas por anúncios e valores investidos em seu impulsionamento são suficientemente preci-

sos, de maneira que possamos avaliar estratégias de precificação e segmentação de conteúdo impulsionado.

Parâmetros que compõem a dimensão de *Acurácia*

P57: A API divide as faixas de impressões por segmento de público em pequenos intervalos que possibilitam identificar, com alguma precisão, tendências e estratégias de segmentação de audiência?*

Verifica se a API do repositório de anúncios da plataforma de rede social disponibiliza o volume de impressões dos anúncios, apresentando dados atualizados e divididos em intervalos com amplitude razoável para retratar as impressões de forma próxima ao número real armazenado pela plataforma. Para pontuar neste parâmetro, volumes de até 1.000 impressões devem ser exibidos em intervalos de 100; entre 1.000 e 10.000, em intervalos de 500; entre 10.000 e 100.000, em intervalos de 1.000; acima de 100.000, em intervalos de 10.000; e acima de 1 milhão, em intervalos de 100.000.

P58: A API divide as faixas de investimento em pequenos intervalos que possibilitam identificar, com alguma precisão, tendências e estratégias de precificação de anúncios?*

Verifica se a API do repositório de anúncios da plataforma de rede social recupera dados atualizados sobre investimentos em anúncios, divididos em intervalos com amplitude razoável para retratar o total de investimento de forma próxima ao número real armazenado pela plataforma. Para pontuar neste parâmetro, investimentos de até R$100 devem ser exibidos em intervalos de R$10; entre R$100 e R$1.000, em intervalos de R$100; entre R$1.000 e R$10.000, em intervalos de R$500; até R$100.000, em intervalos de R$1.000; e acima de R$100.000, em intervalos de R$10.000.

P59: A interface do repositório divide as faixas de impressões por segmento de público em pequenos intervalos que possibilitam identificar, com alguma precisão, tendências e estratégias de segmentação de conteúdo?*

Verifica se a interface do repositório de anúncios da plataforma de rede social recupera o volume de impressões dos anúncios, apresentando dados atualizados e divididos em intervalos com amplitude razoável para retratar as impressões de forma próxima ao número real armazenado pela plataforma. Para pontuar neste parâmetro, volumes de até 1.000 impressões devem ser exibidos em intervalos de 100; entre 1.000 e 10.000, em intervalos de 500; entre 10.000 e 100.000, em intervalos de 1.000; acima de 100.000, em intervalos de 10.000; e acima de 1 milhão, em intervalos de 100.000.

P60: A interface do repositório divide as faixas de investimento em pequenos intervalos que possibilitam identificar, com alguma precisão, tendências e estratégias de precificação de anúncios?*

Verifica se a interface do repositório de anúncios da plataforma de rede social disponibiliza dados atualizados sobre investimentos em anúncios, divididos em intervalos com amplitude razoável para retratar o total de investimento de forma próxima ao número real armazenado pela plataforma. Para pontuar neste parâmetro, investimentos de até R$100 devem ser exibidos em intervalos de R$10; entre R$100 e R$1.000, em intervalos de R$100; entre R$1.000 e R$10.000, em intervalos de R$500; até R$100.000, em intervalos de R$1.000; e acima de R$100.000, em intervalos de R$10.000.

## Composição das notas

Entre os 60 parâmetros, 14 foram considerados essenciais à nossa avaliação por indicarem pontos que possibilitam a realização de análises sistemáticas e metodologicamente rigorosas de dados sobre anúncios. Esses parâmetros foram agrupados em seis critérios especiais que compõem 60% da nota, de forma que cada critério especial corresponde a 10% da nota final. Os outros 46 parâmetros correspondem aos 40% restantes da pontuação total e valem 0,87 ponto cada, no caso de uma avaliação positiva, ou cerca de 0,44, no caso de uma avaliação parcial.

Há anos, a literatura acadêmica especializada já vem apontando para as deficiências e limitações dos repositórios de transparência

de anúncios online (ver Bossetta, 2020; Edelson; Lauinger; McCoy, 2020; Leerssen *et al*., 2019; Santini *et al*., 2024b) e muitos dos anseios expostos nessas pesquisas foram atendidos, inclusive, pela regulamentação de serviços digitais na União Europeia. O DSA exige, por exemplo, que as *big tech* disponibilizem API e interface de usuário de seus repositórios de anúncios, por meio das quais deve ser possível acessar informações de segmentação de peças publicitárias (União Europeia, 2022; Estados Unidos da América, 2023). No mesmo sentido, a Resolução 23.732/2024 do TSE (Brasil, 2024) obriga que provedores de serviços de publicidade online que veiculam anúncios políticos-eleitorais forneçam API e interface semelhantes, porém que sejam navegáveis e buscáveis por diferentes parâmetros, como o nome do anunciante e palavras-chave. Todos estes aspectos foram incorporados à formulação dos critérios especiais.

Para pontuar em quatro dos seis critérios especiais, é necessário que a plataforma atenda, ao menos parcialmente, a todos os parâmetros que o compõem. Se uma plataforma for avaliada de maneira negativa em algum dos parâmetros que formam um critério especial, isso é o suficiente para que ela não receba nenhum dos pontos possíveis. Similarmente, para receber todos os pontos aplicáveis, a plataforma precisa ser avaliada de maneira positiva em todos os parâmetros que compõem um critério especial. Assim, uma avaliação parcial em um dos parâmetros que compõem o critério especial, ao lado de outra positiva, já é suficiente para a atribuição de apenas metade dos pontos esperados.

Além disso, dois critérios especiais apresentam dois determinantes (*D1* e *D2*) que são analisados independentemente para definir sua avaliação final. Um dos determinantes é formado por duas perguntas, enquanto o outro é formado apenas por uma. Se ao menos um determinante tiver avaliação positiva ou parcial, isso é suficiente para que o critério especial receba a pontuação esperada por completo ou pela metade, respectivamente. Nenhum ponto é atribuído na avaliação do critério especial nos casos em que os dois determinantes recebem avaliações negativas.

Dessa forma, a distribuição das pontuações das plataformas que compõem o índice se organiza da seguinte maneira:

10 pontos correspondentes ao Critério Especial #1 ("*A plataforma de rede social oferece API para coletar dados sobre o conteúdo de todos os tipos de anúncios publicados?*"), atingidos por plataformas que permitem o acesso e a recuperação sistemática de dados sobre o conteúdo de todos os tipos de anúncios por meio de sua API.

Para pontuar neste critério especial, é preciso atender positiva ou parcialmente aos parâmetros *P34* ("*A plataforma de rede social disponibiliza API para acessar e coletar dados atualizados sobre todos os tipos de anúncios publicados?*") e *P1* ("*A API fornece dados atualizados sobre o conteúdo do anúncio?*"). Entendemos que dados referentes ao conteúdo das peças são fundamentais para uma análise satisfatória das mesmas utilizando ferramentas externas.

10 pontos correspondentes ao Critério Especial #2 ("*A API da plataforma de rede social fornece dados demográficos e geográficos sobre o público que recebeu o anúncio ou sobre os critérios de segmentação definidos pelo anunciante?*"), atingidos por plataformas que permitem o acesso aos dados de segmentação definidos pelos anunciantes ou às informações do público impactado por meio de uma API.

Para pontuar neste critério a plataforma precisa atender positiva ou parcialmente aos parâmetros *P2* ("*A API retorna dados demográficos atualizados sobre o público para o qual o anúncio foi exibido?*") e *P3* ("*A API disponibiliza dados geográficos atualizados sobre o público para o qual o anúncio foi exibido?*"), que compõem o *D1*, ou ao parâmetro *P4* ("*A API recupera todos os dados sobre a segmentação do público-alvo definida pelo anunciante?*"), que compõe o *D2*. Consideramos que a disponibilização dos critérios de segmentação de público por meio de uma API é a principal maneira de possibilitar um entendimento acerca dos algoritmos de microssegmentação da plataforma, bem como das estratégias dos anunciantes. Também é essencial para se conhecer o público atingido, permitindo a identificação de casos de segmentação discriminatória ou abusiva, por exemplo.

10 pontos correspondentes ao Critério Especial #3 ("*A API da plataforma de rede social permite filtrar os dados por termos de busca e por anunciantes de interesse?*"), atingidos por plataformas que oferecem mecanismos eficientes para localizar e filtrar anúncios por meio de uma API.

Para pontuar neste critério a plataforma precisa atender positiva ou parcialmente aos parâmetros *P40* ("*A API disponibiliza meios para recuperar anúncios a partir de termos de busca?*") e *P51* ("É possível filtrar os dados sobre anúncios na API por página ou perfil anunciante?"). As ferramentas de busca e filtragem disponibilizadas pelas ferramentas de transparência das plataformas de redes sociais não devem limitar a elaboração de desenhos de pesquisa consistentes e, consequentemente, impedir que sejam localizados os anúncios relevantes a uma determinada finalidade.

10 pontos correspondentes ao Critério Especial #4 ("*A plataforma de rede social disponibiliza interface de seu repositório de anúncios, pela qual é possível ter acesso a seu conteúdo e extrair seus dados?*"), atingidos por plataformas que permitem o acesso e a coleta de dados sobre todos os tipos de anúncios por meio da interface do repositório de anúncios.

Para pontuar neste critério, é preciso atender positiva ou parcialmente aos parâmetros *P42* ("*A plataforma de rede social disponibiliza interface do repositório de anúncios para acessar dados atualizados sobre todos os tipos de anúncios publicados?*") e *P43* ("É possível extrair os dados exibidos na interface do repositório?"). Entendemos que não basta apenas disponibilizar e exibir o conteúdo dos anúncios em uma interface web, mas também possibilitar que seus dados possam ser coletados e, então, analisados utilizando ferramentas externas.

10 pontos correspondentes ao Critério Especial #5 ("*A interface do repositório da plataforma de rede social disponibiliza dados demográficos e geográficos sobre o público que recebeu o anúncio ou sobre os critérios de segmentação definidos pelo anunciante?*"), atingidos por plataformas que permitem acesso aos dados de segmentação definidos pelos anunciantes

ou às informações do público impactado por meio da interface do repositório de anúncios.

Para pontuar neste critério a plataforma precisa atender positiva ou parcialmente aos parâmetros *P12* ("*A interface do repositório exibe dados demográficos atualizados sobre o público para o qual o anúncio foi exibido?*") e *P13* ("*A interface do repositório exibe dados geográficos atualizados sobre o público para o qual o anúncio foi exibido?*"), que compõem o *D1*, ou ao parâmetro *P14* ("*A interface do repositório recupera todos os dados sobre a segmentação do público-alvo definida pelo anunciante?*"), que compõe o *D2*.

10 pontos correspondentes ao Critério Especial #6 ("*A interface do repositório da plataforma de rede social permite filtrar os dados por termos de busca e por anunciantes de interesse?*"), atingidos por plataformas que oferecem mecanismos eficientes para localizar e filtrar anúncios por meio da interface de seu repositório de anúncios.

Para pontuar neste critério a plataforma precisa atender positiva ou parcialmente aos parâmetros *P44* ("É possível recuperar, na interface do repositório, anúncios atuais e dados atualizados sobre todos os anúncios por meio de termos de busca?") e *P54* ("É possível filtrar os dados sobre anúncios na interface do repositório por página ou perfil anunciante?").

40 pontos correspondentes ao desempenho da plataforma nos 46 parâmetros restantes, dependentes da soma de pontos obtidos a partir de avaliações positivas e parciais em relação ao total de parâmetros aplicáveis.

Assim, a pontuação final do índice é formalmente representada por:

Em que ***ceTotal*** é o total de critérios especiais atendidos positivamente; ***ceParcial*** é o total de critérios especiais atendidos parcialmente; ***ce*** é o número de critérios especiais aplicáveis[18]; ***pTotal*** é o número de

[18] Na avaliação do Telegram, desconsideramos os seis parâmetros de avaliação que compõem os critérios especiais 3 e 6 e readequamos o restante dos cálculos em torno desta decisão.

parâmetros restantes atendidos positivamente; ***pParcial*** é o número de parâmetros restantes atendidos parcialmente; e ***p*** é o número de parâmetros restantes aplicáveis[19].

## Níveis de transparência de dados

Com base nas pontuações finais e para facilitar a interpretação das pontuações obtidas, classificamos e dividimos as plataformas segundo cinco níveis de transparência de dados:

- **Transparência irrelevante ou nula (0 a 20 pontos):** Não oferecem quaisquer medidas para acesso a dados sobre anúncios veiculados no Brasil, seja por meio de interface de usuário ou de API, ou, quando oferecem, apenas disponibilizam conjuntos de dados desatualizados e com grau de completude muito baixo, inviabilizando quaisquer análises.
- **Transparência precária (21 a 40 pontos):** Ainda que disponibilizem medidas de transparência, apenas permitem o acesso a dados sobre anúncios que ainda estão ativos em um dado momento, sem arquivar peças anteriores, inviabilizando a descoberta de amostras significativas de anúncios que circularam no Brasil. Não publicam relatórios de transparência periódicos sobre suas ações de moderação de publicidade no país.
- **Transparência regular (41 a 60 pontos):** Oferecem API e interface do repositório que permitem a navegação por dados sobre anúncios arquivados que circularam no Brasil, mas apenas de peças consideradas políticas, eleitorais e/ou de relevância social. Não publicam relatórios de transparência sobre suas ações de moderação de publicidade no país.
- **Transparência satisfatória (61 a 80 pontos)**: Além de dados sobre anúncios considerados políticos, eleitorais e/ou de relevância social, também arquivam o conteúdo de anúncios

[19] Na avaliação do Google, desconsideramos dois dos critérios padrão restantes e readequamos o restante dos cálculos em torno desta decisão.

comerciais gerais que circularam no Brasil, embora os dados destes não apresentem completude esperada. Divulgam relatórios de transparência sobre suas ações de moderação no país com alguma periodicidade.

- **Transparência ideal (81 a 100 pontos):** Disponibilizam API e interface do repositório de anúncios robustas, permitindo explorar e coletar dados com completude satisfatória sobre todos os tipos de anúncios que circularam no Brasil. Publicam relatórios de transparência em que detalham ações de moderação de anúncios irregulares realizadas proativamente, a pedidos de governos e da justiça e por denúncias de usuários no país.

## Resultados

Nenhuma plataforma avaliada obteve pontuação satisfatória ou ideal quanto às medidas de transparência e acesso a dados sobre anúncios e à qualidade dos dados retornados. A melhor avaliação foi da Meta, com 49,8 pontos, índice considerado regular. Além dela, apenas Telegram pontua na faixa precária e LinkedIn e Google, na faixa irrelevante. X/Twitter, TikTok, Kwai e Pinterest não oferecem quaisquer medidas de transparência de publicidade no Brasil e, por isso, não pontuam em nenhum parâmetro da avaliação.

A seguir, apresentamos uma visão geral da avaliação de cada plataforma. O detalhamento e as justificativas completas estão disponíveis no site do NetLab UFRJ[20].

## Meta

*Transparência de dados regular*

A Meta, cujo ecossistema de publicidade é formado por Facebook, Instagram, Messenger e Audience Network, somou 49,8 pontos em nossa avaliação, sendo sua transparência de publicidade considerada

[20] Disponível em https://netlab.eco.ufrj.br/itp.

regular. Ela é a empresa que apresenta a melhor transparência de anúncios em plataformas de redes sociais no Brasil.

No país, a Biblioteca de Anúncios da Meta (Meta, [*S.d.*]a) é particularmente útil para a investigação de anúncios que tratam de política, eleições e/ou outros temas de relevância social. Na definição da empresa, estes são "tópicos sensíveis que são fortemente debatidos, podem influenciar o resultado de uma eleição ou resultar/relacionar-se com legislação existente ou proposta", como economia, direitos civis, educação, imigração e armamento (Meta, [*S.d.*]c). Os dados e o conteúdo destes anúncios podem ser visualizados na interface (*P42*) e extraídos tanto pela interface (*P43*) quanto pela API (*P1* e *P34*) do repositório, o que leva à pontuação parcial no Critério Especial #1 e no Critério Especial #4.

Anúncios enquadrados nessa categoria são arquivados por sete anos no repositório e podem ser buscados tanto por palavras-chave (*P40* e *P44*) quanto por páginas anunciantes (*P51* e *P54*), o que leva à pontuação parcial da Meta também no Critério Especial #3 e no Critério Especial #6. Informações de veiculação, como período de circulação (*P8* e *P18*) e dados demográficos (*P2* e *P12*) e geográficos do público atingido (*P3* e *P13*), são disponibilizadas em sua interface de usuário e API, fazendo-a pontuar parcialmente no Critério Especial #2 e no Critério Especial #5.

O grande problema da avaliação da Meta está justamente na distinção entre anúncios comerciais e anúncios políticos, eleitorais e/ou de relevância social. Por conta desse fator, que leva à inconsistência na classificação de diversos anúncios (ver NetLab UFRJ, 2023b), a Meta pontua parcialmente em 21 parâmetros de avaliação, incluindo aqueles que formam cinco dos seis critérios especiais. Anúncios que não são categorizados como políticos, eleitorais e/ou de relevância social só podem ser visualizados na interface de usuário do repositório enquanto estão sendo veiculados (*P42*), mas informações específicas sobre sua veiculação não são públicas. Nem a API nem a interface do repositório de anúncios permitem a extração de dados desses anúncios (*P34* e *P43*). Por fim, a Meta também não publica relatórios de transparência sobre

a moderação de anúncios veiculados nas plataformas do ecossistema da empresa (*P30*, *P31*, *P32* e *P33*).

## Telegram

*Transparência de dados precária*

A pontuação da transparência de anúncios do Telegram de 22,8 pontos é considerada precária. A publicidade veiculada no Telegram é bem específica: de acordo com suas políticas e diretrizes (Telegram, [*S.d.*]a), cada anúncio deve ser constituído de um texto e um botão com link, que deve redirecionar os usuários para canais da plataforma, de modo que links para sites externos não são permitidos (Telegram, [*S.d.*]c). Adicionalmente, os anúncios só podem ser veiculados em canais públicos com mais de 1 mil membros e estão limitados a 160 caracteres com espaço. O Telegram é uma exceção dentre as plataformas analisadas por não oferecer opções de microssegmentação de audiência a anunciantes: no momento em que um anúncio é direcionado a um canal, todos seus membros podem visualizá-lo (Telegram, [*S.d.*]b)[21].

A plataforma permite recuperar informações específicas sobre anúncios na mesma API utilizada para coletar dados gerados por usuários. Os poucos dados disponibilizados incluem o conteúdo dos anúncios e a URL para os canais de redirecionamento, o que faz a plataforma pontuar integralmente no Critério Especial #1. Ainda assim, a API do Telegram só permite a coleta de dados sobre anúncios ativos em um canal conhecido e monitorado pelo pesquisador (*P1* e *P34*), não permitindo que sejam coletados dados históricos de anúncios inativos (*P5*).

---

[21] Para o cálculo final da nota do Telegram, desconsideramos seis parâmetros referentes à completude: *P2* ("*A API retorna dados demográficos atualizados sobre o público para o qual o anúncio foi exibido?*"), *P3* ("*A API disponibiliza dados geográficos atualizados sobre o público para o qual o anúncio foi exibido?*"), *P4* ("*A API recupera todos os dados sobre a segmentação do público-alvo definida pelo anunciante?*"), *P12* ("*A interface do repositório exibe dados demográficos atualizados sobre o público para o qual o anúncio foi exibido?*"), *P13* ("*A interface do repositório exibe dados geográficos atualizados sobre o público para o qual o anúncio foi exibido?*"), e *P14* ("*A interface do repositório recupera todos os dados sobre a segmentação do público-alvo definida pelo anunciante?*"). Os parâmetros foram desconsiderados porque a plataforma não disponibiliza aos anunciantes possibilidades de microssegmentação da audiência. Portanto, a avaliação do Telegram foi baseada em 54 parâmetros e quatro dos seis critérios especiais, que passaram a valer 15 pontos cada.

Assim sendo, a plataforma não possibilita que dados sobre anúncios sejam filtrados de qualquer maneira, o que a faz perder pontos no Critério Especial #2. A plataforma tampouco disponibiliza uma interface de usuário do repositório de anúncios (*P42*), fazendo com que ela não pontue no Critério Especial #3 e no Critério Especial #4.

O Telegram também apresenta um fraco desempenho na dimensão de conformidade, posto que que a documentação de sua API não é disponibilizada nativamente em português (*P29*) e o método de recuperação de dados sobre anúncios não é claramente explicado e exemplificado nela (*P27*). Para piorar, a plataforma não disponibiliza relatórios de transparência sobre a moderação de publicidade no contexto brasileiro (*P30*, *P31*, *P32* e *P33*).

## LinkedIn

*Transparência de dados irrelevante*

Somando 18,3 pontos, a transparência de dados de publicidade do LinkedIn é considerada irrelevante, mesmo que a plataforma ofereça tanto uma interface de usuário (*P42*) quanto uma API (*P34*) para acessar seu repositório de anúncios, visto que ambas as soluções são bem precárias. Primeiramente, a interface do repositório não disponibiliza nenhuma opção para a extração de dados sobre anúncios para a realização de análises externas (*P43*), o que leva ao não cumprimento do mínimo esperado no Critério Especial #4. A API cumpre essa função, permitindo que usuários e pesquisadores coletem os dados dos anúncios disponibilizados no repositório, mas ela não permite a recuperação de dados referentes ao conteúdo textual e visual dos anúncios (*P1*), apenas links para acesso online, o que faz com que a plataforma também não cumpra com o esperado no Critério Especial #1.

A plataforma permite que usuários naveguem por dados de todos os anúncios veiculados nela no período de um ano anterior à consulta. Também permite a busca por anúncios segundo palavras-chave (*P40* e *P44*), podendo-se aplicar um filtro de data (*P10* e *P20*). No entanto, não é possível filtrar os dados a partir da seleção de páginas anunciantes

de interesse, apenas buscar por termos presentes nos nomes das mesmas (*P51* e *P54*), o que faz com que a plataforma não atinja o mínimo para pontuar no Critério Especial #3 e no Critério Especial #6.

Tanto via API quanto via interface do repositório, os dados disponibilizados para cada anúncio são consideravelmente limitados: só é possível ter acesso ao conteúdo das peças e informações sobre a página anunciante e a entidade responsável por seu pagamento. Dados relativos a engajamento (*P9* e *P19*), impressões (*P57* e *P59*), investimento (*P58* e *P60*) e período de veiculação (*P8* e *P18*) não são públicos na versão brasileira da ferramenta. Dados sobre a segmentação demográfica e geográfica do público que visualizou o anúncio (*P2*, *P3*, *P12* e *P13*) também não são disponibilizados, levando a plataforma a não pontuar no Critério Especial #2 e no Critério Especial #5. Vale ressaltar que dados relativos ao volume de impressões, aos critérios de segmentação e ao período de circulação são disponibilizados para anúncios que circularam na União Europeia, a fim de atender ao DSA (LinkedIn, [*S.d.*]b).

Ainda que o LinkedIn apresente um relatório com informações sobre a moderação de conteúdos orgânicos que violaram seus termos de uso e sobre pedidos de moderação e envio de dados feitos por entes governamentais (LinkedIn, [*S.d.*]a), este não indica os motivos da moderação de anúncios na plataforma (*P30*, *P31*, *P32* e *P33*).

## Google

### *Transparência de dados irrelevante*

O Google somou 8,2 pontos[22], com sua transparência de anúncios sendo considerada irrelevante por conta de recentes mudanças que diminuíram a disponibilidade de dados sobre anúncios que circulam

[22] Para o cálculo final da nota do Google, desconsideramos dois parâmetros referentes à completude: *P11* ("*A API sinaliza, de forma clara e inequívoca, se os anúncios foram feitos por anunciantes verificados ou não verificados?*") e *P21* ("*A interface do repositório sinaliza, de forma clara e inequívoca, se os anúncios foram feitos por anunciantes verificados ou não verificados?*"). Os parâmetros foram desconsiderados pelo fato de o Google apenas arquivar em seu repositório de publicidade anúncios impulsionados por anunciantes verificados. Portanto, a avaliação do Google foi baseada em 58 parâmetros ao todo.

em suas plataformas no Brasil. Até o início de 2024, era possível extrair dados sobre anúncios políticos e eleitorais veiculados em suas plataformas por meio da API do Google BigQuery, além de visualizá-los e coletá-los por meio da interface de usuário da Central de Transparência de Anúncios do Google. Contudo, em maio de 2024, a empresa proibiu a veiculação destes anúncios (Waltenberg, 2024), fazendo com que só fosse possível coletar dados sobre anúncios políticos e eleitorais veiculados nas plataformas da empresa no Brasil até o fim de abril anterior. Há evidências, porém, de que anúncios políticos continuam circulando sem a devida moderação e transparência (Santini *et al.*, 2024d). Por isso, a empresa foi mal avaliada em todos os critérios referentes à coleta de dados atualizados sobre anúncios de forma programática, a começar pela disponibilidade de uma API que retorne estes dados (*P34*). Pelo mesmo motivo, o Google não pontua em nenhum dos critérios especiais da avaliação do índice.

Além dos anúncios políticos e eleitorais, a Central de Transparência de Anúncios do Google arquiva, por até um ano após a data final de circulação, peças comerciais impulsionadas por anunciantes verificados. Entretanto, como o Google não arquiva também as peças impulsionadas por anunciantes não verificados, consideramos que seu repositório promove uma medida de transparência insuficiente para investigações sistemáticas, baseada em uma amostra de anúncios inauditável e de cuja representatividade não é possível ter certeza (*P42*). No caso dos anúncios impulsionados por anunciantes verificados e disponibilizados na interface do repositório, não é possível ter acesso a suas informações de veiculação como engajamento (*P9* e *P19*), impressões (*P57* e *P59*) e investimento (*P58* e *P60*), apenas à data final de veiculação e ao conteúdo do anúncio. Assim como outras empresas, o Google não oferece no Brasil as mesmas medidas de transparência e acesso a dados sobre anúncios que oferece em países do Norte Global: na União Europeia, por exemplo, a empresa arquiva todos os anúncios veiculados no bloco, bem como todos os dados exigidos pelo DSA (Richardson; O'Connor, 2023).

Independentemente das limitações, ainda consideramos os parâmetros de avaliação sobre o funcionamento técnico da API, que não dependem de dados atualizados para tanto. Por isso, a plataforma pontua em acessibilidade graças ao acesso gratuito (*P35*) e sem limite de criação de *tokens* da API (*P37*). Além disso, as respostas da API são consistentes (*P47*) e coerentes (*P48*) com os parâmetros utilizados nas requisições. Um dos maiores problemas técnicos na recuperação de anúncios por meio da API do Google BigQuery (*P40*) e da interface de usuário da Central de Transparência de Anúncios do Google (*P44*) é a impossibilidade de se buscá-los segundo palavras-chave. Só é possível encontrar anúncios conforme os nomes com os quais os anunciantes se registraram na rede da empresa (*P51* e *P54*), o que prejudica bastante a identificação de conteúdo de interesse – em especial, a identificação de conteúdo irregular e nocivo.

## X/Twitter

*Transparência de dados nula*

O X/Twitter é uma das quatro plataformas a não pontuar em nossa análise, sendo sua transparência de dados de publicidade nula. Além de não disponibilizar uma API (*P34*) ou interface (*P42*) do repositório de anúncios para coleta e análise de dados sobre a publicidade impulsionada no Brasil, o X/Twitter não disponibiliza publicamente quaisquer relatórios de transparência sobre a remoção de anúncios e a suspensão de anunciantes ilegais, irregulares e/ou abusivos (*P30*, *P31*, *P32* e *P33*). Como forma de atender às demandas impostas pelo DSA, a empresa apenas disponibiliza uma API e uma interface de repositório para anúncios que circularam em países-membros da União Europeia (X/Twitter, [*S.d.*]a).

## TikTok

*Transparência de dados nula*

A transparência de dados de publicidade do TikTok é considerada nula, já que não disponibiliza uma interface de usuário (*P42*) ou API (*P34*) para coleta de dados sobre anúncios exibidos a usuários brasileiros. Enquanto isso, para anúncios veiculados em países da União Europeia, Reino Unido e Suíça, o TikTok disponibiliza uma interface de repositório chamada *Commercial Content Library* (TikTok, [*S.d.*]a), na qual arquiva todos os anúncios que tenham sido vistos pelo menos uma vez e que tenham sido publicados de 01 de outubro de 2022 em diante, também permitindo a recuperação de dados por meio de uma API (TikTok, [*S.d.*]b). Além disso, apesar de disponibilizar relatórios de transparência minimamente detalhados sobre conteúdos orgânicos (TikTok, 2024), o TikTok apenas informa a quantidade total de anúncios removidos globalmente, sem especificar a localização e os motivos da moderação (*P30*, *P31*, *P32* e *P33*).

## Kwai

*Transparência de dados nula*

Como o Kwai também não pontua em nenhuma dimensão, sua transparência de dados de publicidade é considerada nula. A plataforma não disponibiliza API (*P34*) nem interface (*P42*) para acesso e coleta de anúncios no Brasil ou em qualquer outro lugar no mundo. Em abril de 2024, a empresa lançou sua Biblioteca de Anúncios Políticos e Eleitorais (Kwai, [*S.d.*]) no Brasil, por meio da qual era possível visualizar poucas peças relacionadas às eleições gerais de 2022. No entanto, cerca de um mês depois, decidiu proibir a veiculação de anúncios políticos e parou de atualizar o repositório (Nóbrega, 2024). No período em que esteve ativa, a biblioteca não permitia o uso de palavras-chave na busca por anúncios (*P44*), limitando sua pesquisa apenas ao nome com o qual os anunciantes se registraram na plataforma (*P54*).

## Pinterest

*Transparência de dados nula*

Por fim, a transparência de dados de publicidade do Pinterest também é considerada nula pelo fato de não disponibilizar interface de usuário (*P42*) ou API (*P34*) do repositório de anúncios no Brasil. A plataforma permite consultar apenas anúncios que circularam em países da União Europeia em uma interface do repositório (Pinterest, [*S.d.*] a). Mesmo nesses países, o Pinterest não oferece uma forma de acessar o repositório por meio de uma API (Mozilla Foundation; Check First, 2024). Conforme a documentação da plataforma (Pinterest, [*S.d.*] b; Pinterest, [*S.d.*]c), os *endpoints* para acessar o repositório de anúncios estariam disponíveis somente na versão 4 da API de negócios do Pinterest, que foi substituída em 2022 pela versão 5 (Pinterest, 2022) e, em 2024, foi completamente descontinuada, não sendo mais possível utilizá-la (Pinterest, [*S.d.*]d). Além disso, o Pinterest não detalha as medidas de moderação aplicadas a anúncios em seus relatórios de transparência (*P30*, *P31*, *P32* e *P33*), limitando-se a afirmar que "as políticas de anúncios são aplicadas de forma diferente do conteúdo orgânico e não estão incluídas neste relatório de transparência" (Pinterest, [*S.d.*]d, n.p.).

## Boas e más práticas na disponibilização de dados sobre anúncios

Com base em nossas avaliações, apresentamos um panorama de medidas que devem ser amplamente adotadas ou evitadas pelas plataformas de redes sociais para garantir um nível satisfatório de transparência e disponibilização de dados sobre anúncios para fins de pesquisa, proteção dos consumidores e auditoria por agentes do mercado. Neste sentido, é essencial que todas as plataformas de redes sociais que permitem o impulsionamento de publicidade no Brasil ofereçam APIs e interfaces de repositórios de anúncios que permitam a visualização e a recuperação de dados atualizados sobre todas as peças veiculadas.

Hoje, apenas Meta e LinkedIn oferecem tanto uma API quanto uma interface do repositório de anúncios para que qualquer interes-

sado consiga consultar o conteúdo e dados atualizados sobre todos os tipos de anúncios. Porém, caso um anúncio veiculado nas plataformas da Meta não tenha sido classificado como político, eleitoral e/ou de relevância social, ele só pode ser consultado enquanto ainda estiver sendo exibido a usuários, não sendo possível extrair seus dados de forma sistemática, impactando negativamente a capacidade de auditoria por agentes externos.

O Google também permite que seu repositório de anúncios seja acessado por meio de uma API, mas só possibilita a extração de dados sobre anúncios políticos-eleitorais veiculados no Brasil até o fim de abril de 2024. Já a interface de seu repositório permite apenas a consulta do conteúdo de peças impulsionadas por anunciantes verificados no último ano, a contar da data de requisição. Ao deixar dados sobre anúncios veiculados por anunciantes não verificados de fora do repositório, o Google torna o consumidor ainda mais vulnerável a anúncios fraudulentos. O Telegram é outra plataforma que impõe limitações sobre a amostra de anúncios passível de ser conhecida, uma vez que sua API somente dá acesso a anúncios ativos em canais já conhecidos por pesquisadores.

Quaisquer ferramentas disponibilizadas para acessar repositórios de anúncios devem apresentar mínimas condições de navegabilidade e utilização. A Meta e o LinkedIn, por exemplo, permitem que dados sobre quaisquer tipos de anúncios veiculados em suas plataformas sejam buscados por palavras-chave definidas pelo usuário interessado, mas o LinkedIn não permite buscas por anunciantes específicos. No caso do Google, acontece o inverso: os dados sobre anúncios apenas podem ser buscados a partir da seleção de um anunciante de interesse, obrigando usuários a conhecerem de antemão os nomes com que anunciantes se registraram na rede da empresa.

A Meta também disponibiliza uma gama variada de filtros de resultados de buscas: a API e a interface de seu repositório possibilitam que dados atualizados sobre anúncios políticos, eleitorais e/ou de relevância social sejam filtrados segundo a unidade federativa brasileira em que se encontram os usuários para os quais eles foram exibidos. O

Google possibilita uma estratégia de busca parecida, mas novamente esbarra na limitação de que só dados desatualizados sobre anúncios políticos e eleitorais veiculados no Brasil até o fim de abril de 2024 podem ser recuperados sistematicamente. Importante ressaltar, entretanto, que nenhuma plataforma oferece filtros de busca por categorias temáticas definidas pelos anunciantes – no caso da Meta, usuários podem filtrar dados sobre anúncios de acordo com categorias como "moradia", "crédito" e "emprego" na União Europeia e nos Estados Unidos.

A Meta é a única a disponibilizar dados atualizados sobre a segmentação geográfica e demográfica do público que visualiza um determinado anúncio por meio da API e da interface de seu repositório de anúncios. No entanto, estes dados, assim como aqueles relativos aos anunciantes, financiadores e período de impulsionamento, estão apenas disponíveis para anúncios considerados políticos, eleitorais e/ou de relevância social. Nestes casos, a Meta também permite a extração sistemática do conteúdo textual dos anúncios para análises a realização de análises em ferramentas externas, algo igualmente feito pelo Telegram, mas não pelo LinkedIn.

Similarmente, a Meta é a única plataforma analisada que permite a recuperação de dados atualizados sobre o investimento e as impressões recebidas por anúncios políticos, eleitorais e/ou de relevância social por meio da API e da interface de seu repositório de publicidade. Estes dados, contudo, são retornados em faixas de valores muito amplas, sendo insuficientes para compreender estratégias de precificação e de segmentação de público. O cenário é ainda pior quando levamos em consideração os dados de engajamento, visto que nenhuma plataforma permite que seja analisada a quantidade de usuários que, de fato, interagem com seus anúncios, impossibilitando avaliações sobre as atitudes e comportamentos do público impactado.

Por fim, o LinkedIn e a Meta, embora esta apenas no caso de anúncios políticos, eleitorais e/ou de relevância social, se destacam na transparência de anúncios excluídos e moderados, sinalizando a remoção, mas permitindo a visualização do conteúdo das peças em seus re-

positórios de anúncios. Contudo, nenhuma plataforma analisada disponibiliza relatórios de transparência sobre a moderação de anúncios realizada no Brasil, ainda que algumas o façam na Europa ou apresentem dados gerais sobre a moderação de publicidade a nível global.

## Perspectivas e considerações finais

Como nenhuma plataforma atingiu um nível ideal ou sequer satisfatório em nossa análise, nossos resultados evidenciam a urgência de melhorias na transparência da publicidade online no Brasil. A transparência de anúncios da Meta, que obteve a melhor pontuação, foi considerada apenas regular, enquanto a de Telegram, LinkedIn e Google foi considerada precária. Ainda mais grave, quatro plataformas analisadas – X/Twitter, TikTok, Kwai e Pinterest – não pontuaram em nenhum dos nossos parâmetros de análise por não oferecerem nenhuma ferramenta ou mecanismo de transparência sobre anúncios veiculados no Brasil, impedindo por completo a auditabilidade de suas operações comerciais no país.

Apenas Meta, Telegram e LinkedIn permitem a recuperação de dados atualizados dos anúncios que circulam em suas plataformas em algum nível, porém com diferentes restrições, mas somente as duas primeiras disponibilizam uma API e uma interface de seus repositórios de anúncios. É fundamental que seja possível recuperar dados atualizados sobre todos os anúncios veiculados nas plataformas de redes sociais por meio de APIs e interfaces de repositórios que sejam públicas e gratuitas para uso, visando o desenvolvimento de pesquisas sobre os impactos sociais da publicidade. Enquanto as APIs garantem acesso programático aos dados, permitindo que os processos de coleta sejam customizados e automatizados e ganhem escala, a interface de usuário facilita o uso do repositório por qualquer pessoa interessada, mesmo que tenha pouco ou nenhum conhecimento técnico ou de programação. É importante que as orientações de uso das APIs sejam disponibilizadas publicamente, tenham fácil acesso, estejam traduzidas para língua portuguesa e tragam regras claras para sua utilização, além de listarem possíveis erros e ofe-

recerem exemplos representativos e compreensíveis sobre as requisições de dados.

É prejudicial para os consumidores diferenciar anúncios políticos, eleitorais e/ou de relevância social dos demais porque essa classificação por parte das plataformas tem se mostrado imprecisa, arbitrária e ineficiente, consequentemente prejudicando a transparência. Essa diferenciação na transparência dos anúncios políticos frente aos demais tem funcionado mais para que plataformas de redes sociais evitem dar transparência plena aos seus serviços de publicidade do que para evitar a manipulação da opinião pública e proteger o consumidor. A API e a interface do repositório de anúncios da Meta, por exemplo, só disponibilizam o histórico de dados sobre anúncios que tratam de temas considerados políticos, condicionando a capacidade de analisar com profundidade a publicidade distribuída nas plataformas da empresa a uma categorização falha dos anúncios.

Da mesma forma, as políticas de verificação dos anunciantes para venda e autorização de anúncios nas plataformas de redes sociais precisam ser fortalecidas. Estelionatários e anunciantes ilegítimos frequentemente usam perfis falsos para se passar por instituições e figuras públicas conhecidas, enganando consumidores sem precisar comprovar sua identidade previamente. Muitas vezes, as plataformas apenas percebem este problema e os moderam depois da venda e veiculação dos anúncios, quando o estrago já está feito e os consumidores já foram lesados. A ausência de sistemas de controle e verificação dos anunciantes não justifica a falta de transparência na publicidade impulsionada em plataformas de redes sociais. O Google, notadamente, disponibiliza apenas uma amostra insuficiente e arbitrária de anúncios em seu repositório, sem atender adequadamente ao interesse público. Mecanismos de verificação deveriam ser uma forma de proteção e responsabilidade das plataformas frente aos consumidores e anunciantes legítimos, mas, no fim, são usados como artifício para manter a opacidade de suas operações comerciais.

Mesmo as plataformas que disponibilizam ferramentas para acessar seus repositórios de anúncios entregam dados com diversos problemas de qualidade, especialmente de completude. Para garantir a qualidade geral dos dados, é essencial que as plataformas de redes sociais os disponibilizem de forma fidedigna em relação a seus próprios bancos de dados. Entretanto, os principais projetos de regulação de plataformas digitais e de redes sociais em todo o mundo, aprovados ou em debate, não versam explicitamente sobre a qualidade e padronização dos dados sobre anúncios. Nesse sentido, o Brasil tem a oportunidade de aprender com as limitações detectadas em propostas de outros países e se posicionar na vanguarda da discussão sobre a importância da transparência vinculada à qualidade de dados.

Alguns dos problemas mais persistentes na qualidade dos dados sobre anúncios das plataformas envolvem as informações sobre microssegmentação do público. Ainda que as técnicas de microssegmentação sejam o principal diferencial dos anúncios nas plataformas de redes sociais, dados que indiquem as preferências definidas pelos anunciantes são escassos. A Meta é a única plataforma a permitir a recuperação de dados atualizados sobre a segmentação demográfica e geográfica de usuários que visualizaram os anúncios, mas somente nos casos daqueles classificados como políticos, eleitorais e/ou de relevância social. Os dados de perfilamento da audiência devem ser acessíveis, completos e precisos para garantir transparência, auditabilidade e proteção ao consumidor. É o caso da União Europeia, onde as plataformas enquadradas pelo DSA são obrigadas a informar o número exato de usuários impactados por todos os anúncios veiculados nelas e os critérios de segmentação definidos pelos anunciantes para a distribuição dos mesmos.

Atualmente, muito se discute sobre a moderação, pelas plataformas, de conteúdos gerados por usuários e as possíveis implicações destas políticas e práticas para a liberdade de expressão online, enquanto pouco se debate caminhos para a moderação de conteúdo publicitário, fundamental para a proteção dos direitos dos consumidores em rede, e a transparência que deve acompanhá-la. Constatamos que nenhuma

plataforma divulga relatórios de transparência sobre a moderação de anúncios no Brasil, com o fim de detalhar ações para coibir anúncios ilegais, irregulares ou abusivos. É preciso garantir que estes relatórios sejam periodicamente publicados, com dados detalhados sobre o volume de anúncios removidos e anunciantes suspensos, bem como sobre os diferentes tipos de irregularidades identificadas, com o objetivo de proteger os consumidores e os anunciantes legítimos e, ao mesmo tempo, desestimular anunciantes maliciosos. Além disso, é importante diferenciar as ações de remoção de anúncios irregulares que foram definidas pelas próprias plataformas daquelas realizadas a pedido da Justiça ou entes governamentais no Brasil, acompanhadas de suas justificativas e com informações sobre as regiões em que residem os usuários impactados por eles. Idealmente, é necessário que esses relatórios sigam critérios e padrões específicos, para que seja possível compará-los.

É importante ressaltar que o acesso a dados sobre a publicidade nas plataformas de redes sociais é consideravelmente mais restrito no Brasil do que na Europa, sobretudo nos casos de X/Twitter, TikTok, Google e Pinterest, graças aos esforços de regulação de serviços digitais no continente. Por conta de obrigações impostas pelo DSA, estas plataformas disponibilizam repositórios de anúncios, acessíveis via APIs e/ou interfaces de usuário, para a consulta de dados atualizados sobre todas as peças veiculadas em países-membros da União Europeia. Assim sendo, argumentamos que seria possível oferecer uma infraestrutura de transparência robusta, similar a esta, no Brasil e em outros países, mas que as plataformas decidem não fazê-lo por motivos políticos e comerciais.

Por fim, consideramos que os achados aqui reunidos reforçam a ideia de que grandes plataformas digitais e de redes sociais têm se consolidado cada vez mais como alguns dos agentes mais importantes e poderosos no setor de publicidade, mas sem o ônus de transparência e responsabilidade de outros meios de comunicação. Essa assimetria regulatória garante forte vantagem competitiva para as *big tech* e gera assimetrias em relação a seus concorrentes. A falta de transparência das redes sociais no Brasil não é acaso, mas uma decisão estratégica das em-

presas, que diferenciam o tratamento dado a usuários no Sul Global em comparação aos europeus e norte-americanos. E essa decisão dificulta a aplicação de normas e legislações locais, como o Código de Defesa do Consumidor, dentre outras regulamentações.

## Referências


ABAP. Associação Brasileira de Agências de Publicidade. X (ou Twitter) vai permitir a volta de anúncios políticos. *ABAP*, [*S.l.*], 1 set. 2023. Disponível em: https://www.abap.com.br/x-ou-twitter-vai-permitir-a-volta-de-anuncios-politicos/. Acesso em: 1 ago. 2024.

ABRAJI. Democracia pede socorro. *Abraji*, 2022. Disponível em: https://www.abraji.org.br/publicacoes/democracia-pede-socorro. Acesso em: 1 ago. 2024.

ADALYTICS. Are YouTube Advertisers Inadvertently Harvesting Data From Millions of Children?. *Adalytics*, [*S.d.*]. Disponível em: https://adalytics.io/blog/are-youtube-ads-coppa-compliant. Acesso em 31 out. 2024.

ALI, Marwa; HOLLGREN, Mikaela. *"Big brother sees you"*: A qualitative study on users' experiences with targeted advertising on Facebook. [*S.l.*], 2022. Disponível em: https://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-91003. Acesso em: 20 fev. 2025.

ALI, Muhammad.; SAPIEZYNSKI, Piotr; BOGEN, Miranda; KOROLOVA, Aleksandra; MISLOVE, Alan; RIEKE, Aaron. Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Biased Outcomes. ACM ON HUMAN-COMPUTER INTERACTION, [*S.l.*], v. 3, n. CSCW, p. 199:1-199:30, 7 nov. 2019. *Anais* [...]. Nova Iorque: Association for Computing Machinery, 2019. Disponível em: https://doi.org/10.1145/3359301. Acesso em: 1 ago. 2024.

ANANNY, Mike; CRAWFORD, Kate. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, [*S.l.*], n. 20. n. 3, p. 973-989, 2024. Disponível em: https://doi.org/10.1177/1461444816676645. Acesso em: 18 fev. 2025.

ANDREOU, Athanasios; SILVA, Marcio; BENEVENUTO, Fabrício; GOGA, Oana; LOISEAU, Patrick; MISLOVE, Alan. Measuring the Facebook advertising ecosystem. *In*: NETWORK AND DISTRIBUTED SYSTEMS SECURITY SYMPOSIUM: Privacy on the web, fev. 2019, San Diego, CA. *Anais* [...]. [*S.l.*]: NDSS, 2019. Disponível em: https://www.ndss-symposium.org/wp-content/uploads/2019/02/ndss2019_04B-1_Andreou_paper.pdf. Acesso em: 27 maio 2023.

ANPD. Autoridade Nacional de Proteção de Dados. Proteção de Dados Pessoais agora é um direito fundamental. *Autoridade Nacional de Proteção de Dados*, [*S.l.*], 10 fev. 2022. Disponível em: https://www.gov.br/anpd/pt-br/protecao-de-dados-pessoais-agora-e-um-direito-fundamental. Acesso em: 31 out. 2024.

ARMITAGE, Catherine; BOTTON, Nick; DEJEU-CASTANG, Louis; LEMOINE, Laureline. Study on the impact of recent developments in digital advertising on privacy, publishers and advertisers. *European Commission*, 2023. Disponível em: https://op.europa.eu/en/publication-detail/-/publication/8b950a43-a141-11ed-b508-01aa75ed71a1/language-en. Acesso em: 2 ago. 2024.

AROGYASWAMY, Bernard. Big tech and societal sustainability: An ethical framework. *AI & SOCIETY*, [*S.l.*], v. 35, n. 4, p. 829–840, 1 dez. 2020. Disponível em: https://doi.org/10.1007/s00146-020-00956-6. Acesso em: 1 ago. 2024.

BARBIERI, Carlos. *Governança de Dados*: Práticas, conceitos e novos caminhos. Rio de Janeiro: Alta Books, 2019.

BATINI, Carlo; SCANNAPIECO, Monica. *Data Quality Concepts, Methodologies and Techniques*. Nova Iorque: Springer Berlin Heidelberg, 2006. Disponível em: https://link.springer.com/book/10.1007/3-540-33173-5. Acesso em: 1 nov. 2024.

BBC. Twitter to ban all political advertising. *BBC*, [*S.l.*], 31 out. 2019. Disponível em: https://www.bbc.com/news/world-us-canada-50243306. Acesso em: 01 out. 2024.

BECHMANN, Anja. Tackling Disinformation and Infodemics Demands Media Policy Changes. *Digital Journalism*, [*S.l.*], v. 8, n. 6, p. 855–863, 2020. Disponível em: https://doi.org/10.1080/21670811.2020.1773887. Acesso em: 20 fev. 2025.

BEN-DAVID, Anat. Counter-archiving Facebook. *European Journal of Communication*, [*S.l.*], v. 35, n. 3, p. 249–264, 1 jun. 2020. Disponível em: https://doi.org/10.1177/0267323120922069. Acesso em: 1 ago. 2024.

BORGES, Eyder. O Poder do Inventário Estratégico na Mídia Programática!. *redmedia*, [*S.l.*], 21 set. 2023. Disponível em: https://www.redmedia.com.br/inventario/. Acesso em: 01 out. 2024.

BOSSETTA, Michael. Scandalous Design: How Social Media Platforms' Responses to Scandal Impacts Campaigns and Elections. *Social Media + Society*, [*S.l.*], v. 6, n. 2, p. 1–4, 1 abr. 2020. Disponível em: https://doi.org/10.1177/2056305120924777. Acesso em: 1 ago. 2024.

BOUKO, Catherine; VAN OSTAEYEN, Pieter; VOUÉ, Pierre. Facebook's policies against extremism: Ten years of struggle for more transparency. *First Monday*, [*S.l.*], v. 26, n. 9, p. 1–22, 2021. Disponível em: https://firstmonday.org/ojs/index.php/fm/article/view/11705. Acesso em: 1 ago. 2024.

BRASIL. Tribunal Superior Eleitoral. Resolução n.º 23.732, de 27 de fevereiro de 2024. Altera a Res.-TSE nº 23.610, de 18 de dezembro de 2019, dispondo sobre a propaganda eleitoral. Brasília, DF, 27 fev. 2024. Disponível em: https://www.tse.jus.br/legislacao/compilada/res/2024/resolucao-no-23-732-de-27-de-fevereiro-de-2024. Acesso em: 5 ago. 2024.

BRIANT, Emma L.; BAKIR, Vian. *Routledge Handbook of the Influence Industry*. Abingdon: Routledge, 2024. Disponível em: https://www.routledge.com/Routledge-Handbook-of-the-Influence-Industry/Briant-Bakir/p/book/9781032188997. Acesso em: 31 jul. 2024.

BROMELL, David. The Business Models of Big Tech. *In*: BROMELL, David (org.). *Regulating Free Speech in a Digital Age*: Hate, Harm and the Limits of Censorship. Cham: Springer International Publishing, 2022. p.

55–80. Disponível em: https://doi.org/10.1007/978-3-030-95550-2_3. Acesso em: 18 fev. 2025.

BUENO, Thales M; CANAAN, Renan G. The Brussels Effect in Brazil: Analysing the impact of the EU digital services act on the discussion surrounding the fake news bill. *Telecommunications Policy*, [*S.l.*], v. 48, n. 5, p. 102757, 2024. Disponível em: https://www.sciencedirect.com/science/article/pii/S0308596124000545. Acesso em: 20 fev. 2025.

BURNS, Kelli. S. The History of Social Media Influencers. *In*: WATKINS, Brandi (Org.). *Research Perspectives on Social Media Influencers and Brand Communication*. Lanham: Rowman & Littlefield, 2020. Disponível em: https://books.google.com.br/books?id=KXwGEAAAQBAJ. Acesso em 10 set. 2024.

CADE. Conselho Administrativo de Defesa Econômica. *Mercados de Plataformas Digitais*, Brasília, ago. 2023. Disponível em: https://cdn.cade.gov.br/Portal/centrais-de-conteudo/publicacoes/estudos-economicos/cadernos-do-cade/Caderno_Plataformas-Digitais_Atualizado_29.08.pdf. Acesso em: 1 ago. 2024.

CAMPBELL, Colin; GRIMM, Pamela E. The Challenges Native Advertising Poses: Exploring Potential Federal Trade Commission Responses and Identifying Research Needs. *Journal of Public Policy & Marketing*, [*S.l.*], v. 38, n. 1, p. 110–123, 1 jan. 2019. Disponível em: https://doi.org/10.1177/0743915618818576. Acesso em: 1 ago. 2024.

CARAH, Nicholas; HAYDEN, Lauren; BROWN, Maria-Gemma; ANGUS, Daniel; BROWNBILL, Aimee; HAWKER, Kiah; TAN, Xue Y.; DOBSON, Amy; ROBARDS, Brady. Observing "tuned" advertising on digital platforms. *Internet Policy Review*, [*S.l.*], v. 13, n. 2, p. 1–16, 26 jun. 2024. Disponível em: https://policyreview.info/articles/analysis/observing-tuned-advertising-digital-platforms. Acesso em: 1 ago. 2024.

COMISSÃO EUROPEIA. Questions and answers on the Digital Services Act. *Comissão Europeia*, Bruxelas, 22 fev. 2024. Disponível em: https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_2348. Acesso em: 4 abr. 2024.

CONAR. Código Brasileiro de Autorregulamentação Publicitária v. 2021/2022. *CONAR*, São Paulo, 2021. Disponível em: http://www.conar.org.br/pdf/codigo-conar-2021_6pv.pdf. Acesso em: 1 ago. 2024.

CONGER, Kate. Twitter to Relax Ban on Political Ads. *The New York Times*, São Francisco, 3 jan. 2023. Disponível em: https://www.nytimes.com/2023/01/03/technology/twitter-political-ads.html. Acesso em: 1 ago. 2024.

COTTER, Kelley; MEDEIROS, Mel; PAK, Chankyung; THORSON, Kjerstin. "Reach the right people": The politics of "interests" in Facebook's classification system for ad targeting. *Big Data & Society*, [*S.l.*], v. 8, n. 1, p. 1–16, 10 mar. 2021. Disponível em: https://doi.org/10.1177/2053951721996046. Acesso em: 1 ago. 2024.

CRAIN, Matthew. *Profit over Privacy*. Mineápolis: University of Minnesota Press, 2021. Disponível em: https://manifold.umn.edu/projects/profit-over-privacy. Acesso em: 1 ago. 2024.

CURRY, David. TikTok App Report 2025. *Business of Aps*, 18 mar. 2025. Disponível em: https://www.businessofapps.com/data/tiktok-app-report/. Acesso em: 1 abr. 2025.

DANTAS, Dimitrius. Google gastou R$ 837 mil em anúncio para atacar PL das Fake News, admitem plataformas ao STF. *O Globo*, Brasília, 31 maio 2023. Disponível em: https://oglobo.globo.com/politica/noticia/2023/05/plataformas-admitem-ao-stf-que-ataques-do-google-contra-pl-das-fake-news-violaram-suas-regras-de-publicidade.ghtml. Acesso em: 1 ago. 2024.

DATTA, Amit; TSCHANTZ, Michael C.; DATTA, Anupam. Automated Experiments on Ad Privacy Settings. *In*: PRIVACY ENHANCING TECHNOLOGIES SYMPOSIUM, 1., 2015, n. 1, p. 92–112, Filadélfia, EUA. *Anais* [...]. [*S.l.*]: Privacy Enhancing Technologies, 2015. Disponível em: https://petsymposium.org/popets/2015/popets-2015-0007.php. Acesso em: 31 out. 2024.

DE VREESE, Claes; TROMBLE, Rebekah. The Data Abyss: How Lack of Data Access Leaves Research and Society in the Dark. *Political*

*Communication*, [*S.l.*], v. 40, n. 3, p. 356–360, 4 maio 2023. Disponível em: https://doi.org/10.1080/10584609.2023.2207488. Acesso em: 1 ago. 2024.

DOBBER, Tom; KRUIKEMEIER, Sanne; HELBERGER, Nanali; GOODMAN, Ellen. Shielding citizens? Understanding the impact of political advertisement transparency information. *New Media & Society*, [*S.l.*], v. 26, n. 11 p. 6715–6735, 2023. Disponível em: https://doi.org/10.1177/14614448231157640. Acesso em: 1 ago. 2024.

DOMMETT, Katharine.; POWER, Sam. Monitoring digital election campaigns: Assessing the transparency ecosystem in the United Kingdom. *Politics*, [*S.l.*], v. 44, n. 1, p. 119–139, 14 mar. 2024. Disponível em: https://doi.org/10.1177/02633957231156084. Acesso em: 1 ago. 2024.

DOMMETT, Katharine; ZHU, Junyan. What is an online political advert? An interrogation of conceptual challenges in the formation of digital policy response. *Policy & Internet*, Sydney, v. 15, n. 4, p. 713–730, 2023. Disponível em: https://doi.org/10.1002/poi3.350. Acesso em: 1 ago. 2024.

EDELSON, Laura; LAUINGER, Tobias; MCCOY, Damon. A Security Analysis of the Facebook Ad Library. *In*: IEEE SYMPOSIUM ON SECURITY AND PRIVACY (SP), 2020, p. 661–678, São Francisco. *Anais* [...]. [*S.l.*]: IEEE, 2020. Disponível em: https://ieeexplore.ieee.org/document/9152626/authors#authors. Acesso em: 1 ago. 2024.

EDELSON, Laura; CHUANG, Jason; FOWLER, Erika F.; FRANZ, Michael; RIDOUT, Travis N. Universal Digital Ad Transparency. Rochester, NY: Social Science Research Network, 2021. *SSRN Scholarly Paper*. Disponível em: https://papers.ssrn.com/abstract=3898214. Acesso em: 18 fev. 2025.

ESTADOS UNIDOS DA AMÉRICA. Text - S.1876: Platform Accountability and Transparency Act. *118th Congress*, Washington, 6 ago. 2023. Disponível em: https://www.congress.gov/bill/118th-congress/senate-bill/1876/text. Acesso em: 1 ago. 2024.

FALCK, Bruce. Providing more transparency around advertising on Twitter. *X/Twitter*, 28 jun. 2018. Disponível em: https://blog.x.com/en_us/topics/company/2018/Providing-More-Transparency-Around-Advertising-on-Twitter>. Acesso em: 1 ago. 2024.

FERREIRA, Claúdio; DOEDERLEIN, Natalia. TSE define regras para propaganda eleitoral na internet. *Agência Câmara de Notícias*, [*S.l.*], 28 ago. 2018. Disponível em: https://www.camara.leg.br/noticias/544081-tse-define-regras-para-propaganda-eleitoral-na-internet/. Acesso em: 1 ago. 2024.

FERREIRA, Vanessa de M. Porque é necessário que um post publicitário contenha as palavras "publi", "publipost" ou "patrocinado"?. *Jusbrasil*, [*S.l.*], 2022. Disponível em: https://www.jusbrasil.com.br/artigos/porque-e-necessario-que-um-post-publicitario-contenha-as-palavras-publi-publipost-ou-patrocinado/1512310717. Acesso em: 01 out. 2024.

FINGER, Matthias. Algorithms as Public Policy: How to Regulate Them? *In*: FINGER, Matthias. Regulating Digital Platforms. *Network Industries Quarterly*, Lausanne, Suíça, v. 21, n. 4, p. 10–14, 2019. Disponível em: https://cadmus.eui.eu/bitstream/handle/1814/65445/NIQ%20Vol%2021%20-%20Issue%204%20-%20December%202019.pdf. Acesso em: 1 ago. 2024.

FTC. Federal Trade Commission. How to Make Effective Disclosures in Digital Advertising. *Federal Trade Commission*, mar. 2013. Disponível em: https://www.ftc.gov/sites/default/files/attachments/press-releases/ftc-staff-revises-online-advertising-disclosure-guidelines/130312dotcom-disclosures.pdf. Acesso em: 1 ago. 2024.

FTC. Federal Trade Commission. Native Advertising: A Guide for Businesses. *Federal Trade Commission*, 22 dez. 2015a. Disponível em: https://www.ftc.gov/business-guidance/resources/native-advertising-guide-businesses. Acesso em: 1 ago. 2024.

FTC. Federal Trade Commission. Enforcement Policy Statement on Deceptively Formatted Advertisements. *Federal Trade Commission*, 22 dez. 2015b. Disponível em: https://www.ftc.gov/legal-library/browse/com-

mission-enforcement-policy-statement-deceptively-formatted-advertisements. Acesso em: 1 ago. 2024.

FTC. Federal Trade Commission. Google and YouTube Will Pay Record $170 Million for Alleged Violations of Children's Privacy Law. *Federal Trade Commission*, 4 set. 2019. Disponível em: https://www.ftc.gov/news-events/news/press-releases/2019/09/google-youtube-will-pay--record-170-million-alleged-violations-childrens-privacy-law. Acesso em: 31 out. 2024.

FTC. Federal Trade Commission. Who experiences scams? A story for all ages. *Federal Trade Commission*, 9 nov. 2022. Disponível em: https://www.ftc.gov/news-events/data-visualizations/data-spotlight/2022/12/who-experiences-scams-story-all-ages. Acesso em: 1 ago. 2024.

FTC. Federal Trade Commission. 16 CFR Part 255: Guides Concerning the Use of Endorsements and Testimonials in Advertising. *Federal Trade Commission*, 5 jun. 2023. Disponível em: https://www.ftc.gov/legal-library/browse/federal-register-notices/16-cfr-part-255-guides-concerning-use-endorsements-testimonials-advertising. Acesso em: 1 ago. 2024.

FUCHS, Christian. The Google and Facebook Online Advertising Duopoly. *In*: FUCHS, Christian. *The online advertising tax as the foundation of a public service Internet*: A CAMRI extended policy report. Londres: University of Westminster Press, 2018. Disponível em: https://www.jstor.org/stable/j.ctv5vddk0.5. Acesso em: 1 ago. 2024.

FUCHS, Jay. How Facebook Ads Have Evolved [+What This Means for Marketers]. *HubSpot*, [*S.l.*], 11 jun. 2021. Disponível em: https://blog.hubspot.com/marketing/history-facebook-adtips-slideshare. Acesso em: 1 ago. 2024.

FULGÊNCIO, Caio. OOH programática: novas possibilidades na publicidade digital. *Meio e Mensagem*, [*S.l.*], 25 maio 2023. Disponível em: https://www.meioemensagem.com.br/midia/ooh-programatica-novas-possibilidades-para-a-publicidade-digital. Acesso em: 31 jul. 2024.

GILLESPIE, Tarleton. *Custodians of the Internet*: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media. [*S.l.*]: Yale University Press, 2018.

GLOBAL ACTION PLAN. Kids for Sale: Online Advertising and the manipulation of children. Londres: *Global Action Plan*, 2020. Disponível em: https://www.globalactionplan.org.uk/files/kids_for_sale.pdf. Acesso em: 1 ago. 2024.

GLOBAL AD. Kwai em 2024: Um panorama detalhado da plataforma. *Global AD*, [*S.l.*], 22 fev. 2024. Disponível em: https://globalad.com.br/blog/explorando-o-kwai-em-2024-um-panorama-detalhado/. Acesso em: 3 abr. 2024.

GONG. Ads are not labeled as political, and some don't even appear. *Gong*, [*S.l.*], 23 maio 2019. Disponível em: https://gong.hr/en/2019/05/23/ads-are-not-labeled-as-political-and-some-dont-eve/. Acesso em: 1 ago. 2024.

GONZALEZ, Cristiana de O. O modelo de negócio da Google: Entre a eficiência técnico-científica e o imperativo econômico do retorno do investimento extrafiscalidade como instrumento de proteção ambiental no Brasil. *In*: CONGRESSO DE DIREITO DE AUTOR E INTERESSE PÚBLICO, 2012, Florianópolis. *Anais* [...]. Florianópolis: Editora Boiteux, 2012. Disponível em: https://gedai.ufpr.br/wp-content/uploads/2014/07/anais-v-codaip-versao-final-1.pdf. Acesso em: 1 ago. 2024.

GHOSH, Dipayan; SCOTT, Ben. Facebook's New Controversy Shows How Easily Online Political Ads Can Manipulate You. *Time*, [*S.l.*], 2018. Disponível em: https://time.com/5197255/facebook-cambridge-analytica-donald-trump-ads-data/. Acesso em: 20 fev. 2025.

GOEL, Shubhangi. Controversial messaging app Telegram is profitable, says its founder. Here's how it makes money. *Business Insider*, São Francisco, 24 dez. 2024. Disponível em: https://www.businessinsider.com/messaging-app-telegram-profitable-says-durov-2024-12. Acesso em: 1 abr. 2025.

HELBERGER, Natali; HUH, Jisu; MILNE, George; STRYCHARZ, Joanna; SUNDARAM, Hari. Macro and Exogenous Factors in Computational Advertising: Key Issues and New Research Directions. *Journal of Advertising*, [*S.l.*], v. 49, n. 4, p. 377–393, 2020. Disponível em: https://doi.org/10.1080/00913367.2020.1811179. Acesso em: 31 out. 2024.

HELBERGER, Natali; SAMUELSON, Pamela. The Digital Services Act as a Global Transparency Regime. *Verfassungsblog*, [*S.l.*], 2024. Disponível em: https://verfassungsblog.de/the-digital-services-act-as-a-global--transparency-regime/. Acesso em: 20 fev. 2025.

HERMANN, John. Why Nothing on Your Phone Is Safe From Ads. *Intelligencer*, [*S.l.*], 21 ago. 2023. Disponível em: https://nymag.com/intelligencer/2023/08/why-every-tech-company-turns-into-an-ad-company.html. Acesso em: 1 ago. 2024.

HOFFMAN, Jane S. *Your Data, Their Billions*: Unraveling and Simplifying Big Tech. Nova Iorque: Post Hill Press, 2022. Disponível em: https://posthillpress.com/book/your-data-their-billions-unraveling-and--simplifying-big-tech. Acesso em: 1 ago. 2024.

HSU, Tiffany. Twitter's Advertisers Pull Back as Layoffs Sweep Through Company. *The New York Times*, [*S.l.*], 4 nov. 2022. Technology. Disponível em: https://www.nytimes.com/2022/11/04/technology/twitter-advertisers.html. Acesso em: 18 fev. 2025.

ICC. International Chamber of Commerce. ICC Advertising and Marketing Communications Code. *International Chamber of Commerce*, Paris, 2018. Disponível em: https://iccwbo.org/wp-content/uploads/sites/3/2018/09/icc-advertising-and-marketing-communications-code--int.pdf. Acesso em: 1 ago. 2024.

IORY, Nicolas. Rede social X segue o Google e proíbe anúncios políticos na plataforma no Brasil. *O Globo*, São Paulo, 3 maio 2024. Disponível em: https://oglobo.globo.com/politica/noticia/2024/05/03/rede-social--x-segue-o-google-e-proibe-anuncios-politicos-na-plataforma-no-brasil.ghtml. Acesso em: 1 ago. 2024.

IQBAL, Mansoor. Twitter Revenue and Usage Statistics (2025). *Business of Apps*, [*S.l.*], 26 fev. 2025. Disponível em: https://www.businessofapps.com/data/twitter-statistics/. Acesso em: 1 abr. 2025.

JAMISON, Amelia M.; BRONIATOWSKI, David A.; DREDZE, Mark; WOOD-DOUGHTY, Zach; KHAN, DureAden; QUINN, Sandra C. Vaccine-related advertising in the Facebook Ad Archive. *Vaccine*, [*S.l.*], v. 38, n. 3, p. 512–520, 16 jan. 2020. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S0264410X1931446X. Acesso em: 1 ago. 2024.

KEMP, Simon. Digital 2024: Brazil. *Data Reportal*, 23 fev. 2024. Disponível em: https://datareportal.com/reports/digital-2024-brazil. Acesso em: 3 abr. 2024.

KHAN, Lina M.; BEDOYA, Alvaro; SLAUGHTER. Rebecca K. FTC Request For Investigation. Destinatário: *Federal Trade Commission*, 23. ago. 2023. Disponível em: https://fairplayforkids.org/wp-content/uploads/2023/08/FTCRequestForInvestigationAug23.pdf. Acesso em: 31 out. 2024.

KIM, Hwa Y. What's wrong with relying on targeted advertising? Targeting the business model of social media platforms. *Critical Review of International Social and Political Philosophy*, [*S.l.*], v. 0, n. 0, p. 1–21, 29 jan. 2024. Disponível em: https://doi.org/10.1080/13698230.2024.2309047. Acesso em: 1 ago. 2024.

KLEIN, Elana. The Latest Online Culture War Is Humans vs. Algorithms. *Wired*, [*S.l.*], 29 abr. 2024. Disponível em: https://www.wired.com/story/latest-online-culture-war-is-humans-vs-algorithms/. Acesso em: 1 ago. 2024.

KREISS, Daniel; MCGREGOR, Shannon C. The "Arbiters of What Our Voters See": Facebook and Google's Struggle with Policy, Process, and Enforcement around Political Advertising. *Political Communication*, [*S.l.*], v. 36, n. 4, p. 499–522, 2019. Disponível em: https://doi.org/10.1080/10584609.2019.1619639. Acesso em: 20 fev. 2025.

KUAISHOU. Kuaishou Technology Announces Fourth Quarter and Full Year 2024 Financial Results. *Kuaishou*, 25 mar. 2025. Disponível em: https://ir.kuaishou.com/news-releases/news-release-details/kuaishou-technology-announces-fourth-quarter-and-full-year-2024. Acesso em: 1 abr. 2025.

KRUIKEMEIER, Sanne; VERMEER, Susan; METOUI, Nadia; DOBBER, Tom; ZAROUALI, Brahim. (Tar)getting you: The use of online political targeted messages on Facebook. *Big Data & Society*, [*S.l.*], v. 9, n. 2, p. 1-20, jul. 2022. Disponível em: https://doi.org/10.1177/20539517221089626. Acesso em: 1 nov. 2024.

KWAI. Biblioteca de anúncios políticos ou eleitorais. *Kwai*, [*S.d.*]. Disponível em: https://www.kwai.com/business/pt-BR/adstransparency. Acesso em: 1 ago. 2024.

LABORATÓRIO NACIONAL DE COMPUTAÇÃO CIENTÍFICA. Qual a diferença entre dados pessoais e dados sensíveis?. *LNCC*, [*S.l.*], 2024. Disponível em: https://www.gov.br/lncc/pt-br/centrais-de-conteudo/campanhas-de-conscientizacao/campanha-lgpd/2024/qual-a-diferenca-entre-dados-pessoais-e-dados-sensiveis. Acesso em: 18 fev. 2025.

LEERSSEN, Paddy; AUSLOOS, Jef; ZAROUALI, Brahim; HELBERGER, Natali; DE VREESE, Claes H. Platform ad archives: promises and pitfalls. *Internet Policy Review*, [*S.l.*], v. 8, n. 4, out. 2019. Disponível em: https://doi.org/10.14763/2019.4.1421. Acesso em: 9 abr. 2024.

LEERSSEN, Paddy; DOBBER, Tom; HELBERGER, Natali; DE VREESE, Claes. News from the ad archive: how journalists use the Facebook Ad Library to hold online advertising accountable. *Information, Communication & Society*, [*S.l.*], v. 26, n. 7, p. 1381–1400, 26 dez. 2021. Disponível em: https://doi.org/10.1080/1369118X.2021.2009002. Acesso em: 1 ago. 2024.

LINKEDIN. Our Community Report. *LinkedIn*, [*S.d.*]a. Disponível em: https://about.linkedin.com/transparency/community-report. Acesso em: 1 ago. 2024.

LINKEDIN. Parâmetros e impressões de segmentação da Biblioteca de anúncios. *LinkedIn*, [*S.d.*]b. Disponível em: https://www.linkedin.com/help/linkedin/answer/a1620070. Acesso em: 1 ago. 2024.

LOSHIN, David. *Master Data Management*. Burlington: Morgan Kaufmann, 2010.

LUNDEN, Ingrid. U.S. LinkedIn passes $2B in premium revenue in 12 months, with overall revenue up 9% on the year. *TechCrunch*, [*S.l.*], 29 jan. 2025. Disponível em: https://techcrunch.com/2025/01/29/linkedin-passes-2b-in-premium-revenues-in-12-months-with-overall-revenues-up-9-on-the-year/. Acesso em: 1 abr. 2025.

MACNAMARA, Jim. Remodelling Media: The Urgent Search for New Media Business Models. *Media International Australia*, [*S.l.*], v. 137, n. 1, p. 20–35, 1 nov. 2010. Disponível em: https://doi.org/10.1177/1329878X1013700104. Acesso em: 1 ago. 2024.

MAHANTI, Rupa. *Data Quality*: Dimensions, Measurement, Strategy, Management, and Governance. Milwaukee: ASQ Quality Press, 2018.

MAROTTA, Veronica; ABHISHEK, Vibhanshu; ACQUISTI, Alessandro. Online Tracking and Publishers' Revenues: An Empirical Analysis. 2019. [Rascunho] Disponível em: https://www.semanticscholar.org/paper/Online-Tracking-and-Publishers%E2%80%99-Revenues%3A-An--Marotta/bee63f4551c7b6a5a1f07357734a81eab2fec919. Acesso em: 31 jul. 2024.

MCGILVRAY, Danette. *Executing Data Quality Projects*: Ten Steps to Quality Data and Trusted Information (TM). Londres: Academic Press, 2021.

MEDEIROS, Priscila; SALLES, Débora; MAGALHÃES, Thamyres; MELO, Bianca; SANTINI, Rose M. Greenwashing e Desinformação: A Publicidade Tóxica do Agronegócio Brasileiro nas Redes. *Comunicação e Sociedade*, [*S.l.*], v. 45, e024008, p. 1-26, 9 maio 2024. Disponível em: https://doi.org/10.17231/comsoc.45(2024).5417. Acesso em: 1 ago. 2024.

MEDERT, Florian; OTTO, Jan F.; PERCZE; Léna. Enhancing Transparency of Political Micro-targeting on Facebook. Em: BAYER; Judit, GRIMME; Christian. (Orgs.). *Code and Conscience*: Exploring Technology, Human Rights, and Ethics in Multidisciplinary AI Education. Suíça: Springer, 2024. Disponível em: https://link.springer.com/chapter/10.1007/978-3-031-52082-2_4. Acesso em: 31 out. 2024.

MELLO, Patrícia C. TikTok ignorou regra e veiculou anúncios para Lula e Bolsonaro em Portugal. *Folha de S.Paulo*, São Paulo, 14 ago. 2023. Disponível em: https://www1.folha.uol.com.br/poder/2023/08/tiktok-ignorou-regra-e-veiculou-anuncios-para-lula-e-bolsonaro-em-portugal.shtml. Acesso em: 1 ago. 2024.

META. Meta Reports Fourth Quarter and Full Year 2024 Results. *Meta*, 1 fev. 2025. Disponível em: https://investor.atmeta.com/investor-news/press-release-details/2025/Meta-Reports-Fourth-Quarter-and-Full-Year-2024-Results/default.aspx. Acesso em: 1 fev. 2025.

META. Biblioteca de Anúncios. *Meta*, [*S.d.*]a. Disponível em: https://www.facebook.com/ads/library. Acesso em: 1 ago. 2024.

META. O que são os públicos semelhantes do Facebook para anúncios. *Meta*, [*S.d.*]b. Disponível em: https://www.facebook.com/business/help/164749007013531. Acesso em: 1 ago. 2024.

META. Sobre temas sociais. *Meta*, [*S.d.*]c. Disponível em: https://pt-br.facebook.com/business/help/214754279118974. Acesso em: 1 ago. 2024.

META. Sobre os leilões de anúncios. *Meta*, [*S.d.*]d. Disponível em: https://www.facebook.com/business/help/430291176997542. Acesso em: 31 out. 2024.

MICHENER, Greg; BERSCH, Katherine. Identifying Transparency. *Information Polity*, Amsterdã, v. 18, n. 3, p. 233-242, 26 jul. 2013. Disponível em: https://dl.acm.org/doi/abs/10.5555/2659342.2659346. Acesso em: 1 ago. 2024.

MILANO, Silvia; MITTELSTADT, Brent; WACHTER, Sandra. OII | Targeted ads isolate and divide us even when they're not political. *Oxford*

*Internet Institute*, [*S.l.*], 2021. Disponível em: https://www.oii.ox.ac.uk/news-events/targeted-ads-isolate-and-divide-us-even-when-theyre-not-political. Acesso em: 20 fev. 2025.

MIRAGO. Dark Post no Facebook e Instagram: O que é e como usar?. *Mirago*, [*S.l.*], 18 set. 2024. Disponível em: https://www.mirago.com.br/dark-post-facebook/#h-o-que-um-dark-post. Acesso em: 1 out. 2024.

MOZILLA FOUNDATION. Mandating Tools to Scrutinize Social Media Companies. *Mozilla Foundation*, [*S.l.*], [*S.d.*]. Disponível em: https://foundation.mozilla.org/en/campaigns/mandating-tools-to-scrutinize-social-media-companies/. Acesso em: 18 fev. 2025.

MOZILLA FOUNDATION; CHECK FIRST. Full Disclosure: Stress testing tech platforms' ad repositories. *Mozilla Foundation e Check First*, 2024. Disponível em: https://assets.mofoprod.net/network/documents/Full_Disclosure_Stress_Testing_Tech_Platforms_Ad_Repositories_3FepU2u.pdf. Acesso em: 23 set. 2024.

NAPOLI, Philip. *Audience Evolution*: New Technologies and the Transformation of Media Audiences. New York: Columbia University Press, 2010.

NAPOLI, Philip; CAPLAN, Robyn. Why media companies insist they're not media companies, why they're wrong, and why it matters. *First Monday*, [*S.l.*], v. 22, n. 5, 2017. Disponível em: https://firstmonday.org/ojs/index.php/fm/article/view/7051. Acesso em: 1 ago. 2024.

NEKIPELOV, Denis; WANG, Tammy. Inference and auction design in online advertising. *Communications of the ACM*, [*S.l.*], v. 60, n. 7, p. 70-79, 2017. Disponível em: https://dl.acm.org/doi/10.1145/3035966. Acesso em: 31 out. 2024.

NETLAB UFRJ. Irregularidades da propaganda política online durante o 2o turno das Eleições 2022. *NetLab UFRJ*, 29 out. 2022a. Disponível em: https://netlab.eco.ufrj.br/post/voltar-ao-site-irregularidades-da-propaganda-pol%C3%ADtica-online-durante-o-2%C2%BA-turno-das-elei%C3%A7%C3%B5es-2022. Acesso em: 4 abr.. 2025.

NETLAB UFRJ. Irregularidades e opacidade nos anúncios do Google durante as Eleições de 2022. *NetLab UFRJ*, 15 set. 2022b. Disponível em: https://netlab.eco.ufrj.br/post/irregularidades-e-opacidade-nos-anúncios-do-google-durante-as-eleições-de-2022. Acesso em: 1 ago. 2024.

NETLAB UFRJ. A guerra das plataformas contra o PL 2630. *NetLab UFRJ*, 1 maio 2023a. Disponível em: https://netlab.eco.ufrj.br/post/a-guerra-das-plataformas-contra-o-pl-2630. Acesso em: 2 ago. 2024.

NETLAB UFRJ. Anúncios golpistas na biblioteca do Meta Ads: novembro de 2022 a janeiro de 2023. *NetLab UFRJ*, 7 fev. 2023b. Disponível em: https://netlab.eco.ufrj.br/post/anúncios-golpistas-na-biblioteca-do-meta-ads-novembro-de-2022-a-janeiro-de-2023. Acesso em: 1 ago. 2024.

NETLAB UFRJ. Golpes, fraudes e desinformação na publicidade digital desregulada. *NetLab UFRJ*, 20 out. 2023c. Disponível em: https://netlab.eco.ufrj.br/post/golpes-fraudes-e-desinformação-na-publicidade-digital-desregulada. Acesso em: 1 ago. 2024.

NETLAB UFRJ. Anúncios com IA usam imagem de políticos brasileiros para aplicar golpes. *NetLab UFRJ*, 17 jun. 2024a. Disponível em: https://netlab.eco.ufrj.br/post/anúncios-com-ia-usam-imagem-de-políticos-brasileiros-para-aplicar-golpes. Acesso em: 1 ago. 2024.

NETLAB UFRJ. Golpes e falhas sistêmicas: anúncios fraudulentos sobre o Desenrola Brasil e o Voa Brasil seguem circulando após um ano de medida cautelar. *NetLab UFRJ*, 9 out. 2024b. Disponível em: https://netlab.eco.ufrj.br/post/golpes-e-falhas-sist%C3%AAmicas-an%C3%BAncios-fraudulentos-sobre-o-desenrola-brasil-e-o-voa-brasil-seguem-cir. Acesso em: 20 mar. 2025.

NÓBREGA, Liz. Kwai veta anúncios políticos para as eleições de 2024. *desinformante*, [*S.l.*], 16 maio 2024. Disponível em: https://desinformante.com.br/kwai-anuncios-politicos/. Acesso em: 1 ago. 2024.

OCDE. Organização para a Cooperação e Desenvolvimento Econômico. *Facts not Fakes*: Tackling Disinformation, Strengthening Information Integrity. Paris: OECD Publishing, 2024. Disponível em: https://doi.org/10.1787/d909ff7a-en. Acesso em: 1 ago. 2024.

OMS. Organização Mundial da Saúde. Scope and impact of digital marketing strategies for promoting breast-milk substitutes. Geneva: *World Health Organization*, 2022. Disponível em: https://www.who.int/publications/i/item/9789240046085. Acesso em: 31 out. 2024.

O'NEIL, Cathy. *Algoritmos de destruição em massa*. Santo André: Editora Rua do Sabão, 2020.

PAPAKYRIAKOPOULOS, Orestis; HEGELICH, Simon; SHAHREZAYE, Morteza; SERRANO, Juan C. M. Social media and microtargeting: Political data processing and the consequences for Germany. *Big Data & Society*, [*S.l.*], v. 5, n. 2, p. 1-15, 20 nov. 2018. Disponível em: https://doi.org/10.1177/2053951718811844. Acesso em: 1 ago. 2024.

PAUL, Kari. Twitter allows US political candidates and parties to advertise in policy switch. *The Guardian*, [*S.l.*], 30 ago. 2023. Disponível em: https://www.theguardian.com/technology/2023/aug/29/twitter-x-political-ads-us-policy-misinformation. Acesso em: 1 ago. 2024.

PERSHAN, Claire; LESPLINGART, Amaury. Full Disclosure: Stress testing tech platforms' ad repositories. *Mozilla Foundation*, 16 abr. 2024. Disponível em: https://foundation.mozilla.org/en/research/library/full-disclosure-stress-testing-tech-platforms-ad-repositories/. Acesso em: 18 fev. 2025.

PETERSON, Tim. Facebook organic reach is down 52% for publishers' Pages this year. *MarTech*, [*S.l.*], 6 ago. 2016. Disponível em: https://martech.org/facebook-organic-reach-drop-steepens-52-publishers-pages/. Acesso em: 31 jul. 2024.

PINTEREST. Plataforma para Desenvolvedores: Introducing the Pinterest API (v5). *Pinterest*, 12 abr. 2022. Disponível em: https://developers.pinterest.com/blog/2022/04/12/introducing-pinterest-api-v5/. Acesso em: 23 set. 2024.

PINTEREST. Pinterest Announces Fourth Quarter and Full Year 2024 Results, Delivers First Billion Dollar Revenue Quarter. *Pinterest*, 6 fev. 2025. Disponível em: https://investor.pinterestinc.com/news-and-events/press-releases/press-releases-details/2025/Pinterest-Announces-

Fourth-Quarter-and-Full-Year-2024-Results-Delivers-First-Billion-Dollar-Revenue-Quarter/default.aspx. Acesso em: 1 abr. 2025.

PINTEREST. Repositório de Anúncios. *Pinterest*, [*S.d.*]a. Disponível em: https://ads.pinterest.com/ads-repository/. Acesso em: 23 set. 2024.

PINTEREST. Plataforma para Desenvolvedores. API V4: ads_repository. *Pinterest*, [*S.d.*]b. Disponível em: https://developers.pinterest.com/docs/api/v4/#tag/ads_repository. Acesso em: 23 set. 2024.

PINTEREST. Plataforma para Desenvolvedores: Pinterest REST API 5.14.0. *Pinterest*, [*S.d.*]c. Disponível em: https://developers.pinterest.com/docs/api/v5/introduction/. Acesso em: 23 set. 2024.

PINTEREST. Plataforma para Desenvolvedores. *Pinterest*, [*S.d.*]d. Disponível em: https://developers.pinterest.com/. Acesso em: 23 set. 2024.

PODER 360. Google lança ferramenta de transparência política no Brasil. *Poder 360,* [*S.l.*], 24 jun. 2022. Disponível em: https://www.poder360.com.br/poder-eleicoes/eleicoes/google-lanca-ferramenta-de-transparencia-politica-no-brasil/. Acesso em: 1 ago. 2024.

POPIEL, Pawel. The Tech Lobby: Tracing the Contours of New Media Elite Lobbying Power. *Communication, Culture and Critique*, [*S.l.*], v. 11, n. 4, p. 566–585, 1 dez. 2018. Disponível em: https://doi.org/10.1093/ccc/tcy027. Acesso em: 1 ago. 2024.

POCHAT, Victor; EDELSON, Laura; GOETHEM, Tom V.; JOOSEN, Wouter; MCCOY, Damon; LAUINGER, Tobias. An audit of Facebook's political ad policy enforcement. *In*: USENIX SECURITY SYMPOSIUM, 31., ago. 2022, Boston. *Anais* [...]. [*S.l.*]: USENIX Association, 2022. Disponível em: https://www.usenix.org/system/files/sec22-lepochat.pdf. Acesso em: 1 ago. 2024.

REIJMERSDAL, Eva A. van; ROZENDAAL, Esther. Transparency of digital native and embedded advertising: Opportunities and challenges for regulation and education. *Communications*, [*S.l.*], v. 45, n. 3, p. 378–388, 1 set. 2020. Disponível em: https://doi.org/10.1515/commun-2019-0120. Acesso em: 1 ago. 2024.

RIBEIRO, Filipe N.; SAHA, Koustuv.; BABAEI, Mahmoudreza; HENRIQUE, Lucas; MESSIAS, Johnnatan; BENEVENUTO, Fabricio; GOGA, Oana; GUMMADI, Krishna P.; REDMILES, Elissa M. On Microtargeting Socially Divisive Ads: A Case Study of Russia-Linked Ad Campaigns on Facebook. *In*: CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2019, p. 140–149, Atlanta. *Anais* [...]. Nova Iorque: Association for Computing Machinery, 2019. Disponível em: https://doi.org/10.1145/3287560.3287580. Acesso em: 1 ago. 2024.

RICHARDSON, Laurie; O'CONNOR, Jennifer F. Complying with the Digital Services Act. *Google*, 24 ago. 2023. Disponível em: https://blog.google/around-the-globe/google-europe/complying-with-the-digital-services-act/. Acesso em: 1 ago. 2024.

ROSENBERG, Matthew. Ad Tool Facebook Built to Fight Disinformation Doesn't Work as Advertised. *The New York Times*, [*S.l.*], 25 jul. 2019. Disponível em: https://www.nytimes.com/2019/07/25/technology/facebook-ad-library.html. Acesso em: 1 ago. 2024.

SAMBHAV, Kumar; RANGANATHAN, Nayantara. Facebook charged BJP less for India election ads than others. *Al Jazeera*, Nova Delhi, 16 mar. 2022. Disponível em: https://www.aljazeera.com/economy/2022/3/16/facebook-charged-bjp-lower-rates-for-india-polls-ads-than-others. Acesso em: 1 ago. 2024.

SAMSING, Caro. O declínio do alcance orgânico do Facebook e como superar o algoritmo. *Hub Spot*, [*S.l.*], 25 jan. 2018. Disponível em: https://br.hubspot.com/blog/marketing/declinio-alcance-organico-facebook. Acesso em: 31 jul. 2024.

SANTINI, Rose M.; SALLES, Débora; BARROS, Carlos, E.; MARTINS, Bruno M. M.; HADDAD, João, G.; SEADE, Renata; GOMES, Matheus; SOUZA, Lucas. Publicidade online sem lei? Tipos de fraudes e golpes em anúncios digitais. NetLab UFRJ, 7 jun. 2023. Disponível em: https://netlab.eco.ufrj.br/post/publicidade-online-sem-lei-tipos-de-fraudes-e-golpes-em-anúncios-digitais. Acesso em: 1 ago. 2024.

SANTINI, Rose M.; SALLES, Débora; MATTOS, Bruno; BELIN, Luciane L.; CANAVARRO, Marcela; MEDEIROS, Stéphanie; HADDAD, João G.; SILVA, Daphne; SEADE, Renata; DIAS, Bernardo; GOMES, Matheus; YONESHIGE, Bernardo; DAU, Erick; DO CARMO, Victor; LOUREIRO, Felipe. Golpes, Fraudes e Desinformação na Publicidade Digital Abusiva Contra Mulheres. *NetLab UFRJ*, 8 mar. 2024a. Disponível em: https://netlab.eco.ufrj.br/post/golpes-fraudes-e-desinformac-a-o-na-publicidade-digital-abusiva-contra-mulheres. Acesso em: 1 ago. 2024.

SANTINI, Rose M.; SALLES, Débora; MARTINS, Bruno M.; MOREIRA, Alékis; HADDAD, João G. Seeing through opacity: The limitations of digital ad transparency in Brazil. *In*: ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY (FAccT), 7., jun. 2024, Rio de Janeiro. *Anais* [...]. Nova Iorque: Association for Computing Machinery (ACM), 2024b. Disponível em: https://dl.acm.org/doi/10.1145/3630106.3659034. Acesso em: 1 ago. 2024.

SANTINI, Rose. M.; FITZGERALD, James; FILHO, Humberto R.; LOKMANOGLU, Ayse D.; SOMBATPOONSIRI, Janjira; SALLES, Débora. The economy and social impact of platform transparency: As regulation for the construction of inclusive digital markets. T20 Policy Brief: Inclusive Digital Transformation. *G20 Brasil*, 2024c. Disponível em: https://t20brasil.org/media/documentos/arquivos/TF05_ST_06_The_Economic_and_So66cf6bee8d2ed.pdf. Acesso em: 30 ago. 2024.

SANTINI, Rose. M.; SALLES, Débora; MATTOS, Bruno; CANAVARRO, Marcela; HADDAD, João G.; SILVA, Daphne; LOUREIRO, Felipe. Nota técnica: Google Diminui Transparência de Anúncios Políticos no Brasil e Desobedece Resolução do TSE. *NetLab UFRJ*, 12 jul. 2024d. Disponível em: https://netlab.eco.ufrj.br/post/nota-tecnica-google. Acesso em: 1 ago. 2024.

SANTINI, Rose. M.; SALLES, Débora; MATTOS, Bruno; MOREIRA, Alékis; MELLO, Danielle; HADDAD, João G.; DIAS, Bernardo; GOMES, Matheus; BORGES, Amanda; LOUREIRO, Felipe. Danos causados pela publicidade enganosa na Meta: Anúncios fraudulentos promo-

vem desinformação sobre o Pix para lesar cidadãos brasileiros. *NetLab UFRJ*, 4 fev. 2025. Disponível em: https://netlab.eco.ufrj.br/post/danos-causados-pela-publicidade-enganosa-na-meta. Acesso em: 20 mar. 2025.

SCHNAIDER, Amanda. A evolução da "publi": Marcas cocriando com os influenciadores. *Meio e Mensagem*, [*S.l.*], 1 set. 2022. Disponível em: https://www.meioemensagem.com.br/midia/a-evolucao-da-publi-marcas-cocriando-com-os-influenciadores. Acesso em: 31 jul. 2024.

SERPRO. Dados sensíveis — LGPD - Lei Geral de Proteção de Dados Pessoais. *Serpro*, [*S.l.*], [*S.d.*]. Disponível em: https://www.serpro.gov.br/lgpd/menu/protecao-de-dados/dados-sensiveis-lgpd. Acesso em: 18 fev. 2025.

SILVEIRA, Stefanie; MORISSO, João G. O uso de algoritmos na mídia programática. *Parágrafo*, São Paulo, v. 6, n. 1, p. 71–82, 29 jun. 2018. Disponível em: https://revistaseletronicas.fiamfaam.br/index.php/recicofi/article/view/710. Acesso em: 1 ago. 2024.

SILVERMAN, Craig; BENGANI, Priyanjana. Exploiting Meta's Weaknesses, Deceptive Political Ads Thrived on Facebook and Instagram in Run-Up to Election. *ProPublica*, [*S.l.*], 31 out. 2024. Disponível em: https://www.propublica.org/article/facebook-instagram-meta-deceptive-political-ads-election. Acesso em: 31 out. 2024.

SINGH, Shubham. Telegram Users Statistics (2025) – New Global Data. *DemandSage*, [*S.l.*], 18 jan. 2024. Disponível em: https://www.demandsage.com/telegram-statistics/. Acesso em: 3 abr. 2024.

SOSNOVIK, Vera; GOGA, Oana. Understanding the Complexity of Detecting Political Ads. *In*: WEB CONFERENCE 2021 (WWW '21), 2021, p. 2002-2013, Ljubljana, Slovenia. *Anais* [...]. Nova Iorque: Association for Computing Machinery (ACM), 2021. Disponível em: https://hal.science/hal-03450501. Acesso em: 1 ago. 2024.

STATISTA. Digital advertising worldwide - statistics & facts. *Statista*, 2024a. Disponível em: https://www.statista.com/topics/7666/internet-advertising-worldwide/. Acesso em: 1 ago. 2024.

STATISTA. LinkedIn revenue generated in the past 12 months as of June 2024, by business segment. *Statista*, 2024b. Disponível em: https://www.statista.com/statistics/1472741/linkedin-revenue-by-business-segment/. Acesso em: 1 abr. 2025.

STATISTA. Advertising revenue of Google from 2001 to 2024. *Statista*, 2025a. Disponível em: https://www.statista.com/statistics/266249/advertising-revenue-of-google/. Acesso em: 1 abr. 2025.

STATISTA. Revenue of Google from 1st quarter 2008 to 4th quarter 2024. *Statista*, 2025b. Disponível em: https://www.statista.com/statistics/267606/quarterly-revenue-of-google/. Acesso em: 1 abr. 2025.

TELEGRAM. Ad Policies and Guidelines. *Telegram*, [*S.d.*]a. Disponível em: https://ads.telegram.org/guidelines. Acesso em: 1 ago. 2024.

TELEGRAM. Telegram Ad Platform. *Telegram*, [*S.d.*]b. Disponível em: https://ads.telegram.org/. Acesso em: 1 ago. 2024.

TELEGRAM. Telegram Ad Platform Explained. *Telegram*, [*S.d.*]c. Disponível em: https://ads.telegram.org/getting-started. Acesso em: 1 ago. 2024.

TIKTOK. Expanding TikTok's Research API and Commercial Content Library. *TikTok*, 20 jul. 2023. Disponível em: https://newsroom.tiktok.com/en-eu/expanding-tiktoks-research-api-and-commercial-content-library. Acesso em: 1 ago. 2024.

TIKTOK. Biblioteca de anúncios. *TikTok*, [*S.d.*]a. Disponível em: https://library.tiktok.com/ads. Acesso em: 1 ago. 2024.

TIKTOK. Commercial Content API. *TikTok*, [*S.d.*]b. Disponível em: https://developers.tiktok.com/products/commercial-content-api. Acesso em: 1 ago. 2024.

TIKTOK. Acerca da Biblioteca de conteúdos comerciais do TikTok. *TikTok*, [*S.d.*]c. Disponível em: https://library.tiktok.com/faq. Acesso em: 1 ago. 2024.

TRINDADE, Naira. Kwai veta anúncios políticos para eleições 2024. *O Globo*, [*S.d.*], 15 maio 2024. Disponível em: https://oglobo.globo.com/

blogs/lauro-jardim/post/2024/05/kwai-veta-anuncios-politicos-para-eleicoes-2024.ghtml. Acesso em: 1 ago. 2024.

TUFEKCI, Zeynep. Engineering the public: Big data, surveillance and computational politics. *First Monday*, [*S.l.*], v. 19, n. 7, 2 jul. 2014. Disponível em: https://firstmonday.org/ojs/index.php/fm/article/view/4901. Acesso em: 1 ago. 2024.

TUFEKCI, Zeynep. We're building a dystopia just to make people click on ads. Vídeo. 22min45s. Pulicado pelo *Ted Talks*. set. 2017. Disponível em: https://www.ted.com/talks/zeynep_tufekci_we_re_building_a_dystopia_just_to_make_people_click_on_ads?subtitle=en&geo=pt-br&trigger=15s&lng=pt-br. Acesso em: 31 jul. 2024.

TUROW, Joseph. *The Daily You*: How the New Advertising Industry Is Defining Your Identity and Your Worth. New Haven: Yale University Press, 2011. Disponível em: https://yalebooks.yale.edu/book/9780300188011/the-daily-you/. Acesso em: 01 nov. 2024.

TUTTLE, Hilary. Facebook Scandal Raises Data Privacy Concerns. *Risk Management*, [*S.l.*], v. 65, n. 5, p. 6–9, 2018. Disponível em: https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=00355593&v=2.1&it=r&id=GALE%7CA538250056&sid=googleScholar&linkaccess=abs&userGroupName=anon%7E4c85060e&aty=open-web-entry. Acesso em: 20 fev. 2025.

UNESCO. Organização das Nações Unidas para a Educação, Ciência e Cultura. *Guidelines for the governance of digital platforms*: Safeguarding freedom of expression and access to information through a multi-stakeholder approach. Paris: Unesco, 2023. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000387339. Acesso em: 1 ago. 2024.

UNIÃO EUROPEIA. Regulation (EU) 2022/2065 of the European Parliament and of the Council. Digital Services Act. *Official Journal of the European Union*, Bruxelas, 19 out. 2022. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32022R2065. Acesso em: 1 ago. 2024.

UR, Blase; LEON; Pedro G.; CRANOR, Lorrie F.; SHAY, Richard; WANG, Yang. Smart, useful, scary, creepy: Perceptions of online behavioral advertising. *In*: SYMPOSIUM ON USABLE PRIVACY AND SECURITY (SOUPS), 8., 2012, Washington, DC. *Anais* [...]. Association for Computing Machinery: Nova Iorque, 2012. Disponível em: https://doi.org/10.1145/2335356.2335362. Acesso em: 31 out. 2024.

VAN DIJCK, José; NIEBORG, David; POELL, Thomas. Reframing platform power. *Internet Policy Review*, [*S.l.*], v. 8, n. 2, p. 1–18, 2019. Disponível em: https://policyreview.info/articles/analysis/reframing-platform-power. Acesso em: 1 nov. 2024.

VAN LOOY, Amy. *Social Media Management*: Technologies and Strategies for Creating Business Value. Cham: Springer International Publishing, 2016. Disponível em: https://link.springer.com/10.1007/978-3-319-21990-5. Acesso em: 18 fev. 2025.

VARIAN, Hal R. Computer Mediated Transactions. *American Economic Review*, [*S.l.*], v. 100, n. 2, p. 1–10, maio 2010. Disponível em: https://doi.org/10.1257/aer.100.2.1. Acesso em: 1 ago. 2024.

WALKER, Kent. Supporting election integrity through greater advertising transparency. *Google*, 4 maio 2018. Disponível em: https://blog.google/outreach-initiatives/public-policy/supporting-election-integrity-through-greater-advertising-transparency/. Acesso em: 1 ago. 2024.

WALTENBERG, Guilherme. Google veta impulsionamento eleitoral em 2024. *Poder 360*, [*S.l.*], 23 abr. 2024. Disponível em: https://www.poder360.com.br/eleicoes/google-veta-impulsionamento-eleitoral-em-2024-e-pressiona-tse/. Acesso em: 1 ago. 2024.

WANG, Shan. When a Facebook test moves news stories to a separate feed, traffic — and public discourse — are at stake. *Nieman Lab*, [*S.l.*], 26 out. 2017. Disponível em: https://www.niemanlab.org/2017/10/when-a-facebook-test-moves-news-stories-to-a-separate-feed-traffic-and-public-discourse-are-at-stake/. Acesso em: 31 jul. 2024.

X/TWITTER. Ads repository. *X/Twitter,* [*S.d.*]a. Disponível em: https://ads.twitter.com/ads-repository. Acesso em: 1 ago. 2024.

X/TWITTER. Ads transparency. *X/Twitter*, [*S.d.*]b. Disponível em: https://business.x.com/en/help/ads-policies/product-policies/ads-transparency.html. Acesso em: 1 ago. 2024.

X/TWITTER. Political Ads Disclosure. *X/Twitter*, [*S.d.*]c. Disponível em: https://business.x.com/en/help/ads-policies/ads-content-policies/political-content/political-ads-disclosure.html. Acesso em: 1 ago. 2024.

X/TWITTER. Political Content. *X/Twitter*, [*S.d.*]d. Disponível em: https://business.x.com/en/help/ads-policies/ads-content-policies/political-content.html. Acesso em: 1 ago. 2024.

ZALNIERIUTE, Monika. "Transparency-Washing" in the Digital Age: A Corporate Agenda of Procedural Fetishism. *Critical Analysis of Law*, Sydney, v. 8 n. 1, p. 39-53, 2021. Disponível em: https://papers.ssrn.com/abstract=3805492. Acesso em: 1 ago. 2024.

ZUBOFF, Shoshana. Big other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology*, [*S.l.*], v. 30, n. 1, p. 75–89, 2015. Disponível em: https://doi.org/10.1057/jit.2015.5. Acesso em: 18 fev. 2025.

ZUBOFF, Shoshana. "We Make Them Dance": Surveillance Capitalism, the Rise of Instrumentarian Power, and the Threat to Human Rights. *In*: JØRGENSEN, Rikke Frank. *Human Rights in the Age of Platforms*. Cambridge: The MIT Press, 2019. Disponível em: https://direct.mit.edu/books/oa-edited-volume/4531/chapter/202528/We-Make-Them-Dance-Surveillance-Capitalism-the. Acesso em: 18 fev. 2025.

ZUBOFF, Shoshana. *A era do capitalismo de vigilância*. Rio de Janeiro: Editora Intrínseca, 2021.

**Figura 1:** Representação visual da pontuação passível de ser obtida pelas plataformas analisadas, considerando a aplicabilidade de todos os parâmetros de avaliação propostos

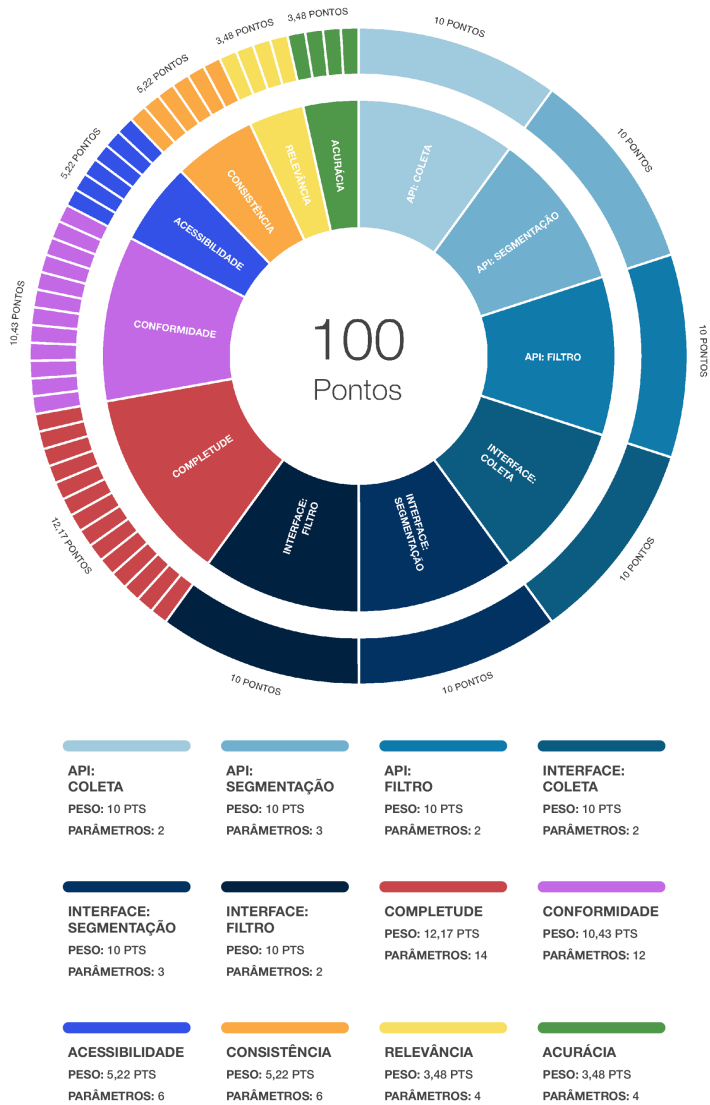


**Figura 2:** Visão geral das avaliações das plataformas, da maior à menor nota obtida

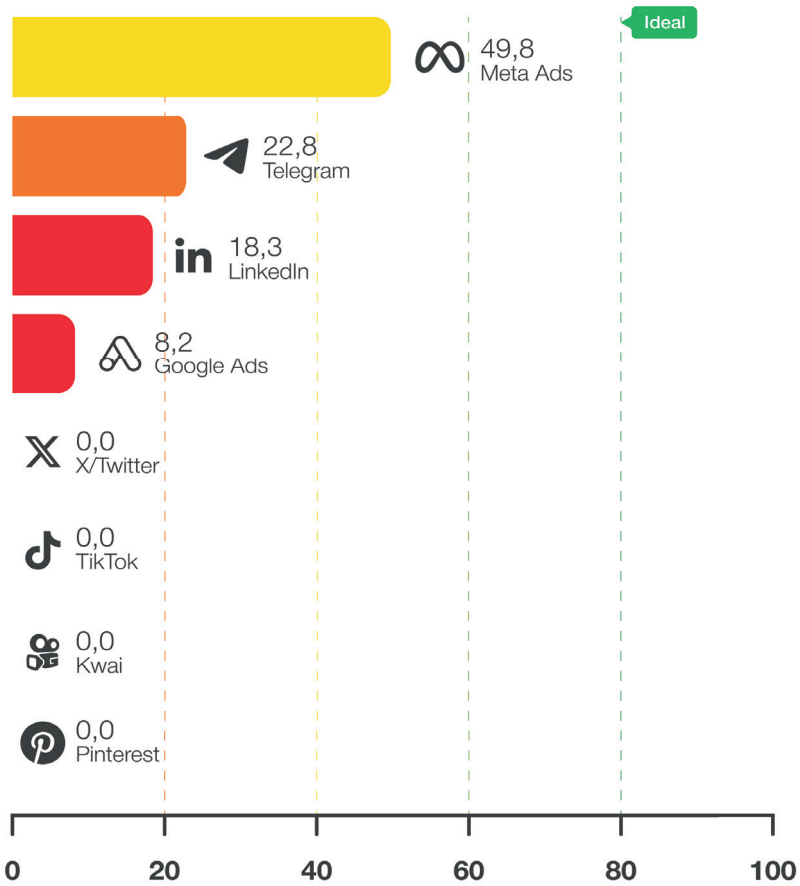


**Figura 3:** Representação visual da nota obtida pela Meta em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas


**Figura 4:** Representação visual da nota obtida pelo Telegram em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

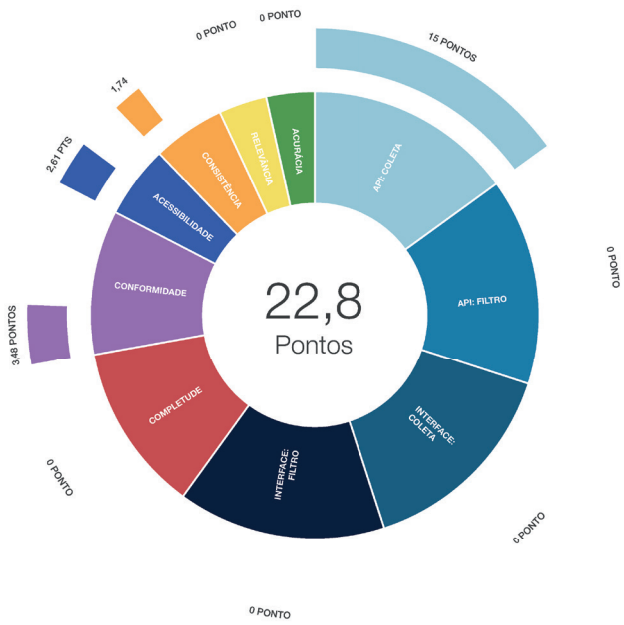


**Figura 5:** Representação visual da nota obtida pelo LinkedIn em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas


**Figura 6:** Representação visual da nota obtida pelo Google em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

**Figura 7:** Representação visual da nota obtida pelo X/Twitter em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas


**Figura 8:** Representação visual da nota obtida pelo TikTok em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

**Figura 9:** Representação visual da nota obtida pelo Kwai em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas


**Figura 10:** Representação visual da nota obtida pelo Pinterest em nossa avaliação, segmentada pelas diferentes dimensões de qualidade analisadas

# Capítulo 3

## Anúncios falsos, seus mecanismos e potenciais danos à sociedade

Os índices de transparência de dados e publicidade desenvolvidos pelo NetLab UFRJ (Santini et al., 2024a; 2024b), já apresentados nos capítulos anteriores, contribuem para a compreensão aprofundada das dinâmicas e características de transparência e acesso a dados das plataformas de redes sociais. Os estudos evidenciam a opacidade dessas redes também no que diz respeito aos anúncios e aos potenciais impactos sobre o mercado de publicidade digital.

Considerando os dados discutidos nos Índices, este capítulo visa apresentar evidências mapeadas em diferentes casos analisados pelo laboratório que ilustram como as plataformas são utilizadas por anunciantes maliciosos para promoção de fraudes em ambientes digitais. Neste capítulo apresentamos as principais conclusões extraídas de seis estudos de caso (NetLab UFRJ, 2024) que examinam e descrevem as características das fraudes e golpes nas redes sociais online que causam prejuízos aos consumidores, às empresas e às instituições no ecossistema de publicidade digital.

Para efeito da pesquisa, foram considerados anúncios fraudulentos os conteúdos pagos e impulsionados que se apropriam de técnicas de engenharia social para vender soluções e informações falsas, buscando obter intencionalmente renda ilegal e causar danos materiais e morais a outras pessoas (Button et al., 2014; Kikerpill; Siibak, 2021; Kotila et

al., 2016). Considerando a centralidade que as plataformas digitais têm assumido no mercado publicitário global, a compreensão desses mecanismos se tornou um passo essencial para que fraudes e golpes possam ser evitados e combatidos e os consumidores, protegidos.

## 3.1. Epidemia de informações falsas e fraudes nas plataformas online

Nos últimos anos, a publicidade passou a representar a principal fonte de receita para gigantes como a Meta e o Google, movimentando centenas de bilhões de reais. Em 2024, a receita da Google foi de 348,1 bilhões de dólares, dos quais 76% são provenientes de publicidade digital. Já a da Meta foi de 164,5 bilhões de dólares, com 98% fruto do mesmo tipo de anúncio (Statista, 2023; Meta, 2024). Só no Brasil, o investimento em publicidade em mídias digitais movimentou R$35 bilhões em 2023, um crescimento de 5% em relação a 2022 (Kantar Ibope Media, 2023).

Anúncios online devem seguir leis e normas de publicidade brasileiras - incluindo as especificidades setoriais previstas na Constituição Federal (Brasil, 1988), como, por exemplo, o artigo 220, §4º, as normas de publicidade do Código de Defesa do Consumidor (Brasil, 1990), como o Artigo 37, §§ 1º e 2º, e a fiscalização do Conselho Nacional de Autorregulamentação Publicitária (Conar). Contudo, a dificuldade de acesso às informações sobre estas negociações tornam o ecossistema de publicidade digital um espaço pouco regulado. Responsabilidades definidas no Código de Defesa do Consumidor para quem "faz ou promove publicidade" (arts. 67 e 68) não são reconhecidas pelas plataformas digitais (Silva, 2022). No entanto, elas são entes ativos na promoção dos anúncios.

Outro aspecto característico deste tipo de publicidade que contribui para seu uso malicioso é a disponibilidade de ferramentas de microssegmentação da audiência, que permitem o direcionamento preciso de anúncios a públicos específicos (Zuboff, 2021). Embora essa tecnologia seja um diferencial competitivo para as *big tech*, ela também contri-

bui com a disseminação de fraudes, como explicamos em mais detalhes no capítulo anterior.

Com isso, as plataformas de redes sociais se tornaram o canal preferido para que estelionatários de todos os tipos encontrem suas novas vítimas. Segundo a UK Finance (2023), principal associação de bancos do Reino Unido, mais de 75% das fraudes financeiras no país se originam em plataformas de redes sociais, formando uma indústria que movimenta mais de 1,2 bilhão de libras anuais. Só nos Estados Unidos, o lucro dos criminosos passou de 42 milhões, em 2017, para 1,2 bilhão de dólares em 2022, segundo a Federal Trade Commission (FTC, 2023), órgão do governo dos Estados Unidos responsável pela proteção dos direitos do consumidor. Um estudo elaborado pela World Federation of Advertisers (WFA, 2016), que assume fazer previsões conservadoras, trabalha com dois cenários para explicar a fraude publicitária: um mais otimista, no qual a fraude em anúncios representa cerca de 10% do mercado digital, e um mais pessimista, em que a fraude em anúncios representa mais de 30% deste mercado. Segundo a projeção mais pessimista, em 2024, a fraude em anúncios movimentou 140 bilhões de dólares (3 vezes mais do que o tráfico de drogas).

As tendências não são muito diferentes no Brasil: mapeamento da OLX em parceria com a AllowMe (UOL, 2023) estima que brasileiros teriam perdido R$551 milhões em 2022 com fraudes online. A pesquisa também indica que uma das modalidades preferenciais de estelionatários na internet é o impulsionamento de anúncios fraudulentos. Além disso, um relatório desenvolvido pela empresa Silverguard (2024), que notifica bancos sobre golpes financeiros a partir das denúncias das vítimas feitas através do serviço "SOS Golpe", mostrou que 79% das fraudes financeiras digitais envolvendo Pix tiveram início nas plataformas da Meta.

Com o objetivo de contribuir com a melhor compreensão dos mecanismos de aplicação de golpes, fraudes e estelionato no ecossistema de publicidade digital, a pesquisa apresentada neste capítulo consiste em um compilado dos resultados de seis estudos de caso (NetLab UFRJ,

2024). Todos eles analisam anúncios fraudulentos e evidenciam como essas práticas exploram lacunas regulatórias e tecnológicas para prejudicar consumidores, reforçando a necessidade de medidas mais robustas para combater fraudes online.

## 3.2. Metodologia

Metodologicamente, a pesquisa se baseou na análise de dados digitais não obstrutivos (Chen, 2017), ou seja, na coleta de dados e informações produzidos e disponibilizados em plataformas digitais de maneira pública e sem a interferência dos pesquisadores no objeto de pesquisa. A análise dos dados adotou métodos digitais (Rogers, 2013) com técnicas computacionais aliadas a análises qualitativas das informações dos anúncios.

Os casos consideraram diferentes intervalos de tempo e trataram de temas diversos, de modo a apresentar a amplitude do uso das plataformas da Meta para aplicação de fraudes financeiras. Foram analisados anúncios veiculados entre 2023 e 2024 no Facebook, Instagram, Messenger e Audience Network - sistema de veiculação de anúncios em outros apps parceiros da Meta, incluindo aplicativos de games, entretenimento e outros formatos.

A opção por analisar a Meta se justifica pois esta era, durante a execução dos estudos, a única plataforma que oferecia um repositório navegável e pesquisável de anúncios no Brasil, por meio de sua Biblioteca de Anúncios (Meta, *[S.d]a.*). Essa ferramenta inclui anúncios políticos, eleitorais e de relevância social, cujos dados eram armazenados por até sete anos. No entanto, conforme apresentado no Capítulo 3, sobre o Índice de Transparência de Publicidade, os critérios de categorização desses anúncios são auto-declarados pelos anunciantes, o que gera inconsistências. Foram considerados dois tipos de anúncios - comerciais e políticos. Como no Brasil não há regulamentação de plataformas digitais, as empresas Meta e Google decidiram disponibilizar repositórios de anúncios separando os anúncios políticos do restante. Porém, em 2024 o Google anunciou que não permitiria anúncios políticos durante as eleições municipais daquele ano (Agência Brasil, 2024).

Teoricamente o objetivo dessa separação é oferecer mais transparência para a publicidade política do que para os demais conteúdos impulsionados. Porém, a interpretação do que é considerado "político" varia significativamente em cada plataforma e em cada país ou região (Leerssen *et al.*, 2019). Como discutimos no capítulo anterior, a literatura propõe diversas possibilidades para distinguir o que são anúncios políticos do que são anúncios comerciais, demonstrando como a falta de consenso favorece ambiguidades na classificação dos conteúdos pagos (Dommett; Zhu, 2023). Essa separação é uma decisão tomada pelas plataformas, que, quando disponibilizam alguma transparência ou acesso a dados, o fazem para uma amostra arbitrária de anúncios com base em uma definição ambígua, estipulada por elas próprias. Sob esse pretexto, as plataformas se eximem de disponibilizar dados sobre todas as peças publicitárias veiculadas. Isto é, ao criarem a ideia de que anúncios "políticos" devem ter mais transparência que outros, acabam por dar um viés político ao direito do consumidor e assim ofuscam questões mais substantivas e fundamentais (Zalnieriute, 2021), que podem colocar os consumidores em risco e causar danos aos usuários.

Levando em consideração as limitações impostas por esta separação, a pesquisa partiu de dados coletados por meio da Biblioteca de Anúncios da Meta com dois métodos distintos: coleta individualizada realizada manualmente pela equipe de pesquisa e/ou raspagem de dados automatizada.

Uma vez coletados, os anúncios foram analisados para identificar indícios de potenciais fraudes, golpes e estelionatos. Para classificação de anúncios fraudulentos, a pesquisa partiu da análise e identificação de infrações de normas jurídicas e códigos autorregulatórios brasileiros estabelecidas sobre consumo, como o Código de Defesa do Consumidor (Brasil, 1990), o Código Penal (Brasil, 1940) e o Conar (2024). Além disso, também foram considerados os aspectos listados na Tabela 1. Em geral, os anúncios fraudulentos apresentavam vários dos problemas listados em uma mesma conta ou post impulsionado, o que contribui com a identificação de conteúdos nocivos na plataforma.

Tabela 1: Critérios de identificação de fraudes em anúncios digitais

| Problema | Elementos indicativos do problema |
|---|---|
| **Uso indevido de imagem e marcas** | • Fotos e vídeos manipulados digitalmente ou via inteligência artificial generativa;<br>• Imagens não autorizadas de pessoas comuns ou de figuras públicas, como artistas, jornalistas e autoridades;<br>• Uso indevido de logotipos ou bancadas de empresas de comunicação, como emissoras de televisão e sites noticiosos;<br>• Uso indevido de logotipos ou fotos de fachadas de órgãos públicos, como a Agência de Vigilância Sanitária, o Ministério da Saúde ou o Supremo Tribunal Federal. |
| **Contas e sites suspeitos ou inautênticos** | • Anunciantes com perfis de Instagram e Facebook recém criados, sem seguidores ou sem engajamento orgânico, ou que buscam se passar por pessoas reais;<br>• Perfis com fotografias retiradas de bancos de imagem;<br>• Múltiplas publicações altamente padronizadas e impulsionadas massivamente;<br>• Direcionamento do usuário para conversas de WhatsApp ou Telegram com números suspeitos;<br>• Direcionamento do usuário para domínios que imitam sites oficiais, com variações sutis de grafia;<br>• Ausência de informações sobre o responsável pelo site, como endereço ou CNPJ, ou registro constando dados falsos. |
| **Solicitações financeiras suspeitas** | • Solicitação de pagamento antecipado para "garantir" um produto;<br>• Pedidos de envio de quantias ou supostas taxas prévias via Pix;<br>• Solicitação de dados pessoais, como CPF ou informações de cartão de crédito. |
| **Conteúdo enganoso ou desinformativo** | • Promessa de resultados milagrosos ou irrealistas;<br>• Preços ou descontos muito abaixo do praticado no mercado;<br>• Uso malicioso do recurso de anúncios dinâmicos, em que um mesmo anúncio exibe conteúdos diferentes de acordo com os usuários que foram impactados pela peça. |

Fonte: Elaboração própria.

Para o desenvolvimento deste estudo, foram coletados dados de anúncios sobre seis casos nas plataformas da Meta. A Tabela 2, abaixo, apresenta os conjuntos de dados analisados neste capítulo, detalhando número de anúncios avaliados, contexto temático da coleta, período em que foram veiculados e tipo de coleta utilizado em cada estudo. Uma vez mapeados os indícios de golpes ou fraudes financeiras, foram analisados aspectos relacionados a estas evidências e à segmentação dos anúncios classificados como políticos, incluindo alcance de audiência, valores investidos, público para o qual o anúncio é segmentado, faixa etária e localização geográfica, considerando que essas informações só estão disponíveis para anúncios políticos.

Tabela 2: Estudos de caso analisados na pesquisa

| Caso analisado | Descrição | Total de anúncios analisados | Período de coleta | Tipo de coleta |
|---|---|---|---|---|
| Caso 1 - Publicidade digital desregulada: Golpes, fraudes e desinformação em anúncios comerciais veiculados nas plataformas da Meta | Levantamento qualitativo voltado a identificar a presença de fraudes na publicidade localizada na Biblioteca de Anúncios da Meta. Identificou as organizações e instituições mais prejudicadas com a falta de transparência na publicidade digital, a saber: instituições financeiras, governo e instituições públicas, empresas de varejo, instituições de ensino e imprensa. | Pesquisa exploratória | Maio a junho de 2023 | Coleta individual via interface da Biblioteca de Anúncios da Meta |
| Caso 2 - O STF e o MJSP em anúncios fraudulentos nas plataformas da Meta: Análise do uso indevido da imagem dos ministros | Anúncios de produtos e serviços que utilizam as imagens do Ministério da Justiça e Segurança Pública (MJSP) e do Supremo Tribunal Federal (STF), assim como seus ministros, para aplicar golpes financeiros e operações de influência que miram as instituições públicas. | 630 anúncios políticos 21 anúncios comerciais | Setembro de 2023 a Fevereiro de 2024 | Coleta automática via Biblioteca de Anúncios da Meta |
| Caso 3 - Rede internacional de fraudes: Páginas de políticos indianos compram anúncios fraudulentos para atingir consumidores brasileiros | Anúncios veiculados pelas páginas oficiais de parlamentares indianos segmentados para usuários brasileiros nas plataformas da Meta. Publicados e impulsionados por Kondeti Chittibabu e Jai Pratap Singh, os anúncios impactaram outros dez países além do Brasil: Ucrânia, Vietnã, Paquistão, Filipinas, Indonésia, Bangladesh, Chile, México, EUA e a própria Índia. As peças não citavam as eleições indianas, que ocorreram entre 19 de abril e 1 de junho de 2024, nem temas relacionados à política. | 530 anúncios políticos | Abril a Junho de 2024 | Coleta automática via Biblioteca de Anúncios da Meta |

| Caso analisado | Descrição | Total de anúncios analisados | Período de coleta | Tipo de coleta |
|---|---|---|---|---|
| Caso 4 - 'Decreto assinado e aprovado!': Como um único perfil veiculou milhares de anúncios com imagens e vozes de políticos brasileiros manipuladas com inteligência artificial para aplicar golpes | Um único perfil, denominado *thiagocampanhola*, veiculou milhares de anúncios com imagens e vozes de políticos brasileiros manipuladas com inteligência artificial para aplicar golpes. O perfil foi o oitavo maior responsável por anúncios políticos nas plataformas da Meta entre os meses de abril e maio de 2024. | 2.137 anúncios políticos | Março a Maio de 2024 | Coleta automática via Biblioteca de Anúncios da Meta |
| Caso 5 - Golpes sobre a Serasa na Meta: Mais anúncios fraudulentos nas plataformas da Meta manipulam imagens e vozes de políticos brasileiros com inteligência artificial | Foram coletados anúncios contendo golpes contra o consumidor usando a falsa narrativa de que a Serasa estaria indenizando clientes por conta de um vazamento de dados que teria ocorrido na plataforma da empresa. | 2.995 anúncios políticos e 72 anúncios comerciais | Março a Maio de 2024 | Coleta individual via interface da Biblioteca de Anúncios da Meta |
| Caso 6 - 'A revelação chocante que pode mudar tudo': Anúncios manipulam discursos de líderes religiosos com inteligência artificial para aplicar golpes em consumidores | Anúncios que utilizam imagens manipuladas de líderes religiosos brasileiros para aplicar golpes financeiros e fraudes, divulgando supostos produtos, serviços e tratamentos a partir das histórias pessoais de padres e pastores e explorando a fé dos seguidores. | 1.850 anúncios comerciais | Julho de 2024 | Coleta individual via interface da Biblioteca de Anúncios da Meta |
| | Total de anúncios coletados no período: | 8.235 anúncios | Maio de 2023 e julho de 2024 | |

Os dados coletados foram analisados e os resultados foram apresentados em seis capítulos do estudo publicado pelo NetLab UFRJ em 2024 (NetLab UFRJ, 2024), como parte do **Observatório da Indústria da Desinformação e seu Impacto nas Relações de Consumo no Brasil**. O primeiro caso buscou explorar quem são os principais prejudicados com a falta de transparência na publicidade digital. Foram observadas e descritas instituições financeiras, públicas e de ensino, empresas de varejo e veículos de imprensa que frequentemente são alvo de fraudes ou têm suas marcas utilizadas para enganar os consumidores. O segundo caso reúne anúncios que exploravam figuras do Ministério da Justiça e Segurança Pública e/ou do Supremo Tribunal Federal (STF). Imagens relacionadas aos ministros eram usadas indevidamente nas plataformas da Meta para atacá-los, comercializar produtos falsos ou promover fraudes. Já o Caso 3 identificou anúncios fraudulentos que segmentam usuários brasileiros, porém foram comprados por páginas de políticos indianos nas plataformas da Meta. O quarto caso, por sua vez, é composto pela veiculação massiva de anúncios comprados por um único perfil, que utilizava imagens e vozes de políticos brasileiros manipulados por meio de inteligência artificial para promover golpes financeiros. Além dos políticos, instituições como o Serasa também são alvos de crimes, como mostram os anúncios coletados e analisados no Caso 5, onde imagens e vozes de autoridades brasileiras foram manipuladas por inteligência artificial para promover fraudes sobre a empresa de avaliação de crédito Serasa. Por fim, o último conjunto de dados coletados para este estudo reúne anúncios fraudulentos que utilizaram e manipularam indevidamente a imagem de lideranças religiosas brasileiras para promover golpes nas plataformas da Meta.

Neste capítulo, apresentamos as principais evidências de golpes e fraudes comuns aos diversos casos e os principais indícios de um padrão de atuação dos golpistas identificados. Serão discutidos 1) quem eram as principais vítimas dessas fraudes; 2) quais foram os recursos utilizados pelos anúncios fraudulentos para enganar os consumidores; 3) quais os principais padrões de uso destes recursos; 4) quais os principais riscos

e prejuízos em potencial oferecidos pela publicidade digital enganosa. Estas evidências são apresentadas no próximo tópico.

## 3.3. Fraudes e publicidade enganosa nas plataformas digitais

Os padrões encontrados na atuação dos criminosos indicam potenciais ações coordenadas na aplicação de golpes nos usuários. Neste tópico, serão discutidos dois aspectos essenciais para a compreensão do ecossistema de publicidade fraudulenta e enganosa. Primeiro, as evidências que indicam que se trata de anúncios fraudulentos e como esses conteúdos se estruturam; em segundo lugar, como os recursos das plataformas são acionados pelos golpistas e as oportunidades que essas ferramentas geram aos usuários mal intencionados para a prática de estelionato.

### *3.3.1 Evidências de Anúncios Fraudulentos*

Cada um dos casos analisados trata de temáticas e personagens distintos, mas adota abordagens, estratégias e narrativas similares para enganar os consumidores e usuários. Em todos os anúncios foram identificados aspectos listados na Tabela 1, o que sugere que há um método sendo aplicado para ampliar o alcance e a eficácia dos anúncios.

#### *Uso indevido de imagem e marcas*

Em todos os casos analisados, foram identificados anúncios com fotos e vídeos editados ou manipulados com inteligência artificial generativa para alterar a imagem e a fala de outros indivíduos. Os principais alvos desta manipulação indevida da imagem pessoal são pessoas públicas com presença midiática ou nas redes sociais. Apresentadores de televisão, artistas, líderes religiosos, políticos e autoridades estão entre as principais figuras reproduzidas. Além disso, logos e fachadas de marcas, empresas e instituições públicas são manipuladas e exploradas pelos golpistas nos anúncios.

Os estelionatários distorcem imagens de autoridades e políticos especialmente nos anúncios analisados nos casos 1, 2, 4 e 5. No caso 1, que mapeou os principais tipos de empresas e instituições utilizadas em golpes nos consumidores, encontramos o uso indevido da imagem de figuras ligadas ao governo federal, além de instituições empresas públicas e privadas, além de instituições e empresas públicas e privadas. No caso 2, 91,4% dos 595 anúncios analisados exploram fotos e vídeos de ministros do judiciário brasileiro, enquanto 56 (8,6%) citam o Ministério da Justiça e Segurança Pública (MJSP) e o Supremo Tribunal Federal (STF) em texto ou fotos. Nos golpes envolvendo a empresa Serasa analisados no estudo 4, dos 2.137 anúncios analisados, 59% reproduziam imagens e áudios manipulados de parlamentares. Os nomes mais acionados foram os dos deputados federais Nikolas Ferreira (PL/MG) (1.254 anúncios) e Sargento Fahur (PSD-PR) (575 anúncios). Vídeos de ambos foram manipulados com inteligência artificial para exibi-los afirmando que a Serasa pagaria até R$30 mil aos brasileiros como suposta indenização por um vazamento de dados ocorrido em 2021. A imagem do presidente Luís Inácio Lula da Silva (PT) também foi utilizada em parte das peças.

O uso criminoso de inteligência artificial para manipular imagens de políticos e interferir nos processos eleitorais nos últimos anos levou o Tribunal Superior Eleitoral (TSE) a definir regras sobre o uso destas ferramentas na geração de conteúdos políticos e eleitorais. A partir de uma resolução de 27 de fevereiro de 2024, o TSE regulamentou no artigo 9º-B (Brasil, 2024b) como o uso de "conteúdo sintético multimídia gerado por meio de inteligência artificial" deve ocorrer. O texto afirma que os responsáveis pela propaganda devem "informar, de modo explícito, destacado e acessível que o conteúdo foi fabricado ou manipulado" por inteligência artificial generativa, além de informar também qual "a tecnologia utilizada". Apesar destas regras terem sido amplamente divulgadas, nos anúncios identificados neste caso não havia qualquer indicação sobre o uso de inteligência artificial generativa. Ou seja, isso indica não só má-fé dos anunciantes que claramente tinham

como objetivo enganar os usuários, mas principalmente falha e/ou falta de controle da empresa Meta em identificar o uso de IA nos anúncios que vendem em suas plataformas, o que desrespeita as normas do TSE.

Além do uso de inteligência artificial para manipular a imagem de políticos e autoridades, artistas e outras personalidades também são alvo. No terceiro caso, onde perfis de parlamentares indianos patrocinaram sistematicamente conteúdos enganosos, cantores, humoristas, apresentadores e até mesmo participantes de reality shows tiveram seus rostos e vozes alterados para vender produtos nas plataformas da Meta. No dia 11 de abril de 2024, o então parlamentar indiano Jai Pratap Singh publicou 58 anúncios idênticos com a cantora brasileira Ana Castela como chamariz para golpes financeiros. Ela, Zezé Di Camargo e Gusttavo Lima apareceram em 153 anúncios (28,8%). Em 370 anúncios (69,8%), os conteúdos traziam imagens associadas a programas de entretenimento da Rede Globo, como o Big Brother Brasil (BBB) de 2024 e o Bate-Papo BBB, além do The Noite, exibido pelo SBT. Entre os perfis mais explorados, destacou-se Leidy Elin, ex-participante do Big Brother Brasil 2024, que apareceu em 46% dos anúncios (244). Outras personalidades da TV também tiveram a imagem utilizada, como os apresentadores Eliana e Celso Portiolli, a atriz e cantora Larissa Manoela, e Key Alves, da edição de 2023 do BBB.

Além de políticos e artistas, os anúncios também utilizaram figuras de líderes religiosos para legitimar seus conteúdos. No estudo 6, em 1.668 peças (90,2% das analisadas), vídeos foram manipulados com uso de inteligência artificial para modificar a voz e a imagem de quatro pastores e três padres: os pastores Cláudio Duarte (477 anúncios), Silas Malafaia (3 anúncios), Rodrigo Silva (4) e um homem desconhecido descrito como pastor (310 anúncios) - ora chamado de Francisco, ora de Enéas - além dos padres Fábio de Melo (333), Reginaldo Manzotti (306) e Marcelo Rossi (123).

As evidências de manipulação e uso indevido de imagem coletadas nestes casos revelaram a ampla utilização de inteligência artificial para manipular vídeos e áudios, criando narrativas falsas e reforçando

a credibilidade dos golpes, prática que vem sendo adotada com frequência por criminosos em ambientes digitais (King et al, 2020; Meireles; Pasito, 2024). Esse uso massivo de tecnologias avançadas sem qualquer tipo de rotulagem ou aviso ao consumidor infringe diretrizes básicas de transparência e ética no ambiente digital (Helmus, 2022). Também acende um alerta sobre o uso ético de IA e indica que estas ferramentas se posicionam cada vez mais como um dos grandes desafios para o combate à desinformação, manipulação de conteúdo e crimes digitais.

Outra estratégia dos golpistas também envolve uso de imagem de terceiros, mas sem inteligência artificial. Pessoas anônimas e modelos cujas fotos estão disponíveis em bancos de imagens também são alvo dos criminosos. Elas são usadas para que eles se passem por médicos, fisioterapeutas, advogados e outros profissionais, reforçando a credibilidade dos perfis falsos. No estudo 4, por exemplo, a página com o maior número de anúncios (809) chamava-se Jessica Santos. Comentários de outros usuários em sua foto de perfil denunciavam o golpe. Além de não apresentar postagens orgânicas, a segunda colocada, Debora Freitas, que publicou 395 anúncios, também foi denunciada por outros usuários que afirmavam se tratar de um perfil enganoso. Ambas utilizavam em seus perfis imagens genéricas de mulheres, extraídas de banco de imagens.

Em relação às marcas e instituições, as imagens também são usadas. Os dados analisados no caso 1 mostram como os anúncios manipulam marcas já consolidadas, com destaque para as empresas de varejo. Estelionatários associavam, por exemplo, notícias sobre o fechamento de lojas de móveis de decoração Tok & Stok à promoção de canais de e-commerce falsos. Cupons de descontos falsos e promoções inverídicas de lojas como Magazine Luiza, Americanas, Renner, Hering e C&A, anunciando promoções ilusórias para enganar usuários, também foram encontrados. Outros exemplos são anúncios de produtos de beleza de marcas como Kérastase, L'Oréal e Avon, que apareceram regularmente em anúncios pagos por páginas falsas. As peças redirecionavam os usuários para sites fraudulentos de todo tipo – desde sites de falsos revende-

dores a páginas que emulavam os portais de grandes empresas do varejo nacional. Nos anúncios, os produtos eram anunciados por preços que variam entre 10% e 16% dos praticados no mercado. Anúncios também prometiam retornos financeiros ilusórios por meio de investimentos via Pix em ações de grandes empresas como a Ambev. Entre as instituições de ensino, anúncios prometiam diplomas e certificados de cursos superiores para estudantes que não haviam concluído a graduação. As certificações, supostamente reconhecidas pelo Ministério da Educação, faziam menções a empresas como a Universidade Estácio de Sá e ao Senac.

Nos casos analisados, os conteúdos que adotavam IA e exploravam a reputação de pessoas ou marcas reconhecidas para ampliar a legitimidade das narrativas nos anúncios fraudulentos eram frequentemente associados a uma segunda estratégia, a de direcionar a sites ou conversas suspeitas, aspecto que será discutido no próximo tópico.

### *Contas e sites suspeitos ou inautênticos*

Os casos analisados ofereceram diversos indícios de que os anúncios fraudulentos utilizavam contas de usuário falsas e redirecionavam o público a links suspeitos. Perfis e páginas do Instagram e Facebook recém-criados, sem seguidores ou sem engajamento orgânico são um exemplo. O uso de identidade falsa descrito anteriormente foi outra evidência mapeada. Nos casos analisados, frequentemente perfis usavam nomes se passando por "doutores" ou "consultores", mas sem fornecer nenhuma comprovação de título. Outras formas de influenciar o comportamento dos consumidores foi o compartilhamento de supostos testemunhos de pessoas que teriam utilizado os serviços ou produtos, mas cujos perfis também traziam os mesmos indícios de inautenticidade. Com frequência, estes perfis sem atividade orgânica publicavam dezenas ou centenas de anúncios idênticos ou com poucas distinções entre si, o que sugere uma ação coordenada de impulsionamento massivo.

A estratégia de utilizar páginas e perfis falsos foi identificada em anúncios coletados em todos os seis casos analisados nesta pesquisa. No

primeiro caso, páginas que se apropriavam indevidamente da imagem da Caixa Econômica Federal e contavam, inclusive, com erros de digitação em seu nome, prometiam empréstimos com aprovação em até 24 horas e de forma facilitada. No caso 2, indícios de comportamento inautêntico estiveram presentes em perfis como Mikaela Queiroz, cuja foto aparecia em outros dois perfis, nomeados como Bianca Soares e Bianca Silva. A página de Mikaela Queiroz, criada poucos dias antes da coleta de dados, publicou 140 anúncios.

Além das contas suspeitas, os usuários também eram direcionados para outros espaços igualmente problemáticos. Entre as características destes sites estão links para conversas automatizadas de WhatsApp ou Telegram, para domínios que imitavam sites reais, com variações sutis de grafia na URL, e para sites sem conteúdo e sem informações sobre o responsável pelo domínio, como endereço ou CNPJ.

No caso 1, um dos sites linkados nos anúncios era a loja de aplicativos do Google, a Google Play Store, onde o público poderia fazer o download dos aplicativos sugeridos pelos anunciantes. Porém, ao serem instalados em celulares e mediante autorização do usuário, os aplicativos poderiam acessar funcionalidades como a localização e o microfone do aparelho. Além disso, conseguiam solicitar dados extras como o número de cartão de crédito e infectar o dispositivo com vírus e/ou *malwares*.

Outra estratégia identificada foi a hospedagem de sites fraudulentos em servidores fora do Brasil ou em servidores que facilitam o anonimato. Isso dificulta a responsabilização dos estelionatários - sem registro no Brasil, é improvável que eventuais denúncias resultem em algum tipo de punição. Entre os principais sites utilizados em fraudes nos anúncios coletados no Caso 2, por exemplo, destacou-se a plataforma Kiwify, especializada na venda de produtos digitais, além de chatbots que simulavam conversas. No caso 3, foram encontrados seis sites que utilizavam o mesmo expediente: uma primeira página simulando veículos de notícia conhecidos, e uma segunda página simulando uma nova empresa de investimentos - ambas no mesmo domínio. Os respon-

sáveis utilizaram a empresa estadunidense Nicenic International Group Co. Limited para terceirizar o registro de todos os sites, mantendo seu anonimato. No estudo 6, 201 anúncios analisados (11%) tinham links que levavam para 10 sites inautênticos. Os proprietários dos domínios usaram empresas nos EUA e na Islândia para realizar o registro. Apenas um dos sites analisados foi registrado no Brasil e vinculado a uma pessoa física. Um aspecto agravante é que os sites tinham dados alarmantes de tráfego. Oito dos 10 sites estavam ativos no momento da análise e receberam mais de 660 mil visitas entre maio e julho de 2024, de acordo com a plataforma de análise SimilarWeb, que monitora tráfego online. Em julho, mês de veiculação dos anúncios, quase todos os sites linkados aos anúncios deste estudo de caso apresentaram expressivo crescimento no volume de acessos.

Outro aspecto preocupante em anúncios que direcionavam a estes domínios e que se apresentou emtodos os casos analisados foi que, em algumas das peças patrocinadas, a URL para a qual o anúncio direcionava era diferente daquela que aparecia descrita no anúncio. Algumas peças veiculadas pelas plataformas da Meta possuíam um botão chamado "Saiba mais", que indicava o endereço de um site e um título. Esse campo deve ser preenchido pelo próprio anunciante no momento da elaboração do anúncio. Em alguns casos, a URL do site sugeria que, ao clicar, o usuário seria redirecionado para a página que aparecia no campo, mas o usuário era direcionado a sites que simulavam a interface dos canais oficiais de veículos midiáticos ou agências governamentais, reproduzindo o menu superior, a estética e o design dos sites originais, dificultando a diferenciação entre o site fraudulento e o verdadeiro. Os conteúdos apresentavam textos extensos e forjados para simularem notícias reais. Esse recurso também foi adotado pelos golpistas no caso 4, em que os anúncios promoviam principalmente três sites: *treinamento.wellitonmonteiro.com*, *noticiasatualizadasonline.online* e *barbaestilo.pepperdigital.com.br*.

Também foram identificados anúncios que apresentavam indevidamente o endereço do site da emissora televisiva SBT para conduzir o usuário para o site do golpe.

Essas práticas evidenciam que a Meta não verifica a veracidade dos conteúdos. Este tipo de exploração de inconsistências da interface é identificada no campo da cibersegurança como "*domain spoofing" ou* "falsificação de domínio em publicidade" (CloudFlare, [s.d.]). Trata-se do que a literatura chama de "*dark pattern*", ou seja, padrões problemáticos na interface de sites e aplicações, que subvertem ou prejudicam a autonomia, a tomada de decisão ou a escolha do consumidor, sendo usados em práticas maliciosas de coação ou manipulação (OECD, 2022).

Além disso, os sites eram diferentes de acordo com a geolocalização. Nos anúncios citados no caso 4, residentes brasileiros e europeus, ao clicarem no mesmo link, eram redirecionados para endereços distintos. Os brasileiros eram encaminhados aos falsos sites de notícias, enquanto os europeus chegavam a uma página da Wikidata sobre a cantora Ana Castela. Trata-se de uma exploração ainda mais sofisticada de *dark patterns* de manipulação da navegação (Mathur et. al, 2021), que segmenta o golpe de acordo com a localização dos usuários.

Nos seis casos também foram encontrados anúncios que levavam a formulários onde o usuário era convidado a inserir dados como nome completo, e-mail e telefone para que a suposta empresa entrasse em contato. Para criar um falso senso de urgência, um aviso indicava que haveria "apenas X vagas restantes" ou que a oferta era válida apenas para os primeiros inscritos. Essas características estão entre os *dark patterns* descritos pela OECD (2022) como *"forced action"* e *"urgency"*. No Caso 3, essa estratégia era frequente. Segundo os próprios sites fraudulentos identificados neste caso, o usuário deveria realizar apenas três passos para receber uma quantia em dinheiro de imediato: 1) Preencher o formulário; 2) Receber uma ligação do "gerente"; 3) Investir um valor inicial para começar a receber o dinheiro prometido. É possível que os estelionatários de fato entrassem em contato com as vítimas com orientações sobre a transferência financeira. Mas os dados também poderiam

ser utilizados para outros tipos de crimes. Diversos golpes tinham como ponto de partida a coleta de dados de usuários vulneráveis.

Outras inconsistências nos sites e contas também sugeriam que os anúncios configuravam golpes. Nos casos 4 e 5, foram identificadas páginas com informações discrepantes - por exemplo, telefone de contato com DDD do estado do Rio de Janeiro e endereço em Rio Branco, no Acre, o que mais uma vez demonstra falta de efetividade nos sistemas de verificação de anunciantes da Meta. No caso 5, apenas uma anunciante havia declarado informações sobre o responsável pelo pagamento dos anúncios. Essas informações são obrigatórias para anúncios categorizados como políticos, eleitorais e/ou de relevância social e incluem dados como telefone, e-mail e endereço.

Sites fraudulentos também tendem a ser efêmeros - os domínios ficam pouco tempo no ar, sendo retirados poucos dias após a publicação ou, em alguns casos, no mesmo dia. No caso 5, entre os 54 sites identificados, 50 já estavam fora do ar no momento da análise. Um dos quatro ainda ativos reproduzia de forma idêntica a identidade visual do site oficial do Ministério Público Federal. E site estava registrado em nome da empresa estadunidense Privacy Protect, um intermediário que vende serviços de anonimato na web. Dois dos sites aos quais os anúncios fraudulentos direcionavam fingiam ser canais do jornalismo da emissora SBT. Neles, os usuários eram redirecionados a sites que simulavam a página da Serasa, onde se solicitavam informações pessoais como o número do CPF. Ambos redirecionavam usuários para sites que se passavam por sites da Serasa, nos quais eram solicitadas informações pessoais, como números de CPF.

Mesmo quando passavam mais tempo no ar, estas páginas tinham comportamentos de audiência pouco usuais. Os seis sites fraudulentos identificados no Caso 3 caso receberam, juntos, mais de 99.746 visitas de usuários durante os quase dois meses de veiculação dos anúncios, segundo dados da plataforma SimilarWeb referentes aos meses de março e abril de 2024 (dados de maio ainda não estavam disponíveis no momento da análise). Por padrão, a SimilarWeb apresenta dados exatos

apenas para sites com mais de 5.000 visitas por mês. Por isso, apenas quatro sites possuíam informações sobre a plataforma de origem do clique. Esses sites também eram inacessíveis por mecanismos de busca, uma característica típica de domínios fraudulentos. Porém, receberam em média mais de 90% dos acessos através das plataformas da Meta, Facebook e Instagram. Mais de 99% das visitas recebidas vinham do Brasil.

Além das inconsistências em contas e sites suspeitos e potencialmente fraudulentos linkados aos anúncios, muitos desses links levam a operações financeiras que prejudicam os consumidores. Elas serão descritas no tópico a seguir.

### *Solicitações financeiras suspeitas*

Embora o foco de alguns dos anúncios fraudulentos analisados fosse obter dados dos usuários, parte significativa explorava estratégias que levavam a prejuízos financeiros imediatos aos usuários. Isso era feito principalmente a partir de solicitações de pagamento antecipado via Pix ou boleto para para "garantir" um produto ou serviço, para assegurar os descontos prometidos ou para aproveitar uma oportunidade única de investimento. Essa estratégia foi identificada em anúncios de todos os casos analisados. No Caso 1, por exemplo, anúncios também prometiam retornos financeiros ilusórios por meio de investimentos via Pix em ações da empresa Ambev.

Uma segunda estratégia é a falsa oferta de empréstimos sem burocracia, voltada especialmente para autônomos ou pessoas sem crédito ativo no mercado. Golpes também recorriam a supostas indenizações do Serasa, com vítimas sendo induzidas a preencher formulários com dados sensíveis e, posteriormente, solicitação de envio de quantias em dinheiro como "taxa" para receber a suposta indenização. Este tipo de fraude financeira, identificado principalmente nos Casos 1 e 2, também afeta a reputação do governo e das instituições públicas. No caso 1, páginas do Facebook recém-criadas e com baixa atividade se apresentavam como supostas representantes de instituições como o Banco do Brasil e

prometiam empréstimos "rápidos e sem burocracias", muitas vezes voltados a servidores públicos. Anúncios prometiam resgate de seguros e as páginas dos anunciantes possuíam comentários de perfis de falsos beneficiários, que comemoravam os empréstimos supostamente recebidos, o que tinha como objetivo dar credibilidade às páginas enganosas. Em meio a comentários feitos por estes perfis falsos, contudo, era possível encontrar denúncias de pessoas reais que haviam sido enganadas pela suposta empresa. Por exemplo, anúncios ofereciam falsos empréstimos em nome do BNDES para pessoas físicas e microempreendedores, também veiculados por páginas recém-criadas no Facebook. Os anúncios divulgavam falsos benefícios nas negociações, como longos prazos para a realização de pagamentos, grandes descontos e outras vantagens mal explicadas. Os anúncios ofereciam a possibilidade de uma "simulação sem compromisso" de um eventual acordo por meio de contatos travados via WhatsApp. Outros anúncios ofereciam empréstimos consignados sob condições especiais para beneficiários de programas como o Bolsa Família. Uma das páginas que oferecia "empréstimos facilitados" e de "liberação rápida" para os beneficiários do programa era um perfil falso associado à Crefisa. Clientes da Caixa Econômica Federal também eram comumente visados pelos enganadores.

Além disso, aplicativos de jogos e investimentos fraudulentos que coletam dados financeiros dos usuários também foram identificados. No caso 1, os usuários eram encaminhados para conversas no WhatsApp ou para a loja de aplicativos da Google, onde era disponibilizado um aplicativo chamado "Pix Trade", que, no momento da análise, já havia sido baixado mais de 1 milhão de vezes. Na página do Pix Trade, porém, foram localizadas avaliações negativas de usuários que efetuaram investimentos e não receberam os valores prometidos.

Essas estratégias que envolvem transações financeiras são uma ameaça crescente aos brasileiros. Em 2024, por exemplo, a pesquisa Panorama Político, do Data Senado, mostrou que 24% dos brasileiros diziam ter perdido dinheiro em crimes cibernéticos no ano anterior, em situações como os golpes digitais (Senado Federal, 2024). Outros

levantamentos e relatórios mostram números similares. Um relatório da empresa de segurança digital NordVPN diz que 30% dos usuários brasileiros de internet já perderam entre R$ 250 e R$ 500 com golpes, em média (Augusto, 2024). Ao adotar identidades visuais que simulam canais oficiais e empregar uma retórica de convencimento, os golpes digitais são cada vez mais exitosos. Essas estratégias de convencimento e os tipos frequentes de conteúdo enganoso identificado nas análises é o tema do tópico a seguir.

*Conteúdo enganoso ou desinformativo*

As escolhas retóricas e narrativas adotadas pelos anunciantes, especialmente quando associadas aos demais problemas listados acima, tornam as evidências de publicidade abusiva e/ou enganosa ainda mais robustas. Entre os aspectos observados estão as promessas irreais de resultados, sejam relacionados à saúde ou estética, seja de ganhos financeiros ou vantagens profissionais.

As narrativas dos anúncios recorrem a problemas comuns a muitos brasileiros, como doenças e angústias compartilhadas por diferentes grupos sociais ou etários. Um dos casos mais emblemático é o sexto estudo, onde anúncios utilizavam as imagens de padres e pastores e incluíam soluções para "homens acima de 45 anos" ou tratavam de doenças como artrite e artrose. Apesar da aplicação de golpes em idosos ser um agravante reconhecido no Código Penal brasileiro (Brasil, 1940, Art. 171), o sistema de publicidade da Meta pode estar ajudando estelionatários a atingirem esse perfil de usuários com seus recursos de microssegmentação. No entanto, a empresa não oferece meios para sua auditabilidade e não é possível verificar o público segmentado ou atingido pelos anúncios, pois estas peças em particular não foram identificadas como políticas.

Para legitimar suas narrativas, os anúncios exibiam falsas "provas" de que os produtos ou serviços oferecidos eram verdadeiros. No Caso 1, golpistas publicavam fotos de extratos bancários falsificados, em que exibiam altas quantias de dinheiro supostamente obtidas por

meio de aplicativos de jogos de azar, os mesmos que o anúncio buscava convencer o usuário a utilizar. Outra evidência neste sentido eram vídeos indicando supostos bens e produtos, como automóveis, adquiridos com a renda obtida com o investimento que o anúncio promovia. Estas peças miravam abertamente pessoas autônomas que tinham dificuldade de comprovar renda e conseguir empréstimos por vias oficiais, oferecendo empréstimos no cartão de crédito com valores que chegavam a R$10 mil. As ofertas eram apresentadas como "fácil, rápida, segura e com ótimas taxas de juros" para clientes que "precisam de dinheiro na hora". Também foram encontrados anúncios sobre uma falsa liberação de dinheiro pelo governo Lula, por conta do "placar do PL 2630". O valor prometido passava de R$ 6,5 mil e era apresentado como uma oferta "para todos os brasileiros".

Descontos irreais ou promoções que parecem "boas demais para serem verdade" também funcionam como um indicativo de golpes financeiros ou estelionato. É o caso dos produtos de beleza de marcas respeitadas ofertados no Caso 1 anunciados por menos de 20% dos valores praticados no mercado em situações normais. Outra evidência identificada nesses dados foram anúncios que prometiam empréstimos consignados "fáceis e sem burocracia" para clientes de diversos bancos. Os anúncios redirecionavam os usuários para sites externos às plataformas da Meta ou para conversas de WhatsApp com supostos consultores bancários. Contudo, as páginas destes "consultores", na maioria das vezes, estavam fora do ar ou, no caso das contas de WhatsApp, não contavam com nenhum tipo de atividade além de imagens de perfil genéricas falsas, dialogando com o aspecto discutido no tópico anterior.

Em geral, os conteúdos também recorrem a chamadas sensacionalistas e atrativas para o consumidor, que sugerem a necessidade de urgência, como "compre agora", "últimas unidades" ou "somente hoje". Essas chamadas, embora sejam corriqueiras em transações comerciais legais, quando associadas aos demais problemas listados acima, podem funcionar como uma *red flag*, ou seja, um sinal de que pode se tratar de golpe. Em todos os casos analisados, os anúncios que utilizavam inteli-

gência artificial para manipular imagens de jornalistas, artistas e outras pessoas famosas também empregavam este tipo de linguagem.

As narrativas falsas e desinformativas usavam trechos de notícias reais, publicadas por veículos de comunicação sérios, e alteravam apenas um segmento para confundir o consumidor. Nos golpes identificados no Caso 6, por exemplo, os anúncios recortavam trechos de vídeos publicados pelos padres e pastores em seus canais próprios e associavam o tema do vídeo ao produto anunciado. Neste conjunto de dados, oito anúncios vendiam soluções milagrosas para perda de peso e destacavam o que chamam de "jornada incrível de emagrecimento" de Marcelo Rossi, explorando um aspecto da trajetória pessoal do padre (Almeida, 2023). Outras sete peças promoviam produtos para zumbido no ouvido. Nos dois casos, os conteúdos exploravam recortes de uma participação do sacerdote no programa Domingão com Huck, da Rede Globo. Outros três anúncios ofertavam supostos tratamentos para ansiedade e depressão, com soluções sem comprovação científica. Nas peças analisadas neste caso, quase 90% dos anúncios apresentavam riscos à saúde dos consumidores. As peças promoviam soluções fáceis para problemas comuns, com destaque para a oferta de medicamentos sem comprovação científica, presente em 1.657 peças identificadas (89,6%). A maioria deles (66,9%) promovia medicamentos para dores diversas. Estes resultados dialogam com outra pesquisa realizada pelo NetLab UFRJ sobre anúncios nas plataformas da Meta, que mostravam que, da publicidade voltada a mulheres analisada no estudo, 79% apresentavam riscos à saúde da mulher (Santini et al, 2024c).

Em algumas situações, os anúncios recorriam à afiliação política ou ideológica do público para vender produtos. No caso 2, um dos conteúdos que utilizavam a imagem dos ministros Dias Toffoli e Gilmar Mendes, divulgado inúmeras vezes, era um suposto livro sobre o ex-presidente Jair Bolsonaro, vendido em diversos sites diferentes. Um deles trazia a capa do livro com o título "Bolsonaro: desvendando a verdade", apresentado como o livro oficial de Bolsonaro e vendido por R$ 22,22 – em referência ao número utilizado pelo ex-presidente durante

a campanha eleitoral de 2022. Em outro anúncio, trechos do Jornal Nacional, da Rede Globo, apareciam editados junto a falas do ministro Alexandre de Moraes sobre liberdade de expressão. Em outro anúncio, que também vendia um livro falso sobre Bolsonaro, as imagens dos ministros e do Jornal Nacional eram utilizadas junto a trechos de uma fala do ex-presidente questionando o processo eleitoral.

Mais do que utilizar uma linguagem apelativa e chamativa, como pode ser comum na publicidade, os anúncios analisados nos seis estudos de caso distorcem a realidade, editam, alteram falas e induzem o consumidor ao erro por meio de narrativas enganosas. Mais do que propagar desinformação, o que este tipo de publicidade faz é recorrer a práticas criminosas para convencer a audiência a clicar, inserir seus dados ou realizar alguma operação financeira. Embora esses anúncios sejam fruto da ação de criminosos, são as plataformas que oferecem as ferramentas e recursos que, com sua finalidade original distorcida, são utilizadas para a aplicação dos golpes, e elas tomam poucas medidas para barrar ou retirar do ar estes conteúdos. No próximo tópico, discutiremos quais são e como são exploradas essas *affordances*.

### *3.3.2 Microssegmentação e conteúdos dinâmicos*

A análise dos dados coletados nos seis casos discutidos neste capítulo elucida quais são as principais ferramentas da Meta exploradas pelos anunciantes que cometem fraudes e estelionato. Duas delas, em especial, chamam a atenção no estudo: a microssegmentação e o uso de anúncios dinâmicos.

Embora seja uma ferramenta importante para anúncios comerciais, a microssegmentação de anúncios também pode ser utilizada por agentes mal intencionados. Esse recurso permite aos golpistas direcionar anúncios às vítimas mais vulneráveis, potencializando os danos financeiros, emocionais e sociais dessas pessoas (Zuboff, 2021). Contudo, conforme apresentado na seção de metodologia, um dos desafios para estudar os impactos da microssegmentação na promoção de golpes e fraudes é a ausência de dados para anúncios não classificados como po-

líticos. A ausência de transparência também impede que se saiba quais são os tipos de dados de usuários utilizados, tampouco como a empresa os utiliza. Isso limita a compreensão do grau de vulnerabilidade e risco a que o público está exposto. Mesmo na Meta, que oferece uma Biblioteca de Anúncios, limitações relacionadas à categorização de conteúdos e inconsistências nos dados disponíveis impossibilitam uma análise completa e detalhada.

Dos seis casos analisados neste capítulo, cinco encontraram conteúdo enganoso e fraudulento em anúncios comerciais, ou seja, não classificados como políticos, o que impede impede o acesso a informações de microssegmentação. Os anúncios que contam com detalhamento de público em aspectos como gênero, faixa etária e geolocalização de faixa etária, bem como valores investidos e alcance das peças, estão nos estudos 2 a 5.

A exploração dos dados de microssegmentação permite identificar quem são as potenciais vítimas deste tipo de golpe, mas também oferece pistas para identificação dos criminosos. Por exemplo, no estudo 3, sobre anunciantes estrangeiros, foi possível avaliar que, para impulsionar as peças publicitárias, foram investidos R$48.610. Os anúncios foram pagos em três diferentes moedas: reais brasileiros, pesos mexicanos e pesos chilenos. Apesar do uso dessas moedas, a Biblioteca de Anúncios da Meta não indicou que qualquer anúncio destas páginas tivesse circulado no Chile ou no México no mesmo período. Com esse valor investido, as peças alcançaram 2.118.916 impressões, entre as quais 58,7% atingiram usuários de São Paulo, Rio de Janeiro, Minas Gerais, Rio Grande do Sul e Bahia. A maior parte dos investimentos dos anúncios em segmentação também foi alocada para atingir usuários nesses mesmos estados, que concentraram 61% dos valores gastos para impulsionamento dos anúncios. Considerando a segmentação por gênero, as mulheres foram o principal público atingido pelos anúncios, com 72% das impressões. Já por segmentação etária, os anúncios foram visualizados principalmente pelo público de 35 a 44 anos. Com este tipo de informação, é

possível desenvolver políticas públicas de conscientização da população e de combate à propagação de fraudes em ambientes digitais.

Outro caso emblemático é o estudo 5. Considerando o número de impressões e os valores gastos com os anúncios categorizados como políticos, foi possível verificar que os estados do Sudeste e Sul foram os principais alvos da publicidade enganosa sobre o Serasa. A soma de investimentos em anúncios direcionados para os sete estados destas regiões foi de R$97 mil, o que corresponde a 54,7% do total investido. Quase na mesma proporção, os anúncios segmentados para estas duas regiões obtiveram 54,5% das 2.150.605 impressões totais alcançadas - somaram 1.171.279 impressões, ou seja, atingiram potencialmente mais de um milhão de pessoas no Brasil. Além disso, os anúncios circularam em pelo menos 1.690 localidades fora do Brasil - ou seja, estados e/ou províncias de diferentes países estrangeiros. O número de impressões realizadas em localização desconhecida também é expressivo, superando as impressões únicas de 11 estados brasileiros. Ao analisar a segmentação por gênero, percebe-se que os homens eram o público predominante dos anúncios, correspondendo a 72% do total de impressões. Em termos de segmentação etária, os anúncios foram vistos principalmente por indivíduos entre 35 e 54 anos, que representaram 60% das impressões totais.

Contudo, a categorização de um anúncio como político nem sempre garante que todos os dados estarão disponíveis. No caso 2, por exemplo, dos 630 anúncios políticos analisados, apenas 220 apresentavam dados de segmentação e 410 tinham dados incompletos. Embora a Meta afirme que é possível consultar os dados demográficos e geográficos de anúncios desse tipo, são desconhecidos os motivos pelos quais essas informações não estavam disponíveis nesses casos.

Já os anúncios dinâmicos são um recurso desenvolvido pela Meta para tornar as peças publicitárias mais eficientes que, no entanto, passou a ser usado pelos golpistas em seu favor. O sistema de publicidade da Meta permite que qualquer anunciante veicule diferentes versões de um mesmo anúncio criando, de forma automática, uma combinação dos

conteúdos personalizada de acordo com o perfil dos usuários impactados pelas peças. Para isso, basta que o anunciante defina um conjunto de diferentes imagens, legendas e outros elementos do anúncio para que os algoritmos da Meta impulsionem as combinações consideradas mais rentáveis para cada tipo de usuário que compõe o público-alvo. Ao clicar na opção "Ver detalhes do anúncio", disponível na interface de usuário da Biblioteca de Anúncios da Meta, é possível visualizar as diferentes versões geradas pelos algoritmos da Meta de um anúncio que utilizou esse serviço.

Esse recurso foi explorado principalmente em dois dos casos analisados. No caso 4, ele apareceu em 1.164 anúncios (38%). As peças traziam imagens variadas - cachorros, gatos, um curso de barbearia e a imagem de uma mulher anônima -, porém, ao selecionar a opção "Ver detalhes do anúncio" na interface de usuário da Biblioteca de Anúncios da Meta, foi possível identificar que esses anúncios utilizavam o serviço de anúncio dinâmico com imagens, legendas e outros conteúdos que tinham o objetivo de aplicar o golpe da falsa indenização da Serasa. Dessa forma, a depender do tipo de usuário impactado pelas peças, uma parte do público-alvo visualizou imagens inofensivas de cachorros, gatos, uma barbearia ou uma mulher anônima e outra parte recebeu o golpe. No Caso 5, o perfil de um anunciante mostrava, além dos golpes usando as imagens dos deputados e a narrativa do falso processo do Serasa, outros 647 anúncios (30,3%) com uma segunda narrativa que mencionava a venda de um suposto curso de barbearia, a depender do tipo de usuário impactado. Essa estratégia de criação de conteúdo pelo sistema da própria Meta pode ser uma forma de burlar a moderação de anúncios da empresa, uma vez que, dos 647 anúncios dinâmicos divulgando golpes no Caso 5, 610 escaparam de qualquer moderação.

Assim, a análise dos anúncios nos seis casos demonstra que os mecanismos de transparência da plataforma possuem lacunas e que os recursos de impulsionamento de conteúdos são usados de maneira mal intencionada. As lacunas estão tanto em classificar como sensíveis os anúncios que tratam de temas sociais, política e eleições, quanto em

disponibilizar as informações desses anúncios quando, de fato, são categorizados dessa maneira. Além disso, o direcionamento dos anúncios com golpes e fraudes a públicos sensíveis, possibilitado pela microssegmentação, é outro aspecto identificado nas análises. Como já vem sendo discutido na literatura sobre o uso de algoritmos e inteligência artificial na publicidade digital, essas ferramentas facilitam que os agentes maliciosos possam inferir informações sensíveis sobre seus alvos ou vítimas (Arsenault, 2020).

## 3.4. Considerações finais

Os seis casos analisados neste capítulo revelam padrões e recursos usados para promover fraudes e enganar consumidores usando as plataformas da Meta. Apontam também as principais pessoas, marcas, empresas e instituições prejudicadas com a circulação desses conteúdos enganosos e os potenciais prejuízos e danos. Somados, os estudos de caso apresentados neste capítulo analisaram 8.235 anúncios fraudulentos que evidenciaram a sofisticação e a organização por trás dessas práticas. Os resultados das análises ilustram como os criminosos exploram de maneira maliciosa os recursos oferecidos pelas plataformas, sua opacidade e a falta de regulamentação do ecossistema publicitário digital.

A análise dos anúncios também evidenciou que muitos dos conteúdos continuaram no ar por períodos extensos. Uma situação emblemática neste sentido se verifica no caso 4. Embora os mais de 2 mil anúncios sejam fraudulentos e tenham sido veiculados pelo mesmo perfil, a Meta informava ter moderado menos da metade dos conteúdos (1.064), por terem infringido "leis locais" e diretrizes sobre propriedade intelectual não especificadas. Em 859 anúncios, a Meta justifica a retirada dos conteúdos por não terem seguido as "políticas sobre propriedade intelectual" da empresa, enquanto apenas um anúncio foi moderado por infringir "Leis Locais". Estes 860 anúncios estavam indisponíveis para os pesquisadores no momento da análise. Outros 133 foram veiculados sem a inserção de um rótulo de conteúdo político e 71 anúncios foram moderados por problemas relacionados aos "Padrões de publicidade" da

Meta, mas, em ambos os casos, ainda era possível acessar seu conteúdo na biblioteca no momento da realização do estudo. Mas a pergunta é: por que a Meta não moderou e tirou de veiculação os outros 1.073 anúncios fraudulentos comprados e veiculados pelo mesmo perfil, com as mesmas características e com o mesmo tipo de conteúdo e de golpe que impactaram milhares de consumidores em suas plataformas? Vale observar que independente da identificação de problemas relacionados aos anúncios, e de moderá-los ou não, a plataforma não deixa de receber as quantias investidas nesse tipo de publicidade enganosa e fraudulenta. A remoção parcial dos conteúdos fraudulentos evidencia que o sistema de controle de anúncios da Meta é falho e apresenta elevado risco aos consumidores, uma vez que os demais anúncios continuaram no ar. Por outro lado, o método de retirada de circulação dos anúncios fraudulentos impede a análise posterior dos pesquisadores, o que representa um problema para as boas práticas de transparência e ao acesso aos dados das plataformas para pesquisa e observação externa.

Algumas das evidências identificadas no estudo sugerem, ainda, a presença de uma rede organizada de anunciantes criminosos com atuação nacional e internacional. É o caso, por exemplo, do alto número de anúncios, muitas vezes feitos por um único anunciante, como no caso 4, e do volume significativo de dinheiro investido, evidenciado principalmente no caso 6. Outro aspecto que indica a coordenação das ações na publicidade enganosa é a existência de inconsistências entre os países e sistemas monetários distintos onde os anúncios são criados e administrados, como explicita o caso 3.

Esses resultados indicam a necessidade de regulamentação mais robusta e de medidas para aumentar a transparência da publicidade online que circula nas redes sociais. Sem a regulamentação e aplicação efetiva de regras que protejam o consumidor, as plataformas continuam sendo um terreno fértil para práticas abusivas que comprometem tanto os direitos dos consumidores quanto a integridade das instituições.

Da mesma maneira, o uso da imagem de políticos e autoridades, como ministros de Estado, bem como de programas e campanhas go-

vernamentais para legitimar produtos falsos e confundir e enganar os consumidores, mapeado nos estudos de caso 1, 4 e 5, é outra evidência da necessidade urgente de se desenvolver formas de combater e desestimular financeiramente a publicidade digital abusiva e fraudulenta.

Este capítulo buscou contribuir para o debate sobre governança digital, oferecendo insumos para a formulação de políticas públicas e estratégias de mitigação que possam proteger os consumidores brasileiros e promover um ambiente digital mais seguro e ético.

No próximo capítulo, apresentamos outros estudos que mostram como a desinformação e a distorção de informações oficiais são usadas para enganar os consumidores nas plataformas digitais causando prejuízos às políticas e instituições públicas.

## Referências

AGÊNCIA BRASIL. Google não permitirá anúncios de políticos nas eleições de outubro. *Agência Brasil*, Brasília, 24 abr. 2024. Disponível em: https://agenciabrasil.ebc.com.br/justica/noticia/2024-04/google-nao-permitira-anuncios-de-politicos-nas-eleicoes-de-outubro. Acesso em: 2 abr. 2025.

ALMEIDA, Diogo. Foco, força e fé: Padre Marcelo Rossi revela rotina que adotou para manter a forma física após empurrão. *G1*, Paraíba, 21 jun. 2023. Disponível em: https://g1.globo.com/pb/paraiba/noticia/2023/06/21/foco-forca-e-fe-padre-marcelo-rossi-revela-rotina-que-adotou-para-manter-a-forma-fisica-apos-empurrao.ghtml. Acesso em: 18 mar. 2025.

ALPHABET. Annual Report Pursuant to Section 13 or 15(d) of the Securities Exchange act of 1934. *Alphabet*, 2024. Disponível em: https://abc.xyz/assets/43/44/675b83d7455885c4615d848d52a4/goog-10-k-2023.pdf. Acesso em: 1 ago. 2024.

ARSENAULT, Amelia. Microtargeting, Automation, and Forgery: Disinformation in the Age of Artificial Intelligence. *Communities & Collections*, u Ottawa, Collections Affaires publiques et internationales - Mé-

moires, 2020. Disponível em: https://doi.org/10.20381/ruor-24728. Acesso em 24 mar. 2025

AUGUSTO, Jean. Golpes financeiros no Brasil: uma realidade comum e perigosa. *NordVPN*, 19 dez. 2024. Disponível em: https://nordvpn.com/pt-br/blog/golpes-financeiros/. Acesso em: 21 mar. 2025.

BRASIL. Lei no 8.078, de 11 de setembro de 1990. Código de Defesa do Consumidor. *Diário Oficial [da] República Federativa do Brasil*, Poder Executivo, Brasília, DF, 11 set. 1990. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 10 dez. 2024.

BRASIL. Lei nº 2.848, de 7 de dezembro de 1940. Código Penal. Presidência da República, *Casa Civil*, Subchefia para Assuntos Jurídicos, Rio de Janeiro, RJ, 7 dez. 1940. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 10 dez. 2024.

BRASIL. Constituição da República Federativa do Brasil de 1988. Brasília, DF: Presidente da República, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm Acesso em: 31 jan. 2025.

BRASIL. Ministério da Saúde. Agência Nacional de Vigilância Sanitária, Anvisa. Resolução da Diretoria Colegiada RDC nº 855, de 23 de abril de 2024. Brasília, DF: *ANVISA*, 2024a. Disponível em: https://anvisalegis.datalegis.net/action/UrlPublicasAction.php?acao=abrirAtoPublico&num_ato=00000855&sgl_tipo=RDC&sgl_orgao=RDC/DC/ANVISA/MS&vlr_ano=2024&seq_ato=000&cod_modulo=310&cod_menu=9434. Acesso em: 10 dez. 2024.

BRASIL. Tribunal Superior Eleitoral. *Resolução nº 23.732*, de 27 de fevereiro de 2024. Brasília, DF: *Tribunal Superior Eleitoral*, 2024b. Disponível em: https://www.tse.jus.br/legislacao/compilada/res/2024/resolucao-no-23-732-de-27-de-fevereiro-de-2024. Acesso em: 10 dez. 2024.

BUTTON, Mark; NICHOLLS, Carol N.; KERR, Jane; OWEN, Rachael. Online frauds: Learning from victims why they fall for these scams. *Australian & New Zealand Journal of Criminology, [S.l]*, v. 47, n. 1, p. 3-19,

2014. Disponível em: https://doi.org/10.1177/0004865814521224. Acesso em: 9 dez. 2024.

CHEN, Vivian. H. H. Unobtrusive Measures in Studying Social Media. In: MATTHES, Jorg. P.; DAVIS, Christine. S.; POTTER, Robert. F. (Eds.). *The international encyclopedia of communication research methods*. Hoboken, NJ: Wiley-Blackwell, 2017. Disponível em: https://doi.org/10.1002/9781118901731.iecrm0257. Acesso em: 10 dez. 2024.

CLOUDFLARE. Página inicial.*[S.d]* Disponível em: https://www.cloudflare.com/pt-br/. Acesso em: 31 jan. 2025.

CONAR, Conselho Nacional de Autorregulamentação Publicitária. Código Brasileiro de Autorregulamentação Publicitária. São Paulo: *Conar*, 2024. Disponível em: http://www.conar.org.br/pdf/Codigo-CONAR-2024.pdf. Acesso em: 10 dez. 2024.

DOMINGOS, Roney. É #FAKE vídeo que diz que governo federal condenou Serasa a pagar indenização de R$ 30 mil. *G1*, *[S.l.],* 30 mar. 2024. Disponível em: https://g1.globo.com/fato-ou-fake/noticia/2024/03/30/e-fake-video-que-diz-que-governo-federal-condenou-serasa-a-pagar-indenizacao-de-r-30-mil.ghtml. Acesso em: 10 dez. 2024.

DOMMETT, Katharine; ZHU, Junyan. What is an online political advert? An interrogation of conceptual challenges in the formation of digital policy response. Sydney: *Policy & Internet*, v. 15, p. 713–730, 2023. Disponível em: https://doi.org/10.1002/poi3.350. Acesso em: 10 dez. 2024.

FAUSTINO, Marco. É falso que Serasa pagará R$ 30 mil a quem teve dados vazados. *Aos Fatos*, *[S.l.],* 8 fev. 2024. Disponível em: https://www.aosfatos.org/noticias/falso-serasa-indenizacao-dados-vazados/. Acesso em: 10 dez. 2024.

FTC - FEDERAL TRADE COMMISSION. FTC issues orders to social media and video streaming platforms regarding efforts to address surge in advertising fraud. *FTC*, *[S.l.],* 2023. Disponível em: https://www.ftc.gov/news-events/news/press-releases/2023/03/ftc-issues-orders-social-media-

-video-streaming-platforms-regarding-efforts-address-surge-advertising. Acesso em: 9 dez. 2024.

HELMUS, Todd C. Artificial intelligence, deepfakes, and disinformation: a primer. Califórnia: *RAND Corporation*, n.p. 2022. Disponível em: https://www.rand.org/pubs/perspectives/PEA1043-1.html. Acesso em: 31 jan. 2025.

KANTAR IBOPE MEDIA. Advertising Intelligence - Projeções IAB. *Kantar Ibope*, 2023. Disponível em: https://iabbrasil.com.br/pesquisa-digital-adspend-2023/. Acesso em: 9 dez. 2024.

KIKERPILL, Kristjan; SIIBAK, Andra. Abusing COVID-19 pan(dem) ic. In: LEIGHTON, Timothy; JAMES, Eleanor (eds.). *COVID-19 in International Media*. London: Routledge, 2021. Disponível em: https://www.taylorfrancis.com/chapters/edit/10.4324/9781003181705-25/abusing-covid-19-pan-dem-ic-kristjan-kikerpill-andra-siibak. Acesso em: 9 dez. 2024.

KING, Thomas C.; AGGARWAL, Nikita; TADDEO, Mariarosaria; FLORIDI, Luciano. Artificial intelligence crime: an interdisciplinary analysis of foreseeable threats and solutions. *Science and Engineering Ethics*, *[S.l.]* v. 26, p. 89–120, 2020. Disponível em: https://doi.org/10.1007/s11948-018-00081-0. Acesso em: 31 jan. 2025.

KOTILA, Mikko; RUMIN, Ruben C.; DHAR, Shailin. The advertising Fraud Council. Compendio sobre el fraude publicitario para inversores en medios. *World Federation of Advertisers*, *[S.l.]*, 2016. Disponível em: https://www.anda.cl/estudios-akc/articulos-wfa/. Acesso em: 9 dez. 2024.

LEERSSEN, Paddy; AUSLOOS, Jef; ZAROUALI, Brahim; HELBERGER, Natali; DE VREESE, Claes. Platform ad archives: promises and pitfalls. *Internet Policy Review*, *[S.l.]* v. 8, n. 4, 2019. Disponível em: https://doi.org/10.14763/2019.4.1421. Acesso em: 10 dez. 2024.

MATHUR, Arunesh; KSHIRSAGAR, Mihir; MAYER, Jonathan. What makes a dark pattern... dark? Design attributes, normative considerations, and measurement methods. In: Proceedings of the 2021 CHI con-

ference on human factors in computing systems. New York: *Association for Computing Machinery*, 2021. p. 1-18.

MEIRELES, Isys G.; PASITTO, Fernando T. Estelionato e suas implicações: o constante crescimento dos golpes virtuais. São Paulo: *Revista Ibero-Americana de Humanidades, Ciências e Educação*, v. 10, n. 11, p. 6303–6316, 2024. Disponível em: https://doi.org/10.51891/rease.v10i11.17063. Acesso em: 10 dez. 2024.

META. Meta Reports Fourth Quarter and Full Year 2023 Results; Initiates Quarterly Dividend. *Meta*, 1 fev. 2024. Disponível em: https://investor.fb.com/investor-news/press-release-details/2024/Meta-Reports-Fourth-Quarter-and-Full-Year-2023-Results-Initiates-Quarterly-Dividend/default.aspx. Acesso em: 1 ago. 2024.

META. Biblioteca de Anúncios. Meta, *[S.d.]*a. Disponível em: <https://www.facebook.com/ads/library>. Acesso em: 1 ago. 2024

MIYASHIRO, Kelly. Fátima Pissarra, da Mynd8, sobre escândalo da Choquei: 'Sofri ameaças'. *Veja - Coluna Tela Plana, [S.l.]*,23 fev. 2024. Disponível em: https://veja.abril.com.br/coluna/tela-plana/fatima-pissarra-da-mynd8-sobre-escandalo-de-choquei-sofri-ameacas. Acesso em: 10 dez. 2024.

MORGADO, Flavio. Golpes financeiros contra pessoas idosas por meio de engenharia social no ambiente digital. São Paulo: *Rev. Longeviver,* Ano VI, n. 24, out./nov./dez. 2024. Disponível em: https://revistalongeviver.com.br/index.php/revistaportal/article/view/17. Acesso em: 10 dez. 2024.

NETLAB UFRJ. *Anúncios falsos, seus mecanismos e potenciais danos aos consumidores brasileiros.* Rio de Janeiro: NetLab/UFRJ, 2024. Disponível em: https://www.netlab.eco.br/observatorio-industria-desinformacao. Acesso em: 21 mar. 2025.

NÚCLEO JORNALISMO. Facebook ainda é usado para comprar e vender armas no Brasil. *Núcleo Jornalismo [S.l.]*, 01 abr. 2021. Disponível em: https://nucleo.jor.br/reportagem/2021-04-01-facebook-comercio-armas/. Acesso em: 10 dez. 2024.

OECD. Dark commercial patterns, OECD Digital Economy Papers, No. 336. Paris: *OECD Publishing*, 2022. Disponível em: https://doi.org/10.1787/44f5e846-en.

ONU MULHERES; PNUD. Prevenir a violência contra as mulheres durante as eleições. Um guia programático. *PNUD*, *[S.l.]*, 2020. Disponível em: https://www.onumulheres.org.br/wp-content/uploads/2021/12/Guia-VCME_web.pdf. Acesso em: 10 dez. 2024.

REUTERS FACT CHECK. Checagem de fatos: Anúncio que promete indenização de R$30 mil do Serasa é golpe. *Reuters Fact Check*, *[S.l.],* 5 mar. 2024. Disponível em: https://www.reuters.com/fact-check/portugues/YMDGWFSSSBI3TMIF7M7K2WQWHU-2024-03-05/. Acesso em: 10 dez. 2024.

RICH, Jessica L. When women are targeted for scams. *Federal Trade Commission*, *[S.l.]* 11 mar. 2016. Disponível em: https://www.ftc.gov/news-events/news/public-statements/when-women-are-targeted-scams. Acesso em: 10 dez. 2024.

ROGERS, Richard. *Digital Methods*. Cambridge, MA: The MIT Press, 2013. Disponível em: https://doi.org/10.7551/mitpress/8718.001.0001. Acesso em: 10 dez. 2024.

SANTINI, R. Marie; SALLES, Débora; BARROS, Carlos Eduardo; MAURICIO, Bruno; MOREIRA, Alékis; DIAS, Bernardo; HADDAD, João Gabriel; GOMES, Matheus. Golpe financeiro através de anúncios no Meta Ads. Rio de Janeiro: *Netlab* - Laboratório de Estudos de Internet e Redes Sociais, Universidade Federal do Rio de Janeiro (UFRJ). 23 abr. 2023. Disponível em: https://netlab.eco.ufrj.br/post/golpe-financeiro-atrav%C3%A9s-de-an%C3%BAncios-no-meta-ads. Acesso em: 10 dez. 2024.

SANTINI, R. Marie; SALLES, Débora; MATTOS, Bruno; CANAVARRO, Marcela; BARROS, Car los E.; MOREIRA, Alékis; MEDEIROS, Priscila; GRAEL, Felipe; FERREIRA, Fernando; MELO, Danielle; BORGES, Marcio; HADDAD, João G.; MURAKAMI, Lucas; SILVA, Daphne; DAU, Erick; LOUREIRO, Felipe. Índice de Transparência de Dados das Plataformas de Redes Sociais. Rio de Janeiro: *NetLab* – Labo-

ratório de Estudos de Internet e Redes Sociais, Universidade Federal do Rio de Janeiro (UFRJ). Publicado em 04 de novembro de 2024a.

SANTINI, R. Marie; SALLES, Débora; MATTOS, Bruno; CANAVARRO, Marcela; BARROS, Carlos E.; MOREIRA, Alékis; GRAEL, Felipe; FERREIRA, Fernando; MELO, Danielle; BORGES, Marcio; CIODARO, Thiago; SANCHOTENE, Nicole; HADDAD, João G.; MURAKAMI, Lucas; SILVA, Daphne; DAU, Erick; LOUREIRO, Felipe. Índice de Transparência da Publicidade nas Plataformas de Redes Sociais. Rio de Janeiro: *NetLab* – Laboratório de Estudos de Internet e Redes Sociais, Universidade Federal do Rio de Janeiro (UFRJ). Publicado em 04 de novembro de 2024b.

SANTINI, R. Marie et al. Golpes, Fraudes e Desinformação na Publicidade Digital Abusiva Contra Mulheres. Rio de Janeiro: *NetLab* - Laboratório de Estudos de Internet e Redes Sociais - Universidade Federal do Rio de Janeiro (UFRJ), 8 mar. 2024c. Disponível em: https://netlab.eco.ufrj.br/post/golpes-fraudes-e-desinformac-a-o-na-publicidade-digital-abusiva-contra-mulheres. Acesso em: 31 jan. 2025.

SENADO FEDERAL. Panorama Político 2024: apostas esportivas, golpes digitais e endividamento. 21ª ed. Instituto de Pesquisa DataSenado. Brasília, DF: Senado Federal, set. 2024. Disponível em: https://www12.senado.leg.br/institucional/datasenado/materias/relatorios-de-pesquisa/golpes-digitais-atingem-24-dos-brasileiros-aponta-21a-edicao-da-pesquisa-panorama-politico. Acesso em: 21 mar. 2025.

SERASA. Golpe do Limpa Nome: conheça e previna-se. *Serasa*, 10 nov. 2023. Disponível em: https://www.serasa.com.br/premium/blog/golpe-limpa-nome/. Acesso em: 10 dez. 2024.

SHUKLA, Vandinika. Deepfakes and Elections: The Risk to Women's Political Participation. *Tech Policy Press*, *[S.l.],* 29 fev. 2024. Disponível em: https://www.techpolicy.press/deepfakes-and-elections-the-risk-to-womens-political-participation/. Acesso em: 10 dez. 2024.

SILVA, Vitor Esmanhotto da. A (in)aplicabilidade do CDC a contratos de prestação de serviço de marketing digital. *Consultor Jurídico*, *[S.l.],* 16 fev. 2022. Disponível em: https://www.conjur.com.br/2022-fev-16/

opiniao-inaplicabilidade-cdc-contratos-marketing-digital/. Acesso em: 10 dez. 2024.

SMITHERS, Rebecca. Women 'most likely to fall for internet scams'. *The Guardian*, 10 nov. 2010. Disponível em: https://www.theguardian.com/money/2010/nov/10/internet-scam-female-victims. Acesso em: 10 dez. 2024.

SOARES, Gabriela. Criminosos dão golpe em bolsonaristas com venda de livros falsos. *Lupa*, 30 set. 2023. Disponível em: https://lupa.uol.com.br/jornalismo/2023/09/30/criminosos-dao-golpe-em-bolsonaristas-com-venda-de-livros-falsos. Acesso em: 10 dez. 2024.

SOARES, Gabriela. Padre Fábio de Melo não divulgou plataforma de investimento em entrevista no The Noite com Danilo Gentili. *Lupa*, 19 jul. 2024. Disponível em: https://lupa.uol.com.br/jornalismo/2024/07/19/padre-fabio-de-melo-nao-divulgou-plataforma-de-investimento-em-entrevista-no-the-noite-com-danilo-gentili. Acesso em: 10 dez. 2024.

SOBIERAJ, S. Bitch, slut, skank, cunt: patterned resistance to women's visibility in digital publics. *Information, Communication and Society, [S.l.]*, 2017. Disponível em: https://doi.org/10.1080/1369118X.2017.1348535. Acesso em: 10 dez. 2024.

STATISTA. Annual advertising revenue of Meta Platforms worldwide from 2009 to 2022. *Statista*, *[S.l.]*, 30 jan. 2023a. Disponível em: https://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/. Acesso em: 9 dez. 2024.

STATISTA. Ad-selling companies U.S. digital ad revenue shares 2021-2026. *Statista*, 2024. Disponível em: https://www.statista.com/statistics/242549/digital-ad-market-share-of-major-ad-selling-companies-in-the-us-by-revenue/. Acesso em: 1 ago. 2024.

UOL. Medo de comprar online? Veja os golpes recentes mais aplicados. *UOL*, *[S.l.]*,2023. Disponível em: https://www.uol.com.br/tilt/noticias/redacao/2023/02/07/medo-de-comprar-online-veja-os-golpes-recentes-mais-aplicados.htm. Acesso em: 9 dez. 2024.

UK FINANCE. Over £1.2 billion stolen through fraud in 2022, nearly 80% via app. *UK Finance [S.l.]*,11 mai. 2023. Disponível em: https://www.ukfinance.org.uk/news-and-insight/press-release/over-ps-12-billion-stolen-through-fraud-in-2022-nearly-80-cent-app. Acesso em: 9 dez. 2024.

VASCONCELLOS, Hygino. Serasa: criminosos criaram mais de 54 mil anúncios sobre falsa indenização. *UOL Notícias*, 29 fev. 2024. Disponível em: https://noticias.uol.com.br/cotidiano/ultimas-noticias/2024/02/29/serasa-criminosos-criaram-mais-de-54-mil-anuncios-sobre-falsa-indenizacao.htm?cmpid=copiaecola. Acesso em: 10 dez. 2024.

WFA - ADVERTISING FRAUD COUNCIL. Compendium of Ad Fraud Knowledge for Media Investors. *WFA*, *[S.l.]*, 2016. Disponível em: https://wfanet.org/knowledge/item/2016/06/03/Compendium-of--ad-fraud-knowledge-for-media-investors. Acesso em: 9 dez. 2024.

ZALNIERIUTE, Monika. "Transparency-Washing" in the Digital Age: A Corporate Agenda of Procedural Fetishism. Critical Analysis of Law, v. 8, n. 1, p. 39–53, 2021. Sydney: *UNSW Law Research Paper*, n. 21-33. Disponível em: https://ssrn.com/abstract=3805492. Acesso em: 10 dez. 2024.

ZUBOFF, Shoshana. *A era do capitalismo de vigilância*. Editora Intrínseca, *[S.l.]*, 2021. Disponível em: https://intrinseca.com.br/livro/a-era-do--capitalismo-de-vigilancia/. Acesso em: 9 dez. 2024.

# Capítulo 4

## Ecossistema de desinformação e fraudes com marcas de programas governamentais

Conforme discutido no capítulo anterior, uma estratégia recorrente empregada por estelionatários para promover fraudes por meio de publicidade online é captar a confiança da população apropriando-se da imagem de personalidades públicas, como celebridades, líderes religiosos, figuras políticas e instituições governamentais. Essa tática visa conferir credibilidade a anúncios enganosos, induzindo consumidores a acreditarem que se trata de informações oficiais e confiáveis. Além disso, o uso de ferramentas de Inteligência Artificial tem aprofundado esse problema, facilitando a adulteração de imagens desses atores e veículos de mídia, tornando ainda mais difícil distinguir esses conteúdos de fontes legítimas.

Essa indústria de anúncios irregulares integra um amplo ecossistema de desinformação, que não se limita à disseminação de informações falsas, mas envolve uma rede complexa de relações, fluxos de atenção e dinâmicas de recepção, combinando diferentes atores e tecnologias e se expandindo continuamente em múltiplos níveis e ramificações (Salles; Martins; Santini, 2024). Argumentamos ao longo dos capítulos que o uso avançado de ferramentas de microssegmentação das plataformas permite a personalização de conteúdos enganosos e a exploração de vulnerabilidades da audiência. Dessa maneira, a publicidade irregular impulsionada pelas plataformas reforça e sofistica o ecossistema de desin-

formação, tendo impactos negativos concretos sobre os usuários, como perdas financeiras diretas e roubo de dados pessoais para uso posterior em novos golpes.

Além destes prejuízos ao consumidor, a desinformação pode causar danos sociais mais amplos, impactando negativamente a divulgação e implementação de políticas públicas, dificultando a comunicação dos governos com a população e inibindo a aderência a benefícios sociais (Cugler; Vaz, 2022). No caso da publicidade irregular nas plataformas digitais, a aplicação de fraudes por meio do uso da imagem de personalidades políticas e de instituições governamentais induz os usuários ao erro, afeta a percepção da população sobre direitos e benefícios oferecidos pelo Estado e compromete a confiança da população no governo, nas instituições públicas e nos seus canais de comunicação oficiais.

A partir desse contexto e das evidências de publicidade irregular nas plataformas da Meta apresentadas no capítulo 4, este capítulo analisa como as políticas públicas se tornaram alvo dessas operações fraudulentas. Esse lucro beneficia não apenas os criminosos que operam os esquemas, mas também as próprias plataformas de redes sociais, que lucram com a distribuição de publicidade enganosa.

## 4.1. Roubo de dados e o aumento do endividamento pelas redes sociais

Pesquisadores e agentes do poder público em diversos países demonstraram que uma das estratégias usadas por estelionatários em fraudes online é o uso indevido do nome e da imagem de instituições governamentais para atrair vítimas e aplicar golpes (Kemp & Pérez, 2023; Goel, 2021; Setera, 2021). Passando-se pelo governo e por organizações legítimas, agentes maliciosos aproveitam-se de audiências vulneráveis a determinadas demandas, necessidades e interesses para divulgar desinformação, promover fraudes financeiras e roubar dados pessoais dos usuários (van den Hout *et al.*, 2022).

No Brasil, a imprensa e entidades governamentais têm alertado que sites e plataformas digitais são frequentemente usados para distri-

buição e veiculação de diversos tipos de fraudes que atingem milhares de usuários utilizando indevidamente a imagem de atores políticos e jornalistas, em muitos casos servindo-se de ferramentas de Inteligência Artificial (Aleixo, 2024; Maia, 2024; Procon-SP, 2024; Secretaria de Comunicação Social, 2024). Baseado em quatro estudos de caso desenvolvidos pelo NetLab UFRJ em 2023 e 2024 (Santini et al., 2023a; Santini et al. 2023b; Salles et al., 2023; NetLab UFRJ, 2024), este capítulo visa abordar este fenômeno no Brasil utilizando dois casos recentes: os programas governamentais Desenrola Brasil e Voa Brasil, lançados em 2023 e 2024, respectivamente.

Em um cenário de endividamento crescente das famílias brasileiras (Abdala, 2023), em 17 de julho de 2023, o governo federal lançou a primeira etapa do programa Desenrola Brasil para facilitar a renegociação de dívidas. Na primeira fase do programa, pessoas que tivessem dívidas bancárias de até R$100 teriam seus nomes automaticamente limpos, enquanto pessoas físicas com renda mensal de até R$20 mil com dívidas de qualquer valor poderiam negociar diretamente com as instituições financeiras (Ministério da Fazenda, 2023). Posteriormente, em outubro de 2023, foi lançada outra fase do programa, destinada à população de baixa renda, com foco em cidadãos com renda de até dois salários mínimos ou inscritos no Cadastro Único (CadÚnico). Essa etapa incluiu, além de dívidas bancárias, débitos de contas de luz e água, mensalidades educacionais e compras no varejo (Barcellos, 2024).

O Voa Brasil, por sua vez, é um programa governamental que busca democratizar o acesso ao transporte aéreo no Brasil, dando direito à compra de passagens de avião por até R$200 por trecho. Inicialmente, o projeto tinha previsão de lançamento para agosto de 2023 (Ribeiro, 2023), mas a primeira fase do programa foi lançada apenas em julho de 2024 (Dantas, 2024). Nos meses que antecederam o lançamento, o programa foi amplamente divulgado pela mídia brasileira, com declarações de autoridades e especulações sobre sua implementação (Sabóia, 2024). Na primeira fase, a iniciativa contemplava aposentados do INSS de todas as faixas de renda, que poderiam adquirir até dois bilhetes por ano.

A segunda fase, prevista para o primeiro semestre de 2025, visa beneficiar estudantes do ProUni e do Pronatec (Portos e Aeroportos, 2024).

A imprensa brasileira noticiou amplamente que esses programas têm sido utilizados por estelionatários, que compram anúncios nas plataformas digitais fingindo ser facilitadores da adesão de tais programas. A Folha de S. Paulo, por exemplo, apontou que o nome do Voa Brasil foi usado em meados de 2023, antes mesmo do lançamento do programa, para aplicar golpes em usuários que procuravam por passagens aéreas (Barboza, 2023). O estudo do NetLab UFRJ (Salles et al., 2023) embasou outras matérias em veículos de grande impacto nacional que reforçavam essas denúncias (Martins; Duarte, 2023), demonstrando como essa publicidade irregular atraía usuários para um ecossistema de desinformação que culmina em golpes financeiros. Ou seja, apesar das evidências e do amplo conhecimento público do problema, a Meta não tomou medidas para removê-los ou impedir sua recorrência.

Neste capítulo, partimos deste diagnóstico inicial para demonstrar como essas fraudes não são apenas irregularidades eventuais, mas que constituem uma indústria de publicidade fraudulenta, cuja consolidação no Brasil tem sido facilitada e alavancada pelas plataformas digitais. Apresentamos evidências, números deste mercado e estratégias de persuasão utilizadas por anunciantes ilegítimos para atingir novas vítimas e aplicar diversos tipos de golpes online. Para isso, exploramos quatro diferentes casos em que detalhamos a veiculação de anúncios fraudulentos que circularam nas plataformas da Meta aproveitando-se do Desenrola Brasil e do Voa Brasil para atingir populações vulneráveis e aplicar golpes.

Como discutido nos capítulos anteriores, a Meta só permite a pesquisadores o acesso a dados retroativos de anúncios classificados como políticos, eleitorais e/ou socialmente relevantes (Meta, s.d.). Por se tratar de programas de governo, entendemos que os anúncios fraudulentos analisados neste capítulo possuem uma natureza explicitamente política e socialmente relevante, devendo, portanto, ser categorizados dessa forma nas plataformas da Meta. Porém, demonstramos que o sistema de

classificação utilizado pela plataforma é falho, facilmente burlado por anunciantes e consequentemente subrepresentado no conjunto de dados que analisamos.

No primeiro estudo, nos debruçamos sobre os anúncios suspeitos sobre o Desenrola Brasil veiculados nas plataformas da Meta. À época da campanha, o NetLab UFRJ divulgou publicamente os resultados da pesquisa, com evidências de que anúncios fraudulentos sobre o Desenrola Brasil circularam entre janeiro e julho de 2023, o que foi amplamente repercutido em diversos veículos de imprensa. O estudo do NetLab UFRJ (Santini et al., 2023a) resultou em uma medida cautelar publicada pela Secretaria Nacional do Consumidor do Ministério da Justiça e Segurança Pública (Senacon/MJSP) em 26 de julho de 2023, que obrigava as plataformas digitais a retirar do ar todos os anúncios fraudulentos referentes ao programa, sob pena de multa diária de R$150 mil em caso descumprimento (Ministério da Justiça e Segurança Pública, 2023). Na mesma data, a Meta alegou que removeria "anúncios enganosos sobre o programa [...] assim que identificados" (Bischoff, 2023).

Diante desses desdobramentos, iniciamos uma segunda busca, que deriva do primeiro estudo e do seu impacto na discussão pública sobre os direitos dos consumidores nas plataformas de redes sociais online. Retomamos a busca por anúncios fraudulentos envolvendo o Desenrola Brasil para verificar se a política de remoção anunciada pela Meta havia de fato sido colocada em prática ou se as fraudes envolvendo o programa do governo havia voltado a circular nas plataformas (Santini et al. 2023b).

Em seguida, reproduzimos os métodos de coleta e análise para investigar se fraudes semelhantes ocorreram utilizando o programa Voa Brasil (Salles et al., 2023) a fim de identificar padrões em anúncios irregulares que se aproveitam de diferentes programas governamentais. Por fim, no último estudo deste capítulo, investigamos se anúncios suspeitos circularam nas plataformas um ano depois da medida cautelar da Senacon (NetLab UFRJ, 2024), a fim de observar o respeito da Meta

em relação à medida cautelar, e a efetividade de seus sistemas de controle de anúncios em suas plataformas de redes sociais.

## 4.2. Procedimentos de pesquisa e linha do tempo dos estudos

Nossa pesquisa sobre publicidade enganosa envolvendo políticas públicas se dividiu em quatro estudos de caso, desenvolvidos em diferentes períodos, entre julho de 2023 e agosto de 2024. Em todos eles, utilizamos a API da Biblioteca de Anúncios da Meta para coletar os dados dos anúncios categorizados como políticos, eleitorais e/ou de relevância social utilizando termos relacionados ao Desenrola Brasil e ao Voa Brasil publicados nos períodos analisados. No caso de anúncios categorizados como políticos, a Meta disponibiliza dados sobre investimento e alcance. Assim, uma vez identificados os anúncios sobre os programas, consolidamos essas informações, bem como o público para os quais eles foram distribuídos.

Ainda na etapa de coleta de dados para as análises, também utilizamos a interface de usuário da Biblioteca para buscar anúncios comerciais sobre os programas governamentais, para verificar eventuais inconsistências de categorização. É importante lembrar que a ferramenta da Meta não permite a coleta automática de dados de anúncios que não são classificados como políticos, eleitorais e/ou de relevância social, que só podem ser vistos enquanto ainda estão sendo exibidos para os usuários. Por isso, coletamos e armazenamos manualmente os anúncios comerciais identificados ao longo dos quatro estudos.

Para a análise qualitativa dos anúncios coletados e a identificação de possíveis irregularidades, incluindo fraudes, golpes e estelionatos contra consumidores nas plataformas, adotamos os mesmos critérios empregados nos estudos apresentados no capítulo 4. Conforme detalhado na Tabela 1 do capítulo 4, a avaliação considerou aspectos formais e de conteúdo, autenticidade das fontes e conformidade com normas legais e diretrizes de autorregulação.

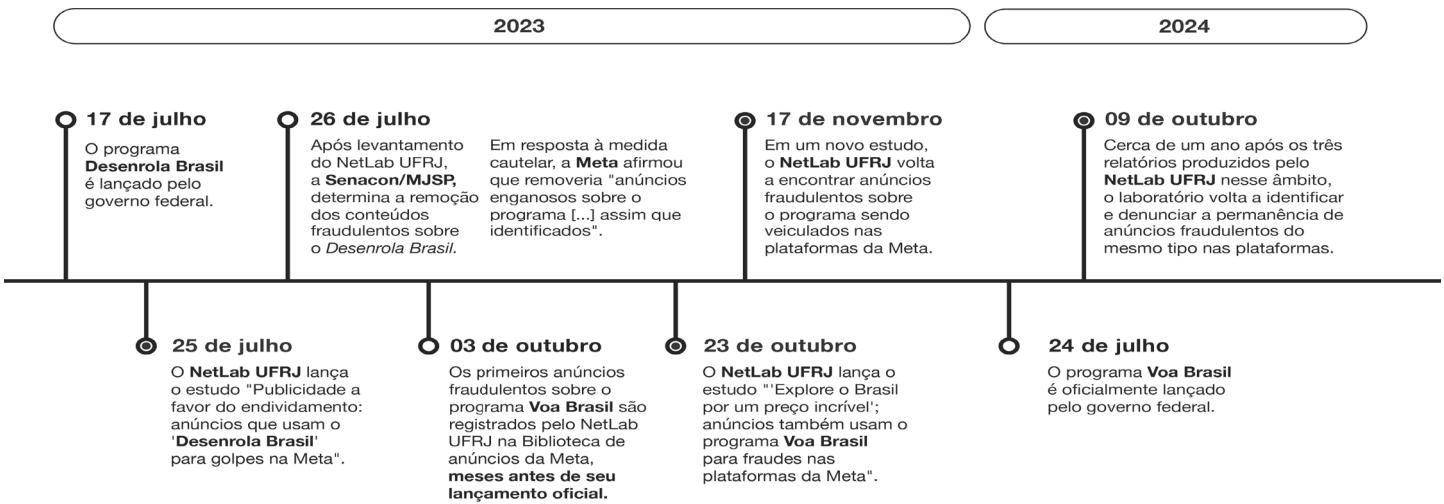
No primeiro estudo, coletamos anúncios com termos relacionados ao programa Desenrola Brasil que circularam nas plataformas da Meta entre os dias 19 e 21 de julho de 2023. Após a medida cautelar e a resposta da Meta alegando que removeria os conteúdos irregulares sobre os programas em suas plataformas, realizamos um segundo estudo, retomando a coleta de anúncios com termos relativos ao Desenrola Brasil publicados nas plataformas da Meta entre 26 de julho e 26 de setembro de 2023.

Após encontrarmos evidências de fraudes nos dois primeiros estudos, repetimos o experimento utilizando termos relativos ao programa Voa Brasil. Para este terceiro estudo, coletamos anúncios políticos e comerciais publicados nas plataformas da Meta entre os dias 18 e 19 de outubro de 2023. Assim como nos estudos sobre o Desenrola Brasil, após a coleta, procuramos por indícios de fraudes e inautenticidade de anunciantes para identificar os anúncios suspeitos na amostra, assim como os links para os quais os usuários foram direcionados.

Por fim, realizamos um quarto estudo cerca de um ano depois da medida cautelar da Senacon. Para esta análise, coletamos dados de anúncios publicados entre 07 de julho e 07 de agosto de 2024 tanto sobre o Desenrola Brasil quanto sobre o Voa Brasil. Nesta busca, não encontramos anúncios políticos relativos aos programas, ou seja, toda a amostra utilizada nesta análise é composta por anúncios comerciais. Deste modo, não foi possível obter dados como valores investidos para impulsionamento de conteúdo nem dados sobre o perfil da audiência dos anúncios. Também utilizamos um algoritmo de Perceptual Hashing e, a partir de uma análise de redes, agrupamos as imagens que apareciam nas peças fraudulentas segundo a similaridade entre elas para identificar os padrões visuais e imagens mais frequentes nos anúncios.

Para visualizar a relação entre os estudos de caso e os acontecimentos mais importantes durante o período das análises, incluindo o lançamento dos programas governamentais, a medida cautelar da Senacon e as respostas da Meta, organizamos esses eventos na linha do tempo, apresentada abaixo.

Figura 1: Eventos chave durante os estudos do Netlab UFRJ sobre anúncios irregulares envolvendo programas governamentais.




Fonte: NetLab UFRJ.

## 4.3. Evidências de anúncios fraudulentos

Nesta seção, apresentamos as principais características dos anúncios fraudulentos que exploram programas governamentais para aplicar fraudes, buscando observar padrões nos anúncios e nas páginas responsáveis por eles, além da resposta da Meta a esse problema. Embora este capítulo busque traçar uma visão geral desse fenômeno, identificando suas dinâmicas, estratégias e recorrências, os detalhes dos estudos e exemplos dos anúncios mencionados foram divulgados publicamente ao longo do projeto e podem ser acessados no site do NetLab UFRJ (Santini et al., 2023a; Santini et al. 2023b; Salles et al., 2023; NetLab UFRJ, 2024).

### *A publicidade a favor do endividamento: Anúncios usam o programa Desenrola Brasil para fraudes nas plataformas da Meta*

Ao buscarmos por anúncios relativos ao Desenrola Brasil entre 19 e 21 de julho de 2023, encontramos 1.048 anúncios suspeitos veiculados por 52 páginas e perfis anunciantes que utilizavam indevidamente o nome e/ou a imagem do programa. Ou seja, em apenas dois dias de coleta, foi possível identificar evidências de como estelionatários exploram as plataformas para promover fraudes em larga escala.

Dos 1.048 anúncios suspeitos que identificamos, 688 foram categorizados como políticos, enquanto 360 não receberam essa classificação. Apenas 153 dos 688 (22%) anúncios fraudulentos categorizados como políticos foram removidos por violarem os Padrões de Publicidade da Meta. Significa dizer que 78% dos anúncios fraudulentos envolvendo os programas sociais do governo circularam nas plataformas de redes sociais sem moderação da Meta. Como só é possível ter acesso aos dados detalhados daqueles anúncios categorizados como políticos, identificamos que os 688 anúncios que receberam esta classificação tiveram até 1,8 milhão de impressões e que até R$87 mil foram investidos para impulsionamento dos conteúdos irregulares. Esses números demonstram o potencial de impacto da maioria dos anúncios dos quais não temos dados de impressões e do número de visualizações pelos usuários.

Nesta amostra, também observamos que as fraudes começaram a circular antes mesmo do lançamento do Desenrola Brasil. Encontramos anúncios veiculados nas plataformas da Meta desde janeiro de 2023, arquivados na Biblioteca de Anúncios e marcados como políticos, eleitorais e/ou de relevância social. Ou seja, apenas os rumores da implementação do programa já foram suficientes para criar interesse na população e oportunidades para estelionatários promoverem as fraudes.

Uma característica comum observada nos anúncios era a apropriação indevida da imagem do governo federal e da Serasa. Identificamos que esse padrão permaneceu mesmo após a publicação de reportagens em jornais e emissoras de TV denunciando esses golpes (Teixeira, 2023; Jornal Nacional, 2023). Uma das páginas anunciantes, chamada "Desenrola Brasil" e que possuía apenas 15 seguidores, foi criada no dia do lançamento do programa e simulava ser um canal oficial. Em poucos dias, a página passou a alegar que o governo federal teria criado um "Feirão Limpa Nome", com negociações "100% online e descontos de até 90% de pagamentos via Pix" para endividados.

Identificamos diversas páginas que utilizavam o nome ou a marca do Desenrola Brasil em seus perfis e fotos. Por exemplo, 10 páginas utilizando esta estratégia veicularam 89 anúncios fraudulentos que não foram categorizados como políticos, eleitorais e/ou de relevância social. O conteúdo desses anúncios incluía o compartilhamento de links que supostamente direcionavam para consultas individuais no WhatsApp e a promoção de sites de orações que prometiam "transformar a vida financeira" de seus seguidores. Veiculados principalmente no Facebook e no Instagram, esses anúncios também reproduziam trechos descontextualizados de telejornais e informações oficiais da Serasa, para aferir legitimidade à fraude pela associação a veículos da mídia tradicional.

Diversos anúncios aproveitaram-se da imagem de veículos jornalísticos, como um anunciante chamado "Informações Diária (sic)" cuja página utilizava uma foto simulando ser o portal G1. As fraudes reproduziam montagens e prints de portais jornalísticos, como o R7 e a Folha de S. Paulo, alegando que o Desenrola Brasil já estaria disponível

e que, ao clicar no link, o usuário encontraria uma solução para "resolver problemas como Nome Sujo e score baixo". Algumas peças também mencionavam uma suposta plataforma para renegociar as dívidas. De fato, embora houvesse a expectativa de lançamento de uma plataforma para realizar essas negociações, o governo só previa implementá-la posteriormente, na segunda fase do programa (Moreno, 2023) – esta fase foi iniciada apenas em outubro de 2023 (Máximo, 2023). A promessa de oportunidades de renegociação ilusórias, como descontos de 99% no pagamento das dívidas, também era frequente nos anúncios.

Na análise do perfil dos anunciantes, identificamos que as páginas responsáveis pelas peças com fraudes possuíam um perfil bastante variado: embora a tendência fosse imitar páginas oficiais do governo federal ou do programa Desenrola Brasil, também encontramos perfis fraudulentos que usavam fotos de terceiros, fingindo ser usuários autênticos – incluindo imagens de influenciadores, pessoas anônimas e falecidas. Por exemplo, um anunciante divulgou o site "Fórum Limpa Nome" ainda no começo de 2023 para se promover por meio das discussões públicas sobre a renegociação de dívidas bancárias. Este perfil promovia anúncios utilizando como foto de perfil uma imagem de Daniele Lyra Nattrodt Barros, vítima de feminicídio no Rio de Janeiro em 2022 (Ventura; Araújo, 2022).

Além da divulgação de conteúdo suspeito, também encontramos anunciantes com outras evidências de comportamento inautêntico, como veiculação de volumes desproporcionais de anúncios em um curto período de tempo e informações desencontradas na descrição do perfil das páginas. Por exemplo, um dos perfis identificados foi responsável por impulsionar 192 anúncios sobre o Desenrola Brasil em um único dia – neste caso específico, é importante frisar que as peças não foram categorizadas como políticas, eleitorais e/ou de relevância social. A página estava registrada na categoria de "artista musical" e utilizava indevidamente a foto de uma influenciadora que atua no Instagram.

Nossa busca também revelou que os mecanismos de autenticação dos anunciantes oferecidos pela Meta apresentavam diversas falhas. En-

contramos anúncios que falsificavam metadados sobre o patrocinador do anúncio, fazendo parecer que o próprio programa Desenrola Brasil (ou seja, o Governo Federal) seria o responsável pelo impulsionamento das peças. Porém, em muitos casos, anúncios políticos não apresentavam nenhuma informação sobre o patrocinador, desrespeitando os próprios termos de uso da plataforma.

Ao analisar os links para os quais os anúncios direcionavam os usuários, encontramos 30 sites suspeitos sobre o Desenrola Brasil. Vários desses sites copiavam o layout dos canais oficiais da Serasa ou exibiam o logotipo da empresa. Nesses sites, o usuário era induzido a achar que estava em um site oficial e estimulado a inserir suas informações pessoais, como o número de CPF, sob o pretexto de realizar uma suposta consulta gratuita de elegibilidade ao programa. Tentando se passar por um canal de comunicação oficial do governo, um dos anunciantes que usava o nome do programa também divulgou um site desenvolvido e hospedado no Google Sites. O endereço contava com vários links de redirecionamento para uma falsa central de atendimento no WhatsApp e exibia os logotipos da Febraban, do Banco Central e do governo federal.

Também encontramos anúncios que redirecionavam para conversas com *chatbots*, simulando um atendimento da Serasa para induzir o usuário a fornecer informações pessoais. Outros sites exibiam vídeos sobre o Desenrola Brasil ou instruções sobre como receber dinheiro para então direcionar o usuário para uma conversa no WhatsApp. Com vídeos longos e sensacionalistas, os golpes buscavam prender a atenção dos usuários, enfatizando ser necessário assistir o conteúdo até o final para "descobrir como solicitar seu dinheiro". Em um dos casos, o site reproduzia uma suposta matéria do Jornal da Noite, da Band, explicando as regras do programa. Ao final dos vídeos, o usuário era levado ao WhatsApp, onde a fraude seria de fato concretizada.

Além disso, nossa análise revelou constantes falhas na moderação da Meta, que levantam dúvidas sobre os critérios adotados pela empresa, a aplicação efetiva de suas próprias políticas e seu real comprome-

timento em coibir esse tipo de golpe. Por um lado, a maior parte dos anúncios mapeados sobre o Desenrola Brasil continuaram no ar, a despeito de seu caráter evidentemente fraudulento e das denúncias públicas, matérias na imprensa e medida cautelar da Senacon/Ministério da Justiça, com apenas uma pequena fração dessa publicidade fraudulenta tendo sido efetivamente moderada pela Meta. Por outro, ironicamente, encontramos anúncios legítimos e sem risco para os consumidores que foram removidos sob a justificativa de violação dos padrões de publicidade da Meta, embora não infringissem os termos da plataforma nem a legislação brasileira. Dois casos emblemáticos foram as remoções dos anúncios com informações legítimas sobre o programa, veiculados pelos deputados federais José Guimarães (PT/CE), líder do governo Lula na Câmara dos Deputados, e do parlamentar Odair Cunha (PT/MG). Um anúncio do jornal local Gazeta da Amazônia, que apenas direcionava para uma matéria que explicava o Desenrola Brasil, também foi derrubado. Apesar dessas falhas, não houve qualquer justificativa por parte da Meta sobre a remoção dessas peças.

Por fim, é importante destacar que a Meta provavelmente utiliza um sistema automatizado para moderação, prática cada vez mais adotada pelas *big tech* (Gorwa; Binns; Katzenbach, 2020). Além disso, mesmo quando há participação humana, a moderação de conteúdo ocorre em condições precárias, levantando questionamentos sobre sua eficácia e as condições de trabalho dos profissionais envolvidos (Roberts, 2016). Nosso estudo demonstra as falhas desse sistema, que evidenciam um modelo de moderação questionável, sem o devido investimento por parte de uma das empresas de tecnologia mais poderosas e ricas do mundo.

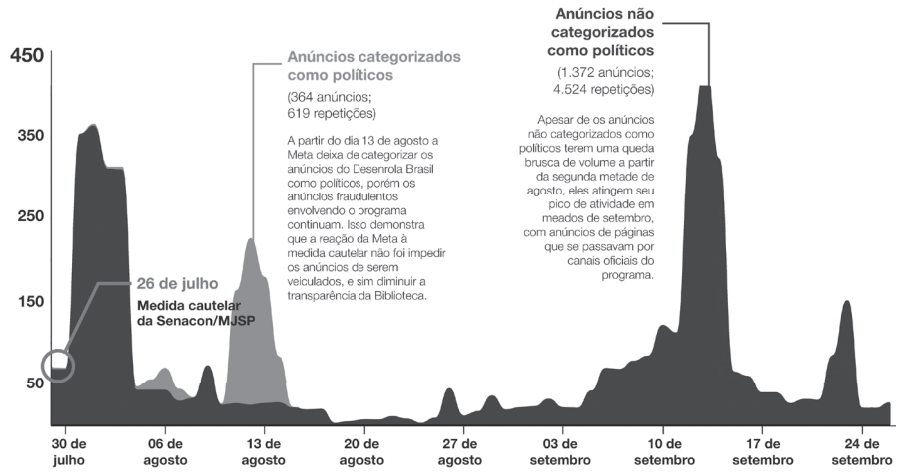
### *'O fim dos seus problemas': Anúncios fraudulentos sobre o Desenrola Brasil seguiram circulando na Meta após a medida cautelar*

Após a medida cautelar da Senacon em 26 de julho de 2023, identificamos 1.817 anúncios fraudulentos que circularam nas plataformas da Meta nos dois meses seguintes à determinação do governo. Nesta amostra, 383 foram categorizados como anúncios políticos, enquanto 1.434 foram veiculados sem qualquer categorização – nem pelos anun-

ciantes, nem pela Meta. Somente 89 dos 383 (23%) dos anúncios fraudulentos categorizados como políticos foram removidos por violarem os Padrões de Publicidade da Meta. Os outros 77% não foram moderados pela plataforma. Além disso, apenas 2 dos 383 (0,5%) dos anúncios políticos irregulares foram categorizados como tal por seus anunciantes – 99,5% foram recategorizados pela Meta.

A partir do dia 13 de agosto, a Meta deixou de categorizar os anúncios sobre o Desenrola Brasil como políticos – os anúncios, porém, continuaram circulando na plataforma. Apesar de ter havido uma queda brusca de volume dos anúncios não categorizados como políticos a partir da segunda metade de agosto, eles atingiram um expressivo pico de atividade em meados de setembro. Como ilustrado na Figura 2, anúncios fraudulentos relacionados ao Desenrola Brasil estiveram ativos todos os dias ao longo dos dois meses analisados; no pico de atividade, cerca de 450 anúncios estiveram ativos em um único dia.

Figura 2: Evolução de anúncios ativos diariamente, nos dois meses seguintes à medida cautelar da Senacon, em 26 de julho de 2023.




Fonte: Gráfico elaborado pelo NetLab UFRJ a partir dos dados obtidos na Biblioteca de Anúncios da Meta.

Assim como no estudo realizado antes da medida cautelar da Senacon, os anúncios faziam promessas irreais e não previstas pelo Desenrola Brasil, como descontos de até 99% das dívidas, parcelas de até

80x sem juros e CPF limpo independentemente do valor da dívida. Os anúncios apareciam frequentemente acompanhados de falsos depoimentos de pessoas que alegavam ter conseguido se livrar das dívidas por meio dos métodos divulgados.

Nossa busca também identificou que 115 páginas e perfis anunciantes foram responsáveis pela veiculação das peças. A maior parte deles não apresentava histórico de atividade orgânica regular nas plataformas da Meta, tendo sido exclusivamente criadas para veicular anúncios. O baixo número de seguidores também foi um padrão recorrente: dos 80 anunciantes com até dez seguidores, 46 não eram seguidos por ninguém. A baixa atividade e poucos seguidores das páginas são características que deveriam facilitar os mecanismos de detecção automática de comportamento inautêntico por parte da Meta. No entanto, a recorrência desse padrão entre os anunciantes sugere que a plataforma não implementa medidas eficazes para coibir esse tipo de prática, permitindo que páginas com indícios de comportamento inautêntico continuem a veicular anúncios. Apesar deste padrão predominar, também identificamos perfis suspeitos com milhares de seguidores; em diversos casos, os seguidores destes anunciantes também apresentavam indícios de comportamento inautêntico.

Os perfis dos anunciantes fraudulentos apresentavam evidências de diversas irregularidades. Uma prática recorrente foi a apropriação da identidade visual da Serasa e da marca do governo: 294 anúncios foram veiculados por anunciantes que utilizavam o nome e a identidade visual do próprio Desenrola Brasil. Símbolos políticos e o uso da imagem de figuras públicas demonstram que não houve escolhas partidárias para a disseminação das fraudes: os anúncios apresentavam, por exemplo, o logotipo do governo de Jair Bolsonaro associado à imagem de Lula e ao nome do programa. Além disso, trechos de participações de lideranças como o presidente Lula e o ministro da Fazenda Fernando Haddad em programas de entrevistas e podcasts foram frequentemente editados e descontextualizados para convencer os usuários de que se tratava de uma publicidade institucional. Os políticos apareciam em vídeos fa-

zendo discursos, assinando documentos, dando entrevistas e em coberturas da imprensa. O deputado federal de oposição Celso Russomanno (Republicanos/SP) também teve sua imagem vinculada a anúncios irregulares, especialmente em uma série de anúncios que se passavam por notícias da GloboNews. Ainda na análise das páginas anunciantes, também observamos comentários feitos por vítimas denunciando os golpes financeiros nessas páginas, que em alguns casos já haviam sido deletadas ou suspensas.

Outro padrão que permaneceu nos anúncios após a medida cautelar foi o uso da linguagem e da imagem de veículos jornalísticos. Um dos anunciantes se passava por um canal oficial da GloboNews em sua foto de perfil e de capa. Neste caso, os estelionatários também enriqueciam o perfil com avaliações falsas e comentários elogiosos, como se a página de fato pertencesse ao Grupo Globo. Outro perfil fingia ser um canal do G1, chegando a apresentar "g1noticias@gmail.com" como endereço oficial. Em outros exemplos da amostra, páginas veiculando conteúdo irregular adotavam tom jornalístico mesmo que não tivessem qualquer ligação com veículos de comunicação.

Também encontramos evidências de publicidade veiculadas por anunciantes cujas fotos de perfil haviam sido geradas por Inteligência Artificial. Em um desses casos, o perfil apresentava comentários suspeitos em diversos idiomas e muitos seguidores, embora tivesse apenas duas publicações em seu histórico. Outro indício de comportamento suspeito era o padrão de publicidade dos perfis: encontramos anunciantes com anúncios ativos sobre outros produtos, como medicamentos sem registro na Anvisa para dores articulares, e anúncios promovendo bolos confeitados, mas que levavam os usuários para sites fraudulentos.

Os anúncios também apresentavam irregularidades como roubo de identidade de terceiros. Uma das páginas utilizava a imagem de Mindy Gale, uma influenciadora com mais de 100 mil seguidores no Instagram; outra utilizava uma foto de Leila Luz, executiva com ampla experiência em multinacionais. Os anunciantes frequentemente apresentavam nomes seguidos de "Dr" ou "Consultora". Em um caso, o

perfil identificado como “Letícia - Consultora” exibia a foto de uma funcionária da Serasa, cujo crachá mostrava o nome “Débora”, distinto do nome usado na tentativa de fraude.

Além disso, a quantidade de anúncios veiculados por cada anunciante era, no geral, muito baixa: 39 anunciantes (34%) impulsionaram apenas um anúncio. Assim, poucos anunciantes veiculavam centenas de anúncios, que costumavam ser idênticos entre si em termos de conteúdo, provavelmente segmentando públicos diferentes. No entanto, não foi possível verificar a segmentação desses anúncios devido à falta de informações detalhadas sobre essas peças na Biblioteca de Anúncios da Meta.

Já a análise das informações dos 383 anúncios categorizados como políticos revelou que até R$39,1 mil foram gastos para impulsionamento das peças, que tiveram até 508,6 mil impressões. Embora a Meta limite o acesso às informações completas sobre os critérios de segmentação utilizados pelos anunciantes, foi possível identificar o uso de segmentação por região, gênero e faixa etária, o que indica tentativas de direcionamento estratégico dos golpes financeiros a públicos específicos. Ainda assim, cerca de um terço desses anúncios não disponibilizava dados de segmentação, o que implica que as estimativas de impressões e gastos por estado que encontramos subdimensionam o problema. Considerando os dados disponíveis, observou-se que os estados com maior número de impressões tendem a coincidir com aqueles que possuem o maior número de habitantes — como São Paulo e Minas Gerais. Uma das exceções foi a Bahia: mesmo com população menor e menos negociações na Serasa em agosto de 2023 do que o Rio de Janeiro (Serasa, 2023), foi o terceiro estado com maior número de impressões.

Já na distribuição por faixa etária, os dados encontrados mostram que o conteúdo foi mais visto por pessoas de 35 a 44 anos (36,1% das impressões) e entre 45 e 54 anos (26,5%) – isso reforça que os anúncios atingiram o principal grupo populacional de inadimplentes no Brasil (Serasa, 2023). Embora a Meta restrinja o acesso às informações sobre segmentação definidas pelo anunciante, na amostra que analisamos

foi possível detectar que 13 das 19 páginas cujas informações estavam disponíveis para anúncios impulsionados no período, responsáveis por 47,3% dos anúncios classificados como políticos, selecionaram o público de 30 anos ou mais para ser impactado pelos anúncios. Por outro lado, a segmentação por gênero foi pouco explorada por anunciantes: os anúncios são apenas um pouco mais vistos por homens (53,8%) que por mulheres (46,2%).

Por fim, exploramos os sites para os quais os anúncios direcionavam os usuários. Encontramos 46 sites nos anúncios sobre o Desenrola Brasil: 1.739 (95,7%) anúncios levavam para 45 sites, enquanto 77 (4,2%) direcionavam para o WhatsApp – 3 (0,1%) não tinham nenhum link associado[23]. Diversos anúncios levavam os usuários para a plataforma *Kiwify*, utilizada para quem vende produtos e serviços digitais. Os estelionatários utilizavam a plataforma para vender supostos métodos garantindo "descontos imperdíveis" na renegociação de dívidas, "aprovação de crédito" e "controle de finanças", muitas vezes passando-se pelo governo federal. A plataforma propunha diferentes formas de pagamento, como cartões de crédito, boleto e Pix. A aprovação de crédito era oferecida pelo site como um "adicional" à compra no site, que variava entre R$29 e R$50.

Nos sites analisados, também encontramos *chatbots* que simulavam a interface do WhatsApp, com falsas conversas que buscavam se passar por órgãos oficiais e a Serasa. Por exemplo, estelionatários forjaram conversas com supostas atendentes virtuais da Serasa, apresentando o espaço como "o canal oficial da Serasa" e alegando garantir a segurança dos dados do cliente. A ideia era que o usuário pensasse estar em um lugar no qual seria possível renegociar suas dívidas. Para prosseguir na "renegociação", os usuários deveriam fornecer informações como nome, número de celular, CPF e e-mail.

Outra prática comum para a aplicação dos golpes foi o uso de sites que imitavam o layout dos canais oficiais da Serasa, a identidade oficial

[23] A soma das porcentagens ultrapassa 100% porque 33 anúncios (1,8%) redirecionavam para mais de um site. Um deles, por exemplo, levava tanto para um site quanto para o WhatsApp.

do governo federal e do Desenrola Brasil para convencer os consumidores. Nesses casos, os usuários eram convencidos a inserir suas informações pessoais, incluindo o número de CPF, sob o pretexto de realizar uma suposta consulta gratuita de elegibilidade ao programa.

As evidências encontradas neste estudo não apenas confirmam os padrões identificados na primeira coleta, que permaneceram essencialmente os mesmos, mas também demonstram que esses indícios, facilmente identificáveis, não foram suficientes para que a Meta aplicasse suas próprias políticas de moderação. Além disso, a continuidade dos golpes mesmo após a medida cautelar da Senacon indica que a empresa não apenas falha em coibir essas práticas, mas também desafia a soberania das leis locais ao permitir a circulação persistente de conteúdos nocivos em seu ecossistema de publicidade, mesmo que isso represente um risco à integridade de seus usuários.

### *'Explore o Brasil por um preço incrível!': Anúncios usam o programa Voa Brasil para promover fraudes nas plataformas da Meta*

Após a coleta e análise dos dados sobre anúncios relacionados ao Voa Brasil entre os dias 18 e 19 de outubro de 2023, identificamos 622 anúncios suspeitos de promover fraudes nas plataformas da Meta. Destes, 205 foram categorizados como políticos: para o impulsionamento dessas peças, os anunciantes investiram cerca de R$30,9 mil, e seus conteúdos foram vistos cerca de 522,7 mil vezes. Apenas 4 dos 205 (2%) anúncios fraudulentos categorizados como políticos foram removidos por violarem os Padrões de Publicidade da Meta. Outros 417 anúncios sobre o Voa Brasil circularam nas plataformas da Meta sem serem devidamente categorizados como políticos.

Para convencer os usuários de que de fato se tratava de uma publicidade do governo, uma estratégia empregada em diversos anúncios era a reprodução de imagens do ex-ministro dos Portos e Aeroportos Márcio França. Eles prometiam uma "aventura única" por meio do Voa Brasil, alegando que já era possível comprar passagens para todo o país por R$200. Para apressar as vítimas e instigá-las a prosseguir na su-

posta oferta, os anunciantes usavam o argumento de que o prazo do programa estaria chegando ao fim e que haveria um número limitado de vagas restantes. Os anúncios também direcionavam os usuários para sites que prometiam descontos em hotéis e sorteios semanais de viagens internacionais.

Os estelionatários impulsionaram anúncios com trechos descontextualizados de vídeos de influenciadores do TikTok e do YouTube, apropriando-se indevidamente de sua imagem e de seus conteúdos, explorando sua credibilidade para enganar os usuários. Por exemplo, alguns anúncios usavam vídeos do influenciador Estevam Pelo Mundo, que tem mais de 2 milhões de inscritos em seu canal no YouTube.

Os anúncios que compõem a amostra deste estudo foram impulsionados por 52 páginas e perfis. Os autores das peças frequentemente utilizavam o nome e a imagem de instituições governamentais, como no caso do perfil "Programa VoaBrasil". Versões sutilmente modificadas do nome do programa ou de órgãos do governo apareceram em 41,4% dos anúncios fraudulentos. O uso indevido da imagem de empresas também se reproduziu no caso de veículos jornalísticos: diversos perfis circulavam simulando ser portais como CNN e G1. Uma das páginas, por exemplo, chamava-se "CNN Notícias" fingindo ser um canal oficial da emissora. No entanto, a página havia sido criada dias antes da veiculação do anúncio e não tinha nenhuma atividade orgânica.

Para enganar os usuários, anunciantes recortavam, adicionavam narrações e descontextualizavam trechos de reportagens de veículos de notícias como CNN e Globo, além de se passarem por publicações de páginas online como a Choquei. Um dos anúncios apresentava um trecho de telejornal da CNN Brasil em que a voz do jornalista Rafael Colombo era manipulada com uma nova dublagem falsa, que alegava ser necessário que os cidadãos pagassem uma taxa para acessar o suposto aplicativo do Voa Brasil. As páginas responsáveis pelos anúncios recorriam a múltiplas estratégias enganosas sem se preocuparem com coerência entre os conteúdos e as marcas utilizadas, combinando identidades visuais e nomes de veículos reconhecidos. Por exemplo, uma página

chamada "G1 Notícias" fingia ser um canal oficial do portal G1. No entanto, ao clicarem no anúncio, os usuários eram redirecionados para sites que emulavam o portal oficial da CNN Brasil e exibiam reportagens manipuladas com informações falsas sobre o programa, para então dar sequência às etapas da fraude. Ao analisarmos as informações sobre esses anúncios, identificamos que as peças foram segmentadas sobretudo para usuários com mais de 45 anos.

Os anunciantes muitas vezes fingiam ser indivíduos comuns, utilizando foto e nomes genéricos. Eles se definiam na plataforma como "criadores de conteúdo", mas não apresentavam publicações nem seguidores. Para enganar as vítimas, esses anunciantes ocultavam os comentários das poucas publicações feitas, muito provavelmente para impedir que novos usuários tivessem acesso a denúncias das vítimas das fraudes.

Enfim, na análise sobre os links para os quais os usuários eram direcionados ao clicarem nos anúncios, encontramos 43 sites. Uma estratégia recorrente era o redirecionamento dos usuários para sites imitando o layout de portais de notícias, principalmente o da CNN Brasil, com vídeos de reportagens manipuladas. Para consumar o golpe, estas páginas contavam com vários links de redirecionamento para canais que se passavam por sites oficiais do governo. Assim como nas análises sobre o Desenrola Brasil, outro padrão destes sites era o uso de *chatbots* customizados que simulavam canais de contato e cadastro oficiais do Voa Brasil para convencer os usuários a enviar informações pessoais.

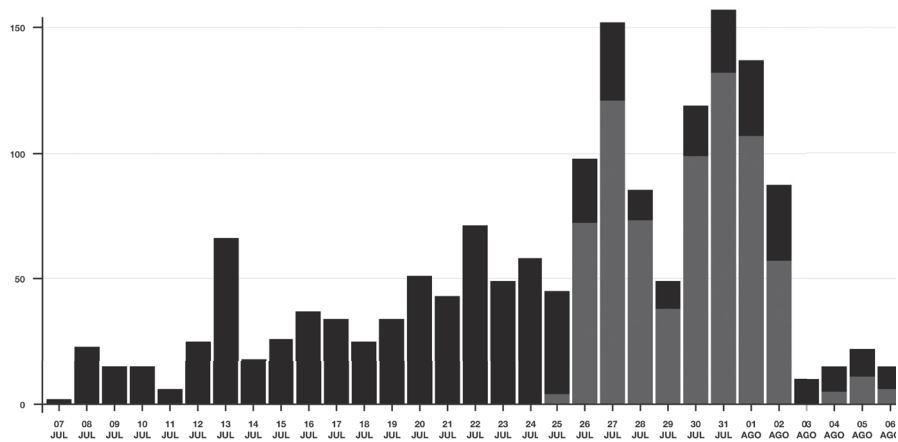
### *Golpes e falhas sistêmicas nas plataformas da Meta: Anúncios fraudulentos sobre o Desenrola Brasil e o Voa Brasil persistem após um ano de medida cautelar*

Neste estudo, retomamos a coleta de anúncios sobre o Desenrola Brasil e o Voa Brasil em 2024, um ano após a medida cautelar da Senacon. Ao todo, nessa nova busca, identificamos 643 novos anúncios fraudulentos utilizando as mesmas táticas para enganar usuários que já haviam sido relatadas em 2023: a exploração e a manipulação de trechos de noticiários de televisão, a falsificação de portais de notícias

e o direcionamento dos usuários para sites falsos de empresas como a Serasa, para citar alguns exemplos. Entre os anúncios, 55% eram sobre o Voa Brasil, enquanto 45% exploravam o Desenrola Brasil para promover fraudes.

Enquanto os anúncios sobre o Desenrola Brasil circularam ao longo de todo o período analisado, começando dia 07 de julho de 2024, o Voa Brasil foi mencionado a partir do dia 25 de julho, logo após seu lançamento. Somados, os dois programas tiveram picos de veiculação de anúncios enganosos na última semana de julho, chegando a ter 157 anúncios exibidos simultaneamente nas plataformas (Figura 3). O volume de anúncios que de fato circularam no período tende a ser maior do que o representado no gráfico abaixo, já que a Biblioteca de Anúncios da Meta não permite o acesso a anúncios não classificados como políticos, eleitorais e/ou de relevância social depois que estes deixam de circular.

Figura 3: Evolução de anúncios ativos diariamente nos meses seguintes à medida cautelar, entre os dias 07 de julho e 06 de agosto de 2024.




Fonte: Gráfico elaborado pelo NetLab UFRJ a partir dos dados obtidos na Biblioteca de Anúncios da Meta.

As publicações foram impulsionadas por 42 páginas e perfis anunciantes, dos quais 30 não receberam qualquer tipo de moderação da Meta. O perfil que mais veiculou anúncios foi responsável, sozinho,

pelo impulsionamento de 142 peças sobre o Voa Brasil. Dentre os dez perfis que mais veicularam anúncios, apenas três haviam sido excluídos pela Meta até nossa análise. A remoção parcial de anunciantes fraudulentos evidencia as inconsistências na moderação de anúncios praticada pela empresa. Outra característica que se repetiu em relação aos estudos anteriores foi a criação das páginas logo antes da veiculação dos anúncios. Algumas dessas páginas eram administradas por usuários na Indonésia, França e Portugal. Dentre estes 42 anunciantes identificados com indícios de comportamento inautêntico, alguns perfis também violavam o direito de uso de marcas registradas, como os logotipos da Serasa e do Nubank. Uma das páginas utilizava o nome "Alfinetei", fingindo ser um conhecido portal de notícias sobre celebridades.

A análise baseada no agrupamento das imagens dos anúncios mostrou que a maior parte das publicações fraudulentas relacionadas ao Desenrola Brasil usavam as marcas de outros órgãos públicos como Procon e Banco Central, e privados como Serasa, SPC e BoaVista. As imagens também incluíam fotos de filas de banco e falsos prints de renegociações de dívidas em aplicativos financeiros. A maioria das peças utilizava fotos genéricas de bancos de imagem, com destaque para casais idosos em golpes do Voa Brasil, uma vez que pensionistas do INSS teriam direito a usufruí-lo.

Para garantir credibilidade às promessas de acesso aos programas, 83,2% dos anúncios irregulares utilizavam a identidade visual e os logotipos do governo federal. As peças sobre o Desenrola Brasil usavam indevidamente artes oficiais feitas pela administração federal para promover o programa, como uma imagem de divulgação da Caixa Econômica Federal. Além de prometerem "limpar o nome" de pessoas endividadas, as peças incitavam os usuários a realizar supostas consultas sobre sua situação financeira, verificar a elegibilidade no Voa Brasil ou comprar passagens aéreas, explorando diferentes estratégias para levá-los a interagir com os anúncios fraudulentos.

Outro padrão que se repetiu neste estudo foi o uso de notícias forjadas para simular o portal G1, presente em 51 anúncios (7,9%). A

marca da TV Globo também foi explorada em 40 peças (6,2%) sobre o Voa Brasil, com destaque para William Bonner que apareceu em 20 anúncios simulando chamadas do Jornal Nacional e da GloboNews. As montagens combinavam chamadas sensacionalistas e dados falsos sobre o programa, como informações sobre um suposto benefício a ser pago aos cidadãos brasileiros para quitar dívidas. Em 11 anúncios (1,7%) sobre o Desenrola Brasil, as peças combinavam vídeos manipulados e descontextualizados, misturando notícias da Rede Record, do G1 e peças do governo federal.

Além disso, 22 anúncios (3,4%) exploravam a marca e o conteúdo da Rede Record, com trechos de notícias retiradas de contexto e conteúdos sensacionalistas sobre os programas federais. O Jornal da Record apareceu em 11 peças (1,7%). Já a imagem do portal R7, do Cidade Alerta e de Luiz Bacci apareceram em um anúncio sobre o Voa Brasil: além de o nome do apresentador ter sido grafado incorretamente, o conteúdo do anúncio era visivelmente manipulado por ferramentas de Inteligência Artificial. Outros veículos de mídia e personalidades da imprensa também foram retratados nas fraudes analisadas, como William Waack, da CNN Brasil. Utilizando imagens de divulgação oficial do governo, 17 anúncios (2,6%) se passavam por publicações feitas pela página "Choquei". Os textos alegavam que o governo teria liberado três meses de seguro-desemprego no valor de R$1.412,90 por mês para que cidadãos quitassem dívidas. As peças direcionavam os usuários para sites externos que solicitavam dados pessoais para suposta liberação do benefício.

Imagens de celebridades também foram exploradas nas fraudes, com destaque para Rodrigo Faro, presente em 26 anúncios (4,2%). Anúncios sobre o Desenrola Brasil usaram a imagem do apresentador para prometer 99% de desconto na renegociação de dívidas. Em outro caso, um vídeo de Celso Portiolli manipulado com Inteligência Artificial induzia aposentados a clicar em um link e supostamente consultar promoções de voos usando seus CPFs.

Além disso, 600 anúncios (93,3%) direcionavam os usuários para 28 sites fraudulentos, e 42 (6,5%) continham links para o WhatsApp[24]. Desses 28 sites, 18 estavam ativos no momento da análise; assim, pudemos identificar os dados de acesso de seis deles, além do registro de quatro. Entre os meses de maio e julho, os seis sites tiveram 37.593 acessos, um número que dá dimensão do alcance e do impacto da circulação desses anúncios, indicando que milhares de pessoas podem ter sido expostas aos golpes. Apenas um site tem registro no Brasil, com dados sobre o responsável, como e-mail e CPF. Outros três sites foram registrados nos Estados Unidos e na Islândia. A falta de dados e verificação facilita a atuação de esquemas fraudulentos, dificultando a responsabilização dos envolvidos.

Além disso, dois sites promovidos em 110 anúncios tinham *chatbots* que solicitavam ao usuário o número de CPF. Ambos estavam relacionados a fraudes e golpes sobre o Desenrola Brasil e faziam referência à *fintech* Acordo Certo, com imagens do logotipo ou da operação "Zero Dívida" da empresa. A maior parte das páginas de *chatbots* identificadas utilizavam o serviço do site *Typebot*.

Uma análise específica dos 354 anúncios analisados relacionados ao programa Voa Brasil revelou que 328 peças (92,6%) direcionavam os usuários para falsos portais contendo o nome do programa. Esses anúncios levavam a 10 sites distintos, dos quais apenas dois estavam ativos no momento da análise. Outros 68 anúncios redirecionavam os usuários para sites que também se apropriavam do logotipo do programa. As páginas simulavam um canal oficial do governo federal e anunciavam que "explorar o Brasil não é um sonho distante". Mimetizando o login do portal gov.br, os sites buscavam roubar dados de usuários por meio de uma técnica conhecida como *phishing*. Essa estratégia explora vulnerabilidades humanas ao manipular mensagens e imitar fontes confiáveis para confundir as vítimas e obter indevidamente suas informações, como senhas, dados bancários, números de cartões de crédito etc. (Lastdrager, 2014; Konji; Iraqi; Jones, 2013).

[24] Apenas um anúncio (0,2%) não apresentava nenhum link associado.

Como evidenciado pelos outros casos aqui apresentados, esse tipo de técnica tem sido comum em golpes promovidos nas plataformas da Meta. No caso de anúncios sobre o Desenrola Brasil, dois sites imitavam canais oficiais da Serasa, reproduzindo elementos da identidade visual da empresa, como cores, imagens e logotipos. Um deles, veiculado em oito anúncios, trazia um botão "Ver ofertas" que redirecionava o usuário para um falso atendimento online. Outro site que mimetizava o canal da Serasa era quase idêntico ao site oficial. Em 17 anúncios, identificamos duas páginas que utilizavam o termo "gov" na URL e o nome do programa Desenrola Brasil. Embora as páginas apresentassem conteúdos quase idênticos, uma delas afirmava ser possível quitar as dívidas por R$47,90, enquanto a outra trazia o mesmo conteúdo, mas com um valor diferente, de R$97,90.

Presente em seis anúncios sobre o Voa Brasil, um dos sites fingia ser a página dos Correios e solicitava que os usuários informassem o CPF para uma suposta verificação de dados. Outros seis anúncios divulgavam um blog com postagens sobre o Voa Brasil, com textos com indícios de terem sido gerados por Inteligência Artificial, prática cada vez mais comum em sites voltados à maximização de receita proveniente de anúncios (Brewster; Fischman; Xu, 2023; Lebow, 2024). Outro anúncio redirecionava para um site com conteúdos em português, inglês e espanhol, promovendo o Voa Brasil usando a imagem oficial do governo. Em outra seção, o mesmo site oferecia produtos suspeitos para emagrecimento, dores e diabetes.

Um site promovido em 50 anúncios sobre o programa Desenrola Brasil incluía um vídeo com trechos de uma matéria do Jornal Nacional sobre acesso a crédito. O vídeo promovia a venda de um suposto guia chamado "O Máximo Score" por R$19,90, associando-o a uma pessoa identificada como Letícia Monteiro, apresentada como analista de crédito de um banco. O site tratava o guia como um "segredo revelado" para aumentar o score de crédito e exibia comentários e testemunhos de supostos clientes satisfeitos. Todos os anúncios foram publicados pelo anunciante "Central Digital Empreendimentos", que usava o logotipo

da Serasa como imagem de perfil e possuía poucas publicações orgânicas.

Uma tática comum era a combinação de informações verdadeiras com dados falsos, explorando a confusão para tornar o golpe mais convincente ou inserindo detalhes enganosos que passavam despercebidos pelos usuários. Por exemplo, 11 anúncios redirecionavam para uma notícia falsa em um site que mimetiza o portal G1. O subtítulo apresentava dados verdadeiros sobre o Desenrola Brasil, como o total de beneficiários e o valor bruto renegociado. O título, porém, alegava que a inscrição no programa garantiria auxílio de R$1.412,90 durante três meses para as pessoas "se organizarem financeiramente". O título utilizava destaques em vermelho e caixa-alta no início de cada palavra, diferente do padrão do G1. A maior parte da falsa matéria copiava uma publicação do Ministério da Fazenda, exibia uma foto do presidente Lula e incluía um passo a passo para realizar o cadastro e receber o "auxílio". Ao clicar em "Eu quero o auxílio", o usuário era redirecionado para uma página que solicitava dados como nome, CPF e chave Pix.

## 4.4. Publicidade fraudulenta sobre políticas públicas e o papel das plataformas digitais

Nossos resultados demonstram que os anúncios explorando programas governamentais seguiram padrões bem estabelecidos, que podem ser observados no perfil dos anunciantes, na estética das peças veiculadas e nas estratégias empregadas para captar a atenção e convencer os consumidores, facilitando a aplicação de golpes financeiros e o roubo de dados pessoais. De modo geral, os perfis anunciantes contavam com pouca ou nenhuma atividade orgânica, não eram seguidos por outros usuários e apresentavam fotos genéricas, muitas vezes oriundas de bancos de imagens públicas, apropriadas de marcas e instituições consolidadas, ou inteiramente geradas por IA. Ou seja, de alguma maneira são ainda toscos, e seus padrões poderiam ter sido facilmente detectados pela Meta.

Os anúncios exploravam argumentos comerciais por meio de diversas táticas de desinformação, evidenciando a ausência de compromisso com a coerência e a tentativa de criar deliberadamente um caos informacional a respeito das políticas públicas. Essa estratégia ocorreu por meio da combinação de informações de diferentes canais de mídia, na associação arbitrária com partidos políticos de todo o espectro político e na tentativa de atingir diversos públicos ao usar imagens de representantes de diversas entidades públicas e privadas.

Também observamos padrões recorrentes nas estratégias para persuadir os usuários e aplicar as fraudes. Entre os comportamentos que se repetiram nestes quatro estudos sobre os programas governamentais, destaca-se a tendência de exagerar os supostos benefícios oferecidos pelo governo para chamar a atenção do público com falsas promessas e soluções fáceis e rápidas sem garantia de retorno e não previstas pelos programas, como zerar dívidas com rapidez e sob o pagamento do menor valor possível. De fato, o nome e a identidade visual do governo brasileiro foram os principais elementos usados em anúncios irregulares e com desinformação. Com falsas informações sobre os programas, incluindo benefícios inexistentes, além de falsos formulários de cadastro e contatos fraudulentos, muitos dos anúncios encontrados carregavam a identidade visual oficial do governo federal, do Desenrola Brasil e do Voa Brasil como forma de aparentar credibilidade.

Além disso, os anúncios exploravam indevidamente a imagem e o nome de entidades do governo federal, figuras políticas e empresas privadas, especialmente a Serasa. Observamos que estelionatários instrumentalizaram o jornalismo para simular que os anúncios irregulares divulgavam informação confiável. Os anúncios veiculados por anunciantes que se passavam por veículos jornalísticos e canais de comunicação oficial do governo tinham potencial para confundir as vítimas desatentas a detalhes como quantidade de seguidores e estilo de conteúdo. Com o intuito de passar credibilidade ao conteúdo impulsionado, os anunciantes fraudulentos comumente empregavam tom jornalístico e apostavam na apropriação da imagem de veículos da mídia profissional,

estratégia já reportada em estudos anteriores sobre anúncios com fraudes e desinformação (Rao, 2022).

Além de se aproveitar da reputação de atores como figuras políticas, influenciadores, canais oficiais e veículos jornalísticos, os anúncios exploraram diferentes práticas para induzir os usuários ao erro, diversas delas amplamente trabalhadas por pesquisadores em estudos sobre fraudes online (Button *et al.*, 2014). Destacaram-se, por exemplo, apelos de urgência, com ênfase em linguagens remetendo a ofertas ou benefícios "exclusivos" com prazo limitado para resposta, recompensas "imediatas", descontos desproporcionais e apelos emocionais, a fim de tornar os indivíduos mais abertos à persuasão. Nos estudos de caso apresentados neste capítulo, esses apelos ainda exploraram vulnerabilidades socioeconômicas para atrair usuários inadimplentes ou de baixa renda, prometendo soluções ao endividamento ou a possibilidade de viajar de avião a baixo custo. Outra característica observada foi a indução ao comprometimento comportamental (Button *et al.*, 2014), isto é, a fraude ocorria por meio de uma série de etapas que pouco a pouco engajavam os usuários e os convenciam a avançar até a fase final do golpe.

Portanto, ao explorar a imagem do governo, suas iniciativas e instrumentalizar empresas, figuras públicas e padrões do jornalismo, as fraudes podem impactar negativamente a reputação de múltiplas instituições e a própria ideia de política pública em si, com consequências para as pessoas individualmente, mas também para as políticas que visam o fortalecimento da ideia de coletividade na sociedade. Primeiramente, esse tipo de publicidade enganosa expõe os consumidores a diversos riscos e agrava a situação de vulnerabilidade de grupos socioeconômicos afetados. Segundo, a disseminação de golpes e esquemas fraudulentos online compromete a implementação de políticas públicas que visam beneficiar a população. Nos dois programas governamentais que analisamos neste capítulo, os anúncios circularam antes mesmo da data de lançamento dos projetos, demonstrando que o ecossistema de desinformação e fraude online prejudica as iniciativas antes mesmo de serem postas em prática. Em terceiro lugar, considerando que as políti-

cas públicas são fragilizadas pelos conteúdos irregulares e que a imagem do governo e de atores políticos é empregada para promover essas fraudes, a credibilidade das próprias instituições governamentais é colocada em risco. Por fim, a confiança na mídia tradicional é fragilizada na medida em que estelionatários apropriam-se dos conteúdos, nomes e marcas de alguns dos principais veículos de comunicação para enganar usuários, que podem acreditar serem vítimas da própria imprensa, degradando sua imagem diante do público.

Nossos dados também mostraram que criminosos se serviram das ferramentas de segmentação da Meta para direcionar anúncios para pessoas consideradas mais suscetíveis a acreditar no conteúdo das peças. Por um lado, nos conteúdos sobre o Desenrola Brasil, usuários entre 35 a 54 anos foram especialmente atingidos. Desta forma, o público impactado com os anúncios é de fato o mais propenso a apresentar maiores taxas de endividamento, que pertence às faixas etárias de 41 a 60 anos e de 26 a 40 anos, segundo a Serasa (2023). Por outro lado, a população idosa foi o principal público-alvo dos anúncios sobre o Voa Brasil. Retratando pessoas idosas com frequência, os conteúdos dos anúncios sugeriram que a segmentação dos criminosos mirava em pessoas dessa faixa etária, que costumam ser vulneráveis a golpes online. Apesar do direcionamento de golpes e furtos a idosos ser um agravante reconhecido no Código Penal brasileiro (Brasil, 1940), e esse grupo social ser considerado juridicamente hipossuficiente (STJ, 2015), o sistema de publicidade da Meta não moderou anúncios fraudulentos que buscavam atingir usuários dessa faixa etária. Em ambos os casos, trata-se de grupos sociais que podem ser altamente impactados por prejuízos financeiros de qualquer porte. Ao permitir que essas peças circulem utilizando técnicas de microssegmentação, o sistema de publicidade da Meta e suas práticas inconsistentes de moderação contribuem para que criminosos atinjam esse perfil de usuário.

De fato, a persistência de anúncios fraudulentos sobre o Desenrola Brasil evidencia não apenas o uso sistemático das plataformas da Meta para atingir pessoas superendividadas, mas também os impactos

da atuação inconsistente e pouco transparente das plataformas digitais, que facilita a disseminação desse tipo de prática. No caso dos anúncios do Desenrola Brasil, nem mesmo a medida cautelar da Senacon foi suficiente para impedir a circulação dos anúncios irregulares nas plataformas. Poucos dias após a decisão, a Meta apenas deixou de categorizar como políticos os anúncios sobre o programa, o que não impedia que eles continuassem circulando em suas plataformas. Isso demonstra que a reação da Meta à medida cautelar não foi impedir a veiculação dos anúncios, e sim diminuir a transparência da Biblioteca. Apesar de a Meta ter declarado que removeria os anúncios fraudulentos envolvendo o Desenrola Brasil, esse tipo de publicidade seguiu sendo veiculado. Assim, a reação da Meta não apenas deixou os usuários vulneráveis a essas campanhas criminosas como também limitou a possibilidade de escrutínio público para denúncia e fiscalização de irregularidades. Ou seja, a opção da empresa foi por menos transparência, e não por remover os anúncios.

Além disso, é importante destacar que a remoção dos anúncios após sua veiculação é ineficaz para conter a proliferação de novos golpes financeiros. Essa medida, além de tardia, não repara os danos causados às vítimas nem impede a recorrência dos golpes, tornando-se insuficiente como estratégia de proteção aos consumidores. Ao mesmo tempo, a Meta retém o valor pago pelos estelionatários e segue lucrando com anúncios irregulares sem cumprir seus próprios termos de uso.

Do ponto de vista comercial, é fundamental destacar que as práticas de moderação de anúncios da Meta não são consistentes e tampouco eficazes a ponto de tornar o ambiente de publicidade de suas plataformas seguro para os consumidores e para anunciantes legítimos. Essa fragilidade configura um risco sistêmico para o mercado publicitário digital, pois compromete a credibilidade do setor. Primeiramente, a exibição de anúncios regulares ao lado de publicidade irregular e fraudulenta compromete a credibilidade do mercado, prejudicando a imagem das marcas legítimas. Em segundo lugar, a falta de clareza nos critérios de moderação prejudica o próprio mercado, lesando anun-

ciantes legítimos que investem em campanhas sem garantia de que seus anúncios serão exibidos. Como demonstramos, a Meta removeu anúncios sem qualquer irregularidade aparente e sem fornecer justificativas, o que reforça a sensação de arbitrariedade nos métodos de moderação e gera incertezas para as campanhas publicitárias em suas plataformas. Ainda assim, é importante lembrar que essas irregularidades só podem ser identificadas porque a Meta oferece um nível mínimo de transparência que permite identificar essas falhas. No entanto, conforme apresentamos no capítulo 3, outras plataformas operam como verdadeiras caixas-pretas, sem qualquer abertura para análise externa independente, o que indica que o problema pode ser ainda mais grave.

Apesar de a Biblioteca da Meta nos fornecer algumas evidências que permitem estimar o tamanho deste mercado e seu público-alvo, a ferramenta disponibilizada pela plataforma também impõe uma série de limitações críticas aos nossos estudos. Por exemplo, apesar de nossa análise detectar a aplicação de critérios regionais para segmentação dos golpes financeiros, não é possível fazer uma análise sistemática de estratégias de segmentação usadas por anunciantes e pela plataforma. Do mesmo modo, nossa pesquisa tem lacunas sobre dados de investimento e alcance em relação a anúncios comerciais, uma vez que a Meta só torna públicos dados de anúncios categorizados como políticos. Entretanto, considerando que centenas dos anúncios encontrados não foram corretamente categorizados, uma parte significativa de nossa amostra carece de dados detalhados sobre seu impacto. Essas limitações indicam que os dados sobre publicidade fraudulenta podem estar subdimensionados, sendo provavelmente muito mais expressivos do que os obtidos nesta pesquisa.

Diante das evidências de anunciantes maliciosos promovendo produtos e medicamentos irregulares, sem qualquer vínculo com os programas governamentais, este capítulo reforça como as plataformas têm se configurado como um ambiente propício para a consolidação de uma indústria da desinformação e da influência – neste caso, às custas da imagem de instituições públicas e privadas e das vulnerabilidades

dos consumidores. Essa indústria movimenta centenas de milhares de reais e articula uma complexa rede de sites, canais de mensageria e publicidade online.

A partir desse estudo empírico, esperamos embasar discussões sobre governança relacionadas à publicidade digital e o modelo de negócios das plataformas digitais. Ao identificar e reunir evidências destas práticas, os resultados de pesquisa aqui apresentados podem contribuir para prevenir crimes contra os consumidores e tornar o ambiente das plataformas mais transparente e seguro para a população brasileira.

## Referências


ABDALA, Vitor. Endividamento atinge 78,3% das famílias brasileiras, diz CNC. Rio de Janeiro: *Agência Brasil*, 04 maio de 2023. Disponível em: https://agenciabrasil.ebc.com.br/economia/noticia/2023-05/endividamento-atinge-783-das-familias-brasileiras-diz-cnc. Acesso em: 12 dez. 2024.

ALEIXO, Isabela. Golpistas usam dados pessoais para se passar por sites do governo. São Paulo: *UOL*, 27 fev. 2024. Disponível em: https://noticias.uol.com.br/confere/ultimas-noticias/2024/02/27/golpistas-que-se-passam-por-sites-do-governo-tem-acesso-a-dados-pessoais.htm. Acesso em: 12 dez. 2024.

BARBOZA, Vinícius. Golpe do Voa Brasil rouba dados e dinheiro de vítimas; saiba como funciona. *Folha de S. Paulo*, *[S.l.]*, 19 out. 2023. Disponível em: https://www1.folha.uol.com.br/mercado/2023/10/novo-golpe-do-voa-brasil-rouba-dados-e-dinheiro-de-vitimas-saiba-como-funciona.shtml. Acesso em: 12 dez. 2024.

BARCELLOS, Thaís. Desenrola chega ao fim alcançando 5 milhões de pessoas na faixa voltada à baixa renda. Brasília: *O Globo*, 18 maio 2024. Disponível em: https://oglobo.globo.com/economia/noticia/2024/05/18/desenrola-chega-ao-fim-alcancando-5-milhoes-de-pessoas-na-faixa-voltada-a-baixa-renda.ghtml. Acesso em: 12 dez. 2024.

BISCHOFF, Wesley. Governo manda Google e Facebook retirarem do ar anúncios falsos do 'Desenrola'. São Paulo: *G1*, 26 jul. 2023. Disponível

em: https://g1.globo.com/politica/noticia/2023/07/26/governo-manda-google-e-facebook-retirarem-do-ar-anuncios-falsos-do-desenrola.ghtml. Acesso em: 12 dez. 2024.

BRASIL. Lei nº 2.848, de 7 de dezembro de 1940. *Código Penal*. Presidência da República, Casa Civil, Subchefia para Assuntos Jurídicos, Rio de Janeiro, RJ. 1940. Disponível em: https://www.planalto.gov.br/ccivil_03/decreto-lei/del2848compilado.htm. Acesso em: 10 dez. 2024.

BRASIL. Lei no 8.078, de 11 de setembro de 1990. *Código de Defesa do Consumidor*. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF. 1990. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm. Acesso em: 10 dez. 2024.

BREWSTER, Jack; FISHMAN, Zack; XU, Elisa. Funding the Next Generation of Content Farms: Some of the World's Largest Blue Chip Brands Unintentionally Support the Spread of Unreliable AI-Generated News Websites. *NewsGuard*, *[S.l.]*, jun. 2023. Disponível em: https://www.newsguardtech.com/misinformation-monitor/june-2023/. Acesso em: 11 dez. 2024.

BUTTON, Mark; NICHOLLS, Carol McNaughton; KERR, Jane; OWEN, Rachael. Online frauds: Learning from victims why they fall for these scams. *Australian & New Zealand Journal of Criminology*, *[S.l.]*, v. 47, n. 3, p. 391-408, 2014. https://doi.org/10.1177/0004865814521224

CONAR, Conselho Nacional de Autorregulamentação Publicitária. *Código Brasileiro de Autorregulamentação Publicitária*. São Paulo: Conar, 2024. Disponível em: http://www.conar.org.br/pdf/Codigo-CONAR-2024.pdf. Acesso em: 10 dez. 2024.

DANTAS, Marina. 'Voa Brasil' vai contemplar aposentados e pensionistas do INSS. *Rádio Senado*, *[S.l.]*, 26 jul. 2024. Disponível em: https://www12.senado.leg.br/radio/1/noticia/2024/07/26/voa-brasil-vai-contemplar-aposentados-e-pensionistas-do-inss. Acesso em: 12 dez. 2024.

GOEL, Rajeev K. Masquerading the Government: Drivers of Government Impersonation Fraud. *Public Finance Review*, *[S.l.]*, v. 49, n. 4, p. 548-572, 2021. https://doi.org/10.1177/10911421211029305.

GORWA, Robert; BINNS, Reuben; KATZENBACH, Christian. Algorithmic content moderation: technical and political challenges in the automation of platform governance. *Big Data & Society*, v. 7, n. 1, 2020. Disponível em: https://doi.org/10.1177/2053951719897945. Acesso em: 12 dez. 2024.

JORNAL NACIONAL. MP-MG envia ofícios para que a Meta bloqueie perfis e anúncios de golpes que envolvem o Desenrola. *Jornal Nacional*, *[S.l.]*, 20 jul. 2023. Disponível em: https://g1.globo.com/jornal-nacional/noticia/2023/07/20/mp-mg-envia-oficios-para-que-a-meta-bloqueie-perfis-e-anuncios-de-golpes-que-envolvem-o-desenrola-brasil.ghtml. Acesso em: 12 dez. 2024.

KEMP, Steven; PÉREZ, Nieves Erades. Consumer Fraud against Older Adults in Digital Society: Examining Victimization and Its Impact. *International Journal of Environmental Research and Public Health*, *[S.l.]*, v. 20, n. 7: 5404, p. 1-17, 2023. https://doi.org/10.3390/ijerph20075404.

KHONJI, Mahmoud; IRAQI, Youssef; JONES, Andrew. Phishing detection: a literature survey. *IEEE Communications Surveys & Tutorials*, v. 15, n. 4, p. 2091-2121, Fourth Quarter 2013. DOI: 10.1109/SURV.2013.032213.00009.

LASTDRAGER, Elmer. E. Achieving a consensual definition of phishing based on a systematic review of the literature. *Crime Science*, v. 3, n. 9, 2014. https://doi.org/10.1186/s40163-014-0009-y.

LEBOW, Sarah. How big of a problem are made-for-advertising websites, and what can advertisers do about them?. *Emarketer*, *[S.l.]*, 26 jun. 2024. Disponível em: https://www.emarketer.com/content/made-for-advertising-websites-advertisers. Acesso em: 11 dez. 2024.

MAIA, Gustavo. Governo alerta para golpe com deepfake de Haddad sobre dinheiro em contas. *Veja*, *[S.l.]*, 22 out. 2024. Disponível em: https://veja.abril.com.br/coluna/radar/governo-alerta-sobre-golpe-com-deepfake-de-haddad-sobre-dinheiro-em-contas. Acesso em: 12 dez. 2024.

MARTINS, Marco Antônio; DUARTE, Helter. Estelionatários anunciam na internet propagandas falsas de programa do governo que não

foi lançado. J*ornal Nacional e G1 Rio*, 16 nov. 2023. Disponível em: https://g1.globo.com/rj/rio-de-janeiro/noticia/2023/11/16/estelionatarios-anunciam-na-internet-propagandas-falsas-de-programa-do-governo-que-nao-foi-lancado.ghtml. Acesso em: 12 dez. 2024.

MÁXIMO, Wellton. Segunda fase do Desenrola exige cadastro no Portal Gov.br. Brasília: *Agência Brasil*, 03 out. 2023. Disponível em: https://agenciabrasil.ebc.com.br/economia/noticia/2023-10/segunda-fase-do-desenrola-exige-cadastro-no-portal-govbr. Acesso em: 12 dez. 2024.

META. Anúncios Relacionados com Questões Sociais, Eleições ou Política. *Meta*, *[S.d.]*. Disponível em: https://transparency.meta.com/en-gb/policies/ad-standards/SIEP-advertising/SIEP/. Acesso em: 12 dez. 2024.

MINISTÉRIO DA FAZENDA. Desenrola Brasil lança Plataforma para Renegociação de Dívidas. *Planalto*. Presidência da República, 09 out. 2023. Disponível em: https://www.gov.br/planalto/pt-br/acompanhe-o-planalto/noticias/2023/10/desenrola-brasil-lanca-plataforma-para-renegociacao-de-dividas. Acesso em: 12 dez. 2024.

MINISTÉRIO DA JUSTIÇA E SEGURANÇA PÚBLICA. Google e Facebook deverão retirar do ar anúncios fraudulentos sobre o programa Desenrola Brasil. *Ministério da Justiça e Segurança Pública*, *[S.l.]*, 26 jul. 2023. Disponível em: https://www.gov.br/mj/pt-br/assuntos/noticias/google-e-facebook-deverao-retirar-do-ar-anuncios-fraudulentos-sobre-o-programa-desenrola-brasil. Acesso em: 12 dez. 2024.

MORENO, Sayonara. Desenrola Brasil: renegociação de dívidas da faixa 2 começa na segunda. Brasília: *Agência Brasil*, 14 jul. 2023. Disponível em: https://agenciabrasil.ebc.com.br/radioagencia-nacional/economia/audio/2023-07/desenrola-brasil-renegociacao-de-dividas-da-faixa-2-comeca-na-segunda. Acesso em: 12 dez. 2023.

NETLAB UFRJ. *Golpes e falhas sistêmicas*: anúncios fraudulentos sobre o Desenrola Brasil e o Voa Brasil seguem circulando após um ano de medida cautelar. Rio de Janeiro: NetLab UFRJ - Laboratório de Estudos de Internet e Redes Sociais, 9 out. 2024. Disponível em: https://netlab.eco.ufrj.br/post/golpes-e-falhas-sist%C3%AAmicas-an%C3%BAncios-frau-

dulentos-sobre-o-desenrola-brasil-e-o-voa-brasil-seguem-cir. Acesso em: 12 dez. 2024.

PORTOS E AEROPORTOS. Maior programa social da aviação brasileira, Voa Brasil tem impulsionado os destinos regionais. *Presidência da República*, 23 set. 2024. Disponível em: https://www.gov.br/portos-e-aeroportos/pt-br/assuntos/noticias/2024/09/maior-programa-social-da-aviacao-brasileira-voa-brasil-tem-impulsionado-os-destinos-regionais. Acesso em: 12 dez. 2024.

PROCON-SP. Alerta: golpistas recriam voz do governador Tarcísio de Freitas para aplicar fraude nas redes sociais. São Paulo: *Procon-SP*, 13 set. 2024. Disponível em: https://www.procon.sp.gov.br/alerta-golpistas-recriam-voz-do-governador-tarcisio-de-freitas-para-aplicar-fraude-nas-redes-sociais/. Acesso em: 12 dez. 2024.

RAO, Anita. Deceptive Claims Using Fake News Advertising: The Impact on Consumers. *Journal of Marketing Research*, *[S.l.]*, v. 59, n. 3, p.534-554, 2022. https://doi.org/10.1177/00222437211039804.

RIBEIRO, Cristiane. Márcio França confirma lançamento do Voa Brasil para agosto. Rio de Janeiro: *Agência Brasil*, 14 jul. 2023. Disponível em: https://agenciabrasil.ebc.com.br/radioagencia-nacional/economia/audio/2023-07/marcio-franca-confirma-lancamento-do-voa-brasil-para-agosto. Acesso em: 12 dez. 2024.

ROBERTS, Sarah. Commercial content moderation: digital laborers' dirty work. *Media Studies Publications*, 2016. Disponível em: https://ir.lib.uwo.ca/commpub/12. Acesso em: 12 dez. 2024.

SABÓIA, Gabriel. 'Voa Brasil': programa de passagens a R$200 começa em fevereiro e pode atingir até 20,6 milhões de pessoas, diz ministro. Brasília: *O Globo*, 09 jan. 2024. Disponível em: https://oglobo.globo.com/economia/noticia/2024/01/09/voa-brasil-programa-de-passagens-a-r-200-comeca-em-fevereiro-e-pode-atingir-25-milhoes-de-pessoas-diz-ministro.ghtml. Acesso em: 12 dez. 2024.

SALLES, Débora; MARTINS, Bruno Maurício; SANTINI, Rose Marie. "Deus, Pátria, Família e Liberdade": a radicalização política no ecossiste-

ma de mídia evangélica digital no Brasil. Niterói: *Mídia e Cotidiano*, v. 18, n. 1, p. 36-63, jan.-abr. 2024. Disponível em: https://periodicos.uff.br/midiaecotidiano/article/view/59933/35992. Acesso em: 03 fev. 2024.

SALLES, Debora; SANTINI, Marie; BARROS, Carlos; MATTOS, Bruno; HADDAD, João; GOMES, Matheus; SEADE, Renata. *"Explore o Brasil por um preço incrível!"*: anúncios que usam o programa Voa Brasil para golpes e fraudes nas plataformas Meta. Rio de Janeiro: NetLab UFRJ - Laboratório de Estudos de Internet e Redes Sociais, 2023. Disponível em: https://doi.org/10.13140/RG.2.2.18684.04483. Acesso em: 12 dez. 2024.

SANTINI, Marie; SALLES, Débora; BARROS, Carlos Eduardo; MARTINS, Bruno Mauricio Mattos; MOREIRA, Alékis; HADDAD, João Gabriel; SEADE, Renata; SOUZA, Lucas. *A publicidade a favor do endividamento*: Anúncios que usam o programa Desenrola Brasil para golpes e fraudes nas plataformas Meta. Rio de Janeiro: NetLab UFRJ - Laboratório de Estudos de Internet e Redes Sociais, 25 jul. 2023a. Disponível em: https://netlab.eco.ufrj.br/post/publicidade-a-favor-do-endividamento-an%C3%BAncios-que-usam-o-desenrola-brasil-para-golpes-na-meta. Acesso em: 12 dez. 2024.

SANTINI, Marie; SALLES, Débora; MARTINS, Bruno Mauricio Mattos; BARROS, Carlos Eduardo; MOREIRA, Alékis; HADDAD, João Gabriel; SILVA, Daphne; SEADE, Renata; SOUZA, Lucas; YONESHIGUE, Bernardo. *'O fim dos seus problemas'*: A permanência de anúncios que usam o programa 'Desenrola Brasil' para golpes e fraudes nas plataformas Meta. Rio de Janeiro: NetLab UFRJ - Laboratório de Estudos de Internet e Redes Sociais, 17 nov. 2023b. Disponível em: https://netlab.eco.ufrj.br/post/o-fim-dos-seus-problemas-a-perman%C3%AAncia-de-an%C3%BAncios-que-usam-o-programa-desenrola-brasil-para-go. Acesso em: 12 dez. 2024.

SECRETARIA DE COMUNICAÇÃO SOCIAL. Cuidado: estelionatários simulam sites e políticas do Governo Federal; saiba como não cair nessa. *Secretaria de Comunicação Social, [S.l.]*, 11 jan. 2024. Disponível em: https://www.gov.br/secom/pt-br/fatos/brasil-contra-fake/noti-

cias/2024/cuidado-estelionatarios-simulam-sites-e-politicas-do-governo-federal-saiba-como-nao-cair-nessa. Acesso em: 12 dez. 2024.

SERASA. *Mapa da inadimplência e renegociação de dívidas*. Serasa, *[S.l.]*, 2023. Disponível em: https://cdn.builder.io/o/assets%2Fb212bb18f00a40869a6cd42f77cbeefc%2Fff9409a38a114135afd16d89734f5b0f?alt=media&token=de5430db-e168-411c-a174-5700f80f8368&apiKey=b212bb18f00a40869a6cd42f77cbeefc. Acesso em: 11 dez. 2024.

SETERA, Kristen. FBI Warns Public to Beware of Government Impersonation Scams. *FBI*, *[S.l.]*, 21 abr. 2021. Disponível em: https://www.fbi.gov/contact-us/field-offices/boston/news/press-releases/fbi-warns-public-to-beware-of-government-impersonation-scams. Acesso em: 12 dez. 2024.

STJ, Superior Tribunal de Justiça. Embargos de Divergência em Resp Nº 1.192.577 - RS (2014/0246972-3). *Superior Tribunal de Justiça, [S.l.]*,2015. Disponível em: https://scon.stj.jus.br/SCON/GetInteiroTeorDoAcordao?num_registro=201402469723&dt_publicacao=13/11/2015. Acesso em: 11 dez. 2024.

TEIXEIRA, Pedro S. Golpistas já distribuem links falsos do Desenrola Brasil no Facebook e WhatsApp. São Paulo: *Folha de S. Paulo*, 18 jul. 2023. Disponível em: https://www1.folha.uol.com.br/mercado/2023/07/golpistas-ja-distribuem-links-falsos-do-desenrola-brasil-no-facebook-e-whatsapp.shtml. Acesso em: 12 dez. 2024.

VAN DEN HOUT, Thijs; WABEKE, Thymen; MOURA, Giovane C.M.; HESSELMAN, Cristian. LogoMotive: Detecting Logos on Websites to Identify Online Scams - A TLD Case Study. In: HOHFELD, Oliver; MOURA, Giovane; PELSSER, Cristel (Eds.). *Passive and Active Measurement*. PAM, *[S.l.]*, 2022. Lecture Notes in Computer Science, vol 13210. Springer, Cham, 2022. https://doi.org/10.1007/978-3-030-98785-5_1.

VENTURA, Giulia; ARAÚJO, Pedro. Mulher morta e esquartejada pelo namorado já teria sido mantida em cárcere privado, diz madrinha. Rio de Janeiro: *O Globo*, 10 dez. 2022. Disponível em: https://oglobo.globo.com/rio/noticia/2022/12/mulher-morta-e-esquartejada-pelo-namorado-ja-teria-sido-mantida-em-carcere-privado-diz-madrinha.ghtml. Acesso em: 12 dez. 2024.

# Capítulo 5

## *Big techs*, direito do consumidor e interesse público

Os estudos apresentados nos últimos dois capítulos evidenciaram que as políticas autorregulatórias adotadas pelas plataformas atualmente têm sido ineficazes no enfrentamento a conteúdos irregulares, incluindo campanhas de desinformação e práticas comerciais que expõem os consumidores a diferentes riscos. Ao permitir e promover a circulação desses conteúdos servindo-se de dados dos próprios usuários, as plataformas não apenas tornam os consumidores vulneráveis a estratégias comerciais enganosas ou fraudulentas, que em diversos casos culminam em perdas financeiras, mas lucram com essa dinâmica, às custas da segurança e do bem-estar da população.

Neste capítulo, abordamos a centralidade do conceito de interesse público na formulação de políticas de governança no mercado midiático, atualizando o debate que tradicionalmente estava focado apenas nos meios de comunicação tradicionais. A partir deste conceito, discutimos a responsabilidade das *big tech* como agentes deste mercado, que operam não apenas como intermediárias, mas influenciam ativamente os fluxos de informação, comportamentos, opiniões e práticas de consumo. Assim, argumentamos que as atuais políticas das plataformas favorecem seus próprios interesses em detrimento dos interesses da população, violando normas locais, ameaçando os direitos dos cidadãos e a própria soberania dos Estados.

## 1. Interesse público como princípio da governança midiática

A literatura científica tem destacado que o conceito de interesse público há décadas ocupa um papel central na discussão sobre governança e na avaliação do desempenho das mídias (Napoli, 2015), sendo um parâmetro essencial para orientar a formulação de políticas públicas e estabelecer limites e expectativas para o funcionamento dos meios de comunicação. O conceito também é considerado fundamental para pautar o debate público e as ações de organizações da sociedade civil engajadas na defesa da integridade da informação, além de orientar a atuação dos profissionais do campo midiático. No entanto, a defesa do interesse público não só é a base para argumentos regulatórios e princípios de conduta profissional, mas deve ser pensado, sobretudo, como expressão das necessidades e expectativas dos usuários (Napoli, 2015).

A partir da necessidade de se defender o interesse público, a ideia de governança tem marcado o debate sobre os processos regulatórios das mídias, sendo utilizada para descrever uma série de mecanismos, leis e diretrizes nacionais e supranacionais pensados para determinar regras, direitos e deveres das plataformas. A organização das empresas de mídia online por meio de um sistema de governança busca englobar diferentes atores, incluindo não apenas formuladores de políticas, mas também partes interessadas da indústria, ONGs, organizações da sociedade civil e até mesmo o público (Napoli, 2015). Portanto, a ideia de governança da mídia pode ser caracterizada como deliberações, processos e resultados regulatórios que ocorrem tanto no âmbito estatal quanto para além dele (Napoli, 2019). O papel das políticas de governança incluem restrições sobre o que as empresas de mídia podem ou devem fazer (por exemplo, em relação a conteúdos de pornografia, violência, drogas etc.) ou exigências e responsabilidades afirmativas para atender as necessidades da população, como padrões mínimos de qualidade de conteúdo noticioso e informativo, padrões de precisão da informação e valores jornalísticos (Napoli, 2015).

Embora a defesa do interesse público seja amplamente reconhecida no campo da governança da mídia tradicional, pesquisadores do

tema apontam que esse conceito aparece nas discussões sobre mídias sociais de maneira mais implícita do que explícita (Napoli, 2015). De modo geral, as políticas de governança e a atuação das plataformas ainda se baseiam num modelo que as caracteriza como meras intermediárias prestadoras de serviços de tecnologia (Napoli; Caplan, 2017; Gillespie, 2010), resultando em uma percepção do interesse público como secundário e quase acessório.

Além disso, nos últimos anos, as plataformas de mídias sociais têm promovido mudanças semânticas na forma como se apresentam ao público, buscando mitigar o desgaste de sua imagem, cada vez mais associada a problemas como a falta de segurança online e os impactos prejudiciais ao bem-estar psíquico e emocional dos usuários. No lugar de "mídia social", passou-se a adotar nomenclaturas mais neutras e convidativas, como "comunidade de comunidades" ou "plataforma de entretenimento" (Santini, 2024). Contudo, a suposta imparcialidade que as plataformas alegam ter obscurece o fato de que elas impactam significativamente o fluxo de informações, influenciam os hábitos de consumo dos usuários e a dinâmica democrática. Ou seja, o poder de curadoria e distribuição de conteúdo online das plataformas de redes sociais as equipara, nesse sentido, aos papéis historicamente desempenhados pela mídia tradicional (Napoli, 2015).

Portanto, apesar de ser publicamente reconhecido que essas plataformas desempenham um papel central na produção, disseminação e consumo de notícias e informações, elas continuam a operar em um quadro regulatório que carece de diretrizes explícitas sobre o que constitui o interesse público no mundo digital, e sobre de que modo elas se enquadram como agentes que deveriam obrigatoriamente atuar em benefício da sociedade, ainda mais se considerarmos a concentração deste mercado em poucas empresas. Ao eximir as plataformas de responsabilidade pelos impactos resultantes de seus sistemas de recomendação de conteúdo e do uso abusivo de dados de usuários, a responsabilização passa a ser do elo mais fraco, que são os consumidores e produtores de conteúdo. Diferentemente das empresas de mídia tradicional, as plata-

formas não estão sujeitas a obrigações legais e éticas de interesse público, o que lhes permite priorizar interesses comerciais enquanto influenciam, de maneira direta e em larga escala, a opinião e o comportamento da população.

Napoli (2019) argumenta que esse modelo reduz o alcance do princípio do interesse público ao privilegiar uma abordagem individualista e obscurecer a autoridade editorial algorítmica. Nesse contexto, as plataformas digitais assumem apenas o papel de facilitadoras, promovem apenas a responsabilidade e a autonomia individuais na produção, disseminação e consumo de informações, e isentam-se do compromisso com a integridade informacional e a defesa do interesse público.

## 2. O interesse público e as políticas das plataformas

Transformações tecnológicas no campo da comunicação e da informação sempre demandaram esforços de inovação regulatória (Obar; Wildman, 2015). No caso das atividades das plataformas, as mídias sociais têm trazido desafios para entidades governamentais e formuladores de políticas públicas em diversas áreas, incluindo questões sobre privacidade, direito autoral, vigilância de usuários, controle abusivo de empregadores sobre seus funcionários, proteção da criança e do adolescente e direito do consumidor, entre outras (Obar; Wildman, 2015). Além disso, questões políticas importantes relacionadas à ascensão das plataformas no mercado digital incluem temas relativos à concentração de mercado, responsabilidade, regulação, diversidade e controle sobre dados gerados pelo uso da mídia em espaços públicos e privados (Bizberge; Mastrini; Gómez, 2023). Apesar disso, em grande parte dos países da América Latina, os debates sobre a regulação da internet permanecem periféricos no cenário político e social (Bizberge; Mastrini; Gómez, 2023), o que levanta preocupações sobre soberania digital, especialmente considerando que o mercado global é amplamente dominado por gigantes tecnológicos dos Estados Unidos (Jiang; Belli, 2025).

Neste contexto, o conceito de transparência emerge como um aspecto central no atual debate sobre a governança das plataformas di-

gitais, pois permite que o público, os governos e outros stakeholders avaliem a performance, políticas e práticas dessas empresas (Urman; Makhortykh, 2023). Por exemplo, pesquisadores frisam a importância de práticas de transparência, isto é, que empresas privadas que detêm plataformas voltadas para o público externo sejam capazes de prover informações internas a respeito de temas de interesse público (Urman; Makhortykh, 2023). No entanto, no Brasil, as iniciativas visando aumentar a transparência das plataformas atualmente enfrentam uma série de obstáculos e desafios, incluindo a limitação nas interfaces dos repositórios de anúncios e resistências corporativas e partidárias.

Enquanto a implementação de medidas regulatórias nacionais e supranacionais não têm avanços significativos ou ainda estão em fase de aprimoramento, as empresas atuam por meio de intervenções específicas, alterando suas políticas internas e termos de uso, resultando num modelo de autogovernança que lhes garante ampla autonomia e flexibilidade (Gorwa, 2019; Ye *et al.*, 2025). Este modelo de autorregulação permite que as plataformas projetem políticas que favorecem seus próprios modelos de negócios e, ao mesmo tempo, cria um ambiente favorável para que que essas empresas implementem políticas alinhadas aos seus próprios termos e princípios de segurança, sem, no entanto, lidar de fato com suas responsabilidades e deveres de prestação de contas (Ye *et al.*, 2025). A ampla liberdade das empresas para definir e aplicar suas próprias políticas frequentemente resulta na violação dessas diretrizes pelas próprias plataformas (Le Pochat et al., 2022). Esse cenário, combinado com o modelo de negócios altamente lucrativo dessas organizações, demonstra como as plataformas acabam se tornando cúmplices de atores maliciosos na medida em que lucram com a distribuição e amplificação de conteúdos nocivos à população (Kim, 2024).

Para sustentar o modelo atual, empresas e atores políticos frequentemente utilizam o argumento da "liberdade de expressão" como estratégia para barrar iniciativas voltadas ao aumento da transparência nas atividades das plataformas, tratando responsabilidade e liberdade de expressão como conceitos incompatíveis. No Brasil, as plataformas

participam ativamente na manipulação do debate público a respeito de questões regulatórias, como no caso da campanha contra o PL2630, quando publicaram anúncios que associavam a proposta à "censura" ou retratavam-na como uma iniciativa que poderia "piorar a sua internet" (NetLab UFRJ, 2023).

O caso emblemático mais recente foi o anúncio de Mark Zuckerberg na primeira semana de 2025 de que a Meta vai acabar com o sistema de checagem de fatos feita por terceiros e substituir por um sistema de notas da comunidade, seguindo o que foi feito no X/Twitter sob o comando de Elon Musk (Tagiaroli, 2025). Sem detalhar as alterações nas políticas, as mudanças previstas pela Meta incluem a redução nas restrições a temas como "imigração e gênero", e a substituição de filtros automatizados por denúncias que devem ser feitas pelos próprios usuários para identificar conteúdos nocivos classificados como "violações de baixa gravidade" (Galf, 2025). A postura da Meta ilustra como o posicionamento público das *big tech* delega aos usuários a tarefa de corrigir informações e filtrar conteúdos tóxicos, estimulando que os usuários trabalhem para as empresas denunciando e organizando conteúdo que não considerem adequado. Em vez de enquadrar a circulação de conteúdo tóxico como um problema estrutural, a empresa aposta em uma abordagem individualista, como se a questão central não fosse a indústria da desinformação em si, mas uma possível falta de letramento digital dos indivíduos.

Na prática, porém, essa política limita ainda mais o conhecimento público sobre as estruturas e os critérios de moderação aplicados, permitindo que as plataformas operem de forma ainda mais opaca e tomem decisões arbitrárias sobre a remoção de conteúdos. Além disso, o discurso de Zuckerberg chama atenção ao argumentar que a medida busca "restaurar a liberdade de expressão nas plataformas" e pelas críticas incisivas a governos, formuladores de políticas públicas e à mídia tradicional, alegando que esses atores "pressionam por censurar cada vez mais" (Tagiaroli, 2025; Spagnuolo, 2025). Por fim, embora Zuckerberg alegue agir para defender "pessoas inocentes" cujos conteúdos são

removidos, a omissão de qualquer compromisso com maior transparência reforça o argumento de que as próprias plataformas operam de forma obscura, sem clareza sobre seus sistemas algorítmicos, equipes de moderadores, regras ou critérios para a moderação ou sua ausência. A reversão da política de checagem de fatos, inicialmente instituída para conter a disseminação de desinformação nas plataformas da Meta, foi apontada pela mídia especializada como um claro indicativo de como a empresa está se reposicionando para o segundo mandato de Donald Trump (Schleifer; Isaac, 2025).

Embora as *big tech* defendam a liberdade de expressão como princípio para garantir que suas atividades ocorram sem escrutínio público, na perspectiva da defesa do interesse público e da democracia, a garantia efetiva da liberdade de expressão só é possível quando acompanhada de transparência e responsabilidade na moderação de conteúdos. Sem a devida transparência, consolida-se, na prática, um monopólio de poucas empresas privadas que dominam o mercado digital, utilizando dados de usuários para moldar a distribuição de conteúdo e maximizar seus ganhos financeiros. Esse controle é exercido sem que a sociedade tenha conhecimento sobre os critérios adotados para a circulação de informações e, frequentemente, sem punição para irregularidades relacionadas ao uso desses dados ou à falta de prestação de contas. A retórica da liberdade de expressão, em última instância, serve para preservar a imunidade dessas empresas frente às leis e regulamentações locais, criando um desequilíbrio entre os direitos dos usuários e o poder de empresas privadas.

## 3. Plataformas Digitais e Violações aos Direitos dos Consumidores

Embora estudiosos e formuladores de políticas públicas enfatizem a necessidade de responsabilizar as plataformas digitais e pressioná-las por maior transparência, a implementação dessas mudanças requer um esforço coordenado entre diversos atores tanto a nível local quanto global. A responsabilidade das plataformas digitais e a necessidade de maior transparên-

cia têm sido amplamente debatidas por estudiosos e formuladores de políticas públicas em diversos países de acordo com suas especificidades locais. Na América Latina, pesquisadores destacam como as assimetrias de poder e desafios históricos da comunicação moldam o debate sobre governança da internet (Bizberge; Mastrini; Gómez, 2023). Assim, práticas abusivas das plataformas e de atores maliciosos que as utilizam para promover conteúdos tóxicos exigem ações capazes de articular desafios mundiais com especificidades locais, como as relacionadas à regulação, estruturas de mercado e cultura institucional (Bizberge; Mastrini; Gómez, 2023).

Apesar da centralidade da discussão sobre regular as plataformas e responsabilizá-las em maior grau por suas atividades, há um problema imediato que não pode ser negligenciado neste debate: as frequentes violações de leis e normas vigentes por conteúdos que circulam nesses ambientes online. Os estudos realizados ao longo do projeto mostram como as plataformas não só se comportam como meras intermediárias, mas recomendam, impulsionam e lucram com conteúdos que infringem suas próprias políticas e diversas leis e normas brasileiras. Essas práticas incluem publicidade enganosa e abusiva, charlatanismo, estelionato, entre outros crimes que compram os serviços de distribuição de publicidade das plataformas para direcionar conteúdo e assim atingir de forma intencional principalmente grupos populacionais vulneráveis.

Considerando que essas atividades infringem regulamentações já vigentes nas leis brasileiras, as evidências que encontramos demonstram que essas normas não têm sido devidamente aplicadas em ambientes digitais e apresentam riscos críticos e diretos à saúde, à integridade e à dignidade dos consumidores brasileiros. O Código de Defesa do Consumidor, por exemplo, estabelece como direitos básicos dos consumidores:

I. *a proteção da vida, saúde e segurança* contra os riscos provocados por práticas no fornecimento de produtos e serviços considerados perigosos ou nocivos;
II. a educação e divulgação sobre o consumo adequado dos produtos e serviços, asseguradas a liberdade de escolha e a igualdade nas contratações;

III. *a informação adequada e clara sobre os diferentes produtos e serviços*, com especificação correta de quantidade, características, composição, qualidade, tributos incidentes e preço, bem como sobre os riscos que apresentem;

IV. *a proteção contra a publicidade enganosa e abusiva, métodos comerciais coercitivos ou desleais*, bem como contra práticas e cláusulas abusivas ou impostas no fornecimento de produtos e serviços; (...)

V. a efetiva *prevenção e reparação de danos patrimoniais e morais*, individuais, coletivos e difusos (...) (Brasil, 1990, grifos nossos).

Considerando que "a atividade de intermediação de conteúdo desenvolvida pelas plataformas de redes sociais as caracteriza como fornecedoras de serviços" (Brasil, 2023) e que a prestação de um serviço falho ou defeituoso representa risco à segurança do consumidor, a veiculação de publicidade abusiva, irregular ou com desinformação representa ameaça à vida, à saúde e à segurança, ferindo um direito básico do usuário e consumidor. Além de representar riscos diretos à segurança do consumidor, o modelo atual de funcionamento das plataformas tem sido criticado por pesquisadores em diversos países devido à violação de direitos básicos relacionados à privacidade, à dignidade e aos direitos fundamentais dos indivíduos (Armitage et al., 2023).

Apesar dos esforços atuais da legislação brasileira, as evidências de publicidade irregular encontradas ao longo dos estudos deste projeto seguem circulando no Brasil sem que haja mecanismos efetivos de fiscalização, documentação e responsabilização de anunciantes fraudulentos e plataformas. Ou seja, isso inclui não apenas quem produz e aplica os golpes, mas também os atores que possibilitam o direcionamento específico às vítimas mais vulneráveis, ampliam a visibilidade desses conteúdos tóxicos e obtêm lucros com essas práticas.

## 4. Medidas e propostas para transparência das plataformas em outros países

Temos demonstrado que as políticas das plataformas têm sido ineficazes e que, frequentemente, as próprias empresas violam as regras estabelecidas por elas mesmas, alimentando um modelo de negócios baseado na exploração dos dados dos usuários e permitindo a promoção recorrente de práticas criminosas sem medidas efetivas para coibi-las. Diante disso, com o objetivo de proteger os direitos dos cidadãos, assegurar políticas de concorrência justas e mitigar os impactos do impulsionamento pago de conteúdo irregular, enganoso e desinformativo, iniciativas regulatórias de diversos países têm imposto normas para que as plataformas forneçam informações claras, completas e de modo acessível sobre publicidade digital – por exemplo, quem paga pelos anúncios, dados demográficos do público-alvo e os critérios de segmentação empregados.

A maioria das propostas regulatórias que apresentaremos a seguir aborda explicitamente a necessidade de mais transparência sobre dados de anúncios, sobretudo aqueles relativos a temas considerados políticos ou de importância nacional. Algumas das primeiras propostas com esta abordagem incluem o Honest Ads Act, nos Estados Unidos, como resposta à revelação da intervenção da Internet Research Agency da Rússia nas eleições de 2016, e o Code of Practice on Disinformation na União Europeia (Leerssen *et al.*, 2019). Mais recentemente, dentre as propostas legislativas que passaram a tratar os posts impulsionados de forma mais rigorosa, destaca-se principalmente o DSA (União Europeia, 2022a), na Europa.

### *Digital Services Act (DSA) – União Europeia (2022)*

O *Digital Services Act* (DSA) busca proteger direitos fundamentais de cidadãos europeus e oferecer "maior controle democrático e vigilância nos sistemas de plataformas" (European Commission, [*S.d.*]d, n.p.). O DSA estabeleceu medidas envolvendo coleta de dados, privacidade, discurso de ódio e desinformação, que passaram a ser obrigatórias, des-

de agosto de 2023, para grandes plataformas e ferramentas de busca que atuam na União Europeia. A lei considera que as grandes plataformas, chamadas de "*very large platforms*" são aquelas usadas por pelo menos 10% dos cidadãos europeus (União Europeia, 2023).

Quanto aos dados de anúncios, as diretrizes do DSA estabeleceram novos padrões de transparência, como a disponibilização de repositório público acessível via interface de usuário e API com todos os anúncios veiculados pelas plataformas, em território europeu, no último ano. O repositório deve incluir, no mínimo: *(i)* todo o universo de anúncios veiculados na plataforma nos 365 dias anteriores; *(ii)* o conteúdo dos anúncios; *(iii)* quem os publicou; *(iv)* quem pagou por eles; *(v)* o período de veiculação; *(vi)* os principais parâmetros de segmentação aplicados; e *(viii)* quantas pessoas de cada Estado europeu os viram (União Europeia, 2022a).

No que diz respeito ao acesso a dados por pesquisadores, o DSA estabelece que as plataformas devem fornecer dados a membros de organizações que conduzam pesquisas científicas alinhadas a sua missão de interesse público (Digital Services Act, [S.d.]a). O DSA determina que os pesquisadores qualificados, que atendam requisitos específicos, solicitem esses dados às plataformas e ferramentas de busca online de grande porte, obrigando essas empresas a fornecê-los gratuitamente dentro de um "prazo razoável". Além disso, também especifica que esses dados devem servir "*exclusivamente* para a realização de pesquisas que contribuam para a detecção, identificação e *compreensão de riscos sistêmicos na União*" (Digital Services Act, [S.d.]b, tradução e grifos nossos). A definição exata de riscos sistêmicos não fica clara na lei, de forma que não é possível saber se são aqueles que impactam no funcionamento de sistemas sociotécnicos ou os riscos causados ou exacerbados por estes sistemas – ou ambos (Sullivan; Pielemeier, 2023). De todo modo, o DSA elenca como exemplos a disseminação de conteúdo ilegal em plataformas e efeitos negativos sobre o exercício de direitos fundamentais, a integridade do processo eleitoral e a proteção da saúde pública (Jahangir; Hendrix, 2024).

Dessa forma, embora o DSA imponha obrigações às plataformas para ampliar a transparência de dados a pesquisadores, ele estabelece um acesso mediado e restrito, vinculado a finalidades específicas previstas em lei, o que limita seu alcance e dá margem para que as plataformas explorem brechas legais. Um exemplo emblemático da não conformidade das plataformas com as determinações do DSA é o caso do X/Twitter, que descumpre diversos requisitos previstos pela regulamentação. Uma avaliação da Comissão Europeia identificou falhas que incluem falta de transparência na publicidade no X e restrições no acesso a dados para pesquisadores (Werner, 2024). Entre os problemas apontados estão a burocracia excessiva e cobrança de taxas desproporcionalmente elevadas para que os pesquisadores tenham acesso à API. Além do X/Twitter, a Comissão Europeia também apontou falhas similares por parte de outras plataformas, incluindo TikTok e Meta (Werner, 2024).

Com isso, desde o final de 2024, a Comissão Europeia tem discutido diretrizes para implementar de modo mais firme as normas do DSA e garantir o acesso a dados de modo mais seguro e eficaz (Vermeulen, 2024). Isso tem se desenvolvido sob a forma de um ato não legislativo que visa complementar a regulamentação vigente com detalhes técnicos ou ajustes específicos, com atualizações que levem em conta transformações decorrentes de progressos técnicos e científicos posteriores à criação da legislação atual. Neste caso, a Comissão Europeia prevê a criação de um inventário de dados, um portal de acesso a dados unificado para pesquisadores, plataformas e formuladores de políticas públicas, além de estender o acesso a organizações de pesquisas fora da União Europeia, desde que desenvolvam pesquisas relativas a riscos sistêmicos no continente (Vermeulen, 2024).

### *Digital Markets Act (DMA) – União Europeia (2023)*

O Digital Markets Act é uma regulamentação complementar ao DSA (European Commission, [*S.d.*]c) que impõe diversas obrigações e restrições às *big tech*. Considerando que as plataformas são "uma passagem obrigatória para usuários e empresas", gerando um contexto de

"quase monopólio" (Ministère de l'économie, 2024), a proposta visa regulamentar a atividade econômica das grandes plataformas digitais para garantir mais igualdade aos mercados digitais na União Europeia.

O DMA é apresentado como uma das primeiras ferramentas regulatórias (European Commission, [*S.d.*]a) que visa abordar as empresas de plataformas com alto impacto em território europeu, sobretudo Alphabet, Amazon, Apple, ByteDance, Meta e Microsoft, de modo a evitar que essas empresas abusem de uma posição dominante no mercado (Vie publique, 2024). Entre as obrigações previstas para ampliar a transparência comercial online e promover práticas de concorrência mais justas, está a exigência de interoperabilidade e facilitação do acesso aos anunciantes a dados sobre os serviços de publicidade nas plataformas. Essas empresas devem disponibilizar, mediante solicitação, os valores pagos por cada serviço e dados de performance publicitária. O DMA também impede as plataformas de usarem dados não públicos (como cliques, pesquisas, visualizações e comandos de voz) sem consentimento. Além disso, proíbe que dados gerados a partir das atividades de um anunciante sejam utilizados pelas próprias plataformas para promover seus produtos, o que caracterizaria como concorrência desleal (European Commission, [*S.d.*]a).

Em termos de transparência, o DMA sublinha o risco da opacidade das operações comerciais, com potenciais danos a anunciantes e agências por conta da falta de informações sobre os efeitos de seus próprios anúncios. Assim, a norma busca aumentar e aprimorar o acesso a dados, obrigando as plataformas a fornecerem gratuitamente informações detalhadas sobre os serviços de publicidade online a agências e empresas que os contratarem. A exigência de transparência é destacada como um elemento que exerce pressão externa sobre os controladores de acesso para evitar que a definição de perfis exaustivos de consumidores se consolide como prática predominante no setor, considerando que potenciais novos operadores ou empresas em estágio inicial não possuem acesso ao mesmo volume de dados nem à mesma escala de exaustividade (União Europeia, 2022b).

*Code of Practice on Disinformation – União Europeia (2018)*

Lançado em 2018 e atualizado em 2022, o Código de Conduta proposto pela União Europeia definiu que plataformas online, associações comerciais e os principais atores do setor de publicidade se comprometessem a combater a desinformação e melhorassem suas próprias políticas (European Commission, [*S.d.*]b).

A versão de 2018 propunha uma autorregulação e fazia com que os signatários se engajassem com medidas como identificação e transparência de anúncios políticos. Em 2022, com o DSA, o código foi atualizado e reforçado. No lugar da autorregulação, foi adotado um modelo de corregulação das grandes plataformas orientado pelas normas do próprio DSA. Além disso, as medidas com as quais os signatários devem se comprometer passaram a ser mais detalhadas e sofisticadas. Entre elas, a desmonetização de conteúdo desinformativo, o aumento da transparência em anúncios políticos, o aprimoramento do acesso e da qualidade dos dados, além da criação de um Centro de Transparência (European Commission, [*S.d*]b.; European Commission, 2022a).

Diferentemente do DSA, que se caracteriza como um quadro legislativo, o Código de Conduta é uma iniciativa de cooperação visando combater o problema da desinformação online, que depende de compromissos voluntários de signatários, como plataformas online e anunciantes, e não implica em penalidades legais. Entre os signatários da versão mais recente do documento (European Commission, 2022b), destacam-se diversas plataformas, como Meta, Google, TikTok, Microsoft e Twitch, entre outros.

*Online Safety Act – Reino Unido (2023)*

O Online Safety Act é um conjunto de leis proposto no Reino Unido visando a proteção dos usuários de mídias sociais e a criação de um ambiente online seguro, principalmente para as crianças (Gov.UK, [*S.d.*]). A medida é uma iniciativa para combater a disseminação de conteúdo potencialmente danoso online, seja ele orgânico ou impulsionado mediante pagamento. As leis buscam impor diversas obrigações às

plataformas e torná-las mais responsáveis pela segurança dos usuários. Estão submetidos às leis diversos tipos de sites, aplicativos, plataformas de redes sociais, aplicativos de mensageria, plataformas de compartilhamento de vídeos, fóruns, sites de armazenamento e compartilhamento de arquivos e aplicativos de relacionamento (Gov.UK, [*S.d.*]).

A proposta visa mitigar os impactos negativos dos algoritmos nas recomendações de conteúdo, como o aumento da exposição de conteúdos ilegais ou prejudiciais, sejam eles anúncios ou posts orgânicos. Uma vez identificados os riscos, os provedores de serviços devem reformular o design, as funcionalidades e os algoritmos para garantir a segurança dos usuários, sob pena de multas substanciais e outras sanções, incluindo a proibição de atividades comerciais e anúncios online. Além disso, as empresas deverão publicar relatórios anuais de transparência sobre a segurança online e o impacto dos algoritmos, fornecendo dados essenciais para jornalistas, formuladores de políticas públicas, pesquisadores, investidores e anunciantes. O *Office of Communications* (Ofcom), responsável pela aplicação da lei, poderá solicitar informações adicionais sobre a aplicação de políticas e tecnologias de moderação (Gov.UK, [*S.d.*]; Harling; Henesy; Simmance, 2023).

## 5. Considerações finais

Embora as plataformas de redes sociais se apresentem como meras intermediárias entre usuários, essas empresas devem ser caracterizadas como atores relevantes na indústria da informação e da comunicação. A circulação de publicidade abusiva, irregular ou com desinformação não apenas lesa individualmente os consumidores por meio de diversos tipos de fraudes, colocando em risco sua dignidade e integridade física e mental, mas tem também consequências coletivas mais amplas, como a desconfiança quanto às políticas públicas e a perda de credibilidade de anunciantes legítimos e veículos jornalísticos (Zeng; Kohno; Roesner, 2021).

Além disso, essas plataformas têm uma participação ativa na sociedade e no mercado caracterizada por uma série de contradições. Ao

mesmo tempo em que afirmam defender ideais de igualdade e diversidade, operam em um sistema de alta concentração de mercado e de práticas de concorrência desproporcionais; alegam defender o interesse público, mas se organizam de modo corporativo e priorizam seus lucros; seus efeitos aparentam ser locais, mas o alcance e impacto de suas operações são globais; sugerem dar voz à população para substituir o poder centralizado dos governos, "empoderando" usuários, mas faz isso por meio de uma estrutura altamente centralizada e opaca (Bizberge; Mastrini; Gómez, 2023; van Dijck; Poell; de Waal, 2018).

Atualmente, parte considerável dos parâmetros de transparência e governança que não são atendidos pelas plataformas na operação brasileira são ofertados em outros países por estas mesmas empresas, evidenciando que a assimetria regional pela ausência desse tipo de iniciativa no Brasil é uma decisão política e não técnica. A pouca transparência e as brechas na regulamentação das atividades dessas empresas têm deixado lacunas e dificultado a fiscalização, permitindo que anunciantes explorem um ambiente desregulado e promovam produtos e serviços por meio de anúncios nocivos ou influenciadores.

Esses impactos reforçam a necessidade de regulamentações eficazes e a adoção de mecanismos que imponham obrigações às plataformas, viabilizando uma fiscalização pública rigorosa e responsabilização dos diversos atores que lucram com inúmeras práticas irregulares que têm se perpetuado nos ambientes digitais. Para coibir práticas online que violem os direitos dos consumidores e garantir que as plataformas sejam seguras para os usuários, é preciso assegurar a auditabilidade desses espaços, não só sobre os conteúdos irregulares que circulam neles, mas sobre sua própria arquitetura e funcionamento. Na medida em que o poder das plataformas impacta diretamente as dinâmicas de mercado, o acesso à cultura e à informação, as práticas políticas locais e as democracias ao redor do mundo, autores apontam para a necessidade de debater como os Estados podem atuar para garantir maior soberania digital, considerando a necessidade de uma atuação colaborativa entre organizações, instituições e governos (Bizberge; Mastrini; Gómez, 2023).

Além disso, para lidar com os problemas trazidos pelas práticas abusivas de atores maliciosos e das próprias plataformas, a literatura científica aponta para a importância de se revisitar o conceito de interesse público, frequentemente marginalizado nas discussões sobre a governança do setor de mídia digital (Napoli, 2019).

Considerando que a tecnologia sempre avançará mais rapidamente do que a esfera do direito, as leis e propostas de regulamentação em outros países que apresentamos neste capítulo são pontos de partida para futuros desdobramentos e atualizações. A experiência com iniciativas como o DSA mostra que mesmo medidas regulatórias já em vigor também podem demandar aprimoramentos, seja para lidar com lacunas existentes ou para assegurar maior eficácia na aplicação das normas. Por exemplo, um estudo sobre a transparência das plataformas antes e depois da regulamentação demonstrou que de fato houve avanços significativos nos relatórios de transparência das plataformas. No entanto, os pesquisadores também observaram disparidades na granularidade, consistência e padronização das informações entre as plataformas para que haja análises e comparações mais eficazes entre as práticas de moderação de conteúdo nas plataformas (Zornetta, 2024). Assim, o caso do DSA é exemplar pois, apesar dos progressos alcançados, ainda persistem lacunas que demandam esforços permanentes de especialistas, pesquisadores, formuladores de políticas públicas e organizações da sociedade civil comprometidos com a promoção da integridade informacional nos ambientes online.

Ainda que futuramente haja uma eventual regulamentação das plataformas digitais no Brasil, é preciso ter em vista que novas ferramentas e usos da tecnologia podem surgir, apresentando novos problemas e demandando novas investigações e descobertas. Por isso, é necessário assegurar meios de fiscalização públicos sustentados por parâmetros formulados com bases científicas para garantir o cumprimento das leis brasileiras e identificar novos problemas e possíveis estratégias para burlar as legislações locais e supranacionais. O papel dos pesquisadores neste contexto é participar no monitoramento contínuo

destes ambientes, trabalhando em conjunto e apoiando o poder público para garantir o cumprimento e o fortalecimento de políticas públicas. Para as pesquisas acadêmicas, o acesso aos dados digitais permite não apenas identificar, descrever e compreender fenômenos e problemas sociais relevantes, mas também formular recomendações embasadas para lidar com eles e mitigá-los. Assim, as pesquisas são essenciais para informar políticas públicas e de governança fundamentadas em análises sobre infraestrutura, economia política e campanhas de desinformação nas plataformas digitais.

## Referências


ARMITAGE, Catherine.; BOTTON, Nick; DEJEU-CASTANG, Louis; LEMOINE, Laureline. Study on the impact of recent developments in digital advertising on privacy, publishers and advertisers. Brussels: *European Commission*, 2023. Disponível em: https://op.europa.eu/en/publication-detail/-/publication/8b950a43-a141-11ed-b508-01aa75ed71a1/language--en. Acesso em: 2 ago. 2024

BIZBERGE, Ana; MASTRINI, Guillermo; GÓMEZ, Rodrigo. Discussing internet platform policy and regulation in Latin America. United Kingston: *Journal of Digital Media & Policy*,v. 14, Issue Emerging Debates on Internet Platform Policy and Regulation in Latin America, Jun. 2023, p. 135 - 148, 2023. https://doi.org/10.1386/jdmp_00118_2.

BRASIL. Ministério da Justiça e Segurança Pública. Portaria do Ministro nº 351, de 14 de julho de 2023. *Diário Oficial da União*: seção 1, Brasília, DF, 17 jul. 2023. Disponível em: https://www.gov.br/mj/pt-br/centrais-de-conteudo/publicacoes/categorias-de-publicacoes/portarias/portaria-do-ministro_plataformas.pdf/view. Acesso em: 10 jan. 2025.

BRASIL. Lei nº 8.078, de 11 de setembro de 1990. *Código de Defesa do Consumidor*. Diário Oficial [da] República Federativa do Brasil, Poder Executivo, Brasília, DF.Disponível em: https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm#:~:text=O%20fornecedor%20n%C3%A3o%20poder%C3%A1%20colocar,periculosidade%20

%C3%A0%20sa%C3%BAde%20ou%20seguran%C3%A7a. Acesso em: 6 jan. 2025.

DIGITAL SERVICES ACT. The final text of the Digital Services Act (DSA). *Preamble 91-100*. Digital Services Act (DSA), [*S.d.*]a. Disponível: https://www.eu-digital-services-act.com/Digital_Services_Act_Preamble_91_to_100.html. Acesso em: 04 fev. 2025.

DIGITAL SERVICES ACT. The final text of the Digital Services Act (DSA). *Article 40, Data access and scrutiny*. Digital Services Act (DSA), [*S.d.*]b. Disponível em: https://www.eu-digital-services-act.com/Digital_Services_Act_Article_40.html. Acesso em: 04 fev. 2025.

EUROPEAN COMMISSION. The 2022 Code of Practice on Disinformation | Shaping Europe's digital future. *European Commission*, 2022a. Disponível em: https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation. Acesso em: 1 ago. 2024.

EUROPEAN COMMISSION. Signatories of the 2022 Strengthened Code of Practice on Disinformation. *European Commission*, 2022b. Disponível em: https://digital-strategy.ec.europa.eu/en/library/signatories-2022-strengthened-code-practice-disinformation. Acesso em: 1 ago. 2024.

EUROPEAN COMMISSION. About the Digital Markets Act. *European Commission*, *[S.d.]*a. Disponível em: https://digital-markets-act.ec.europa.eu/about-dma_en. Acesso em: 1 ago. 2024.

EUROPEAN COMMISSION. A strengthened EU Code of Practice on Disinformation. *European Commission*, *[S.d.]*b. Disponível em: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/new-push-european-democracy/protecting-democracy/strengthened-eu-code-practice-disinformation_en. Acesso em: 1 ago. 2024.

EUROPEAN COMMISSION. The Digital Markets Act: ensuring fair and open digital markets. *European Commission*, [*S.d.*]c. Disponível em: <https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en>. Acesso em: 1 ago. 2024.

EUROPEAN COMMISSION. The EU's Digital Services Act. *European Commission*, *[S.d.]*d. Disponível em: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en. Acesso em: 1 ago. 2024.

GALF, Renata. Anúncio da Meta indica embate contra regulação e mudança sobre conteúdos políticos e minorias. São Paulo: *Folha de S. Paulo*, 07 jan. 2025. Disponível em: https://www1.folha.uol.com.br/poder/2025/01/anuncio-da-meta-indica-embate-contra-regulacao-e-mudanca-sobre-conteudos-politicos-e-minorias.shtml. Acesso em: 07 jan. 2025.

GILLESPIE, Tarleton. The politics of 'platforms'. *New Media & Society*, *[S.l.]*, v. 12, n. 3, p. 347–364, 9 fev. 2010. Disponível em: https://journals.sagepub.com/doi/10.1177/1461444809342738. Acesso em: 1 ago. 2024.

GORWA, Robert. What is platform governance? London: *Information, Communication & Society*, v. 22, n. 6, p. 854-871, 2019. Disponível em: https://doi.org/10.1080/1369118X.2019.1573914. Acesso em: 06 jan. 2024.

GOV.UK. Online Safety Act: explainer. *Department for Science, Innovation & Technology [S.d.]*. Disponível em: https://www.gov.uk/government/publications/online-safety-act-explainer/online-safety-act-explainer. Acesso em: 1 ago. 2024.

HARLING, Anne-Sophie; HENESY, Declan; SIMMANCE, Eleanor. Transparency Reporting: The UK Regulatory Perspective. Stanford: *Journal of Online Trust and Safety*, v. 1, n. 5, p. 1-8, 30 jan. 2023. Disponível em: https://tsjournal.org/index.php/jots/article/view/108. Acesso em: 1 ago. 2024.

JAHANGIR, Ramsha; HENDRIX, Justin. Understanding Systemic Risks under the Digital Services Act. *Tech Policy Press*, *[S.l.]*, 15 set. 2024. Disponível em: https://www.techpolicy.press/understanding-systemic-risks-under-the-digital-services-act/. Acesso em: 05 fev. 2025.

JIANG, Min; BELLI, Lucca. Contesting Digital Sovereignty: Untangling a Complex and Multifaceted Concept. In: JIANG, M.; BELLI, L. (Eds.). *Digital Sovereignty in the BRICS Countries*: How the Global South and Emerging Power Alliances Are Reshaping Digital Governance. Cambridge; New York; Port Melbourne; New Delhi; Singapore: Cambridge University Press, 2025, p. 1-38. https://doi.org/10.1017/9781009531085.

KIM, Hwa Y. What's wrong with relying on targeted advertising? Targeting the business model of social media platforms. *Critical Review of International Social and Political Philosophy*, *[S.l.],* v. 0, n. 0, p. 1–21, 29 jan. 2024. Disponível em: https://doi.org/10.1080/13698230.2024.2309047. Acesso em: 14 jan. 2025.

LEERSSEN, Paddy; AUSLOOS, Jef.; ZAROUALI, Brahim.; HELBERGER, Natali; VREESE, Claes de. Platform ad archives: promises and pitfalls. *Internet Policy Review*, *[S.l.]*, v. 8, n. 4, p. 1–21, 9 out. 2019. Disponível em: https://policyreview.info/articles/analysis/platform-ad-archives-promises-and-pitfalls. Acesso em: 1 ago. 2024.

LE POCHAT, Victor; EDELSON, Laura; GOETHEM, Tom. V.; JOOSEN, Wouter; MCCOY, Damon; LAUINGER, Tobias. An audit of Facebook's political ad policy enforcement. In: USENIX Security Symposium, 31., ago. 2022, Boston: *Anais* [...] USENIX Association, *[S.l.]*, 2022. Disponível em: https://www.usenix.org/system/files/sec22-lepochat.pdf. Acesso em: 14 jan. 2025.

MINISTÈRE DE L'ÉCONOMIE DES FINANCES ET DE LA SOUVERAINETÉ INDUSTRIELLE ET NUMÉRIQUE. Numérique : le règlement sur les marchés numériques (DMA) entre en application. *Ministère de l'économie*, 5 mar. 2024. Disponível em: https://www.economie.gouv.fr/actualites/numerique-reglement-sur-les-marches-numeriques-dma-applicable. Acesso em: 1 ago. 2024.

NAPOLI, Philip M. *Social Media and the Public Interest*: Media Regulation in the Disinformation Age. Nova York: Columbia University Press, 2019.

NAPOLI, Philip M. Social media and the public interest: Governance of news platforms in the realm of individual and algorithmic gatekee-

pers. *Telecommunications Policy*, *[S.l.]*, v. 39, n. 9, out. 2015, p. 751-760, 2015. Disponível em: https://doi.org/10.1016/j.telpol.2014.12.003. Acesso em: 19 dez. 2024.

NAPOLI, Philip M; CAPLAN, Robyn. Why media companies insist they're not media companies, why they're wrong, and why it matters. *First Monday*, *[S.l.]*, v. 22, n. 5, 2017. Disponível em: https://firstmonday.org/ojs/index.php/fm/article/view/7051. Acesso em: 1 ago. 2024.

NETLAB UFRJ. A guerra das plataformas contra o PL 2630. Rio de Janeiro: *NetLab* - Laboratório de Estudos de Internet e Redes Sociais - Universidade Federal do Rio de Janeiro (UFRJ). 1 maio 2023. Disponível em: https://netlab.eco.ufrj.br/post/a-guerra-das-plataformas-contra-o--pl-2630. Acesso em 16 dez. 2024.

OBAR, Jonathan A.; WILDMAN, Steven. Social media definition and the governance challenge: An introduction to the special issue. Rochester: *Telecommunications Policy*, v. 39, n. 9, p. 745-750, 2015. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2647377. Acesso em: 07 jan. 2025.

SANTINI, Rose M. Desilusão tecnológica, regulação e o futuro das mídias sociais. *Fast Company*, *[S.l.],* 30 dez. 2024. Disponível em: https://fastcompanybrasil.com/coluna/desilusao-tecnologica-regulacao-e-o-futuro-das-midias-sociais/. Acesso em: 08 jan. 2025.

SCHLEIFER, Theodore.; ISAAC, Mike. Meta's move effectively puts an end to its longstanding fact-checking program. *The New York Times*, 07 jan. 2025. Disponível em: https://www.nytimes.com/live/2025/01/07/business/meta-fact-checking#meta-fact-checking-conservative-views. Acesso em: 07 jan. 2025.

SPAGNUOLO, Sergio. Com fim de checagem, Zuck arrisca transformar redes da Meta num X mais caótico. *Núcleo*, *[S.l.],* 07 jan. 2025. Disponível em: https://nucleo.jor.br/raiox/2025-01-07-com-fim-de-checagem--zuck-arrisca-transformar-redes-da-meta-num-x-mais-caotico/. Acesso em: 07 jan. 2025.

SULLIVAN, David; PIELEMEIER, Jason. Unpacking "Systemic Risk" Under the EU's Digital Service Act. *Tech Policy Press*, *[S.l.],* 19 jul. 2023. Disponível em: https://www.techpolicy.press/unpacking-systemic-risk--under-the-eus-digital-service-act/. Acesso em: 05 fev. 2025.

TAGIAROLI, Guilherme. Meta vai acabar com checagem de fatos nos EUA e copiar modelo usado pelo X.São Paulo: *UOL*, 07 jan. 2025. Disponível em: https://www.uol.com.br/tilt/noticias/redacao/2025/01/07/meta-fim-checagem-notas-da-comunidade.htm. Acesso em: 07 jan. 2025

UNIÃO EUROPEIA. Regulation (EU) 2022/2065 of the European Parliament and of the Council. Digital Services Act. Bruxelas: *Official Journal of the European Union*, 19 out. 2022a. Disponível em: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A32022R2065. Acesso em: 1 ago. 2024.

UNIÃO EUROPEIA. Regulamento (UE) 2022/1925 do Parlamento Europeu e do Conselho. Regulamento dos Mercados Digitais. Bruxelas: *Official Journal of the European Union*, 14 set. 2022b. Disponível em: https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32022R1925. Acesso em: 1 ago. 2024.

UNIÃO EUROPEIA. Digital Services Act takes effect for large online platforms. *European Union*, *[S.l.]*, 25 ago. 2023. Disponível em: <https://data.europa.eu/en/news-events/news/digital-services-act-takes-effect-large-online-platforms. Acesso em: 2 ago. 2024.

URMAN, Aleksandra; MAKHORTYKH, Mykola How transparent are transparency reports? Comparative analysis of transparency reporting across online platforms. *Telecommunications Policy*,*[S.l.]*, v. 47, n. 3, abr. 2023. Disponível em: https://doi.org/10.1016/j.telpol.2022.102477. Acesso em: 19 dez. 2024.

VAN DIJCK, José; POELL, Thomas; DE WAAL, Martijn. 'The platform society as a contested concept'. In: *The Platform Society*: Public Values in a Connective World. Nova York: Oxford University Press, 2018, pp. 7–30.

VERMEULEN, Mathias. Reading the European Commission's Proposed Implementation of DSA Article 40: Six Initial Observations on a New Framework for Research Data Access. *Tech Policy Press*, *[S.l.]*, 29 out. 2024. Disponível em: https://www.techpolicy.press/reading-the-european-commissions-proposed-implementation-of-dsa-article-40-six-initial-observations-on-a-new-framework-for-research-data-access/. Acesso em: 04 fev. 2024.

VIE PUBLIQUE. DMA: le règlement sur les marchés numériques ou Digital Markets Act. *Vie Publique*, *[S.l.]*, 13 maio 2024. Disponível em: https://www.vie-publique.fr/eclairage/284907-dma-le-reglement-sur-les-marches-numeriques-ou-digital-markets-act. Acesso em: 1 ago. 2024.

WERNER, Jeremy. EU Commission Identifies Preliminary Breaches of Digital Services Act by X (formerly Twitter). *Babl*, 07 dez. 2024. Disponível em: https://babl.ai/eu-commission-identifies-preliminary-breaches-of-digital-services-act-by-x-formerly-twitter/. Acesso em: 04 fev. 2025.

YE, Zhen; HUANG, Qian; KRIJNEN, Tonny. Douyin's playful platform governance: Platform's self-regulation and content creators' participatory surveillance. *International Journal of Cultural Studies*, *[S.l.]*, v. 28, n. 1, p. 80-98, 2025. Disponível em: https://doi.org/10.1177/13678779241247065. Acesso em: 06 jan. 2025.

ZENG, Eric; KOHNO, Tadayoshi; ROESNER, Franziska. What makes a "bad" ad? user perceptions of problematic online advertising. New York: *Association for Computing Machinery*, Article No.: 361 p. 1-24, . 2021. Disponível em: https://dl.acm.org/doi/abs/10.1145/3411764.3445459 Acesso em: 4 fev. 2025.

ZORNETTA, Alessia. Is The Digital Services Act Truly A Transparency Machine? *Tech Policy Press*, *[S.l.]*, 11 jul. 2024. Disponível em: https://www.techpolicy.press/is-the-digital-services-act-truly-a-transparency-machine/. Acesso em: 08 jan. 2024.

# Conclusão

## Da cultura maker à cultura faker: a normalização da fraude e da publicidade enganosa em ambientes digitais

A história mostra que há uma correlação entre o surgimento de uma tecnologia e a imediata inovação nos métodos dos crimes (McGuire, 2017). Inúmeras fraudes foram perpetradas com o auxílio de inovações tecnológicas de sua época. Por exemplo, a invenção da prensa por Gutenberg em 1440 revolucionou o mundo da comunicação mas também abriu oportunidade para novas formas de golpes e falsificações, inundando o mercado de livros com cópias piratas e obras adulteradas de editores da época (Whittaker, 2024). Da mesma forma, houve um aumento de grupos falsificadores de dinheiro em Londres desde pelo menos o começo do século XVIII quando a moeda passou a ser impressa – muitos deles com conhecimento técnico avançado. No mundo contemporâneo, a maioria das fraudes ocorrem online. No Brasil, a maioria dos golpes contra o consumidor começam em alguma das plataformas da empresa Meta, dona do Facebook, Instagram e Whatsapp (Silverguard & SOS Golpes, 2024).

Em seu influente livro "Os Meios de Comunicação como Extensões do Homem", publicado em 1964, Marshall McLuhan argumenta que a tecnologia estende a agência humana ao expandir as oportunidades do que seria capaz sem o seu uso. Portanto, no caso dos crimes online, a tecnologia amplifica as capacidades do fraudador contemporâneo que conseguem usar os dados pessoais dos usuários para distribuir de forma certeira e microssegmentada os anúncios fraudulentos. E o serviço que as plataformas oferecem para esses anunciantes é justamente conseguir

atingir as vítimas ideias sob o menor custo, ou seja, impactando somente aqueles usuários que demonstraram interesses e necessidades específicas, o que por si só já justificaria a necessidade de uma regulamentação.

Além de ampliar as capacidades, a tecnologia também pode resultar em "amputações" de vários tipos. No contexto das plataformas digitais, os usuários possivelmente sacrificam algumas faculdades mentais e cognitivas, como concentração, memória e tomada de decisão, em favor da conveniência e da velocidade das transações oferecidas pelos sistemas de recomendação online. Nesse contexto, a teoria da extensão também pode ser útil para embasar a necessidade de proteção do consumidor nas redes sociais sob o ponto de vista do aumento de sua vulnerabilidade em relação às fraudes. Em vez de os consumidores visitarem uma loja física tradicional para inspecionar um produto e verificar se o que pretendem comprar realmente existe e tem qualidade suficiente, os consumidores usam plataformas de vendas online. Assim, amputam seus sentidos e estão mais dispostos a "confiar" que o vendedor é legítimo e não pretende fraudá-los, criando as condições de possibilidade para o engano em um ambiente de completa assimetria de informação e opacidade. Dito de outra maneira, a ignorância do consumidor é um fator crucial para a eficácia da publicidade enganosa em escala massiva.

A arquitetura tecnológica e o modelo de negócios das redes sociais online propiciou muitas oportunidades para fraudadores. Por um lado, assistimos a um fenômeno de digitalização e industrialização de antigas fraudes (Button & Cross, 2017) porém com uma novidade: a lógica do estelionato e da falsificação é incorporada ao coração do modelo de negócios das empresas mais poderosas do mundo atual, as *big tech*, que passam a lucrar com isso. Pesquisas acadêmicas têm mostrado como os fraudadores têm buscado se afastar dos crimes presenciais, e como esses comportamentos ilícitos têm migrado para plataformas de redes sociais, de comércio eletrônico e de mensageria (Gillespie, 2017). Por outro lado, os estelionatários conseguiram criar uma nova modalidade de fraude que não existia antes, que é a promoção de produtos e serviços falsos feitos "sob medida" (Button & Cross, 2017; McGuire, 2018). A oportunidade

de comprar anúncios microssegmentados para atingir consumidores com vulnerabilidades específicas e usar sistema de pagamentos instantâneos como o PIX para lavar dinheiro roubado de vítimas em redes sociais são dois exemplos notáveis dessas novas formas de "crime organizado digital". Com isso, vimos surgir a indústria da desinformação digital estruturada a partir da arquitetura fragmentada da internet (Benkler, 2006), que facilita tanto a distribuição de conteúdos fraudulentos como o anonimato de seus agentes.

Após anos estudando anúncios em plataformas digitais, coletamos inúmeras evidências que mostram que o setor de publicidade digital está repleto de fraudes, golpes e manipulação. Durante as últimas décadas assistimos a desidratação financeira e queda de credibilidade do jornalismo profissional e de fontes tradicionais de informação. Enquanto isso, o modelo de negócios da desinformação online financiado por publicidade gera cada vez mais dinheiro. Ou seja, a publicidade digital está sendo instrumentalizada para vender e ao mesmo tempo financiar a desinformação.

A natureza predatória da desinformação vem criando um caos sem precedentes no ecossistema de comunicação e informação no mundo, no qual a publicidade enganosa se tornou onipresente. E o mercado de mídia, os anunciantes legítimos, mas especialmente, os consumidores e usuários estão cada vez mais vulneráveis a todo tipo de manipulação neste ambiente online sem lei. Porém, mais do que um fenômeno tecnológico, político ou econômico, estamos diante de uma transformação sociocultural. O *zeitgeist* de nossa época é marcado pelo que estamos chamando de "cultura faker" - engajamento acima de tudo e monetização acima de todos - na qual fraudes, golpes, estelionato e a comercialização de produtos e serviços falsos passam a ser hipernaturalizados. Ou seja, argumentamos que estamos vivenciando uma mudança subjetiva onde a normalização da mentira e do charlatanismo vai anestesiando qualquer capacidade crítica e de indignação da população diante da desonestidade e da trapaça no cotidiano dos conteúdos consumidos online.

Entretanto, é necessário produzir ignorância de forma ininterrupta para evitar a revolta contra as fraudes e assim assimilar a cultura *faker*. A agnotologia revela como a ignorância sobre determinados assuntos é fabricada, mantida e disseminada intencionalmente no tecido social por interesses políticos, econômicos e culturais. O estudo da ignorância de Robert N. Proctor e Londa Schiebinger (2008) revela não apenas a ausência de conhecimento, mas também os complexos mecanismos cognitivos, sociais e políticos que a produzem, a mantêm e a exploram. E a história vem mostrando a importância dos agentes econômicos e empresas em disseminar desinformação, esconder informação ou criar dúvidas sobre seus serviços, produtos ou práticas para proteger seus lucros - como por exemplo, a indústria do tabaco negou por décadas os danos do cigarro assim como as *big tech* vem negando a onipresença da publicidade fraudulenta em suas plataformas.

De fato, a ignorância e a opacidade são parte constituinte do setor de publicidade online. Os anunciantes, principais clientes desse mercado, sabem muito pouco sobre o que acontece com seus recursos econômicos e simbólicos nas plataformas digitais. Os anunciantes legítimos, desde grandes marcas até pequenos negócios, desconhecem onde seus anúncios estão sendo distribuídos online, se ou quando seus investimentos acabam financiando a desinformação. Por exemplo, o anúncio pode estar ao lado de um conteúdo tóxico ou fraudulento, confundindo o consumidor, ou mesmo ser veiculado dentro de um vídeo com informações falsas, o que pode comprometer a credibilidade da própria publicidade. Uma vez que os anunciantes desconhecem os contextos em que a publicidade aparece para seus públicos-alvo, não conseguem dimensionar como isso afeta a percepção negativa de sua marca, de seus produtos e de sua reputação. Isso ocorre porque a publicidade microssegmentada online não é pública, mas sim personalizada para cada usuário. Logo, a publicidade nas redes sociais não é verificável ou auditável por agentes externos.

Os únicos dados e métricas que os anunciantes possuem sobre a distribuição e performance de seus anúncios são fornecidos pelas próprias plataformas que os vendem. Portanto não sabemos o quão enviesados ou

viciados essas medições podem estar, considerando que são as mesmas empresas que comercializam os anúncios, criam suas métricas de performance e controlam os preços do mercado por meio de leilões que elas mesmas administram e que são completamente opacos. As *big tech* trabalham para tornar a compra de anúncios cada vez mais opaca e consequentemente mais cara, sem transparência de dados. A ideia de valor aqui está diretamente ligada à falta de transparência. Ou seja, quanto mais opaco o processo, mais superestimados podem ser os preços. O preço dos anúncios é definido individualmente por cada plataforma digital em leilões operados por elas mesmas, que são completamente obscuros e suscetíveis a diversos tipos de manipulação e viés dado que são as próprias empresas de redes sociais que controlam todas as pontas da cadeia de valor setorial (Van Looy, 2016).

Do ponto de vista do consumidor, a publicidade direcionada fornece rios de dados às *big tech,* porém as pessoas não sabem o que é feito com suas informações pessoais e comportamentais. Além disso, os anúncios personalizados exploram a vulnerabilidade das pessoas, e isso se torna mais grave em países onde há mais desigualdade e injustiça social, como é o caso do Brasil. Os usuários são classificados, categorizados e rankeados em centenas de modelos algorítmicos com base em suas preferências e padrões de interesse. Isso estabelece uma base poderosa para campanhas publicitárias legítimas, mas também abastece seus primos mais predatórios, os anúncios fraudulentos. Os anunciantes maliciosos pagam para as plataformas identificarem com precisão as carências e vulnerabilidades das pessoas e se aproveitam disso para explorar comercialmente, vendendo promessas falsas e exageradas por meio da publicidade digital. O resultado é a perpetuação de desigualdade e injustiça social, especialmente quando se trata de vulnerabilidades financeiras e de saúde. Segundo Cathy O'Neil (2020, p. 112) "em qualquer lugar onde houver a combinação de ignorância e grande carência, provavelmente veremos anúncios predatórios". Ou seja, a publicidade enganosa se concentra nos mais necessitados, porém em uma escala sem precedentes. Isso faz com que as políticas públicas voltadas para as camadas mais pobres da população sejam os

temas mais comumente instrumentalizados nesses anúncios, como vimos nos estudos apresentados neste livro.

Os anúncios fraudulentos exploram a necessidade das pessoas por empréstimos, perdão de dívidas e programas de assistência social, deixando os cidadãos que já tinham necessidades financeiras soterrados em montanhas de dívidas ainda maiores. Na maioria das vezes, os alvos não têm ideia de como foram enganados porque a publicidade online é camuflada no meio do conteúdo orgânico, é opaca, não é pública, e desaparece depois de ser vista. As vítimas raramente descobrem como foram escolhidas ou como os recrutadores sabiam tanto sobre elas. Aqui encontramos mais uma instrumentalização da produção e disseminação da ignorância que gera dinheiro a diferentes agentes econômicos.

Durante a pesquisa, encontrando as evidências dos estudos aqui apresentados, nos questionamos: a sociedade deveria aceitar que as *big tech* ganhem dinheiro com fraudes e crimes como algo normal ou se indignar? Sabemos que essa questão não pode ser tratada somente do ponto de vista da teoria do direito ou da economia. O campo da comunicação tem muito a contribuir enquanto área de conhecimento especializado sobre as mídias e a publicidade, embasando instrumentos regulatórios e políticas públicas com evidências empíricas sobre o fenômeno, e esse foi o nosso objetivo neste livro.

Existem também outras questões que contribuem para o aumento da vulnerabilidade e risco dos usuários nas plataformas digitais. Por exemplo, os processos de verificação das plataformas sobre seus "clientes" - sejam eles usuários (audiência) ou anunciantes - são muito limitados e insuficientes. As plataformas costumam aceitar qualquer tipo de novo usuário como meio de aumentar o máximo possível o valor de sua audiência, principalmente quando utilizam um modelo de negócios *freemium*. Ou seja, é possível criar infinitas contas gratuitamente sem quaisquer procedimentos formais de verificação, abrindo espaço para uma imensidão de bots e páginas falsas. Esses perfis inautênticos são utilizados para comprar anúncios fraudulentos, falsificar a popularidade de conteúdos e manipular as métricas de audiência em determinados conteúdos. No caso do mo-

delo de negócio das *big tech* baseado em publicidade, há um utilitarismo ainda pior em relação à falta de verificação. As plataformas online costumam aceitar qualquer tipo de anunciante sem verificar a autenticidade de quem está comprando e pagando pelos anúncios. Para piorar, a dinâmica de precificação e mensuração de performance dos anúncios fraudulentos são extremamente opacos. A verificação dos perfis deveria ser feita de modo a assegurar a legitimidade do usuário, porém protegendo o sigilo e proteção dos dados pessoais, e a verificação dos anunciantes deveria ser mais rigorosa para dar a segurança aos consumidores.

Assim as *big tech* viabilizaram uma lógica de mercado onde ganham dinheiro com todo tipo de publicidade, tanto com a legítima como com a tóxica e fraudulenta, monetizando qualquer conteúdo e audiência. Na verdade, a desinformação em forma de anúncio gera uma dupla fonte de receita para as plataformas. Primeiro, essas empresas ganham com os anunciantes fraudulentos. Segundo, também lucram com a necessidade crescente de grandes marcas e pequenos negócios de anunciar para tentar salvar sua reputação e divulgar suas verdadeiras mensagens, disputando espaço e audiência com a desinformação sobre si mesmas nas mesmas plataformas digitais. Portanto, argumentamos que as plataformas se beneficiam da publicidade enganosa por dois motivos:

1. **Para gerar receita:** Um componente chave da economia da fraude online é que os fraudadores precisam gastar dinheiro para ganhar dinheiro. Muitos sites de comércio eletrônico fraudulentos investem e reinvestem os fundos roubados de suas vítimas em campanhas de publicidade nas redes sociais, como forma de atrair tráfego para seus sites. Isso também faz com que empresas, marcas e instituições que estão sendo falsificadas tenham que investir cada vez mais em anúncios nas plataformas online também, para tentar divulgar os produtos legítimos e ocupar mais espaço que os estelionatários que os emulam.
2. **Para diminuir custos:** É simplesmente mais fácil e barato para as plataformas digitais ignorar casos de fraude e dizer que é um fenômeno residual para não ter que investir. Para o combate às fraudes é preciso investir constantemente em desenvolvimento

> tecnologia e em departamentos robustos de *Trust & Safety* para combate a abusos e crimes contra os direitos dos consumidores, que deve ser composto por pessoas preparadas, e não somente sistemas ou algoritmos generalistas. Ou seja, estamos falando de equipes de desenvolvedores de inovação e grandes times de moderação, investigação e pesquisa sobre conteúdos problemáticos.

Considerando que há uma onda contínua de criminalidade parasitária online, e que as plataformas online ganham dinheiro com anúncios fraudulentos, é imoral e preocupante que essas empresas continuem menosprezando o problema da publicidade fraudulenta e que esse tema seja tratado com pouca transparência. Quando há moderação de anúncios, isso é feito de forma aleatória e inconsistente, como mostramos nos capítulos 3 e 4 deste livro. Na maioria dos casos é necessário uma pessoa treinada para validar que uma fraude está ocorrendo, mesmo que ela tenha sido inicialmente identificada por uma máquina. Não basta simplesmente que as *big tech* retirem conteúdos fraudulentos depois destes terem sido veiculados, pois nesse momento o estrago já foi feito (e a plataforma já ganhou dinheiro). Além disso, sabemos que esse tipo de anúncio se multiplica em uma velocidade na qual em poucas horas o mesmo conteúdo reaparece.

A normalização da publicidade enganosa e do charlatanismo está criando uma cultura da fraude e do golpe financeiro que começa a afetar não somente a economia global, mas também a política. As próprias políticas públicas passam a precisar da publicidade nas redes sociais para competir com a publicidade fraudulenta, e o cidadão vê lado a lado uma publicidade autêntica e uma falsa, sem nenhuma diferença estética, semântica ou de sinalização. Esse fenômeno também tem minado o próprio ambiente da mídia e a confiança no setor de publicidade que vai se tornando sinônimo de enganação e manipulação. Além disso, a publicidade online começa a misturar vendas de produtos, fraudes e política, como mostramos nos dois estudos empíricos apresentados neste livro.

Independente do entendimento jurídico se o fato das *big tech* ganharem dinheiro com fraudes e golpes em anúncios é crime, um tipo de corrupção ou não, deveríamos enquanto sociedade entender que não se

trata de falha de mercado ou externalidade negativa, mas sim de uma crise ética e moral do tecnocapitalismo em que vivemos. Independente da tipificação criminal ou não, estamos diante de empresas de tecnologia que não tem pudor em usar os dados pessoais dos seus usuários para distribuir golpes e fraudes para os mais vulneráveis, que serão ao final roubados e perderão dinheiro. Independente da regulamentação ou não das plataformas, vivemos uma crise cultural, ética e moral onde banalizamos profundamente a trapaça, a enganação e a mentira como negócio, e as consequências disso ainda estão por ser estudadas.

Por fim, entendemos que as fraudes com anúncios em plataformas de redes sociais vão muito além dos prejuízos econômicos e riscos aos consumidores: este problema deve ser pensado, sobretudo, como uma questão cultural com implicações profundas. Essas plataformas, que figuram entre as empresas mais poderosas do mundo, lucram diretamente com a circulação de anúncios enganosos e fraudulentos. Ao não fiscalizarem, não permitirem o escrutínio público e nem punirem essas práticas de forma efetiva, as plataformas contribuem para a naturalização de formas ilegais de obtenção de renda e lucro. Isso significa que o crime passa a ser incorporado à cultura: quando os principais atores do mercado digital normalizam a receita obtida por meio de fraudes e publicidade enganosa, essas práticas deixam de ser vistas como exceção e passam a ser percebidas como parte do funcionamento "normal" do sistema. Ainda que haja um arcabouço legal destinado a coibir tais práticas, a carência de fiscalização e punição efetivas revela que o cerne do problema é, sobretudo, cultural. Quando as normas são ignoradas e as fraudes naturalizadas, as regras se tornam inúteis diante da "cultura faker" que valida a propaganda enganosa e o ganho ilícito à custa dos cidadãos.

## Referências

BENKLER, Yochai. *The Wealth of Networks*: How Social Production Transforms Markets and Freedom. Yale University Press, 2006.

BUTTON, Mark; CROSS, Cassandra. *Cyber Frauds, Scams and their Victims*. 1º Edição. Londres Nova York: Routledge, 2017.

GILLESPIE, Alisdair A. The Electronic Spanish Prisoner: Romance Frauds on the Internet. *The Journal of Criminal Law*, [*S.l.*], v. 81, n. 3, p. 217–231, 2017. Disponível em: https://doi.org/10.1177/0022018317702803. Acesso em: 2 maio 2025.

MCGUIRE, Michael R. Into the Web of Profit: Understanding the Growth of the Cybercrime Economy. *Bromium*, 2018. Disponível em: https://www.bromium.com/wp-content/uploads/2018/05/Into-the-Web-of-Profit_Bromium.pdf. Acesso em: 2 maio 2025.

MCGUIRE, Michael R. Technology crime and technology control: Contexts and history. *In*: The Routledge Handbook of Technology, Crime and Justice. [*S.l.*]: *Routledge*, 2017.

MCLUHAN, Marshall. *Os Meios de Comunicação Como Extensões do Homem*. 1º Edição. São Paulo: Cultrix, 2012.

O'NEIL, Cathy. *Algoritmos de Destruição em Massa*. Tradução: Rafael Abraham. 1º Edição. Santo André, São Paulo: Editora Rua do Sabão, 2021.

PROCTOR, Robert N.; SCHIEBINGER, Londa (org.). *Agnotology*: The Making and Unmaking of Ignorance. 1º Edição. Stanford, California: Stanford University Press, 2008.

SILVERGUARD; SOS GOLPE. Estudo Golpes com Pix 2024. *Silverguard*, 2024. Disponível em: https://static1.squarespace.com/static/672922c4b034ed7793cab948/t/673368ac9f58876b76a71175/1731422402579/Estudo+Golpes+com+Pix+2024_Silverguard_SOSGolpe.pdf. Acesso em: 2 maio 2025.

VAN LOOY, Amy. Social Media Management: Technologies and Strategies for Creating Business Value. Cham: *Springer International Publishing*, 2016. (Springer Texts in Business and Economics). Disponível em: https://link.springer.com/10.1007/978-3-319-21990-5. Acesso em: 2 maio 2025.

WHITTAKER, Jack M. The Complex Relationship Between Fraud and Technology - Should We Ignore or Regulate Online Platforms? 12º Edição. *Public Sector Counter Fraud Journal*, [*S.l.*], v. 1, n. 12, p. 21–22, 2024. Disponível em: https://philarchive.org/rec/WHITCR-7. Acesso em: 2 maio 2025.

# Sobre o Netlab UFRJ

O NetLab é um laboratório de pesquisa da Escola de Comunicação da Universidade Federal do Rio de Janeiro (ECO - UFRJ). Desde 2013, nos dedicamos a estudar o ecossistema de mídia digital no Brasil. Nos últimos anos, nossa agenda de pesquisa tem focado na indústria da desinformação digital e seus impactos econômicos e sociais no país. Nosso laboratório é formado por uma equipe multidisciplinar com mais de 40 colaboradores de diferentes áreas - Ciência de Dados, Ciência da Informação, Comunicação Social, Sociologia, Ciências Políticas, Engenharia e Computação.

A missão do NetLab UFRJ é produzir pesquisas, baseadas em conhecimento e métodos científicos, que tenham impacto social. Por meio das nossas parcerias com instituições públicas, privadas e com a sociedade civil, queremos contribuir no combate à indústria da desinformação no Brasil - que inclui campanhas de manipulação, uso de IA para apropriação de marcas e imagens, falsificação de audiências, publicidade enganosa etc. Para isso, combinamos abordagens metodológicas tradicionais das ciências sociais — como análises críticas e qualitativas de dados — com métodos digitais e computacionais, incluindo o desenvolvimento de tecnologias, diagnósticos e soluções nas áreas de ciência de dados e inteligência artificial. No âmbito do desenvolvimento tecnológico, focamos na criação e uso de modelos de machine learning, deep learning e inteligência artificial generativa. Atuamos com Processamento de Linguagem Natural (PLN), utilizando tanto métodos com-

putacionais tradicionais quanto modelos de linguagem de larga escala (LLMs), além de realizar análises de redes complexas e de dados visuais por meio de técnicas de análise e processamento de imagens.

Além disso, produzimos análise sobre o comportamento das audiências online, dieta midiática e pesquisas sobre os usos sociais das plataformas digitais para embasar novos métodos de governança e marcos regulatórios que visem tornar o consumo e a distribuição de mídia nas plataformas digitais no Brasil mais seguros para os consumidores e cidadãos. Nos dedicamos também à ampla divulgação científica de nossas pesquisas, pois acreditamos na importância da popularização da ciência para qualificar o debate público e assim contribuir para tornar o ecossistema de mídia no país mais saudável e confiável para a população.

Contato:
netlab.eco.ufrj.br/
netlab@eco.ufrj.br

net lab Laboratório de Estudos de Internet e Redes Sociais

UFRJ UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

**AUTORES**

Alékis Moreira
Arthur Mendes
Bernardo Dias
Bruno Mattos
Carlos Eduardo Barros
Danielle Pinho
Daphane Silva
Débora Salles
Felipe Grael
Fernando Ferreira
João Gabriel Haddad
Lucas Murakami
Luciane Belin
Marcela Canavarro
Marcio Borges
Matheus Gomes
Nicole Sanchotene
Priscila Medeiros
R. Marie Santini
Thiago Ciodaro
Vinicius B. Scortegagna

MINISTÉRIO DA JUSTIÇA E SEGURANÇA PÚBLICA

BRASIL

FDD
Fundo de Defesa de Direitos Difusos

net lab
Laboratório de Estudos de Internet e Redes Sociais

UFRJ
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO

*Editora Sulina*