

A Inteligência Artificial como recurso de desinformação: conflitos éticos e a responsabilização das *Big Techs*¹

Maria Iasmin Pereira DINIZ²

Cláudio MESSIAS³

Universidade Federal de Campina Grande - UFCG

RESUMO

O trabalho versa sobre a necessidade de estruturar, desde já, uma discussão ética, no campo das ciências da comunicação, sobre o uso de inteligência artificial na geração de conteúdos de desinformação que contenham trechos ou partes de informação. A circulação de *deepfake* é fomentada por plataformas geradas, mantidas e dominadas pelas denominadas *Big Techs*, que almejando o lucro ignoram as liberdades individuais dos sujeitos da audiência, seus usuários. É analisado o caso em que Mark Zuckerberg, CEO da Meta, anunciou, em janeiro de 2025, o fim do monitoramento de conteúdos.

PALAVRAS-CHAVE: Inteligência artificial; *Big Techs*; Desinformação; Ética na comunicação; *Deepfake*

INTRODUÇÃO

O célere desenvolvimento das tecnologias de inteligência artificial (IA's) tem causado grandes impactos no campo da comunicação, especialmente, quando tratamos da produção e consumo de informações. No entanto, ao mesmo tempo que potencializam diversas áreas do conhecimento, elas também passam a ser utilizadas como ferramentas para articulação de desinformação. Tal situação faz emergir a demanda de um olhar reflexivo e questionador diante de uma problemática que impõe grandes desafios, sobretudo no que diz respeito à responsabilidade das denominadas *Big Techs*⁴ que, muitas vezes, se posicionam de maneira omissa ou mesmo conivente.

Primordialmente, analisa-se como estes sistemas inteligentes evoluíram de tal forma que, para além dos textos, tornam-se, também, capazes de construir imagens,

¹ Trabalho apresentado no Grupo de Trabalho Desinformação, educação midiática e plataformas digitais, evento integrante da programação do 25º Congresso de Ciências da Comunicação na Região Nordeste, realizado de 26 a 28 de junho de 2025.

² Estudante de Graduação, 6º semestre, do bacharelado Comunicação Social/Educomunicação na Universidade Federal de Campina Grande-PB, e-mail: iasminpdiniz@gmail.com.

³ Orientador do trabalho. Doutor em Ciências da Comunicação pela ECA/USP, docente no bacharelado Comunicação Social/Educomunicação na Universidade Federal de Campina Grande, PB. E-mail: claudio.messias@professor.ufcg.edu.br.

⁴ Denominação, em língua inglesa, para conglomerados empresariais da indústria produtora de dispositivos eletrônicos e de plataformas digitais que estabelecem uma nova forma de hegemonia na indústria cultural, estabelecendo padrão de comportamento de consumo na audiência.

áudios e vídeos hiper-realistas dentro de um processo veloz e imediato. Um exemplo disso é o *deepfake*, uma técnica de inteligência artificial que trabalha na manipulação de vozes e rostos para simular falas, expressões e comportamentos de pessoas reais. O método é capaz de criar um conteúdo, por vezes, tão convincente que acaba por dificultar a distinção entre o falso e o real.

O *deepfake* é utilizado, essencialmente, em contextos eleitorais e jornalísticos com o objetivo de simular declarações falsas ou comprometer a imagem de figuras públicas, comportando-se como uma ferramenta potente na geração de desinformação em larga escala e suscitando uma necessária reflexão sobre os limites éticos dessas tecnologias. Giselle Beiguelman (2021, p. 128) argumenta que personalidades públicas são alvos frequentes de deepfakes devido à grande quantidade de imagens disponíveis online, facilitando o aprendizado algorítmico para simular suas falas e expressões.

Sobre o que expõe Beiguelman recordamos o ocorrido em 2022, com a apresentadora Renata Vasconcellos, jornalista que compõe a bancada do Jornal Nacional, na Rede Globo, ocasião em que a profissional foi vítima da primeira *deepfake* registrada durante as eleições presidenciais daquele ano. No vídeo manipulado, a jornalista anunciava uma suposta pesquisa de intenção de votos na qual o candidato à reeleição da época, Jair Bolsonaro (PL), aparecia liderando com 44% das intenções, à frente do também candidato Luiz Inácio Lula da Silva (PT), que surgia em segundo lugar, com 32%. O conteúdo, todavia, se tratava de uma versão adulterada da edição do Jornal Nacional que foi ao ar no dia 15 de agosto de 2022. Naquela ocasião, a apresentadora anunciara os dados reais da pesquisa do Ipec⁵, que, na verdade, apontava Lula na liderança da corrida presidencial, com Bolsonaro em segundo lugar. As porcentagens foram mantidas, alterando-se apenas a ordem dos candidatos, o que evidenciava uma clara intencionalidade na distorção da informação.

Está posto, portanto, que há ausência de verdade objetiva em relação a um fato divulgado. A verdade central encontra-se em conteúdo incondizente ao fato anunciado enquanto notícia. Trata-se, portanto, de desinformação, que por si só, traz consigo uma série de dilemas éticos que nesse trabalho trazemos à arena das discussões. Levantar, pois, uma discussão no âmbito do fenômeno das tecnologias de inteligência artificial é entender, pensamos, a necessidade de novas reflexões acerca da forma como é

⁵ Instituto de pesquisas que monitora opinião pública a pedido de grupos empresariais e de mídia, com registro, no caso de 2022, junto à Justiça Eleitoral.

estruturada a desinformação na contemporaneidade. Nesse contexto, torna-se inevitável refletir, também, sobre o papel das grandes empresas de tecnologia que, ao mediar os fluxos informacionais, influenciam diretamente na forma como esses conteúdos circulam, além de impactarem seu alcance e legitimação.

A postura adotada pelas *Big Techs* revela, entendemos, uma irresponsável omissão, em um ambiente onde a ética não assume o protagonismo. Em vez disso, os cliques, alimentados pelo ódio de conteúdos polarizadores e sensacionalistas, tornam-se o principal foco de plataformas que priorizam o lucro e o engajamento em detrimento da integridade informacional.

A comprovação do exposto é Mark Zuckerberg, CEO da Meta⁶, que anunciou, em janeiro de 2025, uma série de mudanças nas políticas da empresa, dona de redes sociais populares como o Instagram e o Facebook. Entre elas, estava a substituição dos verificadores de fatos por um sistema de notas da comunidade, semelhante ao já aplicado no X, antigo Twitter. No novo modelo, a veracidade das informações passa a ser avaliada pelos próprios usuários da plataforma, em substituição aos profissionais da comunicação. No mesmo sentido, foi declarado um abrandamento dos filtros de moderação e a ampliação da exibição de conteúdo político.

Diante do exposto, é possível perceber que ambas as medidas comunicadas giram em torno de políticas que não almejam uma simples inovação, mas sim uma adaptação às exigências do capitalismo digital na busca pelo lucro. Evgeny Morozov (2018) traz uma elucidação pertinente, sobre isso, ao postular que

O problema não são as *fake news*, e sim a velocidade e facilidade de sua disseminação, e isso acontece principalmente porque o capitalismo digital faz com que seja altamente rentável – veja o Google e o Facebook – produzir e compartilhar narrativas falsas que atraem cliques (MOROZOV, 2018, p. 184).

Habitualmente, publicações que protagonizam e fomentam a polarização e levantam discussões caracterizadas por polêmicas e pontualmente definidas por uma versão paralela de agenda setting tendem, verificamos, a engajar com mais facilidade. Da mesma forma, conteúdos noticiosos com ausência de verdade acabam se destacando, impulsionados pela lógica algorítmica das plataformas digitais que se eximem de suas responsabilidades éticas.

⁶ Big tech que faz a gestão, no planeta, de plataformas como Facebook, Instagram e Whatsapp, de amplo uso por uma audiência composta por todas as idades.

O usuário-consumidor, por sua vez, inserido na lógica da “Sociedade em Rede” descrita por Manuel Castells (2005) - na qual os fluxos de comunicação e informação são mediados por tecnologias – encontra-se imerso em uma dinâmica de urgência social. Sob essa perspectiva, os conteúdos são produzidos, processados e compartilhados com tamanha velocidade que, muitas vezes, até mesmo a inexistência da veracidade passa despercebida.

À medida que o algoritmo da IA se aperfeiçoa, as informações são concebidas de maneira cada vez mais automatizada, acometendo os leitores e usuários a uma verdadeira tempestade de dados comunicacionais, frequentemente dispersos e desprovidos de narrativa. Esse fenômeno é enfatizado por Byung-Chul Han (2023, p. 20), que argumenta que “a crise narrativa da modernidade se deve ao fato de que o mundo está inundado de informações. O espírito da narração está sendo sufocado pela enxurrada de informações”.

Ante ao exposto, torna-se indispensável uma reflexão crítica acerca da forma como se estrutura a geração de conteúdos informacionais em um contexto marcado pelo protagonismo das tecnologias de inteligência artificial. Para além disso, é preciso que a problemática da desinformação seja colocada em pauta, mas dessa vez, sem desconsiderar todos os agentes envolvidos em sua reprodução – das grandes corporações tecnológicas até os usuários comuns.

METODOLOGIA E FUNDAMENTAÇÃO TEÓRICA

Este trabalho surge sob a perspectiva de provocações relacionadas à atual conjuntura da comunicação social, imersa em um corpo social cada vez mais mediado pelas tecnologias, especialmente as IA's. Para tanto, compreendeu-se a necessidade de realização de uma análise crítica no que tange a forma como a desinformação tem sido gerada e disseminada nesse contexto, bem como o posicionamento das *Big Techs* diante desse cenário.

À luz disso, a pesquisa foi desenvolvida a partir de uma abordagem qualitativa de caráter exploratório, uma escolha metodológica que se justifica pelo desejo de compreender de maneira mais profunda os aspectos comunicativos da contemporaneidade. A metodologia qualitativa, conforme os postulados de Marconi &

Lakatos (2008, p. 269), visa uma análise detalhada e interpretativa de hábitos, atitudes e tendências, o que é essencial para investigar práticas comunicacionais emergentes.

Seguindo essa lógica, foi empreendido um estudo de caso como estratégia metodológica na qual foram analisados dois episódios emblemáticos: o caso do *deepfake* envolvendo a jornalista Renata Vasconcellos, nas eleições presidenciais brasileiras de 2022, e o anúncio feito por Mark Zuckerberg, em 2025, acerca das mudanças nas políticas de moderação da Meta. Ambos os casos contribuem para compreensão contextualizada dos desafios éticos e comunicacionais impostos pelas dinâmicas digitais vigentes, bem como a postura adotada pelas grandes empresas de tecnologia em meio a essas circunstâncias.

Outrossim, recorreu-se à realização de uma revisão bibliográfica com o intuito de fundamentar teoricamente as discussões propostas, articulando os postulados de autores que permeiam os estudos da informação na contemporaneidade, inteligência artificial e *Big Techs*, a fim de sustentar as análises desenvolvidas ao longo do trabalho.

Dentre os principais referenciais teóricos, destaca-se Byung-Chul Han (2023), cuja crítica à crise narrativa que permeia o corpo social oferece mecanismos para compreender a posição do usuário diante da disseminação intensa de informações. O autor observa que “a comunicação é cada vez mais controlada de fora para dentro. Ela parece obedecer a um processo automático, maquinal, controlado por algoritmos, e do qual não somos conscientes” (HAN, 2023, p. 28-29), apontando para o esvaziamento da experiência comunicacional que se vê operada pela engrenagem digital.

Nessa mesma direção, Manuel Castells (2005) contribui com uma leitura aprofundada dando suporte para a apreensão das transformações estruturais da produção da informação no cenário da “Sociedade em Rede” e que, assim como Han (2023), identifica a centralidade das tecnologias da informação na reorganização das dinâmicas sociais, mudando a forma como os indivíduos recebem e processam as mensagens.

Complementando essa discussão, Giselle Beiguelman (2021) chama atenção para a problemática da *deepfake* - um dos mecanismos de inteligência artificial articulados para a manipulação da informação – ao evidenciar a engrenagem algorítmica por trás da ferramenta. Para tanto, a autora afirma que o vídeo modificado

É imagem produzida algoritmicamente, sem mediação humana no seu processamento, que utiliza milhares de imagens estocadas em bancos de dados para aprender os movimentos do rosto de uma pessoa,

inclusive os labiais e suas modulações de voz, para prever como ela poderia falar algo que não disse (BEIGUELMAN, 2021, p. 137).

Complementando, Evgeny Morozov (2018) oferece contribuições significativas ao levantar críticas às plataformas gerenciadas pelas gigantes da tecnologia, as quais adotam uma conduta de isenção de responsabilidade no tocante à propagação de conteúdos com ausência de verdade. Esses conteúdos, por sua natureza, tendem a atrair o engajamento lucrativo, que se revela como o principal interesse dessas empresas, sobrepondo-se à ética.

CONSIDERAÇÕES PRELIMINARES

O trabalho versa sobre um cenário em que é possível perceber que a desinformação, na contemporaneidade, não é um fenômeno isolado, mas um efeito direto de uma lógica algorítmica que potencializa suas ações. Os casos analisados indicam como as tecnologias de Inteligência Artificial atuam não só na automação de conteúdos, como também na produção de conteúdos com ausência de verdade, logo, de desinformação.

Nesse contexto, torna-se evidente a postura adotada pelas *Big Techs*, que apesar de sua capacidade tecnológica, priorizam o lucro acima da ética, demonstrando uma conduta de isenção de responsabilidade diante do papel de suas plataformas na disseminação de conteúdos distorcidos.

Por fim, reforça-se a necessidade de colocar em perspectiva os mecanismos de regulação e responsabilização dessas plataformas. Nesse processo, é indispensável refletir criticamente acerca do papel desempenhado pelos sistemas inteligentes, os quais, muitas vezes, passam a ser utilizados de maneira dissociada da ética. Diante disso, impõe-se o desafio de construir uma cultura digital alicerçada em valores éticos, priorizando a educação midiática, a transparência e a responsabilidade.

REFERÊNCIAS

- BEIGUELMAN, G. **Políticas de imagem: vigilância e resistência na dadosfera**. São Paulo: Ubu Editora, 2021.
- CASTELLS, M. **A sociedade em rede**. São Paulo: Paz e terra, 2005.
- HAN, B-C. **A crise da narração**. Rio de Janeiro: Editora Vozes, 2023.
- MARCONI, M. A., LAKATOS, E. M. **Metodologia Científica**. 5. ed. 2. reimpr. São Paulo: Atlas, 2008.
- MOROZOV, E. **Big tech: A ascensão dos dados e a morte da política**. São Paulo: Ubu Editora, 2018.