

Crise de equidade na governança das plataformas digitais no mega ciclo de 2024 e nas eleições municipais do Brasil

Tatiana Dourado
Bruna Martins
Nina Santos
Rodrigo Carreiro
Carla Vreche

Resumo

Este artigo tem o objetivo de situar o papel das plataformas na chamada crise de equidade global da governança digital de plataformas digitais, apresentar os principais riscos colocados para o mega ciclo eleitoral, com especial atenção ao caso das eleições municipais do Brasil, e discutir os pontos de interseção entre a concepção de integridade eleitoral e a de integridade de informação. Interessamos, portanto, refletir sobre os aspectos-chave do tratamento assimétrico fornecido por essas empresas para países não falantes da língua inglesa e sobre as condições necessárias para restaurar o ambiente digital, tornando-o cada vez mais seguro para os usuários, regido por princípios universais de direitos humanos e ligado à proteção do Estado democrático de direito. Por fim, este artigo lista 7 recomendações para a agenda da governança digital no contexto de realização da Cúpula de Líderes do G-20 no Brasil.

Abstract

The aim of this article is to situate the role of platforms in the so-called global equity crisis in the digital governance of digital platforms, to present the main risks posed to the mega-electoral cycle, with special attention to the case of Brazil's municipal elections, and to discuss the points of intersection between the concept of electoral integrity and that of information integrity. We are therefore interested in reflecting on the key aspects of the asymmetrical treatment provided by these companies to non-English-speaking countries and on the conditions needed to restore the digital environment, making it increasingly secure for users, governed by universal human rights principles and linked to the protection of the democratic rule of law. Finally, this article lists 7 recommendations for the digital governance agenda in the context of the G-20 Leaders' Summit in Brazil.

Introdução

As plataformas digitais de mídias sociais vêm ganhando predominância em nível crescente e têm mudado constantemente os paradigmas dos espaços cívicos e as lógicas que estruturam os processos de formação da opinião pública globalmente. Baseados principalmente nos Estados Unidos (EUA), estes intervenientes beneficiam-se da negligência dos reguladores na abordagem do impacto da tecnologia, sejam em países regidos por formas de governos autocráticas ou democráticas, possibilitando novas experiências de recrudescimento da concentração do poder político, de arroubos autoritários e de extremismo violento na atualidade (TUCKER et al., 2017).

Os vazamentos de informações empresariais, a atuação concertada da sociedade civil especializada e os recentes relatórios das plataformas elaborados por exigência do Regulamento de Serviços Digitais, no âmbito da União Europeia (UE), revelam investimento mínimo na atividade de moderação de conteúdo fora da língua inglesa. Em 2023, na UE,

havia 41 profissionais dedicados à moderação de conteúdo no idioma português no X, contra 2.294 para língua inglesa, 81 alemã, 52 francesa e 20 espanhola (Twitter, 2023). No Facebook e Instagram, durante esse mesmo ano, o serviço para a região era prestado por 58 pessoas proficientes em português contra 226 em alemão, 242 em francês, 163 em espanhol e 106 em inglês (FACEBOOK, 2023).

Estes números contrastam com a quantidade de usuários ativos no território regulado da UE – em média, 100 milhões por mês no X, 225 milhões por mês no Facebook e 250 milhões por mês no Instagram (MARTINS, 2023) – e lançam luz para o quão defasados estão os mercados de países em desenvolvimento, tal é o caso também do Brasil. Em 2024, o país sediará uma das quase setenta eleições que acontecerão em todo o mundo, um mega ciclo eleitoral que tem sido objeto de preocupação em razão da dimensão potencialmente extensa e severa de danos online e da reconhecida negligência das plataformas digitais em fazer cumprir e aprimorar suas próprias políticas e diretrizes de forma também global. O Facebook, que informa estar disponível em mais de 100 idiomas, possui equipe de moderação de conteúdo capaz de compreender apenas 50 idiomas e suas ferramentas de detecção automatizadas contra discurso de ódio estão disponíveis somente para 30 idiomas (FACEBOOK, 2020). Adicionalmente, informações reveladas pelo Facebook Papers, em 2021, evidenciaram estas discrepâncias, como o fato de 84% dos investimentos alocados para conter mis- e desinformação¹ no Facebook e Instagram terem sido usados exclusivamente nos Estados Unidos (ZAKRZEWSKI et al., 2021).

A partir deste quadro, este artigo tem o objetivo de situar o papel das plataformas na chamada crise de equidade global da governança de plataformas digitais, apresentar os principais riscos colocados para o mega ciclo eleitoral, com especial atenção ao caso das eleições muni-

1 Considerar misinformation (informação incorreta) como informação equivocada e falsa, cujo engano é parte da forte crença depositada pelo indivíduo, e desinformação como fluxo de mensagens interessado com objetivo de influenciar outros e de gerar enganos deliberados sobre o entendimento da realidade social (DOURADO, 2021).

cipais do Brasil, e discutir os pontos de interseção entre a concepção de integridade eleitoral e a de integridade de informação. Interessa-nos, portanto, refletir sobre os aspectos-chave do tratamento assimétrico fornecido pelas plataformas digitais para países não falantes da língua inglesa e sobre as condições necessárias para restaurar o ambiente digital, tornando-o cada vez mais seguro para os usuários, regido por princípios universais de direitos humanos e ligado ao Estado democrático de direito. Por fim, este artigo lista 7 recomendações para a agenda da governança digital.

Crise de equidade global na governança de plataformas digitais

Assim como as plataformas digitais de mídias sociais têm ocupado um espaço central no dia a dia de usuários, crescentemente, têm também afetado muitos contextos eleitorais. Ao mesmo tempo em que são vias para acesso a conteúdos com qualidade informativa e ampliação do alcance a uma pluralidade de vozes, são também arenas que favorecem a circulação rápida, frequente e em escala de discursos de ódio, perigosos e conspirações prejudiciais (ver MARWICK e LEWIS, 2017). Neste sentido, a mesma infraestrutura que suporta a mobilização democrática pode ser manipulada por atores mal-intencionados para desestabilizar a integridade eleitoral e as democracias.

É diante desse contexto desafiador que o papel das políticas de moderação e dos termos de uso das plataformas em equilibrar proteção de dados, transparência, liberdade de expressão e contenção de campanhas de desinformação vem sendo discutido amplamente pela comunidade acadêmica, sociedade civil e governos em todo o mundo. Essa crescente relevância se deve a diversos efeitos percebidos em eleições anteriores, como a proliferação de notícias fraudulentas (fake news), o aumento da polarização política e o aumento do poder das plataformas na sociedade (MARRET, 2022; OVERGAARD e COLLIER, 2023). O cenário, portanto, denota uma preocupação recorrente com a integridade da informação e

o seu impacto na manutenção de ambientes digitais mais democráticos e saudáveis, principalmente em período eleitoral.

Políticas de integridade eleitoral

Plataformas digitais têm criado mecanismos específicos de contra-golpe a problemas decorrentes de proliferação de mis- e desinformação em eleições, o que podemos chamar de políticas de integridade eleitoral. Estes documentos abrangem vários aspectos relativos à moderação de conteúdo, como parcerias com programas de verificação de fatos, redução de alcance, provimento de contexto informativo a usuários, interação entre moderação humana e de Inteligência Artificial (IA), incorporação de padrões de direitos humanos às políticas da plataforma, e entendimento sobre liberdade de expressão para aplicação global (ver KATZENBACH et al, 2023).

Processos eleitorais passados, entretanto, demonstram que essas políticas são insuficientes, assim como as medidas e as cooperações implementadas para coibir, com rigor, a instrumentalização técnica da disseminação artificial e a amplificação prejudicial de mensagens que visam manipular o eleitor (ver BALAKHONSKAYA et al., 2020; CARREIRO, ALMADA, DOURADO, 2024). No Brasil, ainda que, desde 2018, a Meta (controladora do Facebook, Instagram, Whatsapp e, a partir de 2023, também Threads), Alphabet (controladora do Google e YouTube) e Twitter (hoje, X) tenham adotado providências nesse sentido, o problema não só permanece ativo como ganhou novos contornos na medida em que novas práticas são adotadas e novas plataformas de alcance popular, a exemplo de Kwai e TikTok, passam a moldar a forma como as pessoas interagem e constroem suas dietas informativas digitalmente.

Um dos principais obstáculos observados no atual cenário é a negligência com que as plataformas encaram os mercados do Sul-Global. Em muitos casos, estas empresas criam orientações que são válidas somente para os EUA, não tendo aplicação estendida para eleições em países com tensões diversas. Ou seja, o terreno segue sem moderação de conteúdo

suficiente ou qualquer outra ação destinada a conter sublevações democráticas e violência política durante eleições – como ocorrido no Brasil em 2022, parte considerável das políticas de integridade eleitoral foi replicada sem adaptações às particularidades do contexto nacional (SAD, 2022).

A presença de redes de contas e a difusão extenuante de mensagens anti-urna eletrônica, maior narrativa usada para desinformação em 2018, 2020 e 2022, e pró-intervenção militar, com releitura do Artigo 142 da Constituição Federal, são exemplos de tópicos que, no Brasil, atraem apoio popular para tentativas de ruptura democrática (ver Linha do Tempo – Instituto Democracia em Xeque, 2024). De acordo com o InternetLab, políticas de plataformas que visem proteger a integridade eleitoral brasileira não podem prescindir de agir contra alegações infundadas de fraude eleitoral, controle da circulação de anúncios eleitorais, estabelecimento de regras para moderação ágil e precisa de posts violativos publicados por figuras públicas e proibição de conteúdos contendo incitação e defesa da violência contra a ordem democrática (BORGES, BRITO CRUZ, CINTRA, 2023). As limitações das políticas das plataformas ficam evidentes, portanto, ao não considerarem aspectos específicos dos contextos nacionais e de eleições de nações socialmente distintas, politicamente tensas e economicamente desiguais.

Um outro problema a ser enfrentado é a falta de uniformidade das regras, termos de uso e políticas de integridade eleitoral entre diferentes plataformas. Além do fato desses documentos serem transportados quase que integralmente a diferentes realidades das mais diversas nações, como já abordado, há uma considerável confusão de escopo entre cada uma dessas orientações. No X, por exemplo, há uma política de integridade cívica, outra para comportamento abusivo e sobre propagação de ódio, além de inúmeras diretrizes de atuação e moderação em casos de comportamento inautêntico, desinformação política, fraude eleitoral, dentre outros – todas com implicações no período eleitoral. No Youtube não é diferente, pois também existem políticas contra desinformação, sobre conteúdo violento e contra discurso de ódio etc. Isso

revela o quão fragmentadas estão as diretrizes de orientação de uso e moderação, o que atinge diretamente tanto o modo como os usuários utilizam as plataformas quanto a forma como as *big techs* organizam-se internamente para estruturar seus modos de atuação para conter abusos contra os sistemas eleitoral e democrático. Levantamentos anteriores já haviam destacado que nem políticas nem processos parecem bem estruturados pelas plataformas digitais para lidar com moderação de conteúdo relativa à desinformação (BARBOSA, HELENA, VALENTE, 2020).

Os problemas se aprofundam ao considerar que contas oficiais de políticos e candidatos a cargos eletivos não estão sujeitas a estas políticas, mesmo que violem as diretrizes da comunidade criadas pelas próprias empresas, aspecto transposto para um modelo de regulação e considerado ponto crítico do Projeto de Lei (PL) 2630/2020 (ITS RIO, 2022). Plataformas digitais têm sido, todavia, vastamente utilizadas como meios propagadores de peças de desinformação como fake news e discurso violento ou de ódio por parte, justamente, dessas figuras públicas e políticas². São essas contas que, reiteradamente, amplificam ideários e publicações que violam regras e termos de uso, ameaçando a integridade eleitoral. Para isso, políticos e candidatos ancoram-se na perspectiva de que não serão moderados por uma espécie de imunidade oferecida pelas próprias plataformas.

Por fim, o cenário brasileiro é particularmente preocupante se comparado com o volume de recursos e esforços que vêm sendo direcionados para proteger o sistema eleitoral e a democracia norte-americana em 2024. Os EUA são o país que receberá a maior parte do investimento da Meta em moderação de conteúdos, ferramentas, tecnologia de IA e outras ferramentas internas e externas. Como em 2020, em 2024, o Facebook não permitirá anúncios políticos na última semana da campanha eleitoral nos EUA. Da mesma forma, anúncios de órgãos midiáticos controlados por Estados serão bloqueados somente quando mira-

2 Ver <https://noticias.uol.com.br/colunas/juliana-dal-piva/2023/08/08/1-a-cada-5-disseminadores-de-fake-news-sao-politicos-metade-deles-e-do-pl.htm>

rem indivíduos nos EUA (META, 2023). Ou seja, apesar do alcance global, as plataformas tendem a ser influenciadas pela pressão dos EUA ao conceber as suas políticas globais de moderação de conteúdo. “Conteúdo violativo que não seja em inglês tem menos probabilidade de ser visto por moderadores humanos para revisão; portanto, alguém que usa uma plataforma de mídia social em um idioma diferente do inglês pode ter maior probabilidade de ser exposto a desinformação prejudicial” (Mozilla, online). Dado que as externalidades negativas da utilização das redes sociais afetam muitos países, é importante que essas empresas definam critérios para reduzir as disparidades e promover uma maior equidade de investimento em todo o mundo.

Problemas de transparência

Nos últimos anos, Twitter, YouTube e TikTok desenvolveram programas voltados para a comunidade científica, incluindo APIs específicas para pesquisadores acadêmicos, mas essas ferramentas têm sido encerradas ou alteradas substancialmente, passando de gratuitas para pagas, a preços considerados excessivamente caros. Outras plataformas, como Meta, com Facebook e Instagram, dependem exclusivamente de aplicativos de terceiros, como o Crowd Tangle, cujo desmonte já vinha sendo realizado ao longo dos anos (MENDOZA, 2022) até o recente anúncio de que será definitivamente descontinuado em agosto de 2024³⁴. Mais do que nunca, o acesso a dados de plataformas digitais, que são também ambientes de pesquisa, tem sido dificultado ou frontalmente impedido, tornando este obstáculo para o desenvolvimento de pesquisas sobre eleições nos ambientes digitais um problema mais recente.

3 Important update to CrowdTangle, March 2024. (2024, March). CrowdTangle Help Center. Retrieved from <https://help.crowdtangle.com/en/articles/9014544-important-update-to-crowdtangle-march-2024>

4 Em substituição, a Meta firmou acordos com uma instituição acadêmica dos Estados Unidos responsável por centralizar a concessão de credenciais para acesso a uma biblioteca e a API de conteúdo (SOARES, 2024).

Diante deste cenário, diversas entidades têm buscado reforçar a importância de existir políticas de transparência que abranjam acesso a dados para desenvolvimento de pesquisas e para escrutínio público por parte da mídia e da sociedade civil. Em carta aberta contendo dezenas de signatários, a Fundação Mozilla enfatizou o prejuízo que o encerramento do CrowdTangle trará para os esforços de prevenção de problemas, tais como incitação à violência, desinformação política ou assédios contra candidatas mulheres nas eleições nos EUA, Brasil, Índia e outros países⁵. Somado a isso, a falta de regularidade e de padronização dos relatórios de transparência e das informações fornecidas nestes documentos para mercados não regulados, como o Brasil, agravam o problema. Nesse país, por exemplo, o X não mais fornece informações sobre contas suspensas por decisão judicial versus taxa de resposta da plataforma, isso depois de ter sido adquirido por Elon Musk (PINHO, 2024).

Além disso, muitas plataformas baseiam-se no fato de serem empresas privadas, de lidarem com dados privados, de estarem sediadas nos Estados Unidos e de supostas limitações técnicas para evitar responder a decisões de tribunais judiciais locais. Em 2022, repetidamente, o Telegram descumpriu ordens da Justiça brasileira, mesmo após aplicação de multa e ordem de bloqueio do serviço⁶. Em geral, mesmo as *big techs* que têm acordos com sistemas judiciais apresentam queixas subsequentes para o descumprimento de ordem judicial. As alegações que apresentam à Justiça são muitas vezes opacas e não estão sujeitas a qualquer tipo de sanção. Em 2018, por exemplo, tanto no Brasil quanto na Índia, diante de problemas crescentes decorrentes de desinformação online, o WhatsApp alegou uma série de dificuldades técnicas na reso-

5 Open letter to Meta: support CrowdTangle through 2024 and maintain CrowdTangle approach. (2024). Mozilla Foundation. Disponível em <https://foundation.mozilla.org/en/campaigns/open-letter-to-meta-support-crowdtangle-through-2024-and-maintain-crowdtangle-approach/>

6 Assinala-se que o pedido de bloqueio do serviço dividiu a opinião de especialistas por razões que estão apresentadas em <https://www.jota.info/coberturas-especiais/liberdade-expressao/stf-telegram-bloqueio-especialistas-18032022>

lução de problemas e na remoção de conteúdos ou usuários da plataforma, que foram rebatidas por uma série de especialistas. Em episódio mais recente, o próprio dono do X atacou o judiciário brasileiro, como será detalhado a seguir.

Por fim, na hora de definir marcos regulatórios, é igualmente importante considerar o ecossistema local de produção e consumo de conteúdo em plataformas digitais. Invariavelmente, esta visão baseia-se fortemente na percepção destas empresas por parte dos usuários norte-americanos. Como resultado, os serviços de mensagens privadas, como o WhatsApp ou o Telegram, têm sido fortemente negligenciados nas políticas regulatórias. Mesmo quando se trata de diretrizes nacionais, há uma tendência para tratar as redes sociais e os serviços móveis de mensagens instantâneas como a mesma coisa, quando, na verdade, as redes de comunicação privadas têm grandes idiosincrasias. Os MIMS, por exemplo, são fortemente permeados por uma opacidade, característica relacionada à sua moderação, mas também à própria circulação do conteúdo. O usuário final não consegue identificar onde seu conteúdo circula, quem já o visualizou e quem o encaminhou para quem, pois esses metadados são omitidos em favor do sistema de criptografia ponta a ponta. No entanto, embora esse tipo de encriptação seja útil para a comunicação *peer-to-peer*, quando se trata de grupos de discussão, estes serviços devem ser regulamentados como ferramentas de transmissão e não como redes privadas. A confusão entre estes dois modos de funcionamento provoca um afrouxamento das regras do MIMS, o que resulta na facilitação da propagação da desinformação e do discurso de ódio através de grupos de discussão. É fundamental reforçar a política de transparência das plataformas, especialmente dos serviços de mensagens privadas.

Plataformas digitais e desinformação no caso das eleições municipais do Brasil

Em 2024, as eleições no Brasil abrangerão a escolha de pessoas para os cargos de prefeito e vereador em 5.568 municípios espalhados

por todo o país. O contexto será especialmente desafiador pela evolução das tecnologias de inteligência artificial e seus usos, dada a incerteza sobre como se dará a aplicação das resoluções aprovadas pelo Tribunal Superior Eleitoral (TSE) em fevereiro deste ano, em particular a que aborda o tema da propaganda eleitoral, e pela ofensiva da extrema direita, consubstanciada pelo apoio de Elon Musk (dono do X) e de lideranças internacionais, contra as Cortes Superiores do Brasil. Este cenário de permanente tensão agita comunidades polarizadas e (hiper)partidárias mantidas em plataformas digitais que, por sua vez, dão vazão a campanhas de desinformação que inundam a internet com mensagens inflamatórias, inverídicas e prejudiciais à integridade do processo eleitoral.

No início do mês de abril, o episódio intitulado Twitter Files, no qual Micheal Shellenberger e outros autores, a partir do acesso a documentos internos do Twitter concedidos por Elon Musk, revelaram supostos arbítrios contidos em decisões judiciais sobre remoção de conteúdo nas eleições de 2022 – o que é considerado falacioso ou sem prova de autenticidade por parte da opinião pública⁷ -, elevou o grau de hostilidade já reiteradamente direcionados ao TSE e ao Supremo Tribunal Federal (STF) entre parlamentares bolsonaristas e seus apoiadores. O cenário agravou-se ainda mais depois de declarações, no X, de Elon Musk, dirigidos pessoalmente ao ministro Alexandre de Moraes, atual presidente do TSE, episódio que passou a ser entendido como abuso de poder econômico e ataque à soberania nacional por especialistas e autoridades⁸. Esta situação ensejou a decisão do presidente da Câmara, o deputado Arthur Lira (PP-AL), de criação de um grupo de trabalho para debater um novo projeto de lei, atravessando a discussão já em andamento do PL 2630/2020, considerada a proposta de regulação de plataformas

7 Ver O que são os Twitter Files, vazados pelo próprio Musk, e como eles chegaram ao Brasil <https://www.aosfatos.org/bipe/twitter-files-brasil-musk/>

8 Ver Elon Musk e a soberania econômica e nacional <https://oglobo.globo.com/opiniao/artigos/coluna/2024/04/elon-musk-e-a-soberania-economica-e-nacional.ghtml>

digitais em estágio mais avançado no Brasil. Esse ato foi fortemente repudiado pela sociedade civil⁹.

O caso levantou diversas discussões e controvérsias, mas um dos seus resultados mais diretos foi o vertiginoso ganho de visibilidade de figuras públicas da extrema direita brasileira a partir de manifestações em suas redes sociais, algumas das quais foram compartilhadas pelo próprio Musk, e a projeção internacional da acusação de “censura” e de “violação de direitos humanos” no Brasil¹⁰. No Instagram e no YouTube, por exemplo, a média de interações por postagem realizada por contas classificadas como extrema direita chegou a 9K e a 8K, respectivamente, e foram baseadas, principalmente, em críticas inflamadas contra Alexandre de Moraes; na defesa de que o Brasil não é mais democracia, e sim uma ditadura disfarçada; e na deferência a Elon Musk como defensor da liberdade (BERNARDI et al., 2024).

Em ano eleitoral, esse caso pode acirrar de forma muito perigosa a animosidade de fatias da população que aderem a ideias populistas de direita contra os poderes constituídos e contra a institucionalidade democrática, tendo em vista o histórico de forte envolvimento dos indivíduos com alegações de fraude nas urnas eletrônicas (desde 2014), com a insurreição de 8 de janeiro de 2023 em Brasília e em apoio ao plano de golpe de Estado no Brasil. Estudos têm informado que níveis elevados de percepção de polarização política, social e afetiva impulsionam o alcance de peças usadas para desinformação nas mídias sociais e os efeitos disso na sociedade, principalmente em períodos de crises e em eleições (TUCKER et al., 2018). Neste sentido, alinhado a este e a outros casos que podem vir a surgir antes do início das campanhas eleitorais oficiais, em

9 Ver Aprovação do PL 2630 é fundamental para regular plataformas e defender a democraciabrasileira <https://direitosnarede.org.br/2024/04/09/aprovacao-pl2630-fundamental-para-regular-plataformas-e-defender-democracia-brasileira/>

10 Ver em Deputados bolsonaristas viajam a Bruxelas para denunciar ‘tirania’ de Lula e ‘censura’ de Moraes <https://www.msn.com/pt-br/noticias/other/deputados-bolsonaristas-viajam-a-bruxelas-para-denunciar-tirania-de-lula-e-censura-de-moraes/ar-BB1l1bhP>

16 de agosto, entende-se que as eleições municipais de 2024 enfrentarão outros desafios relacionados às dinâmicas de campanhas de desinformação, que são:

Fragmentação e novos recursos no ecossistema de plataformas: Estão disponíveis no mercado um maior número de plataformas digitais, a serem usadas de forma significativa por campanhas eleitorais oficiais e não-oficiais durante as eleições, do que nos pleitos anteriores. Desde 2022, somaram-se à Meta (Facebook, Instagram, WhatsApp), Alphanet (Google e YouTube), ao X (então Twitter) e Telegram as plataformas de vídeos curtos (TikTok, Kwai), de compartilhamento de vídeos (Rumble), plataformas alt-tech (BitChute, Gab, Gettr, Parlet, Truth Social), outros mensageiros (Signal), Threads (aplicativo de conversas de texto adjacente ao Instagram, Meta), entre outras.

A existência de todas essas opções não significa que elas sejam efetivamente relevantes ou que seus usos afetem direta ou indiretamente o processo e a integridade eleitorais. Um ecossistema de plataformas expandido, contudo, alerta para maior possibilidade de articulação de redes politicamente radicais e extremistas, que se valem de espaços digitais com menos ou com nenhuma política de moderação de conteúdo para conectar conjuntos abrangentes de comunidades heterogêneas e para distribuir conteúdos de teor violento e dotado de discurso de ódio, o que poderia vir a ser objeto de remoção por violação de políticas em outras plataformas, por exemplo.

Um segundo ponto a ser observado é que, em 2023, o WhatsApp, X, além do Facebook e Instagram, lançaram novos recursos que alteraram consideravelmente o modo como os usuários conectam-se com grupos de interesse, organizando a distribuição de mensagens e colaborando com a finalidade de ampliar o acesso à cultura da correção de posts virais ou controversos. As ferramentas serão testadas nas eleições de 2024, e há riscos relacionados a um novo patamar de transmissão em larga escala de mensagens entre um enorme conjunto de comunidades digitais afins, ao reforço do potencial de viralização de algumas peças de

conteúdo usado por campanhas de desinformação, à facilidade de manipulação de informação e de mídias, principalmente imagens e vídeos, para distribuição hipersegmentada de anúncios; e de ocorrência de erros fruto das dinâmicas de correção colaborativa (Tabela 1).

TABELA 1. Recursos que serão testados no Brasil nas eleições municipais de 2024

	Plataforma	Recurso	Descrição
Viralização	WhatsApp	Canais	Serviço de transmissão unilateral de mensagem para usuários previamente inscritos, que detém um único dono e até 16 admins. É possível compartilhar e reagir a publicações, além de votar em enquetes.
		Comunidades	Serviço que congrega até 50 grupos e que permite envio de mensagens para até 2 mil membros e aos grupos de avisos. Usuários podem participar de enquetes, reagir, compartilhar arquivos e realizar chamadas de voz.
Manipulação da informação	Facebook e Instagram	Ferramenta de criação de anúncios com uso de IA generativa	Ferramenta permite uso de IA generativa para criação de anúncios, com possibilidade de geração de plano de fundo, expansão de imagem e, para a língua inglesa, variação de texto.
Ação corretiva e legitimação de erros	X	Notas da Comunidade	Programa que formaliza, por meio de inscrição e aprovação, usuários para redigirem e avaliarem notas, que são anexadas ao post para oferecer mais contexto à informação disseminada. Pelas regras, a inserção de referência nas notas resulta de acordo entre os usuários inscritos e de perspectivas divergentes ¹ , e o algoritmo de avaliação é passível de ser inspecionado ² .

Fonte: Adaptado de *O Globo* e de sites da Meta, WhatsApp e X.

Tecnologias de inteligências artificiais: A IA generativa reposicionou a inteligência artificial como uma das principais ameaças à integridade eleitoral pela maneira facilitada e acelerada como *deep fakes* podem ser criadas e espalhadas. As *deep fake* de áudio e vídeo, prática que usa a IA para sintetizar vozes e rostos e produzir conteúdos novos e não reais, tem sido crescentemente utilizada no Brasil e no mundo. Desde o pe-

11 Ver em <https://communitynotes.x.com/guide/pt/contributing/diversity-of-perspectives>

12 Ver em <https://communitynotes.x.com/guide/en/under-the-hood/download-data>

ríodo eleitoral de 2022, diversos conteúdos passaram a ser manipulados e distorcidos por meio dessa tecnologia. Com a função de criar vídeos e áudios realistas (mas não reais), essa técnica de manipulação faz com que pessoas apareçam em vídeo e em áudio fazendo e dizendo coisas, inclusive, inverídicas. Os graves impactos da *deep fake* em eleições já foram registradas, segundo a imprensa, em outras partes do globo, como em Bangladesh, Argentina, Reino Unido, Estados Unidos e Coréia do Sul, mas elas também já demonstraram sua força durante o processo eleitoral de 2022 no Brasil, na produção de conteúdo falso sobre as pesquisas de intenção de voto. Notícias na imprensa brasileira ainda dão conta de que existe suspeita de alteração de áudios por essa técnica já no atual período pré-campanha de 2024. Além de confundir os cidadãos, elas poderão impactar negativamente o processo, servindo para espalhar misoginia, racismo e ódio contra grupos sociais historicamente afetados, causando um aprofundamento do problema da violência política, por exemplo. O baixo custo e a capacidade de produção de campanhas de desinformação em grande escala a partir de ferramentas de IA generativa poderão afetar não apenas o direito à informação, mas também à privacidade, como destacado pelo Alto Comissário dos Direitos Humanos em discurso na Universidade de Stanford¹³.

Eleições presidenciais dos Estados Unidos: A cada quatro anos, as campanhas eleitorais dos Estados Unidos e do Brasil ocorrem concomitantemente. Apesar da tradição apontar que, na corrida municipal, eleitores voltam-se essencialmente para em assuntos de nível local e que afetam diretamente o cotidiano dos cidadãos em suas cidades, desde 2020, tem-se dado mais atenção mediática às eleições norte-americanas e observado um incremento da interferência de valores e temas usados pelo trumpismo na agenda de canais dogmáticos e hiperpartidários do Brasil. Em 2020, por exemplo, as alegações falaciosas de fraude eleitoral

13 Disponível em: <https://www.ohchr.org/en/statements-and-speeches/2024/02/human-rights-must-be-core-generative-ai-technologies-says-turk>

foram o primeiro tópico de repercussão entre os bolsonaristas na onda de ataques ao TSE, cujo sistema foi alvo de hackers, e dos clamores para implementação de voto impresso para as próximas eleições, no período de tramitação da então PEC do Voto Impresso (135/19), que também foram bastante volumosos (RUEDIGER, GRASSI, 2020).

É de se notar que as raízes do bolsonarismo têm o trumpismo como principal referência do conservadorismo radical na atualidade por força da influência do seu principal ideólogo, Olavo de Carvalho, que tinha Donald Trump como a grande liderança da frente contra o “globalismo comunista” (SINGER et al., 2020). Em 2024, a campanha presidencial de Trump persiste na acusação de interferência eleitoral, perseguição política e persecução judicial; ela ainda investe na comunicação que associa os Democratas a criminosos, terroristas, narcotraficantes, gangsters e corruptos; explora o sentimento pró-Israel; e destaca nomes como o de Elon Musk¹⁴. Embora este pacote de referências e discursos possam não ter, de todo, utilidade para campanhas à Prefeitura e à Câmara Municipal nas eleições brasileiras, elas são parte do sistema de crenças dos eleitores que se alinham a nomes ligados ao bolsonarismo e, assim, podem ser levadas em conta para sublinhar diferenças entre candidaturas nas principais capitais do país.

Recomendações

Integridade eleitoral e integridade da informação nas plataformas digitais

Desde 2023, a Organização das Nações Unidas vem pautando o termo integridade da informação como espécie de princípio normativo de governança digital, a fim de reposicionar a agenda que passou a ser construída, principalmente a partir de 2016, sobre a interferência nega-

14 Tópicos destacados a partir da observação, em abril de 2024, da página oficial do republicano Donald Trump no Truth Social.

tiva de campanhas de desinformação, operações de influência, manipulação de mídia, discurso de ódio e ecossistemas hiperpartidários da direita radical na esfera pública e nas democracias. Tal termo tem sido cada vez mais discutido e utilizado, sobretudo por organismos internacionais e por governos, como o do Brasil. No documento da ONU que dá notoriedade global ao conceito de forma vinculada às discussões regulatórias das plataformas digitais, integridade da informação é situada como “acurácia, consistência e confiabilidade da informação [que] está ameaçada pela desinformação, pela informação incorreta e pelo discurso de ódio”.

Em fevereiro de 2022, o Programa das Nações Unidas para o Desenvolvimento (PNUD) havia publicado o documento “Information Integrity: Forging a Pathway to Truth, Resilience and Trust”, em que conceituou integridade da informação como sendo “determinada pela precisão, consistência e confiabilidade do conteúdo, dos processos e sistemas de informação para manter um ecossistema de informação saudável”. Por um lado, assinala-se que nem sistemas de mídia, que diferem significativamente contexto a contexto, nem internet e plataformas digitais, cujos usos e os efeitos nas sociedades relacionam-se a múltiplos fatores combinados, são capazes de garantir o fornecimento de informação íntegra nos termos almejados. Por outro, nota-se também que, tendo sido os processos eleitorais os momentos mais afetados por campanhas de desinformação e abusos online, o bom funcionamento democrático presume a existência de eleitores bem informados. Neste sentido, integridade da informação e integridade eleitoral são discussões que podem ser entendidas como interligadas.

Esta última, entretanto, possui base mais ampla, que envolve uma série de requisitos para a realização de eleições íntegras. Entre os requisitos considerados estão aqueles relativos ao ambiente comunicacional. Para Frederico Alvim, ao discutir-se integridade eleitoral, deve-se considerar a qualidade da informação e a relação que ela estabelece com a liberdade de sufrágio e a igualdade de oportunidades nas eleições. Aqui, não é considerada a integridade da informação per si, mas como um

determinado ambiente comunicacional pode influenciar a integridade eleitoral (ALVIM, 2015). De modo específico, integridade eleitoral está relacionada também à sensação de segurança e confiança do cidadão. Para que possa participar da vida pública a partir da apresentação de candidatura ou do ato de votar, entende-se como essencial que o ambiente seja equilibrado e que as pessoas não tenham medo de se manifestar.

Como discutido ao longo do artigo, entende-se que, a despeito de antagonismos sociais e de disputas políticas serem partes naturais e legítimas da convivência democrática, ambientes informativos e processos comunicacionais seguros e protegidos, hoje, dependem de um novo ator, que são as plataformas digitais, particularmente as *big techs*. Apesar disso, essas empresas não têm visado a equidade de investimento, e de forma proporcional ao tamanho do mercado, em moderação de conteúdo, avaliação de riscos e mitigação de danos fora da língua inglesa e do território dos Estados Unidos. Isto enseja a reflexão sobre a existência de uma crise de equidade global. Ao tratarem certos países de forma secundária, principalmente os situados no Sul Global, plataformas digitais tornam-se agentes ativos de riscos impostos a estas democracias, particularmente expostos por atores da direita radical transnacional organizada.

Conclusão e Recomendações

Diante do quadro apresentado, propõem-se, nesta seção, recomendações para a (re)construção da agenda de governança de plataformas digitais de mídias sociais que abordam 7 eixos: padrões de transparência, abordagens regulatórias, políticas digitais, integridade eleitoral, participação digital inclusiva, capacitação digital e mitigação dos danos decorrentes de campanhas de desinformação. Assinala-se como pontos prioritários:

1. Chamada para Padrões de Transparência: Estabelecer padrões de transparência, que considerem nível e qualidade das informações apresentadas, para as plataformas digitais, de modo a aumentar a

capacidade de supervisão do funcionamento e das atividades desses sistemas e oportunizar processos de responsabilização das *big techs* (*tech accountability*);

2. Clareza dos frameworks regulatórios: Estabelecer padrões orientadores para a definição de princípios centrados na participação do cidadão e papéis e responsabilidades para todas as partes interessadas, incluindo empresas de tecnologia, órgãos governamentais, órgãos eleitorais, setor privado e sociedade civil, promovendo a colaboração e a complementaridade do modelo regulatório adotado.
3. Desenvolvimento de políticas digitais abrangentes e equitativas: Formular políticas digitais que abranjam a gestão de dados, a segurança cibernética e a interoperabilidade de sistemas e serviços digitais, assim como garantir o investimento equitativo em recursos e ferramentas de moderação de conteúdos, dando prioridade à cobertura proporcional, incluindo a melhoria das capacidades de detecção automatizada para vários idiomas.
4. Aperfeiçoamento e aplicação das políticas de integridade eleitoral: Implementar políticas de integridade eleitoral consistentes, pesquisáveis e comparáveis, adaptadas aos contextos regionais e locais, baseadas na devida diligência e em avaliações de risco sistêmico.
5. Promoção da cultura de participação digital inclusiva: Facilitar o diálogo regular e inclusivo com a sociedade civil, com as comunidades científicas e as universidades no desenvolvimento e avaliação de estratégias de governação digital, dando prioridade aos interesses e necessidades dos cidadãos.
6. Investir na capacitação digital: Alocar recursos para a capacitação e a formação das partes interessadas na utilização eficaz das tecnologias digitais, garantindo competência e proficiência.
7. Mitigação colaborativa da desinformação: Promover a colaboração entre plataformas e canais para reduzir a propagação de peças de conteúdo e práticas de desinformação, promovendo a transparência e a responsabilização na disseminação extenuante de mensagens interessadas.

Referências

ALVIM, Frederico. Integridade eleitoral: significado e critérios de qualificação. **Revista Ballot**, v1, n2, 2015.

BALAKHONSKAYA, L. V., GAVRA, D. P., BERESNEVA, I. V., & SADIROVA, K. K. Markers of manipulation in the electoral communication in a digital environment. In: IEEE (Ed.), **2020 IEEE Communication Strategies in Digital Society Seminar (ComSDS)**, p. 91-98, April 2020. IEEE.

BARBOSA, B., MARTINS, H., & VALENTE, J. **Pesquisa Fake News: Como as plataformas enfrentam a desinformação**. São Paulo: Intervezes, 2020. Retrieved from <https://app.rios.org.br/index.php/s/JDWtwTS2nBmpQso>

BERNARDI, Ana Julia; CHIODI, Alexsander; BORGES, Tiago. COSTA, Andressa. **Justiça nas Redes - Especial Musk vs Judiciário**. São Paulo: Instituto Democracia em Xeque, 2024.

BORGES, E. (coord); BRITO CRUZ, F e CINTRA, A. M, “Compromisso com a democracia: Integridade eleitoral e o Estado Democrático de Direito nas políticas de plataformas digitais”, *Diagnósticos & Recomendações nº10*. São Paulo: InternetLab, 2023.

CARREIRO, Almada, & DOURADO. (2024, prelo). Ataques e desinformação on-line contra o processo eleitoral de 2022 no Brasil: os memorandos de entendimento firmados entre TSE e big techs. In: **Eleições 2022**. Rio de Janeiro: Instituto Nacional de Ciência e Tecnologia em Democracia Digital.

COUTO, M., & AZEVEDO, F. Do WhatsApp ao YouTube: Novas funções e mudanças na transparência de plataformas desafiam eleições municipais. In **Sonar: A escuta das redes**, January 2024. Disponível em <https://oglobo.globo.com/blogs/sonar-a-escuta-das-redes/post/2024/01/do-whatsapp-ao-youtube-novas-funcoes-e-mudancas-na-transparencia-de-plataformas-desafiam-eleicoes-municipais.gh>t

DOURADO, T. **Fakenews: quando mentiras viram fatos políticos**. Porto Alegre: Zouk, 2021. SHARING ACTIONS ON STOPPING HATE. In Facebook Business Newsroom July 1 2020. <https://www.facebook.com/business/news/sharing-actions-on-stopping-hate>

FACEBOOK. (2023). DSA Transparency Report Oct 2023. Retrieved from <https://transparency.fb.com/sr/dsa-transparency-report-oct2023-facebook/>

INSTITUTO DEMOCRACIA EM XEQUE. **Linha do tempo de ameaças golpistas no Governo Bolsonaro e propostas para defesa da democracia brasileira**. São Paulo: Instituto Democracia em Xeque, 2024.

ITS RIO. 9 pontos de atenção sobre o PL das Fake News (PL 2630/20).

KATZENBACH, C., REDEKER, D., MAGALHÃES, J. C., KOPPS, A., SÜHR, T., CAPLAN, R., YILMA, K. **Policing platforms: addressing power and inequalities in platform policies.** AoIR Selected Papers of Internet Research, 2022. <https://doi.org/10.5210/spir.v2022i0.12959>

MARRET, C. The impact of social media on the polarization of the political debate: three case studies in the European context, 2022. doi: 10.11606/d.101.2022.tde-03112022-123930

MARTINS, L. Plataformas revelam número de usuários na União Europeia. Núcleo Jornalismo, February 17, 2023. <https://nucleo.jor.br/curtas/2023-02-17-plataformas-revelam-numero-de-usuarios-na-uniao-europeia/> How Meta is Planning for Elections in 2024. In Facebook Newsroom, November, 2023. <https://about.fb.com/news/2023/11/how-meta-is-planning-for-elections-in-2024/>

MARWICK, A., & LEWIS, R. *Media Manipulation and Disinformation Online*. Data & Society Research Institute, 2017.

MENDOZA, G. Disinformation campaigns: Rappler denied CrowdTangle data access. **Rappler**, July 11, 2022. Retrieved from <https://www.rappler.com/technology/social-media/disinformation-crowdtangle-data-access/>

OVERGAARD, C.; COLLIER, J. In different worlds: The contributions of polarization and platforms to partisan (mis)perceptions. **New Media & Society**, 2023. Doi: 10.1177/14614448231176551

PINHO, A. X reduziu transparência sobre remoção de contas após ser comprada por Musk. **Folha de São Paulo**, April 10, 2024. Retrieved from <https://www1.folha.uol.com.br/poder/2024/04/x-reduziu-transparencia-sobre-remocao-de-contas-apos-ser-comprada-por-musk.shtml>

RUEDIGER, M. A., & GRASSI, A. (Coord.). O ecossistema digital nas eleições municipais de 2020 no Brasil: o buzz da desconfiança no sistema eleitoral no Facebook, YouTube e Twitter [Policy paper]. Rio de Janeiro: FGV DAPP, 2020.

SALA DE ARTICULAÇÃO CONTRA DESINFORMAÇÃO. O papel das plataformas na proteção da integridade eleitoral em 2022. São Paulo: SAD, 2022. Retrieved from https://drive.google.com/drive/u/1/folders/1y_krbuskNGBH3yoNSOO6zhzOvnoWtE5i

Trained for Deception: How Artificial Intelligence Fuels Online Disinformation. (online). In Mozilla Foundation. <https://foundation.mozilla.org/en/campaigns/trained-for-deception-how-artificial-intelligence-fuels-online-disinformation/recommendations/>

TUCKER, J. A., THEOCHARIS, Y., ROBERTS, M. E., & BARBERÁ, P. From liberation to turmoil: Social media and democracy. **Journal of Democracy**, 28(4), 46, 2017.

X. Twitter transparency report, 2023. Retrieved from <https://transparency.twitter.com/dsa-transparency-report.html#>

SINGER, A., DUNKER, C., ARAÚJO, C., CARVALHO, L., PAULANI, L., BRAGA, R., & SAFATLE, V. Por que assistimos a uma volta do fascismo à brasileira. *Ilustríssima. Folha de S. Paulo*, June 20, 2020. <https://www1.folha.uol.com.br/ilustrissima/2020/06/por-que-assistimos-a-uma-volta-do-fascismo-a-brasileira.shtml>

SOARES, M. Faltam análise de dados nas eleições. *Desinformante*, March 19, 2024. Retrieved from <https://desinformante.com.br/faltam-analise-dados-eleicoes/>

TUCKER, J. A., GUESS, A., BARBERÁ, P., VACCARI, C., SIEGEL, A., SANOVICH, S., ... & NYHAN, B. Social media, political polarization, and political disinformation: A review of the scientific literature. *Political Communication*, 35(4), 609-632, March 19, 2018. <https://doi.org/10.1080/10584609.2018.1473676>

ZAKRZEWSKI, C., VYNCK, G., MASIH, N., & MAHTANI, S. Facebook faces an India reckoning over misinformation, hate speech. *The Washington Post*, October 24, 2021. Retrieved from <https://www.washingtonpost.com/technology/2021/10/24/india-facebook-misinformation-hate-speech/>

Tatiana Dourado · Doutora em Comunicação (UFBA, 2020), pesquisadora pós-doc no âmbito do projeto MediaTrust.Lab, do LabCom – Comunicação e Artes, Universidade da Beira Interior, e pesquisadora associada ao INCT.DD. Membro do Instituto Democracia em Xequê.

Bruna Martins · Bacharel em Direito (Uniceub, 2013) e Gerente de Campanhas Global na Digital Action. Membro da Coalizão Direitos na Rede.

Nina Santos · Doutora em Ciências da Comunicação e da Informação (Université Panthéon-Assas, 2019), pesquisadora do INCT.DD e diretora do Aláfia Lab.

Rodrigo Carreiro · Doutor em Comunicação (UFBA, 2018), pesquisador do INCT.DD e diretor do Aláfia Lab.

Carla Vreche · Doutora em Ciência Política (Unicamp, 2023) e assessora do Programa de Fortalecimento do Espaço Democrático da Conectas Direitos Humanos.