

I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL III

PAULO CEZAR DIAS

VALTER MOURA DO CARMO

FERNANDO GALINDO AYUDA

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

Diretoria - CONPEDI

Presidente - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - FMU - São Paulo

Diretor Executivo - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

Vice-presidente Norte - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

Vice-presidente Centro-Oeste - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

Vice-presidente Sul - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

Vice-presidente Sudeste - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

Vice-presidente Nordeste - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

Representante Discente: Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

Conselho Fiscal:

Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão

Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais

Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte

Prof. Dr. Fernando Passos - UNIARA - São Paulo

Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo

Secretarias

Relações Institucionais:

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal

Comunicação:

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

Relações Internacionais para o Continente Americano:

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Prof. Dr. Felipe Chiarello de Souza Pinto - UPM - São Paulo

Relações Internacionais para os demais Continentes:

Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul

Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI - Santa Catarina

Educação Jurídica

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - PR

Prof. Dr. Rubens Beçak - USP - SP

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - MS

Eventos:

Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo

Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina

Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

Comissão Especial

Prof. Dr. João Marcelo de Lima Assafim - UFRJ - RJ

Profa. Dra. Maria Creusa De Araújo Borges - UFPB - PB

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - MG

Prof. Dr. Rogério Borba - UNIFACVEST - SC

I61

Inteligência Artificial: Desafios da Era Digital III [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Fernando Galindo Ayuda, Paulo Cezar Dias, Valter Moura do Carmo. – Florianópolis: CONPEDI, 2025.

Inclui bibliografia

ISBN: 978-65-5274-097-7

Modo de acesso: www.conpedi.org.br em publicações

Tema: Inteligência Artificial e Sustentabilidade na Era Transnacional

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Internacionais. 2. Inteligência Artificial. 3. Desafios da Era Digital. I
International Experience Perugia – Itália. (1: 2025 : Perugia, Itália).

CDU: 34



Conselho Nacional de Pesquisa
e Pós-Graduação em Direito Florianópolis
Santa Catarina – Brasil
www.conpedi.org.br

I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL III

Apresentação

APRESENTAÇÃO

O I International Experience – Perúgia – Itália foi realizado nos dias 28, 29 e 30 de maio de 2025, com o tema "Inteligência Artificial e Sustentabilidade na Era Digital". O Grupo de Trabalho (GT) "Inteligência Artificial: Desafios da Era Digital III" ocorreu nos dias 29 e 30 de maio, nos períodos vespertinos, na Universidade de Perúgia.

O GT destacou-se não apenas pela qualidade dos trabalhos apresentados, mas também pelo nível acadêmico dos autores — doutores, mestres, professores pesquisadores e seus alunos pós-graduandos. O evento também proporcionou um importante espaço de interlocução internacional, contando com a participação de renomados juristas e professores de instituições estrangeiras, como os Professores Doutores Roberto Cippitani (Universidade de Perúgia) e Fernando Galindo (Universidade de Zaragoza – Espanha), que enriqueceram os debates e contribuíram para o sucesso da atividade.

Foram apresentados 15 (quinze) artigos, os quais foram objeto de intenso debate presidido pelos coordenadores e enriquecido pela participação ativa do público presente na Faculdade de Direito de Perúgia – ITÁLIA.

A apresentação dos trabalhos permitiu discussões atualizadas e profícuas sobre temas como inteligência artificial, uso de dados pessoais, dever de informação, riscos e interações tecnológicas. As abordagens trataram dos desafios enfrentados pelas diversas linhas de pesquisa jurídica no estudo do futuro da regulação no Brasil, dos abusos relacionados à inteligência artificial e das possíveis soluções para a proteção de dados em um mundo globalizado.

As temáticas incluíram: tecnologias relacionadas a fake news, deepfakes e bots; compliance; a consideração do elemento humano na aplicação da I.A. nas decisões judiciais; a inteligência artificial como ferramenta de proteção no sistema de justiça criminal; o consentimento informado e o uso de dados pessoais; regulamentação e governança da I.A.; precarização do governo digital e aplicação da inteligência artificial em distintos setores jurídicos.

A seguir, apresenta-se a relação dos trabalhos que compõem este Grupo de Trabalho, acompanhados de seus respectivos autores:

1. CAPACIDADE ARTIFICIAL DAS MÁQUINAS E A EXIGÊNCIA DE TRANSFORMAÇÕES NA MANEIRA DO SABER DE PROFISSIONAIS, de Fernanda Conceição Pohlmann.
2. AI, VOCÊ ESTÁ AÍ? O PANORAMA JURÍDICO RELATIVO À (AUTO) IDENTIFICAÇÃO DA INTELIGÊNCIA ARTIFICIAL, de Gabriel Siqueira Eliazar de Carvalho, André Fortes Chaves e Marcello Silva Nunes Leite.
3. DEMOCRACIA EM REDE: O PAPEL DA INTELIGÊNCIA ARTIFICIAL E DOS ALGORITMOS NA LIBERDADE DE EXPRESSÃO E NO PLURALISMO POLÍTICO, de Kennedy da Nobrega Martins, Alexandre Manuel Lopes Rodrigues e Jadgleison Rocha Alves.
4. INTELIGÊNCIA ARTIFICIAL E DIREITOS FUNDAMENTAIS: DESAFIOS E TENSÕES NA ERA DIGITAL, de Jesualdo Eduardo de Almeida Junior e Gustavo Roberto Dias Tonia.
5. INTELIGÊNCIA ARTIFICIAL E DEMOCRACIA: O PERIGO DA MANIPULAÇÃO DE INFORMAÇÕES, de Claudia Maria da Silva Bezerra e Luiz Eduardo Simões de Souza.
6. INFLUÊNCIAS DO REALISMO JURÍDICO E O USO DA INTELIGÊNCIA ARTIFICIAL NA ELABORAÇÃO DE DECISÕES JUDICIAIS NO BRASIL: VIESES COGNITIVOS E HEURÍSTICAS NO PROCESSO DECISÓRIO, de Kerry Barreto, Fausto Santos de Moraes e Júlia Regina Bassani Caus.
7. CRITÉRIOS QUANTITATIVOS PARA A MENSURAÇÃO DE RESULTADOS NO JUÍZO 100% DIGITAL: RISCOS PARA A QUALIDADE DA PRESTAÇÃO JURISDICIONAL NO BRASIL, de Orides Mezzaroba, José Renato Gaziero Cella e Lia Loana Curial Oliva.
8. AS PROVAS DIGITAIS NO PROCESSO CIVIL E O (DES)CABIMENTO DA CADEIA DE CUSTÓDIA, de Jesualdo Eduardo de Almeida Junior e Gustavo Roberto Dias Tonia.
9. A REVOLUÇÃO DA INTELIGÊNCIA ARTIFICIAL NOS GABINETES JUDICIAIS: EFICIÊNCIA COM GARANTIAS CONSTITUCIONAIS, de Lisbino Geraldo Miranda do Carmo, Deise Neves Nazaré Rios Brito e Jimmy Souza do Carmo.

10. GENEALOGIA E INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL PARA ELABORAÇÃO DE UM ONOMÁSTICO DOS IMIGRANTES ITALIANOS QUE DESENVOLVERAM O SUL DO ESTADO DE SANTA CATARINA DE 1877 A 1897, de Júlio Cesar Cancellier de Olivo.

11. A REDE-LAB COMO INOVAÇÃO NA POLÍTICA ANTILAVAGEM DE CAPITAIS NO BRASIL, de Lorryne Souza Galli e Matheus Felipe de Castro.

12. ARMAS AUTÔNOMAS LETAIS: OS IMPACTOS DA INTELIGÊNCIA ARTIFICIAL PARA OS DIREITOS HUMANOS E SUA CONSEQUENTE REGULAMENTAÇÃO, de Alexandre Gonçalves Ribeiro e Renata Mantovani de Lima.

13. A INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA ESSENCIAL NA ELUCIDAÇÃO DE CRIMES SEXUAIS PRATICADOS COM VIOLÊNCIA CONTRA A MULHER, de Eneida Orbage de Britto Taquary e Catharina Orbage de Britto Taquary Berino.

14. A INTELIGÊNCIA ARTIFICIAL NO DIREITO PENAL: AVANÇOS, DESAFIOS E IMPACTOS NA INVESTIGAÇÃO E NO SISTEMA JUDICIAL, de Eneida Orbage de Britto Taquary, Bianca Cristina Barbosa de Oliveira e Tiago de Lima Mascarenhas Santos.

15. ENTRE CÓDIGOS E DIREITOS: UMA ANÁLISE CONSTITUCIONAL DA PROTEÇÃO DE DADOS PESSOAIS NO CONTEXTO DA INTELIGÊNCIA ARTIFICIAL, de Lisbino Geraldo Miranda do Carmo, Deise Neves Nazaré Rios Brito e Paulo Henrique da Silva Costa.

Por fim, os organizadores e coordenadores do Grupo de Trabalho "Inteligência Artificial: Desafios da Era Digital III" parabenizam e agradecem aos autores pelos valiosos trabalhos apresentados, cuja leitura certamente contribuirá para o aprofundamento do debate acadêmico e científico na área.

Prof. Dr. Fernando Galindo - Universidad de Zaragoza - Espanha

Prof. Dr. Valter Moura do Carmo – PPGPJDH - ESMAT e UFT

Prof. Dr. Paulo Cezar Dias – Centro Universitário Eurípides de Marília - SP

INTELIGÊNCIA ARTIFICIAL E DEMOCRACIA: O PERIGO DA MANIPULAÇÃO DE INFORMAÇÕES

ARTIFICIAL INTELLIGENCE AND DEMOCRACY: THE DANGER OF INFORMATION MANIPULATION

Claudia Maria Da Silva Bezerra ¹
Luiz Eduardo Simões de Souza ²

Resumo

A ascensão da inteligência artificial (IA) tem transformado profundamente a comunicação digital, influenciando a produção, disseminação e consumo de informações. No entanto, a IA também tem sido utilizada como ferramenta poderosa na manipulação da opinião pública. Tecnologias como fake news, deepfakes e bots automatizados impulsionam a disseminação de desinformação em larga escala, afetando processos eleitorais, polarização social e confiança nas instituições democráticas. A crescente sofisticação dessas técnicas levanta desafios regulatórios e éticos, exigindo respostas de governos e empresas de tecnologia. Diversas iniciativas emergiram para enfrentar essa ameaça, incluindo regulações internacionais, como o Digital Services Act da União Europeia, e projetos legislativos nacionais, como o PL das Fake news no Brasil. Entretanto, equilibrar a liberdade de expressão e moderação de conteúdo gera debates sobre riscos de censura e efetividade das políticas adotadas. Paralelamente, a IA também tem sido empregada no combate à desinformação, com ferramentas de fact-checking automatizado, detecção de deepfakes e moderação algorítmica. Este artigo examina os impactos da manipulação informacional pela IA na democracia, analisando suas consequências políticas e sociais, tentativas de regulamentação e o potencial da própria IA como solução. A metodologia baseia-se em revisão bibliográfica de estudos acadêmicos, relatórios internacionais e legislações emergentes. Argumenta-se que mitigar os danos da manipulação digital exige uma abordagem multidimensional, combinando políticas regulatórias eficazes, transparência algorítmica, educação midiática e desenvolvimento de tecnologias antifraude. A governança digital dependerá da capacidade das sociedades de adaptar-se aos desafios da IA, assegurando seu uso para fortalecer a democracia, e não para subvertê-la.

Palavras-chave: Inteligência artificial, Fake news, Deepfakes, Democracia digital, Regulação da informação

¹ Pós-doutoranda em Direito PPGDIR-UFMA. Doutora e Mestre em Administração - UNINOVE. Editora Associada RIAE. Líder Sustentabilidade SINGEP/UNINOVE/SP. Líder Gestão Socio-ambiental/ODS-EMPRAD/FEA-USP. Pesquisadora NEDC/UFMA. Professora IDEA-DIREITO – São Luís/MA. E-mail: profa.claudiamsbezerra@gmail.com

² Pós-doutor em relações internacionais - INEST/UFF, Doutor e Mestre em História Econômica - USP. Editor-chefe da Revista REPHE. Líder do GEEPHE. Professor Associado da Universidade Federal do Maranhão (UFMA).

Abstract/Resumen/Résumé

The rise of artificial intelligence (AI) has profoundly transformed digital communication, influencing the production, dissemination, and consumption of information. However, AI has also been used as a powerful tool for manipulating public opinion. Technologies such as fake news, deepfakes, and automated bots drive large-scale disinformation dissemination, affecting electoral processes, social polarization, and trust in democratic institutions. The increasing sophistication of these techniques raises regulatory and ethical challenges, demanding responses from governments and technology companies. Various initiatives have emerged to address this threat, including international regulations such as the European Union's Digital Services Act and national legislative projects like Brazil's Fake news Bill. However, balancing freedom of expression and content moderation generates debates about censorship risks and the effectiveness of adopted policies. At the same time, AI is also being employed to combat disinformation, with tools for automated fact-checking, deepfake detection, and algorithmic content moderation. This article examines the impacts of AI-driven information manipulation on democracy, analyzing its political and social consequences, regulatory attempts, and AI's potential as a solution. The methodology is based on a literature review of academic studies, international reports, and emerging legislation. It is argued that mitigating the harms of digital manipulation requires a multidimensional approach, combining effective regulatory policies, algorithmic transparency, media literacy, and the development of anti-fraud technologies. The future of digital governance will depend on societies' ability to adapt to AI challenges, ensuring its use strengthens democracy rather than undermines it.

Keywords/Palabras-claves/Mots-clés: Artificial intelligence, Fake news, Deepfakes, Digital democracy, Information regulation

1. INTRODUÇÃO

"A mentira pode dar a volta ao mundo enquanto a verdade ainda calça os sapatos."

— *Mark Twain*

"As pessoas simplesmente não sabem o que está acontecendo, e nem sabem que não sabem."

— *Noam Chomsky*

A ascensão da inteligência artificial (IA) tem transformado significativamente a maneira como a informação é produzida, distribuída e consumida. Enquanto a IA tem o potencial de democratizar o acesso ao conhecimento, sua utilização também tem facilitado a propagação de desinformação em larga escala, comprometendo a integridade dos processos democráticos (de Carvalho Júnior; Carvalho; De Sousa, 2024). A manipulação da opinião pública por meio de *fake news*, *deepfakes* e *bots* automatizados tem se tornado uma preocupação global, à medida que governos, pesquisadores e a sociedade civil buscam estratégias para mitigar seus impactos (Tucker et al., 2018). Exemplos recentes incluem o uso da IA para espalhar notícias falsas durante as eleições nos Estados Unidos, no Brasil e na União Europeia, demonstrando como essa tecnologia pode ser explorada para enfraquecer a confiança nas instituições democráticas (Bradshaw & Howard, 2019).

O avanço das técnicas de IA generativa, como redes neurais profundas e aprendizado de máquina, permitiu a criação de *deepfakes* altamente realistas, capazes de falsificar discursos de líderes políticos e personalidades públicas. Essas tecnologias têm sido utilizadas para desacreditar adversários políticos, gerar pânico social e até mesmo influenciar mercados financeiros (Chesney & Citron, 2019). Além disso, o uso de *bots* automatizados tem impulsionado campanhas de desinformação em redes sociais, onde algoritmos de recomendação favorecem conteúdos polarizadores, amplificando narrativas falsas e manipulativas (Zollo et al., 2017). Diante desse cenário, surgem desafios éticos e regulatórios sobre quem deve ser responsabilizado pela disseminação desse tipo de conteúdo: as empresas de tecnologia, os governos ou os próprios usuários?

Os impactos da desinformação baseada em IA vão além do ambiente digital, afetando a tomada de decisões políticas e a estabilidade democrática. Estudo realizado pelo MIT Media Lab mostrou que notícias falsas se espalham seis vezes mais rápido do que informações verdadeiras no Twitter (Vosoughi, Roy & Aral, 2018). Esse fenômeno ocorre porque conteúdos

sensacionalistas e polarizadores geram maior engajamento, sendo priorizados pelos algoritmos das plataformas. Em resposta, países como a União Europeia vêm adotando medidas regulatórias, como o *Digital Services Act* (DSA), que obriga grandes plataformas a implementar mecanismos para limitar a disseminação de *fake news* (European Commission, 2022). No entanto, o desafio persiste: como equilibrar a moderação de conteúdo sem ferir princípios fundamentais, como a liberdade de expressão?

Diante da crescente influência da IA na manipulação da informação e seus impactos na democracia, este artigo tem como objetivo analisar os principais desafios decorrentes desse fenômeno, com ênfase em três aspectos: (i) o papel das plataformas digitais na disseminação de desinformação, (ii) as respostas regulatórias adotadas por diferentes países e (iii) as possibilidades e limitações da própria IA na contenção desse problema.

Embora a literatura acadêmica já aborde amplamente os impactos da IA na disseminação de *fake news* e o papel dos algoritmos na personalização da informação, observa-se uma lacuna no que diz respeito à eficácia das regulações implementadas e à viabilidade da própria IA como uma ferramenta de mitigação da desinformação. O presente estudo busca preencher esse gap, explorando as limitações das abordagens regulatórias e investigando como soluções tecnológicas podem contribuir para conter a manipulação informacional sem comprometer direitos fundamentais, como a liberdade de expressão.

Para isso, o presente estudo adota uma abordagem qualitativa, exploratória e jurídico-descritiva, com o objetivo de compreender os impactos da inteligência artificial na manipulação de informações e seus reflexos na democracia. A escolha dessa abordagem justifica-se pela complexidade do fenômeno estudado, que envolve dimensões tecnológicas, sociais, políticas e jurídicas que não podem ser adequadamente mensuradas por métodos quantitativos (Creswell; Poth, 2016; Gustin; Dias; Nicácio, 2020).

A revisão bibliográfica foi selecionada como principal técnica de pesquisa, considerando a necessidade de fundamentação teórica sólida para análise do tema (Mezzaroba; Monteiro, 2017). A coleta de dados foi realizada em bases científicas renomadas, como *Scopus*, *Web of Science* e *SciELO*, garantindo a credibilidade das fontes utilizadas. Além disso, foram analisados relatórios de organismos internacionais e legislações emergentes que tratam da regulação da IA e da desinformação digital, permitindo uma visão global sobre as respostas institucionais ao problema.

Para a análise de conteúdo, adotou-se o método de Bardin (2015), que permite a identificação de categorias temáticas recorrentes nos textos analisados. Esse método foi

escolhido por sua capacidade de oferecer uma visão abrangente dos impactos da IA na democracia, permitindo o agrupamento de informações em categorias analíticas que facilitam a interpretação dos dados (Queiroz; Feferbaum, 2021).

Este estudo foi motivado pela urgência de compreender os impactos da IA na democracia, especialmente em um contexto onde eleições ao redor do mundo são diretamente influenciadas pelo fluxo de informações falsas. Além disso, a pesquisa contribui para o debate sobre como equilibrar inovação tecnológica e proteção da integridade eleitoral, um dos desafios mais prementes da era digital. A busca por um modelo regulatório eficaz, que limite os efeitos da manipulação informacional sem cercear o debate público e a liberdade de expressão, torna-se essencial para garantir a sustentabilidade dos regimes democráticos diante da rápida evolução das tecnologias digitais.

2. A INTELIGÊNCIA ARTIFICIAL E A PRODUÇÃO DE CONTEÚDO MANIPULATIVO

A Inteligência Artificial (IA) tem revolucionado a maneira como o conteúdo é criado e disseminado na internet, tornando-se uma ferramenta poderosa tanto para o acesso à informação quanto para a produção de desinformação em larga escala (Robl Filho; Marrafon; Medón, 2022). Com o desenvolvimento de modelos avançados de aprendizado de máquina, como o *GPT-4*, *DALL·E* e *Stable Diffusion*, tornou-se possível gerar textos, imagens, vídeos e áudios extremamente realistas, muitas vezes indistinguíveis de conteúdos autênticos. Esse avanço trouxe benefícios significativos, como a automação da produção de conhecimento e a personalização de serviços digitais. No entanto, também abriu caminho para novas formas de manipulação da opinião pública, permitindo a criação e disseminação de *fake news*, *deepfakes* e campanhas de desinformação direcionadas (Rugge, 2020).

Entre as ferramentas mais preocupantes da atualidade estão os *deepfakes*, uma tecnologia baseada em redes neurais profundas que permite a substituição do rosto e da voz de uma pessoa em vídeos, gerando conteúdos altamente convincentes. Essa tecnologia já foi usada para imitar discursos de líderes políticos, fabricar escândalos e manipular a percepção pública sobre eventos históricos (RAMOS-ZAGA, 2024). Um dos casos mais emblemáticos ocorreu durante as eleições presidenciais dos Estados Unidos em 2020, quando vídeos falsos de políticos foram amplamente disseminados em redes sociais, confundindo eleitores e aumentando a polarização (Chesney & Citron, 2019). No Brasil, *deepfakes* também foram

utilizados durante campanhas eleitorais recentes, demonstrando o potencial dessa tecnologia para comprometer a integridade dos processos democráticos (Pimenta et al., 2024).

Além dos *deepfakes*, *bots* automatizados e redes de desinformação têm sido amplamente utilizados para impulsionar conteúdos falsos. *Bots* são programas que imitam interações humanas em redes sociais e são frequentemente empregados para amplificar artificialmente determinadas narrativas, criando uma falsa sensação de consenso. Em plataformas como *Twitter* (atualmente *X*), *Facebook* e *TikTok*, estudos mostram que esses sistemas são responsáveis por até 70% do compartilhamento de *fake news* durante períodos eleitorais (Shao et al., 2018). A combinação entre algoritmos de recomendação e o uso estratégico de IA faz com que notícias falsas sejam promovidas de maneira direcionada, influenciando diretamente a percepção pública e até mesmo decisões políticas (Zollo et al., 2015).

A manipulação da informação por IA também ocorre por meio da personalização algorítmica, em que sistemas automatizados direcionam conteúdos específicos para diferentes grupos populacionais com base em seus perfis de consumo de informação. Essa abordagem tem sido explorada por campanhas políticas e organizações interessadas em moldar a opinião pública. O escândalo da *Cambridge Analytica*, que envolveu o uso de dados do *Facebook* para manipular eleitores nos Estados Unidos e no Reino Unido, evidenciou como a IA pode ser utilizada para criar mensagens altamente segmentadas e persuasivas (Cadwalladr & Graham-Harrison, 2018). Esse tipo de manipulação reduz a diversidade informacional e contribui para a formação de bolhas ideológicas, onde os usuários são expostos apenas a conteúdos que reforçam suas crenças preexistentes.

Outro desafio preocupante é o impacto da IA na descredibilização da mídia tradicional e das instituições democráticas. À medida que conteúdos falsificados se tornam mais sofisticados, cresce a dificuldade de distinguir entre fatos e manipulações digitais. Esse fenômeno tem levado ao aumento da desconfiança nas fontes jornalísticas legítimas, favorecendo teorias da conspiração e discursos anticientíficos. O relatório do *Reuters Institute Digital News Report* (2022) apontou que mais de 50% dos entrevistados em diversos países disseram ter dificuldades em identificar notícias confiáveis devido à proliferação de desinformação *online*. Esse cenário é particularmente perigoso para democracias frágeis, onde a instabilidade política pode ser amplificada por campanhas de desinformação orquestradas (Newman et al., 2022).

Diante dessa realidade, a questão da responsabilidade pela produção e disseminação de desinformação por IA se torna central no debate regulatório. Governos, plataformas digitais e a própria sociedade enfrentam o desafio de equilibrar a moderação de conteúdo com a proteção da liberdade de expressão. A implementação de mecanismos de transparência, como a exigência de identificação de conteúdos gerados por IA e a fiscalização de campanhas de desinformação, tem sido discutida globalmente. No entanto, enquanto medidas eficazes não são implementadas, a manipulação da informação por meio da IA continuará a representar um risco significativo para a estabilidade democrática e a confiança pública nos meios de comunicação.

3. O IMPACTO DA MANIPULAÇÃO DIGITAL NA DEMOCRACIA

A disseminação de desinformação por meio de inteligência artificial tem erosionado a confiança nas instituições democráticas e influenciado diretamente processos eleitorais em diversas partes do mundo. A proliferação de *fake news* e *deepfakes* compromete o direito dos cidadãos à informação verídica, essencial para a tomada de decisões políticas fundamentadas. Em eleições recentes nos Estados Unidos, Brasil e União Europeia, campanhas de desinformação coordenadas por *bots* e algoritmos geraram crises institucionais ao espalhar teorias conspiratórias e notícias falsas sobre candidatos e processos eleitorais (Bradshaw & Howard, 2019). A manipulação digital afeta não apenas a percepção dos eleitores sobre os candidatos, mas também mina a legitimidade dos resultados eleitorais, incentivando narrativas de fraude e instabilidade política.

A preocupação com a manipulação digital por IA também é discutida por Harari (2017), que alerta para o fato de que algoritmos avançados podem conhecer os cidadãos melhor do que eles próprios, permitindo a personalização extrema de discursos políticos. Essa hiperpersonalização torna os indivíduos mais vulneráveis à persuasão digital, dificultando o livre-arbítrio e tornando os processos eleitorais menos transparentes e previsíveis. A longo prazo, essa dinâmica pode levar a uma democracia artificialmente controlada, na qual decisões políticas são direcionadas por mecanismos que exploram vieses cognitivos e emocionais da população.

Além do impacto eleitoral, a manipulação digital intensifica a polarização ideológica e o extremismo político. O uso de IA para promover conteúdos altamente segmentados cria bolhas informacionais, onde usuários são expostos apenas a narrativas que reforçam suas crenças preexistentes (Pariser, 2012). Esse fenômeno reduz o debate democrático saudável e incentiva a radicalização, pois os indivíduos deixam de ter contato com perspectivas

divergentes. Plataformas como *Facebook*, *Twitter (X)* e *YouTube* utilizam algoritmos que priorizam conteúdos de alto engajamento, muitas vezes favorecendo informações sensacionalistas e descontextualizadas, que se espalham mais rapidamente do que conteúdos verificáveis (Vosoughi, Roy & Aral, 2018). Dessa forma, a IA não apenas facilita a disseminação da desinformação, mas também alimenta a polarização social ao reforçar visões extremas.

Outro aspecto crítico é o impacto da manipulação digital na confiança pública nos meios de comunicação e na ciência. O avanço das tecnologias de *deepfake* e a capacidade da IA de gerar textos convincentes desafiam a credibilidade do jornalismo tradicional (Botelho; Nöth, 2021). Quando notícias falsas se tornam indistinguíveis de conteúdos legítimos, a sociedade passa a questionar a veracidade de todas as informações, criando um ambiente de ceticismo generalizado (Newman et al., 2022). Esse efeito é explorado por campanhas políticas que buscam desacreditar veículos de imprensa e instituições acadêmicas, promovendo discursos anticientíficos e deslegitimando evidências empíricas.

A desinformação sistêmica prejudica a credibilidade do conhecimento científico e do jornalismo investigativo, o que pode comprometer a governança pública e as políticas baseadas em evidências (Brennen et al., 2020), como aconteceu com a desinformação relacionada à pandemia de COVID-19, por exemplo, que mostrou como a IA pode ser usada para espalhar teorias conspiratórias e minar a adesão a políticas de saúde pública (Brennen et al., 2020). Além disso, Harari (2018) argumenta que a crescente dependência da IA em processos políticos pode levar ao enfraquecimento das instituições democráticas, tornando o controle da informação uma ferramenta de poder ainda mais concentrada nas mãos de grandes corporações e governos autoritários.

Além da polarização e da crise de confiança, a manipulação digital também enfraquece a governança democrática ao dificultar a implementação de políticas públicas eficazes. Líderes políticos e tomadores de decisão enfrentam desafios crescentes ao tentar combater narrativas falsas que desinformam a população e reduzem o apoio a iniciativas institucionais. A disseminação de *fake news* sobre temas sensíveis, como mudanças climáticas, vacinação e segurança pública, leva a uma sociedade menos informada e mais resistente a políticas baseadas em evidências científicas (Lewandowsky et al., 2017). A IA, nesse sentido, tem sido uma ferramenta tanto para governos autoritários que controlam a narrativa política quanto para grupos extremistas que buscam desestabilizar democracias emergentes.

Diante desse cenário, governos e organismos internacionais vêm tentando responder ao desafio da manipulação digital com medidas regulatórias e iniciativas de transparência. A União Europeia implementou o *Digital Services Act* (DSA), que exige que grandes plataformas monitorem e removam desinformação sistemática, além de exigir mais transparência nos algoritmos de recomendação (European Commission, 2022). No Brasil, o Projeto de Lei das *Fake news* busca responsabilizar plataformas e provedores por conteúdos impulsionados de forma artificial. No entanto, essas iniciativas enfrentam resistência tanto de grupos que defendem a liberdade de expressão irrestrita, quanto das próprias *Big Techs*, que alegam dificuldades em controlar o conteúdo publicado por bilhões de usuários sem comprometer a privacidade e o direito à informação.

Além da União Europeia e do Brasil, outros países também têm implementado medidas regulatórias para conter os impactos da manipulação digital na democracia. Nos Estados Unidos, a discussão sobre a regulação da desinformação está centrada na revisão da *Section 230* do *Communications Decency Act*, que atualmente protege plataformas digitais da responsabilidade sobre o conteúdo publicado por seus usuários (Leary, 2018). Há um debate crescente sobre a necessidade de maior transparência nas operações dos algoritmos de recomendação e moderação de conteúdo (Kosseff, 2019). Além disso, foram instauradas investigações pelo Congresso sobre o papel das big techs na disseminação de desinformação eleitoral (Howard, 2020).

No Reino Unido, o *Online Safety Act* propõe tornar as plataformas digitais responsáveis por remover conteúdos prejudiciais, incluindo desinformação sistemática e manipulação digital (Coe, 2022). A legislação exige que empresas como Google, Meta e Twitter adotem medidas preventivas contra fake news e estabeleçam diretrizes de moderação mais rigorosas (UK Parliament, 2022).

A Austrália implementou o *News Media and Digital Platforms Mandatory Bargaining Code*, que exige que grandes plataformas digitais negociem compensações financeiras com veículos de imprensa pelo uso de seus conteúdos. Essa medida visa fortalecer a imprensa tradicional e reduzir a dependência do público de fontes informacionais não verificadas, combatendo indiretamente a proliferação de fake News (Treasury, 2022).

No Canadá, o governo propôs a *Online Harms Bill*, que prevê sanções para empresas de tecnologia que não adotem medidas eficazes contra desinformação e discursos de ódio online. A regulamentação enfatiza a transparência algorítmica e exige relatórios periódicos das plataformas sobre como lidam com conteúdos prejudiciais (Heritage, 2024).

A Alemanha, por sua vez, foi pioneira com a implementação da *NetzDG* (*Netzwerkdurchsetzungsgesetz*), lei que obriga redes sociais a removerem rapidamente conteúdos ilegais, incluindo desinformação eleitoral e discursos de ódio. As plataformas que não cumprirem as diretrizes podem enfrentar multas significativas (Claussen, 2018).

Essas iniciativas refletem diferentes abordagens para combater a manipulação digital e evidenciam a crescente preocupação global com os impactos da desinformação na governança democrática. No entanto, o desafio persiste: equilibrar a proteção da liberdade de expressão com a necessidade de conter a disseminação de conteúdos falsos e manipulativos.

É evidente que a democracia digital enfrenta um dilema: como garantir o direito à liberdade de expressão ao mesmo tempo em que se combate a desinformação e a manipulação política por meio da IA? A resposta a essa questão exige uma abordagem equilibrada que combine educação midiática, transparência algorítmica e regulamentação eficiente. É essencial que governos, empresas de tecnologia e a sociedade civil colaborem para desenvolver estratégias que protejam a integridade da informação sem comprometer direitos fundamentais. O avanço da IA impõe desafios inéditos para a governança democrática, e a maneira como esses desafios forem enfrentados nas próximas décadas definirá o futuro da participação política em um mundo digitalmente interconectado.

4. MEDIDAS REGULATÓRIAS E JURÍDICAS NO COMBATE À MANIPULAÇÃO DA INFORMAÇÃO

A crescente ameaça da desinformação impulsionada por inteligência artificial tem levado governos e organizações internacionais a desenvolver medidas regulatórias e jurídicas para conter os impactos dessa manipulação na sociedade. A União Europeia tem sido pioneira nesse campo, com a implementação do *Digital Services Act* (DSA), que estabelece diretrizes rigorosas para grandes plataformas digitais no combate à disseminação de conteúdos falsos e manipulativos. Essa legislação exige que empresas como Google, Meta e Twitter (X) adotem mecanismos de transparência em seus algoritmos de recomendação, monitorem redes de desinformação e removam conteúdos comprovadamente falsos (European Commission, 2022). Além disso, o Código de Práticas contra a Desinformação da UE obriga plataformas a relatar regularmente as medidas adotadas para conter campanhas de manipulação digital, promovendo maior *accountability* no ecossistema informacional.

No Brasil, o combate à desinformação ganhou força com a tramitação do Projeto de Lei das *Fake news* (PL 2630/2020), que busca regular a atuação de plataformas digitais na

moderação de conteúdos e no combate à desinformação impulsionada por *bots* e redes automatizadas. O projeto propõe regras para a rastreabilidade de mensagens virais em aplicativos de mensagens, como *WhatsApp* e *Telegram*, além da obrigação de que plataformas sinalizem conteúdos impulsionados por meio de inteligência artificial. No entanto, o PL enfrenta forte resistência de grupos que alegam riscos à liberdade de expressão, bem como da indústria de tecnologia, que argumenta que a regulação excessiva pode impactar a inovação e a privacidade dos usuários (Baldissera e Fortes, 2021). A aprovação do projeto continua sendo debatida, refletindo o dilema entre controle da desinformação e preservação de direitos fundamentais.

Nos Estados Unidos, a abordagem regulatória tem sido menos centralizada, com iniciativas individuais de estados e esforços de pressão sobre as *Big Techs* para que adotem políticas internas mais rigorosas contra desinformação. A *Section 230 do Communications Decency Act*, que protege plataformas digitais de responsabilidade pelo conteúdo postado por seus usuários, tem sido um dos pontos mais controversos no debate regulatório. Críticos argumentam que essa lei permite que plataformas lucrem com a desinformação sem serem responsabilizadas, enquanto defensores afirmam que sua revogação poderia comprometer a liberdade de expressão e levar à censura excessiva (Kosseff, 2019). Em meio a essa disputa, grandes empresas como Meta e Google têm investido em ferramentas automatizadas de detecção de *fake news* e parcerias com agências de checagem de fatos, mas essas medidas ainda são limitadas diante da sofisticação das novas tecnologias de IA generativa.

Uma tendência emergente na regulação da IA e da desinformação é a exigência de transparência algorítmica e rastreabilidade de conteúdo. Muitos países estão discutindo mecanismos que obriguem plataformas a revelar como seus algoritmos priorizam conteúdos e quais medidas são adotadas para reduzir a disseminação de informações falsas (Gorwa, 2019). Além disso, propostas como a identificação obrigatória de conteúdos gerados por IA têm ganhado espaço, especialmente diante do crescimento dos *deepfakes* e textos sintéticos, que tornam cada vez mais difícil distinguir o que é real do que é manipulado. No entanto, há desafios significativos na implementação dessas medidas, pois empresas de tecnologia resistem a revelar detalhes de seus algoritmos, alegando questões de sigilo comercial e riscos de exploração por agentes mal-intencionados.

Outra frente de combate envolve o reforço da educação midiática e digital, promovendo a conscientização da sociedade sobre os riscos da manipulação algorítmica e fornecendo ferramentas para que os cidadãos possam identificar e combater a desinformação.

Países como Finlândia e Alemanha implementaram programas educacionais para ensinar pensamento crítico e análise de fontes desde o ensino fundamental, capacitando novas gerações a lidarem com a era digital (Lewandowsky et al., 2017). No Brasil, iniciativas como o EducaMídia, voltado para a alfabetização midiática de estudantes e professores, são passos importantes na construção de um público mais preparado para enfrentar a manipulação digital. A longo prazo, o fortalecimento da educação midiática pode ser uma das estratégias mais eficazes para reduzir a vulnerabilidade da sociedade às campanhas de desinformação.

Embora regulações e medidas legais sejam fundamentais, o combate à manipulação digital exige cooperação global entre governos, setor privado e sociedade civil. A desinformação é um fenômeno transnacional, e sua propagação ocorre em redes descentralizadas que ultrapassam fronteiras nacionais. Dessa forma, acordos internacionais para harmonizar legislações e estabelecer padrões globais de controle da IA são essenciais para enfrentar esse desafio. Modelos como o *AI Act* da União Europeia, que busca estabelecer normas para o desenvolvimento e uso responsável da IA, podem servir de referência para outras jurisdições. O futuro da democracia digital dependerá da capacidade dos atores globais de criar um ecossistema informacional equilibrado, onde a inovação tecnológica possa coexistir com a proteção dos direitos democráticos e da verdade factual.

5. INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA CONTRA *FAKE NEWS*

Se, por um lado, a inteligência artificial tem sido amplamente utilizada para a produção e disseminação de desinformação, por outro, ela também pode ser uma aliada no combate às *fake news*. Avanços recentes em aprendizado de máquina e processamento de linguagem natural têm permitido o desenvolvimento de ferramentas automatizadas capazes de detectar padrões de desinformação, verificar fatos e identificar *deepfakes*. Empresas de tecnologia e organizações de checagem de fatos têm investido em sistemas que utilizam IA para analisar conteúdos em tempo real e sinalizar possíveis violações às diretrizes das plataformas digitais (Horne et al., 2019). O grande desafio, no entanto, está na eficácia e imparcialidade desses sistemas, que devem equilibrar a moderação de conteúdo com a proteção da liberdade de expressão.

Uma das principais aplicações da IA no combate às *fake news* é a verificação automática de fatos (*fact-checking* automatizado). Modelos avançados de Processamento de Linguagem Natural (NLP), como os desenvolvidos pelo Google e pela OpenAI, podem cruzar informações com bancos de dados confiáveis para avaliar a veracidade de uma afirmação em questão de segundos (Thorne & Vlachos, 2018). Além disso, redes neurais treinadas em grandes

volumes de notícias podem identificar padrões textuais típicos de desinformação, como manchetes sensacionalistas, uso excessivo de apelos emocionais e falta de fontes verificáveis. Embora essas tecnologias estejam em evolução, ainda enfrentam desafios, como a necessidade de bases de dados amplas e constantemente atualizadas para garantir maior precisão.

Além do *fact-checking*, a IA tem sido utilizada para detectar *deepfakes* e conteúdos manipulados. Ferramentas como *Deepfake Detection Challenge*, da Meta, e o *Deepfake-O-Meter*, desenvolvido por pesquisadores da Universidade de Stanford, utilizam aprendizado profundo para analisar inconsistências visuais e sonoras em vídeos suspeitos. Essas tecnologias conseguem identificar padrões anômalos, como piscadas artificiais, movimentos labiais desalinhados e sombras incoerentes, características comuns em vídeos sintéticos gerados por IA (Verdoliva, 2020). Apesar dos avanços, os desenvolvedores dessas tecnologias enfrentam uma corrida constante contra os criadores de *deepfakes*, que aprimoram continuamente seus algoritmos para tornar as falsificações mais convincentes e difíceis de detectar.

Outro uso promissor da IA é na moderação automatizada de conteúdos em redes sociais. Algoritmos treinados para identificar padrões de *fake news* podem atuar na sinalização e remoção de conteúdos enganosos, reduzindo a viralização de desinformação antes que ela alcance grandes audiências. O *Twitter* (atualmente X), por exemplo, tem utilizado IA para marcar postagens potencialmente falsas com alertas de contexto adicional, enquanto o *YouTube* aprimorou seus algoritmos para limitar a disseminação de teorias conspiratórias. No entanto, essas abordagens ainda enfrentam críticas devido à falta de transparência e possibilidade de viés algorítmico, levando a preocupações sobre censura e impactos desproporcionais em determinados grupos ideológicos (Gorwa, 2019).

A colaboração entre IA e equipes humanas também tem se mostrado uma estratégia eficaz para combater a desinformação. Plataformas como Facebook e Google têm investido em parcerias com agências de checagem de fatos e organizações de jornalismo independente, onde a IA atua como um filtro preliminar, identificando conteúdos suspeitos que posteriormente são analisados por especialistas (Graves & Cherubini, 2016). Essa abordagem híbrida reduz os erros de moderação exclusivamente algorítmica, garantindo que conteúdos legítimos não sejam removidos indevidamente. Além disso, a inteligência artificial pode auxiliar pesquisadores e jornalistas ao fornecer análises automatizadas sobre tendências de desinformação, permitindo um monitoramento mais eficiente do ecossistema digital.

Apesar dos avanços, a IA ainda não é uma solução definitiva contra *fake news*. O combate à desinformação exige um esforço multidimensional, que combine tecnologia,

regulação, transparência e educação midiática. Sem uma governança adequada, o uso indiscriminado de IA na moderação de conteúdos pode gerar novos problemas, como viés algorítmico e limitação da liberdade de expressão. Portanto, é fundamental que políticas públicas e diretrizes éticas acompanhem o desenvolvimento dessas tecnologias, garantindo que a IA seja utilizada como um instrumento de fortalecimento da democracia e da informação de qualidade, e não como uma ferramenta para restringir indevidamente o debate público.

6. CONSIDERAÇÕES FINAIS

A ascensão da inteligência artificial trouxe oportunidades sem precedentes para o acesso à informação e o desenvolvimento tecnológico. No entanto, ao mesmo tempo em que aprimorou a disseminação do conhecimento, a IA também se tornou uma ferramenta poderosa na produção de desinformação em larga escala. O uso de algoritmos para gerar *fake news*, *deepfakes* e campanhas de manipulação digital tem afetado diretamente processos democráticos, alimentando polarizações, desacreditando instituições e influenciando eleições ao redor do mundo. A sofisticação dessas tecnologias levanta questões relevantes e urgentes sobre responsabilidade, regulamentação e os limites entre liberdade de expressão e moderação de conteúdo, tornando essencial um debate equilibrado sobre o papel da IA na sociedade contemporânea.

Nesse contexto, diversos países têm buscado medidas regulatórias para conter os impactos da manipulação digital, como o *Digital Services Act* (DSA) da União Europeia e o Projeto de Lei das *Fake news* no Brasil. Essas iniciativas propõem transparência algorítmica, rastreamento de conteúdos gerados por IA e maior responsabilidade das plataformas digitais. No entanto, a implementação dessas regulamentações enfrenta desafios significativos, como a resistência das *Big Techs* e as preocupações legítimas sobre eventuais impactos na liberdade de expressão. A busca por um modelo regulatório eficaz deve equilibrar a necessidade de combate à desinformação com a preservação de direitos fundamentais, garantindo que o combate às *fake news* não se transforme em um mecanismo de censura ou controle excessivo sobre o fluxo informacional.

Ao mesmo tempo, a própria inteligência artificial pode ser uma aliada na luta contra a desinformação. Ferramentas de *fact-checking* automatizado, detecção de *deepfakes* e moderação algorítmica já estão sendo desenvolvidas para identificar e conter conteúdos enganosos antes que eles se tornem virais. No entanto, nenhuma solução tecnológica é infalível, e a IA precisa ser combinada com esforços humanos e institucionais para garantir maior eficácia

e justiça na moderação de conteúdos. Além disso, o fortalecimento da educação midiática e digital surge como uma estratégia essencial para capacitar os cidadãos a reconhecerem e lidarem com a manipulação informacional, reduzindo a vulnerabilidade da sociedade ao impacto das *fake news*.

Como limitação, destaca-se a dificuldade de mensuração do impacto real das campanhas de desinformação geradas por IA, uma vez que a disseminação desses conteúdos ocorre de maneira descentralizada e muitas vezes oculta em redes sociais e aplicativos de mensagens. Além disso, a evolução constante das tecnologias de IA gera desafios na atualização de abordagens regulatórias, tornando algumas análises rapidamente obsoletas.

Para pesquisas futuras, sugere-se a realização de estudos empíricos que avaliem, por meio de experimentos e análises de *big data*, o real alcance das campanhas de desinformação e os efeitos das regulamentações na redução da manipulação informacional. Ademais, investigações qualitativas podem explorar como cidadãos percebem e reagem a conteúdos gerados por IA, especialmente no contexto político e eleitoral.

Assim, conclui-se que diante da rápida evolução das tecnologias de IA, a sociedade enfrenta um desafio contínuo na busca por soluções que preservem a integridade da informação sem comprometer valores democráticos. O futuro da governança digital dependerá da colaboração entre governos, empresas de tecnologia, instituições acadêmicas e sociedade civil para criar um ecossistema informacional mais transparente, seguro e confiável. O avanço da inteligência artificial é irreversível, mas sua aplicação pode ser guiada por princípios que priorizem a ética, a veracidade e a proteção da democracia. A forma como esses desafios forem enfrentados nos próximos anos definirá não apenas o futuro da informação digital, mas também os rumos das sociedades democráticas no século XXI.

REFERÊNCIAS

BALDISSERA, Wellington Antonio; FORTES, Vinícius Borges. Regulação das fake news: um dilema diante do direito à liberdade de expressão. **Direito e Desenvolvimento**, v. 12, n. 1, p. 18–36, 2021. DOI: 10.26843/direitoedesenvolvimento.v12i1.957. Disponível em: <https://periodicos.unipe.br/index.php/direitoedesenvolvimento/article/view/957>. Acesso em: 14 mar. 2025.

BARDIN, L. **Análise de Conteúdo**. 1ª edição ed. [s.l.] Edições 70, 2015.

BOTELHO, T. H. F.; NÖTH, W. Deepfake: Inteligência Artificial para discriminação e geração de conteúdos. **TECCOGS: Revista Digital de Tecnologias Cognitivas**, n. 23, 2021.

BRADSHAW, S.; HOWARD, P. N. **The global disinformation order: 2019 global inventory of organised social media manipulation**. Oxford Internet Institute, 2019. Disponível em: <https://demtech.oii.ox.ac.uk/research/posts/the-global-disinformation-order-2019-global-inventory-of-organised-social-media-manipulation>. Acesso em: 08/03/2025.

BRENNEN, J. S.; SIMON, F.; HOWARD, P. N.; NIELSEN, R. K. **Types, sources, and claims of COVID-19 misinformation**. Reuters Institute, 2020. Disponível em: <https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation>. Acesso em: 08/03/2025.

CADWALLADR, C.; GRAHAM-HARRISON, E. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*, 17 mar. 2018. Disponível em: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>. Acesso em: 08/03/2025.

CHESNEY, R.; CITRON, D. Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. **Foreign Affairs**, v. 98, n. 1, p. 147-155, 2019.

CLAUSSEN, V. Fighting hate speech and fake news. The Network Enforcement Act (NetzDG) in Germany in the context of European legislation. **Media Laws**, v. 3, n. 3, p. 110–136, 2018.

COE, P. The Draft Online Safety Bill and the regulation of hate speech: have we opened Pandora's box? **Journal of Media Law**, v. 14, n. 1, p. 50–75, 2 jan. 2022.

CRESWELL, J. W.; POTH, C. N. **Qualitative inquiry and research design: Choosing among five approaches**. [s.l.] Sage publications, 2016.

DE CARVALHO JÚNIOR, O. L.; CARVALHO, S. S.; DE SOUSA, B. A. Democracia em xeque: inteligência artificial e Deep Fake. **Revista JRG de Estudos Acadêmicos**, v. 7, n. 15, p. e151743–e151743, 2024.

EUROPEAN COMMISSION. **The Digital Services Act package**. 2022. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R2065>. Acesso em: 08/03/2025.

GORWA, R. What is platform governance? **Information, Communication & Society**, v. 22, n. 6, p. 854-871, 2019.

GUSTIN, M. B. DE S.; DIAS, M. T. F.; NICÁCIO, C. S. **(Re)pensando a Pesquisa Jurídica: Teoria e Prática**. 5. ed. São Paulo: Almedina, 2020.

GRAVES, L.; CHERUBINI, F. The rise of fact-checking sites in Europe. **Reuters Institute for the Study of Journalism**, 2016. Disponível em:

<https://reutersinstitute.politics.ox.ac.uk/our-research/rise-fact-checking-sites-europe>. Acesso em: 08/03/2025.

HERITAGE, C. **Proposed Bill to address Online Harms**. Disponível em: <<https://www.canada.ca/en/canadian-heritage/services/online-harms.html>>. Acesso em: 14 mar. 2025.

HORNE, B. D.; KHEDR, S.; ADALI, S. Sampling the news producers: A large news and feature data set for the study of the complex media landscape. **Proceedings of the 13th International AAAI Conference on Web and Social Media**, p. 249-258, 2019.

HOWARD, P. N. **Lie machines: How to save democracy from troll armies, deceitful robots, junk news operations, and political operatives**. [s.l.] Yale University Press, 2020.

KOSSEFF, J. **The twenty-six words that created the internet**. Cornell University Press, 2019.

LEARY, M. G. The indecency and injustice of section 230 of the Communications Decency Act. **Harv. JL & Pub. Pol’y**, v. 41, p. 553, 2018.

LEWANDOWSKY, S.; ECKER, U. K. H.; COOK, J. Beyond misinformation: Understanding and coping with the "post-truth" era. **Journal of Applied Research in Memory and Cognition**, v. 6, n. 4, p. 353-369, 2017.

MEZZAROBBA, O.; MONTEIRO, C. S. **Manual de Metodologia da Pesquisa no Direito**. 7ª edição ed. São Paulo: Saraiva, 2017.

NEWMAN, N.; FLETCHER, R.; SCHULZ, A.; ANDI, S.; NIELSEN, R. K. **Reuters Institute Digital News Report 2022**. Reuters Institute for the Study of Journalism, 2022. Disponível em: <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2022>. Acesso em: 08/03/2025.

PARISER, Eli. **The filter bubble: How the new personalized web is changing what we read and how we think**. Penguin Books, 2012.

PIMENTA, Ricardo; REGLY, Tainá; SILVA, Anna Karla S; BELFORT, Hugo. A MECÂNICA DA DESINFORMAÇÃO ELEITORAL: fake news e o paralelo com as “trend topics” das redes sociais em 2022. **Encontros Bibli** 29, 2024. <https://doi.org/10.5007/1518-2924.2024.e100310>.

QUEIROZ, R. M. R.; FEFERBAUM, M. **Metodologia da Pesquisa em Direito: técnicas e abordagens para elaboração de monografias, dissertações e teses**. 2. ed. São Paulo: Saraiva, 2021.

RAMOS-ZAGA, F. Deepfake: Análisis de sus implicancias tecnológicas y jurídicas en la era de la Inteligencia Artificial. **Derecho global. Estudios sobre derecho y justicia**, v. 9, n. 27, p. 359–387, 2024.

ROBL FILHO, I. N.; MARRAFON, M. A.; MEDÓN, F. A inteligência artificial a serviço da desinformação: como as deepfakes e as redes automatizadas abalam a liberdade de ideias no debate público e a democracia constitucional e deliberativa. **Economic Analysis of Law Review**, v. 13, n. 3, p. 32–47, 2022.

RUGGE, Fabio. AI in a Contested Cyberspace in RUGGE, F. (ed.) **AI in the Age of Cyber-Disorder**. ISPI and The Brookings Institution, 2020.

SHAO, C., CIAMPAGLIA, G.L., VAROL, O. et al. The spread of low-credibility content by social bots. **Nature communications**, v. 9, n.1, p. 4787, 2018.
<https://doi.org/10.1038/s41467-018-06930-7>.

THORNE, J.; VLACHOS, A. Automated fact-checking: Task formulations, methods and future directions. **Proceedings of the 27th International Conference on Computational Linguistics**, p. 3346-3359, 2018.

TREASURY, D. OF THE. **News Media and Digital Platforms Mandatory Bargaining Code - The Code's first year of operation** | **Treasury.gov.au**. text. Disponível em: <<https://treasury.gov.au/publication/p2022-343549>>. Acesso em: 14 mar. 2025.

TUCKER, J. A.; GUESS, A.; BARBERA, P.; SIEGEL, A. Social media, political polarization, and political disinformation: A review of the scientific literature. **Political Science Quarterly**, v. 133, n. 3, p. 611-643, 2018.

UK PARLIAMENT. **Parliamentary Bills - UK Parliament**. Disponível em: <<https://bills.parliament.uk/>>. Acesso em: 14 mar. 2025.

VERDOLIVA, L. Media forensics and deepfakes: An overview. **IEEE Journal of Selected Topics in Signal Processing**, v. 14, n. 5, p. 910-932, 2020.

VOSOUGHI, S.; ROY, D.; ARAL, S. The spread of true and false news online. **Science**, v. 359, n. 6380, p. 1146-1151, 2018.

ZOLLO, F. et al. Emotional dynamics in the age of misinformation. **PLoS ONE**, v. 12, n. 9, 2017.