XIII ENCONTRO INTERNACIONAL DO CONPEDI MONTEVIDÉU - URUGUAI

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

EDSON RICARDO SALEME

EUDES VITOR BEZERRA

CINTHIA OBLADEN DE ALMENDRA FREITAS

Copyright © 2024 Conselho Nacional de Pesquisa e Pós-Graduação em Direito

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
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

Membro Nato – Presidência anterior Prof. Dr. Raymundo Juliano Feitosa – UNICAP – Pernambuco

D597

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

[Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Edson Ricardo Saleme, Eudes Vitor Bezerra, Cinthia Obladen de Almendra Freitas – Florianópolis: CONPEDI, 2024.

Inclui bibliografia
ISBN: 978-85-5505-990-2
Modo de acesso: www.conpedi.org.br em publicações
Tema: ESTADO DE DERECHO, INVESTIGACIÓN JURÍDICA E INNOVACIÓN

1. Direito – Estudo e ensino (Pós-graduação) – 2. Governo digital. 3. Novas tecnologias. XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU (2: 2024 : Florianópolis, Brasil).

CDU: 34



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

XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

Apresentação

O conjunto de pesquisas que são apresentadas neste livro faz parte do Grupo de Trabalho de "DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II", ocorrido no âmbito do XIII Encontro Internacional do CONPEDI, realizado entre os dias 18, 19 e 20 de setembro de 2024, na cidade de Montevidéu, Uruguai, promovido pelo Conselho Nacional de Pesquisa e Pós-Graduação em Direito – CONPEDI e que teve como temática central "Estado de Direito, Investigação Jurídica e Inovação".

Os trabalhos expostos e debatidos abordaram de forma geral distintas temáticas atinentes DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS, especialmente relacionadas aos principais desafios que permeiam a tecnologias jurídica, passando pela inteligência artificial, demais meios digitais, também apontando para problemas emergentes e propostas de soluções advindas de pesquisas em nível de pós-graduação, especialmente, Mestrado e Doutorado.

Os artigos apresentados no Uruguai trouxeram discussões sobre: Tecnologias aplicáveis aos tribunais, Governança digital e governo digital, Função notarial e novas tecnologias, Exclusão digital derivando tanto para exclusão social quanto para acesso à justiça, Eleições, desinformação e deepfake, cidades e TICs. Não poderiam faltar artigos sobre privacidade e proteção de dados pessoais, com atenção aos dados sensíveis, consentimento e LGPD, liberdade de expressão, censura em redes sociais, discriminação, herança digital, microtrabalho e o trabalho feminino, uso de sistemas de IA no Poder Judiciário e IA Generativa.

Destaca-se a relevância e artigos relacionados ao tema de Inteligência Artificial, tratando de vieses algorítmicos e do AI Act. E, ainda, aplicação de sistemas de IA ao suporte de pessoas com visão subnormal. Para além das apresentações dos artigos, as discussões durante o GT foram profícuas com troca de experiências e estudos futuros. Metodologicamente, os artigos buscaram observar fenômenos envolvendo Direito e Tecnologia, sem esquecer dos fundamentos teóricos e, ainda, trazendo aspectos atualíssimos relativos aos riscos que ladeiam as novas tecnologias, destacando os princípios e fundamentos dos direitos fundamentais

Considerando todas essas temáticas relevantes, não pode ser outro senão de satisfação o sentimento que nós coordenadores temos ao apresentar a presente obra. É necessário, igualmente, agradecer imensamente aos pesquisadores que estiveram envolvidos tanto na confecção dos trabalhos quanto nos excelentes debates proporcionados neste Grupo de Trabalho. Por fim, fica o reconhecimento ao CONPEDI pela organização e realização de mais um relevante evento internacional.

A expectativa é de que esta obra possa contribuir com a compreensão dos problemas do cenário contemporâneo, com o a esperança de que as leituras dessas pesquisas ajudem na reflexão do atual caminhar do DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS.

Prof. Dr. Edson Ricardo Saleme (UNISANTOS)

Prof. Dr. Eudes Vitor Bezerra (PPGDIR – UFMA)

Profª. Dra. Cinthia Obladen de Almendra Freitas (PPGD - PUCPR)

CAOS INFORMACIONAL E DEEPFAKES NAS ELEIÇÕES BRASILEIRAS EM 2024: A DESINFORMAÇÃO COMO TÁTICA POLÍTICA

INFORMATIONAL CHAOS AND DEEPFAKES IN THE BRAZILIAN ELECTIONS OF 2024: DISINFORMATION AS A POLITICAL TACTIC

Cesar Beck [1]
Thami Covatti Piaia [2]
Juliana Oliveira Sobieski [3]

Resumo

Esse artigo tem como objetivo central analisar os possíveis impactos das deepfakes para as eleições municipais brasileiras de 2024. Deepfakes são mídias sintéticas, de áudio e vídeo, sintetizadas artificialmente por meio de técnicas digitais, que simulam conteúdos reais com alta sofisticação. Essa prática tem sido utilizada como uma das formas de desinformação política neste século, por manipular rostos e vozes de pessoas reais com fins escusos. Dessa forma, questiona-se em que medida esse fenômeno pode influenciar o processo eleitoral brasileiro e analisa estratégias que visam coibir o caos informacional gerado pelo grande fluxo de informações falsas e falsificadas visto nas eleições nos últimos anos. A pesquisa reflete acerca do amplo fenômeno da desinformação, os tipos de informações envolvidos e analisa os mecanismos de prevenção e mitigação dessa problemática. Conclui-se que as ações para preservar a lisura do processo eleitoral depende da participação conjunta de diversos atores sociais e organizações, uma vez que o fenômeno é multifacetado. Entende-se que tal problemática pode resultar em muitos danos para a democracia, no Brasil e no mundo.

Palavras-chave: Deepfakes, Eleições brasileiras, Desinformação, Desordem informacional, Novas tecnologias

Abstract/Resumen/Résumé

The main objective of this article is to analyze the possible impacts of deepfakes on the 2024 Brazilian municipal elections. Deepfakes are synthetic audio and video media, artificially synthesized using digital techniques that simulate real content with high sophistication. This practice has been used as one of the forms of political disinformation in this century, by manipulating the faces and voices of real people for shady purposes. It questions the extent to

[1] Universidade Regional Integrada do Alto Uruguai e das Missões, Santo Ângelo, RS, Brasil. Advogado; Mestre em Direito. Pesquisador Acadêmico e Comentarista Político. Lattes: http://lattes.cnpq.br/6488052912983557. E-mail: ocesarbeck@gmail.com

[2] Universidade Regional Integrada do Alto Uruguai e das Missões, Santo Ângelo, RS, Brasil. Advogada; Doutora em Direito. Pesquisadora Acadêmica. Lattes: http://lattes.cnpq.br/9871466774374427. E-mail: thamicovatti@san.uri.br

[3] Universidade Atitus. Advogada; Mestranda em Direito. Pesquisadora Acadêmica. Lattes: http://lattes.cnpq.br /9401654264861549E-mail: julianasobieski.js@gmail.com

which this phenomenon can influence the Brazilian electoral process and analyzes strategies aimed at curbing the informational chaos generated by the large flow of false and falsified information seen in elections in recent years. The research reflects on the broad phenomenon of disinformation and the types of information involved and analyzes the mechanisms for preventing and mitigating this problem. It concludes that actions to preserve the fairness of the electoral process depend on the joint participation of various social actors and organizations since the phenomenon is multifaceted. It is understood that this problem can result in a great deal of damage to democracy in Brazil and around the world.

Keywords/Palabras-claves/Mots-clés: Deepfakes, Brazilian elections, Disinformation, Information disorder, New technologies

INTRODUÇÃO

Em 2018, um vídeo[1] viral do ex-presidente americano Barack Obama foi divulgado no Youtube e concentrou rapidamente quase 8 milhões de visualizações. O título do vídeo, que utiliza uma técnica de caça-cliques - por ter apelo emocional e senso de urgência -, era "você não vai acreditar no que Obama diz neste vídeo!". No material, Obama olha diretamente para a câmera. O cenário parece ser o Salão Oval da Casa Branca. No vídeo, Obama diz que há uma era atual na qual os inimigos podem fazer parecer que alguém disse algo, em qualquer momento, mesmo que a pessoa nunca tenha dito nada. E exemplifica: eles poderiam me fazer dizer algo como "Trump é um imbecil!" e, por fim, afirma que nunca diria isso, pelo menos não em um discurso público.

O vídeo em questão foi uma criação do diretor de cinema Jordan Peele. Toda a mensagem do vídeo, na verdade, foi dita pelo próprio diretor, que utilizou posteriormente um programa de edição para manipular a imagem de Obama e fazer parecer que ele havia pronunciado as palavras. Segundos após a imagem de Obama, Jordan Peele se revela e aponta os riscos atuais dos usos de tecnologia de manipulação audiovisual e reflete sobre a necessidade de desconfiança e análise crítica sobre os conteúdos divulgados na internet, além de apontar a importância de consumir fontes confiáveis de informação (Benevides, 2018).

Em junho de 2022, circularam vídeos[2] em que o presidente Luiz Inácio Lula da Silva aparecia sendo vaiado nas cidades de Uberlândia e em Garanhuns, durante a sua agenda de campanha política. As vaias, no entanto, não eram reais. O presidente cumpria parte de sua agenda política nas cidades, mas nos vídeos originais não havia qualquer hostilidade por parte do público. O vídeo falso foi compartilhado por milhares de usuários no TikTok.

Em janeiro de 2024, o pré-candidato à prefeitura de Maringá pelo Partido Progressista Silvio Barros (PP-RJ), teve um áudio[3] atribuído a ele, no qual ele supostamente informava que desistiu de concorrer às eleições municipais. No mesmo áudio, ele declara apoio ao seu adversário de candidatura. O arquivo em questão, compartilhado diversas vezes por pessoas

[1] BENEVIDES, Bruno. **Diretor Jordan Peele faz vídeo com Obama para denunciar notícias falsas**. 18 abr. 2018. Disponível em: https://www1.folha.uol.com.br/mundo/2018/04/diretor-jordan-peele-faz-video-com-obama-para-denunciar-noticias-falsas.shtml. Acesso em: 2 fev. 2024.
[2] COELHO, Gabriela; HIRABAHASI, Gabriel. Cármen Lúcia manda remover vídeos contra Lula com “deep fake” | CNN Brasil. 28 out. 2022. Disponível em: https://www.cnnbrasil.com.br/politica/carmen-lucia-manda-remover-videos-contra-lula-com-deep-fake/. Acesso em: 2 fev. 2024.
[3] GÓES, Bruno. **Casos de ‘deep fake’ chegam à Justiça Eleitoral, que cobra WhatsApp**. 24 jan. 2024. Disponível em: https://oglobo.globo.com/politica/noticia/2024/01/24/casos-de-deep-fake-chegam-a-justica-eleitoral-que-cobra-whatsapp.ghtml. Acesso em: 2 fev. 2024.

pelo Whatsapp, tem o timbre, o tom de voz e o jeito de falar do político. O candidato lidera as pesquisas locais das eleições municipais de 2024.

O prefeito de Manaus, David Almeida (AVANTE-AM), teve um áudio atribuído a ele no final de 2023, que dizia que os professores da rede pública são vagabundos, além de outras ofensas semelhantes. O áudio, que tinha uma voz muito parecida com a dele, circulou em diversos grupos nas redes sociais. Essa ação veio após o anúncio de que o pagamento de um abono para a categoria dos professores não seria feito. Ainda no áudio, há uma mensagem de que o prefeito não pagou o FUNDEB, mas que o povo iria esquecer.

Todos os arquivos digitais mencionados acima, e disseminados pelas redes sociais para milhares de pessoas acessarem e compartilharem, trata-se de *deepfakes*. Falsificações de áudio e vídeo de alto nível, que visam atribuir a determinadas pessoas mensagens que elas não emitiram. Os episódios brasileiros citados têm em comum a desinformação como tática para difamar determinada figura política ou reduzir o seu alcance. Eleições não são arenas neutras, nunca foram. Dentre as táticas utilizadas com o objetivo de atingir os adversários, as notícias falsas ou a exposição de comportamentos moralmente comprometedores - devido às compreensões sociais sobre o que é aceitável - são utilizadas como instrumentos de guerra.

O que há, no cenário atual, é a elevação dos fenômenos de difamação a patamares problemáticos, posto que há a facilidade no compartilhamento de informações e a horizontalização das fontes de acesso a informações, o que pode incluir, também, conteúdo fraudulento. Na arena política, uma notícia falsa - ou parcial - é sempre um meio de descredibilizar o outro. Reflete-se, sobretudo, acerca das ameaças que tais armas políticas podem significar para as eleições brasileiras de 2024, especialmente porque tais táticas foram utilizadas anteriormente, de forma massiva, no Brasil e no Mundo. Suas consequências para a democracia são inúmeras.

Os olhos humanos foram destacados, na história das representações, como símbolo de constatação da verdade. É comum que as pessoas utilizem expressões como "eu vi com os meus próprios olhos", para confirmar que determinada coisa tem o status da veracidade. Essa foi uma forma confiável de, se não atestar que algo é verídico, no mínimo constituir uma espécie de prova. Mas e quando os olhos humanos não são mais suficientes para confirmar a verdade de uma notícia? E se os olhos pudessem ser traídos por uma simulação, uma construção quase perfeita que à primeira vista pode nublar o discernimento até mesmo dos mais atentos?

Mulholland e Oliveira (2021) questionaram anteriormente a incapacidade dos olhos de serem elementos críveis de testemunho no âmbito do ecossistema digital. Os exemplos utilizados nas primeiras páginas retratam fielmente a tranquilidade com a qual é possível manipular a imagem de alguém atualmente e fazê-lo proferir discursos reprováveis. Cada vez mais, essas técnicas adquirem características hiper-realistas, o que acirra os debates sobre a sua utilização em diversos meios, especialmente no âmbito político. A popularização do termo *fake news*, como sinônimo genérico para práticas de desinformação, foi um resultado de sua utilização massiva em eleições. Embora se compreenda que a dissimulação e as campanhas de difamação sejam dados antigos no campo político, se antes havia uma dinâmica de difamação de um para muitos, num ciclo de divulgação mais restrito, hoje a mesma notícia pode percorrer o mundo inteiro rapidamente, pela capilaridade da comunicação humana atual e pela facilidade de falsificação de informações no cenário atual.

Dessa maneira, esse artigo tem como objetivo central refletir acerca dos impactos das *deepfakes* para as eleições municipais em 2024. Questiona-se em que medida o fenômeno já conhecido na sociedade brasileira pode influenciar o processo eleitoral e, de maneira complementar, analisar como as estratégias institucionais visam coibir a desinformação como tática política de alta sofisticação. Como objetivos específicos, pretende-se conceituar a desinformação como fenômeno; explorar as nuances do termo genérico "*fake news*"; analisar as práticas atuais de combate das *deepfakes* nas eleições brasileiras e refletir sobre a alfabetização digital como um dos mecanismos de mitigação da problemática.

O trabalho está dividido em três capítulos e percorre, principalmente, o caminho metodológico da revisão teórica crítica, além de explorar os recursos das notícias jornalísticas e políticas como suporte à análise empírica da problemática levantada. O primeiro capítulo discute o fenômeno da desinformação, o surgimento das *fake news* e de sua importância política, assim como a sua evolução por meio das *deepfakes*. Esse capítulo visa explorar conceitualmente o tema e retratar as controvérsias e confusões no uso do termo *fake news* como sinônimo de desinformação.

O segundo capítulo apresenta os problemas das *deepfakes* para as eleições brasileiras municipais em 2024 e a ameaça que tais usos representam para manipular o pleito eleitoral, tendo consequências perversas para a democracia. Esse capítulo reflete acerca do potencial lesivo das *deepfakes* para as escolhas dos eleitores, para a vida das pessoas atingidas e para o contexto político geral. O terceiro capítulo explora o panorama atual de combate das *deepfakes* no Brasil, maneiras de detecção e arquitetura de mitigar os efeitos do fenômeno, assim como

debate a singularidade da alfabetização digital como instrumento de pavimentação de debates eleitorais justos.

1 DESINFORMAÇÃO, *FAKE NEWS* E *DEEPFAKES* NO ÂMBITO POLÍTICO

Os avanços tecnológicos exponenciais no início desse século consolidaram um patamar de disseminação de informações nunca visto. Antes da virada do século, um ecossistema da informação evoluía em um ritmo significativo, porém mais lento, o que permitia que a sociedade se adaptasse gradativamente aos avanços tecnológicos. Entre a invenção da impressa e o desenvolvimento da fotografia, por exemplo, vivenciaram-se quatro séculos. Nas últimas décadas, a humanidade criou a internet, os smartphones e as redes sociais e transformaram o ambiente de informação em uma coisa integrada. O mundo nunca esteve tão interligado, tanto do ponto de vista físico quanto virtual. As pessoas podem se mover a uma velocidade inacreditável, assim como as informações, e a arquitetura desse trânsito permite um fluxo livre de informações para as pessoas com acesso à internet (Van Raemdonck; Meyer, 2022).

Tal avanço trouxe oportunidades incríveis e desafiadoras. A acessibilidade informacional permitiu uma troca enorme de informações e contribuiu para colaborações internacionais, como foi o caso da luta contra o coronavírus. A partilha intensa de dados foi essencial para mapear o vírus, além de possibilitar usos estratégicos da tecnologia para benefício social. Por outro lado, fenômenos como a desinformação se difundem rapidamente, por meio de mecanismos técnicos e cognitivos que facilitam o seu enraizamento (Van Raemdonck; Meyer, 2022).

A sociedade contemporânea tem acompanhado, nas últimas décadas, a maior revolução tecnológica já vista na história, com uma evolução que, além de possibilitar o acesso amplo a informações e a capacidade de conexão mundial, teve também um grande impacto no comportamento das pessoas, pela própria mudança dos modos de vida com a formulação, ampliação e evolução da internet. As modificações cosmológicas como efeito das tecnologias de comunicação digital alteraram a estrutura das organizações e instituições sociais, o desenvolvimento de novos meios de comunicação e formas diversas de exercício da própria política nesse contexto (Mulholland; Oliveira, 2021).

Com o surgimento de uma mídia ampla, independente e horizontalizada - pelo menos em tese, como resultado do acesso à internet - muitos processos sociais complexos coexistem e modificaram a vida das pessoas, como cidadãos, na esfera pública. Se em décadas passadas os jornais televisivos e impressos eram algumas das fontes centrais de informações políticas, hoje

é possível acessar informações da esfera política de uma série de fontes, pouco organizadas, que alimentam cotidianamente as redes sociais e possuem uma capacidade de difusão nunca vista (Colomina, 2022).

A dinâmica social e política se movimenta em face dessas modificações, uma vez que os cidadãos não ocupam mais espaços passivos como espectadores, pois tornaram-se protagonistas nesse processo de consumir informações, mesmo que isso signifique, no contexto atual, receber, processar e compartilhar diversos tipos de informações falsas. A aparente horizontalização da comunicação política por meio da internet é apenas um elemento, um pouco controverso, porque os fenômenos sociais recentes relacionados à política mundial têm demonstrado que, embora seja fácil dizer qualquer coisa na internet, é particularmente difícil ser ouvido (Hindman, 2008). Dessa maneira, as informações repercutem nas redes sociais por meio de bolhas e embora o compartilhamento orgânico seja um dado relevante para a discussão da desinformação, na prática, há uma rede profunda e complexa de pessoas e contas automatizadas que disparam informações falsas nas redes sociais, em perfis que por vezes parecem fidedignos, mas representam o império da desinformação utilizado para difamar pessoas, inclusive os políticos.

O ambiente digital transformou-se num potente instrumento de socialização e exposição de opinião para os indivíduos, mas abriu a porta para um tipo de anonimato - temporário - muito perigoso do ponto de vista social, pois utilizam-se desses perfis falsos para disseminar ódio, cometer crimes diversos e disseminar conteúdo falso. O problema das notícias falsas, nesse contexto, é que elas se espalham muito fácil e largamente, com muito mais persuasão do que informações categoricamente verdadeiras e não maliciosas. Isso porque, junto a essas notícias, utilizam-se táticas diversas para captar a atenção das pessoas que lerão. Ênfase e exagero são partes fundamentais das operações caça-cliques: criam uma espécie de urgência e espetáculo que mexem psicologicamente com as pessoas, como se elas não pudessem seguir adiante sem acessar a informação (Van Raemdonck; Meyer, 2022).

Isso porque, no âmbito da mudança dos modos de vida, há também uma série de operações neurológicas, do funcionamento comportamental e psíquico do ser humano, que se movimentou para comportar a quantidade de informações e possibilidades de acesso que as novas configurações sociais possuem. Tais modificações emitem uma sequência de estímulos, contínuos e ininterruptos, posto que é possível navegar infinitamente nas redes sociais atualmente, que afetam o sistema dopaminérgico dos indivíduos (Van Raemdonck; Meyer, 2022).

As notícias falsas estimulam zonas emocionais dos indivíduos que torna o compartilhamento, mesmo equivocado, uma prática comum, diante da necessidade emocional de fazer parte. Explica-se: as redes sociais elevam o senso de comunidade, uma vez que é possível não apenas se conectar com pessoas íntimas da esfera pessoal, mas com uma rede de pessoas que compartilham valores e ideologias. A desinformação utiliza esse desejo de conexão para espalhar conteúdo malicioso porque, a partir da necessidade de participar, as pessoas compartilham e comentam as informações como parte dessa dinâmica, da arquitetura das plataformas, que conduz à necessidade de ter determinado comportamento para ser visto (Wardle; Derakhshan, 2017).

Quanto maior o compartilhamento, eleva-se também o sentimento de pertencimento. Precisamente por isso que, na ânsia de participar de suas comunidades online, indivíduos divulgam informações equivocadas e descontextualizadas, mesmo que sem a intenção direta de desinformar. Tal dado evoca, igualmente, a arquitetura das redes e das aplicações, que requerem a participação das pessoas como requisito para permanecer, para serem vistos. É o caso, por exemplo, do modelo de curtidas, comentários e compartilhamentos. Nessa dinâmica da participação e da urgência por compor o ciclo de informações, os conteúdos emocionais se espalham amplamente (Wardle; Derakhshan, 2017).

Embora pareça uma digressão, reflete-se que a forma da arquitetura ou *design* das redes – em constante evolução – requer uma presença onipresente e excessiva das pessoas. Há que se considerar, no fenômeno da desinformação, que essa necessidade de presença desencadeada pela arquitetura das aplicações, contribui para que as pessoas tenham gatilhos imediatos com informações que parecem "bombásticas", mesmo que elas sejam falsas. A necessidade de participar, a economia massiva por atenção ininterrupta e as práticas de vigilância são comportamentos estimulados pela própria arquitetura desse ecossistema digital. Assim, tal padrão é um dado importante quando se trata do compartilhamento acrítico de desinformação por pessoas comuns, que acreditam fielmente na veracidade das informações (Sarlet, 2021).

O atual ecossistema de informação, conceituado nesse trabalho como caos informacional, em alusão à noção empregada por Shick (2020), recebe tal denominação em função da poluição, amplitude e perigo das informações nesse ambiente contemporâneo. As sociedades globais vivem, de maneira geral, uma sobrecarga de informações que altera significativamente a percepção e o comportamento de indivíduos, além de ter que lidar com uma onda ininterrupta que desafia os limites do que a capacidade crítica pode processar.

É fundamental destacar que esse caos não é estático, é um fenômeno potente e mutável, que se move à medida que se aperfeiçoam, também, as técnicas para produzir, criar e disseminar informações. A sobrecarga informacional produz diversos fenômenos e tem sido utilizada, também, no campo político. A guerra de narrativas tem sido ampliada no campo político como uma das formas de confundir e contrapor outras opiniões e percepções. É essa operação que Estados, organizações, instituições e pessoas empregam para influenciar desde questões geopolíticas até as vidas individuais (Shick, 2020).

As informações, neste século, podem ser tão perigosas e letais quanto mísseis e bombas. A desinformação é uma tática fundamental para o arsenal de armas políticas e tem sido utilizada globalmente nas guerras entre nações e de maneira local, nas guerras eleitorais internas. Sua ameaça reside na instabilidade que informações confusas podem gerar para a democracia, por criar polarizações políticas extremas e prejudicar o discernimento dos cidadãos que recebem uma ampla e desordenada gama de elementos para decidir. Uma vez que a guerra de informações produz alterações na percepção dos indivíduos sobre o fenômeno da disputa do poder, ela constitui uma ameaça aos processos democráticos (Colomina, 2022).

O que é mais importante nessa questão, no entanto, é que a desinformação está a serviço de um confronto tecnológico e digital que polariza as agendas internacionais e nacionais. A capacidade ofensiva das informações como armas, no entanto, deve-se menos ao conteúdo das mensagens e mais ao poder que as redes sociais têm para tornar essas mensagens virais e garantir a penetração delas na compreensão dos cidadãos. Muito embora o conteúdo seja muito relevante para a análise, ele amplifica a sua capacidade de dano pela forma como as redes sociais transformam rapidamente qualquer coisa em verdades bombásticas (Colomina, 2022).

Em termos geopolíticos, a lógica militar e a transformação tecnológica se uniram no espaço digital, de maneira que a internet se tornou um campo crucial de desestabilização de oponentes. Por vezes, no campo político, tem se construído ofensivas cibernéticas com informações diversas que visam alterar a percepção das pessoas e minar a realidade. Seja em guerras militares tradicionais, ou em guerras eleitorais, o conflito se alastra para as redes sociais, de forma que a arena política não se resume mais apenas à arena pública, presencial, dos debates, dos comícios e das agendas de campanha. Todas essas etapas das guerras eleitorais se transformaram em elementos híbridos, sendo que a disputa por narrativas em meios virtuais garantiu um patamar prioritário em função da capacidade que o ambiente digital tem de amplificar mensagens, especialmente as mensagens falsas e falsificadas (Van Raemdonck; Meyer, 2022).

Essa lógica de desestabilização democrática e da utilização da desinformação como tática política e militar se intensificou na última década, especialmente na guerra da Ucrânia, no que se refere à transmissão em *tempo real* da guerra, narrada de forma fragmentada. Há uma estratégia central na guerra de narrativas que é a fantasia da verdade. Ela se baseia na ideia de que é possível transmitir veracidade por meio de informações fragmentadas, de recortes de vídeos e de imagens que visam "retratar" a devastação em determinado contexto, no que tange às guerras militares. A mediação das guerras por meio das redes sociais utiliza muitos subelementos para moldar a narrativa global sobre determinado fenômeno. No campo das guerras eleitorais, a distorção também teve um papel fundamental, como foi o caso das eleições estadunidenses de 2016 (Colomina, 2022).

A desinformação é uma arma que pode ser utilizada em diversos contextos. No campo político, tem os seus usos centrais em contexto geopolítico, de guerras militares híbridas, e em disputas eleitorais. Em termos genéricos, entende-se que ela é utilizada para desestabilizar os espaços cívicos em diversos níveis, seja local, regional ou nacional, mas o seu trunfo central não é a natureza do conteúdo divulgado, mas a compreensão do poder que a interconexão digital tem no ecossistema político contemporâneo. Por isso, a desinformação não é uma mera campanha de criação e disseminação de notícias falsas, simplesmente, mas um fenômeno acoplado, *sine qua non*, à arquitetura dos ecossistemas digitais, que elevam a publicização das informações a escalas globais (Van Raemdonck; Meyer, 2022).

Embora não seja o objetivo central deste artigo, é essencial posicionar a maneira como a interconexão digital transformou as relações sociais e como modificou os equilíbrios de poder, tanto no que se refere às grandes potências, quanto aos novos atores das relações políticas, sejam estatais, não estatais ou privados. Por isso, a desinformação não pode ser estudada como um fenômeno alheio aos fatores sociopsicológicos, tecnológicos e dos estímulos próprios à hiperconectividade dessa era. Também por isso a desinformação é uma ameaça híbrida que se difere muito da propaganda política, uma tática muito tradicional de distorção, ou utilização estratégica, de informações verdadeiras para prejudicar os oponentes. A propaganda tem bases factuais que são apresentadas de forma tendenciosa para obter um resultado específico (Van Raemdonck; Meyer, 2022).

O primeiro movimento de combate ao fenômeno da desinformação deve incluir uma profunda discussão conceitual sobre cada subfenômeno incluído no tema. Essas diferenciações permitem visualizar como a falsificação de informações, em conjunto com a evolução tecnológica, avançou de forma rápida e consciente por parte de organizações e indivíduos que

promovem a desinformação de forma sistemática. Para esse capítulo, convém apresentar uma diferenciação entre *fake news*, além de refletir acerca da desinformação como um amplo fenômeno.

O termo *fake news*, utilizado como sinônimo de desinformação, é inadequado para descrever o complexo sistema de caos informacional visto no cenário político atual. Com isso, é fundamental apresentar uma estrutura conceitual que diferencie as diversas formas de desordem informacional, incluindo três tipos diferentes de informação: a desinformação, a informação incorreta e a má informação. A informação incorreta, embora falsa, não tem a intenção central de prejudicar algo ou alguém. Ela parte de conexões falsas e evoca contextos equivocados, mas não há prejuízo estimado. Seu compartilhamento pode desencadear ondas de difamação de uma pessoa, mas a sua intenção não é prejudicial (Wardle, 2017).

O elemento "intenção" é fundamental para entender que a desinformação é um fenômeno esperado, consciente e calculado, mas não sempre. A desinformação constrói falsos, manipulados e fabricados contextos. Ela é produzida deliberadamente para causar danos. A má informação é um tipo prejudicial, mas inclui informações genuínas. Seu uso é amplo no campo político, porque permite confrontar os valores e comportamentos do opositor perante os seus eleitores. Normalmente, informações verdadeiras são utilizadas no campo político para expor o outro a um patamar da vergonha, seja por confrontar sua ideologia ou por expor, de forma vexatória, questões de âmbito privado (Wardle; Derakhshan, 2017).

Acerca da intenção como elemento central na definição da desinformação, enfatiza-se que mesmo um conteúdo manipulado seja recebido genuinamente pelo indivíduo que o compartilha, essa prática também é desinformação. Muito embora se entenda que os cenários mais críticos se trata de campanhas intencionais de distribuição de conteúdo malicioso e falso, uma informação espalhada sem a intenção de gerar danos também pode repercutir de forma perigosa (Hasen, 2022).

Em uma alusão explicativa às categorias do direito penal, para fins didáticos, pode-se dizer que a má informação ou informação equivocada trata-se de uma conduta culposa, associada ao indivíduo em função de sua imprudência ou negligência, mas sem a consciência do efeito que pode causar. Em face desse comportamento desatencioso, é possível que outrem seja atingido, mas o agente de distribuição da mensagem não opera em um sentido intencional, o que incorre em uma culpa. O conteúdo falso e malicioso, por outro lado, é caracterizado como dolo, uma vez que há intenção e consciência da conduta e a finalidade é, de fato, causar danos.

Nesse sentido, ambas as espécies produzem efeitos práticos, mas a informação maliciosa contém, como elemento subjetivo, a intenção do agente de produção e de distribuição.

O ecossistema da desinformação inclui três elementos centrais: o tipo de conteúdo criado e compartilhado; as motivações envolvidas no compartilhamento e as formas de divulgações deste. Também por isso esse ecossistema se divide em três elementos: os agentes que criam, as mensagens e quem interpreta as informações. A importância dessa dissecação do fenômeno reside no fato de que quem produz uma mensagem, por exemplo, não necessariamente a distribui. A análise da desinformação tem que, por necessidade, refletir profundamente sobre quem são os agentes que produzem essas informações e quais são as motivações por trás, sobretudo porque essas motivações definem as subcategorias da informação. Num debate sobre a fabricação consciente de notícia falsa, por exemplo, entende-se que a sofisticação das técnicas utilizadas será maior do que da informação equivocada, porque a segunda não tem como motivação provocar o dano (Wardle, 2017).

O termo *fake news*, que pode ser traduzido literalmente como notícias falsas, tem acoplado a uma expressão pouco rigorosa uma série de práticas relacionadas ao fenômeno da desinformação. Por isso, por mais que o tema tenha se popularizado nos últimos anos, especialmente no contexto político, a ausência de definição incorre em alguns equívocos. Caracterizar o que seria *fake news*, de maneira genérica, é difícil, porque seus usos têm sido vastos e até mesmo com significado esvaziado. Por isso, muitos autores e figuras têm rejeitado o uso do conceito, pela compreensão de que é possível fabricar significados muito confusos para esta expressão (Nonô, 2022).

A cobertura midiática o tem utilizado nos últimos anos como sinônimo de publicações virais baseadas em relatos fictícios, que são feitas para parecerem reportagens críveis. Alguns estudos o definem como artigos de notícias que são intencionalmente falsos e visam enganar os eleitores. Ao longo da última década, o termo serviu para descrever diversos conteúdos, mesmo aqueles sem intenção de danos. O que é comum entre todos os tipos de informação é o verniz de legitimidade que utilizam para serem críveis, pois assumem algum símbolo de credibilidade para parecerem informações reais. A expressão se popularizou durante a campanha presidencial dos Estados Unidos em 2016 e desde então tem sido utilizada como termo guarda-chuva para o fenômeno da desordem informacional. Seu uso, embora popularizado, é desaconselhado porque o termo mais confunde do que explica (Tandoc; Lim; Ling, 2017).

Como um termo que abrange todas as formas de notícias incorretas, equivocadas e mal-intencionadas, tende a dissolver o debate mais sério sobre as campanhas intencionais de notícias falsas. Há, no bojo das *fake news*, notícias reais que não merecem a atenção pública, por exemplo, mas que integram a categoria genérica pela falta de rigor conceitual. Entende-se que a popularização do termo gerou certa confusão sobre as formas de desinformação que efetivamente causam riscos para vidas individuais e para a democracia. A expressão tem sido utilizada como sinônimo de sátira, paródia, manipulação, publicidade, fabricação e uma série de produções informacionais diferentes relacionadas à política. Seu uso foi apropriado também por políticos para rebater qualquer conteúdo crítico, inclusive fidedigno, que eles considerassem desagradáveis. Com isso, tornou-se um mecanismo para reprimir, minar e contornar também a imprensa livre que cobre as ações políticas de maneira crítica (Wardle; Derakhshan, 2017).

Wardle (2017) empreendeu, assim como outros autores, um esforço para definir o atual ecossistema da informação, de maneira que o uso do termo *fake news* não tivesse que surgir genericamente com "aspas" para identificar o desconforto e a inadequação em utilizá-lo. A razão pela qual esse termo parece genérico é que se trata de um fenômeno maior do que notícias, simplesmente, mas de um ecossistema informacional, assim como "falso" não descreve adequadamente a complexidade dos tipos de desinformação e como esses diferentes tipos são articulados nas práticas políticas e interferem na compreensão dos cidadãos. A autora descreve pelo menos sete tipos de informações problemáticas que compõem o ecossistema da desinformação, sobretudo porque essa definição tem como objetivo jogar luz especialmente para as práticas intencionais e coordenadas de falsificação política de notícias. As sátiras ou paródias não têm intenção de causar danos, mas podem causar zombaria ou enganar os cidadãos, posto que as interpretações dos conteúdos em ambientes virtuais são polissêmicas. O conteúdo enganoso utiliza informações para enquadrar um tópico ou um indivíduo. O conteúdo impostor pode causar prejuízos significativos, porque reproduz e falsifica a aparência de fontes críveis de notícias - como um jornal popular - para disseminar notícias falsas.

Do ponto de vista contextual, a desinformação tem três tipos: falsas conexões, quando o apelo inicial, seja do título, das imagens ou dos subtextos destoam completamente do conteúdo central da informação; falso contexto, quando um conteúdo genuíno é compartilhado em um contexto falso e fabricado; e o conteúdo manipulado, quando informações textuais ou imagéticas são manipuladas para emitir mensagens falsas. O principal foco desse artigo é o conteúdo fabricado, quando há manipulação total de uma informação, que é criada do zero, com

o intuito de causar danos. Essa modalidade de informação tem um potencial significativo de danos e nos últimos anos se tornou mais popular, em função da sofisticação dos métodos de fabricação no ecossistema informacional. As notícias falsas antes eram estáticas, editadas muitas vezes em baixa qualidade, o que permitia identificar mais rapidamente sua falsidade (Wardle; Derakhshan, 2017).

Até pouco tempo, manipular os meios de comunicação, seja em fotos, vídeos ou áudios era considerado como um domínio técnico especializado ou da ordem de pessoas com muitos recursos. Ao citar "manipulação digital", imaginava-se um estúdio de Hollywood na criação de efeitos especiais para um novo filme. Porém, a democratização da tecnologia tornou essa manipulação acessível a todos. A inteligência artificial permitiu que essa manipulação gerasse formas de comunicação completamente sintéticas - e/ou falsas (Shick, 2020). As *deepfakes* são uma evolução das *fake news,* no sentido de que as *fake news* se travam de conteúdo estático produzido com o intuito de criar contextos falsos, especialmente por meio de imagens. As *deepfakes*, por outro lado, representam uma nova forma de manipular materiais visuais e audiovisuais, por meio de inteligência artificial generativa. Essas tecnologias podem ser utilizadas para alterar imagens, vídeos e áudios reais, de forma realista e de difícil detecção, ou podem ser usadas para criar um conteúdo completamente do zero.

O termo *deepfake* se refere a uma enorme gama de falsificações hiper-realistas de materiais fotográficos, de áudio e audiovisuais realizadas digitalmente. Trata-se da junção entre os termos *deep learning*, um ramo do aprendizado de máquina que realiza aprendizado profundo, e *fake*, que deu origem à expressão. Essas técnicas utilizam algoritmos de aprendizado profundo para inserir rostos e vozes em gravações de áudio e vídeo reais, ou para sintetizar desde o zero os materiais, por meio da junção de elementos - dados anteriores - que permitem a criação de personificações realistas. O resultado atualmente é tão verossímil, que se torna impossível distinguir, à primeira vista, o que é real do que foi sintetizado digitalmente (Mulholland; Oliveira, 2021).

Na definição simplificada de Shick (2020), uma *deepfake* é uma mídia sintética, o que inclui imagens, áudio e vídeo, manipulada ou totalmente gerada por meio de inteligência artificial. Os avanços recentes da inteligência artificial podem gerar meios de comunicação completamente sintéticos; atualmente, a tecnologia já é acessível a ponto de que qualquer pessoa com um smartphone pode produzir alguns tipos de *deepfakes* mais simples, mas também é possível gerar conteúdos, com mais habilidade e esforço, particularmente críveis. Os usos iniciais das *deepfakes* tiveram sentido pornográfico. Se popularizou em 2017, na rede social

Reddit, após um usuário usar a expressão para postar vídeos pornográficos criminosos alterados digitalmente com imagens de pessoas públicas e celebridades, um fenômeno conhecido como *deep porn*. Para alimentar o algoritmo de sintetização, o usuário usava inúmeras imagens e vídeos das pessoas para ensinar a tecnologia a imitar as expressões da pessoa em questão e sobrepunha o seu rosto ao de uma atriz de filme pornô. A técnica é a mesma, mas os usos se modificaram muito.

Há utilizações inofensivas, criativas e artísticas das *deepfakes*, mas era esperado que as utilizações maliciosas e perturbadoras se tornassem quantitativamente expressivas e problemáticas como questão social. Já o era uma vez que expunha pessoas em conteúdo pornográfico nunca realizado por estas, mas as suas utilizações se ampliaram, inclusive na arena política. No capítulo seguinte, trata-se dos problemas das *deepfakes* em termos tecnológicos e políticos, assim como a sua relevância para as eleições brasileiras de 2024.

2 POR QUE TEMER AS *DEEPFAKES* NAS ELEIÇÕES MUNICIPAIS DE 2024?

Os avanços de novas tecnologias, notadamente em inteligência artificial nos últimos anos, foram muito úteis para o surgimento de *deepfakes*. A inteligência artificial, de maneira geral, possui desenvolvimentos e ramificações diversas: o aprendizado de máquina, do inglês *machine learning*, que pode a grosso modo ser descrito como um subconjunto da inteligência artificial que treina as máquinas para aprenderem por meio da alimentação de alto volume, variedade, velocidade e veracidade de dados armazenados em *big datas*, *data lakes* e *cloud computing*; as Redes Adversárias Generativas, ou GAN's, que são arquiteturas de redes neurais profundas, compostas por duas redes, que disputam entre si, para gerar novos dados autênticos a partir de um conjunto de dados de treino. Na competição entre as duas redes neurais, elas criam o melhor resultado imagético. Na criação de *deepfakes*, as Redes Adversárias Generativas são utilizadas para criar vídeos ou imagens hiper-realistas (Caldera, 2019).

Nessa operação, em rápida explicação, utiliza-se dados previamente existentes, como imagens e vídeos, mas substitui-se os componentes sintéticos dos dados, pois o programa de computador encontra os pontos e gera um novo conteúdo, calculado de forma autônoma para gerar as imagens e vídeos mais realistas. No caso dos áudios falsificados, o usuário precisa apenas inserir um material prévio do indivíduo para que o programa sintetize a voz da pessoa, para criar um áudio realista com um discurso inédito. É importante destacar que esses usos são perigosos especialmente porque não são mais uma "caça ao tesouro". Estão abertos ao público

em softwares acessíveis e qualquer pessoa com pequeno conhecimento prático tecnológico pode usar para o que quiser (Caldera, 2019).

Dois elementos são a chave para o sucesso que as *deepfakes* têm representado e para a sua popularidade em usos políticos maliciosos. O primeiro elemento é a credibilidade. O impacto das *deepfakes* é significativo porque, mesmo que a confiança em materiais fotográficos tenha diminuído significativamente nas últimas décadas graças às tecnologias de edição de imagens, a evidência fotográfica ainda possui grande valor social e recebe muita credibilidade. O sistema visual do cérebro pode ser alvo de percepções errôneas, mas as pessoas tendem a confiar mais em vozes e em vídeos do que numa simples imagem estática. Ainda que hoje seja mais difícil confiar "em seus próprios olhos", essa ainda é uma fonte de crença muito grande para as pessoas. De outra maneira, as falsificações de baixa qualidade deram lugar a mídias sintéticas complexas e sofisticadas (Kietzmann *et al.*, 2020).

Há poucos anos era facilmente perceptível que o conteúdo de determinado vídeo havia sido modificado, porque o cérebro humano estranha traços que considera "bizarros", o que era comum nessas execuções. Porém, poucos anos depois, com a célere evolução e popularização dessa inteligência artificial generativa que faz uso do aprendizado profundo, do inglês *deep learning*, via GANs, está mais difícil distinguir vídeos autênticos e conteúdos audiovisuais sintéticos. Esse é o primeiro elemento relevante. O segundo certamente é a acessibilidade. As ferramentas não são mais de domínio exclusivo de técnicos especializados. Indivíduos sem qualquer qualificação podem manipular mídias facilmente, por meio de seus smartphones e as aplicações usadas para tal têm devolvido resultados impressionantes. A junção da credibilidade e da acessibilidade são uma receita perigosa para o sucesso das *deepfakes*. É preciso encarar a magnitude que esse fenômeno tem e que pode ter, porque cada vez mais mídias sintéticas serão produzidas, com mais riscos e danos do que benefícios, já que os usos escusos são numerosos (Kietzmann *et al.*, 2020).

Em usos eleitorais, estima-se que as *deepfakes* podem ser utilizadas como armas. O seu uso malicioso é uma estratégia de desinformação, Shick (2020) adota o uso consciente do dano como qualificador para determinar a *deepfake*, em contraposição a outros meios de comunicação sintéticos que servem a outros objetivos. Na perspectiva de Chesney e Citron (2019), há fundamentos tecnológicos que explicam o problema do uso de *deepfakes*. A primeira delas é o nível de realismo que é possível alcançar, com as técnicas atuais, na produção desse conteúdo. A tecnologia é vanguardista e os resultados costumam ser aparentemente verossímeis.

Porém, eles sempre ocorrem sem o consentimento das pessoas que integram o conteúdo criado sinteticamente.

Esse modo de produção ocorre em um momento perigoso, no qual a atenção dos cidadãos está voltada para as mais diversas fontes de informação, não apenas às organizações de comunicação social de confiança. Por isso, conteúdos falsificados podem - e chegam - a milhares de pessoas em poucos minutos, essa capacidade de audiência perturba o campo político, cabendo à pessoa atingida justificar-se quanto à falsidade da informação. Até esse momento, muitas pessoas já viram, comentaram e compartilharam o conteúdo, alçando-o a um patamar de verdade antes mesmo da conferência das informações. O dano, portanto, já foi causado e pode ser possível revertê-lo apenas parcialmente (Chesney; Citron, 2019).

Mesmo as *cheapfakes*, termo que designa as modificações sintéticas empobrecidas, que podem ser facilmente percebidas como editadas, são compartilhadas e interpretadas como verdadeiras, apesar da manipulação audiovisual grotesca. Uma parte dos usuários das redes sociais acredita nessas montagens, ao ponto de que sites de verificação de conteúdo precisam esclarecer cotidianamente que se trata de montagens. Um aspecto importante dessa questão, que torna as *deepfakes* preocupantes nas eleições, é a crescente qualidade dos vídeos e imagens criados artificialmente. Essa verossimilhança construiu um cenário no qual a mídia sintética é crível; como consequência, o seu impacto é ampliado (Mulholland; Oliveira, 2021).

Três subfenômenos alimentam a dinâmica que pode tornar as *deepfakes* virais e danosas nas eleições. São eles a cascata de informação, a atração das pessoas por informações bombásticas e negativas e os filtros bolha, que ajudam a explicar o porquê das falsificações profundas se tornarem tão perigosas. No que se refere às cascatas de informação, ela é um mecanismo confuso: as pessoas deixam de prestar atenção às próprias informações e confiam no que os outros dizem, pois presumem que estão corretos, mesmo que isso contradiga as suas informações prévias sobre determinado tema. Em certa altura, além de abandonar as próprias informações e confiar em terceiros, elas transmitem o que os outros pensam. Essa rede é uma narrativa que atinge credibilidade à medida que é disseminada. A ideia popularmente conhecida como "disse me disse" se tornou, efetivamente, um fenômeno da informação, em que a cascata se fortalece. As mídias sociais são arquitetadas para estimular essa cascata, uma vez que espalham infinitamente conteúdos de terceiros que são considerados como confiáveis. Se em determinado contexto a cascata de informação pode servir à democracia, quando amplifica vozes silenciadas sistematicamente, em outros ela pode significar a disseminação de desinformação (Chesney; Citron, 2019).

CONSIDERAÇÕES FINAIS

Cumpre destacar ao final que o fenômeno da desordem informacional possui diversas facetas e vastos efeitos em contexto eleitoral. A disseminação de *deepfakes* pode comprometer a integridade do pleito e minar a confiança que os cidadãos possuem nas instituições democráticas. A popularização da manipulação sintética de conteúdo audiovisual, com o intuito de difundir desinformação, produz graves efeitos ao influenciar negativamente, por meio de materiais falsos, a opinião pública. Informações adequadas, verdadeiras e críveis, de interesse público, devem ser valorizadas neste contexto. O que se entende nessa questão é que, embora a desinformação seja um fenômeno histórico-político antigo, com diversos exemplos ao longo da história, o ineditismo deste século tem relação com a dinamicidade do fenômeno, com a criação real de um caos informacional, especialmente pela capacidade de disseminação rápida de campanhas de desinformação.

Do ponto de vista sociopolítico, as *deepfakes* podem alimentar a polarização e distorcer a realidade para criar uma atmosfera política de incerteza. A evolução abrupta tecnológica sem regulação não pode representar um ataque aos fundamentos da democracia, por isso a manutenção da estabilidade política é fundamental nesse tema. O potencial que campanhas de manipulação tem para abalar a opinião pública é significativo. Portanto, é fundamental a adoção de medidas enérgicas pelo Tribunal Superior Eleitoral, que envolva todos os atores relevantes, as plataformas, os políticos, a justiça, o Estado e a população. Medidas de prevenção e mitigação de danos devem ser dinâmicas e proativas, como é o próprio fenômeno, que carece de um fortalecimento legislativo contra a manipulação via inteligência artificial, o desenvolvimento de tecnologias de detecção de *deepfakes*, de maneira que seja possível acompanhar o estado da arte da tecnologia, além da conscientização pública em torno da credibilidade das informações. Uma abordagem ampla perpassa pelo reconhecimento do fortalecimento dos meios de comunicação tradicionais, que devem contribuir para a checagem de informações e fortalecimento de uma mídia confiável.

As *deepfakes* são evoluções das *fake news* e trazem dinamismo ao fenômeno da desinformação. Elas podem ser utilizadas para distorcer a realidade e criar narrativas que interessem apenas àqueles que financiam, produzem e distribuem esse conteúdo. O risco mais relevante é afetar a capacidade dos cidadãos de tomar decisões políticas e participar plenamente do processo democrático. A subversão da verdade é um dado que torna as *deepfakes* uma ameaça imperativa. Somente por meio de um esforço conjunto e coordenado se pode encarar esse desafio e preservar a lisura do processo eleitoral e da democracia como um todo. Desafios

como esse, e ainda maiores, podem se tornar comuns com a sofisticação das tecnologias, das técnicas e dos usos de inteligência artificial generativa. Por isso, deve-se reconhecer a magnitude do fenômeno e criar estratégias tanto de proteção quanto de correção e mitigação dos danos.

O estabelecimento de diretrizes claras e o avanço legislativo – como o projeto de lei que visa criar novas regras para a moderação de conteúdos nas redes sociais, criar garantias e responsabilidades aos usuários e também versa sobre a remoção de publicações ou a suspensão de contas, a PL 592/2023 de autoria do senador Jorge Seif (PL-SC) assim como o projeto de lei 2630/2020, que trata do combate às fake news, sob a relatoria do deputado federal Orlando Silva (PCdoB-SP) e por fim o debate em plenário no senado federal do Marco Civil da Inteligência Artificial –, são fundamentais para fortalecer os mecanismos de defesa contra a manipulação sintética de materiais digitais. As leis podem tanto proibir explicitamente os usos de *deepfakes* integralmente quanto penalizar rigorosamente os usos para fins maliciosas. Mas urge a necessidade de discussões legislativas que visem equilibrar o impacto da desordem informacional nesse século, para que não afetem negativamente as eleições brasileiras e internacionais no ano de 2024.

REFERÊNCIAS BIBLIOGRÁFICAS

BENEVIDES, Bruno. Diretor Jordan Peele faz vídeo com Obama para denunciar notícias falsas. **Folha de São Paulo**, 18 abr. 2018. Disponível em: https://www1.folha.uol.com.br/mundo/2018/04/diretor-jordan-peele-faz-video-com-obama-para-denunciar-noticias-falsas.shtml. Acesso em: 2 fev. 2024.

CALDERA, Elizabeth. "Reject the evidence of your eyes and ears": deepfakes and the law of virtual replicants. **Seton Hall Law Review**, vol. 50, article 5, 2019.

CHESNEY, Robert; CITRON, Danielle Keats. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. **SSRN Electronic Journal**, 2019. Disponível em: https://doi.org/10.2139/ssrn.3213954. Acesso em: 29 fev. 2024.

COLOMINA, Carmen. Words as weapons: from disinformation to the global battle for the narrative. In: BARGUÉS, Pol.; BOUREKBA, Moussa; COLOMINA, Carmen (eds.). **Hybrid threats, vulnerable order**. Cidob Report, 08, Barcelona, 2022, p. 15-22.

DIAZ RUIZ, Carlos; NILSSON, Tomas. EXPRESS: Disinformation and Echo Chambers: How Disinformation Circulates in Social Media Through Identity-Driven Controversies. **Journal of Public Policy & Marketing**, p. 074391562211038, 16 maio 2022. Disponível em: https://doi.org/10.1177/07439156221103852. Acesso em: 20 fev. 2024.

HASEN, Richard. **Cheap speech**: how disinformation poisons our politics - and how to cure it. New Haven: Yale University Press, 2022.

HINDMAN, Matthew. **Myth of Digital Democracy**. New Jersey: Princeton University Press, 2008.

JANSEN, Sue Curry; MARTIN, Brian. The Streisand Effect and Censorship Backfire. **International Journal of Communication**, v. 9, p. 16, fev. 2015. Disponível em: https://ijoc.org/index.php/ijoc/article/view/2498 Acesso em: 28 fev 2024.

KIETZMANN, Jan et al. Deepfakes: Trick or treat? **Business Horizons**, v. 63, n. 2, p. 135-146, mar. 2020. Disponível em: https://doi.org/10.1016/j.bushor.2019.11.006. Acesso em: 21 fev. 2024.

LESSIG, Lawrence. **Code and other laws of cyberspace**. New York: Basic Books, 1999.

MAGRANI, Eduardo. **Democracia conectada**: a internet como ferramenta de engajamento político democrático. Curitiba: Juruá, 2014.

MULHOLLAND, Caitlin; OLIVEIRA, Samuel Rodrigues. Uma Nova Cara Para a Política? Considerações sobre Deepfakes e Democracia. **Direito Público**, v. 18, n. 99, 28 out. 2021. Disponível em: https://doi.org/10.11117/rdp.v18i99.5773. Acesso em: 29 fev. 2024.

NONÔ, Manuella da Silva. Guerra eleitoral, fake news e a tentativa de regulamentar o uso da internet. **Cadernos Aslegis**, Brasília, n. 63, p. 163-183, 2º semestre 2022.

O'CONNOR, Jennifer Flannery; MOXLEY, Emily. Our approach to responsible AI innovation. **Youtube Blog**, 14 nov. 2023. Disponível em: https://blog.youtube/inside-youtube/our-approach-to-responsible-ai-innovation/. Acesso em: 28 fev. 2024.

ORTELLADO, Pablo; RIBEIRO, Marcio Moretto. Polarização e desinformação online no Brasil. **Análise**, 44,3-8, 2018. Disponível em: https://library.fes.de/pdf-files/bueros/brasilien/14629.pdf. Acesso em: 29 fev. 2024.

PRADO, Magaly Parreira do. Deepfake de áudio: manipulação simula voz real para retratar alguém dizendo algo que não disse. **TECCOGS - Revista Digital de Tecnologias Cognitivas**, nº 23, jan./jun. 2021, p. 45-68.

SARLET, Gabrielle Bezerra Sales. A Inteligência artificial no contexto atual: uma análise à luz das neurociências voltada para uma proposta de emolduramento ético e jurídico. **Direito Público**, v. 18, n. 100, 27 jan. 2022. Disponível em: https://doi.org/10.11117/rdp.v18i100.5214. Acesso em: 29 fev. 2024.

SHICK, Nina. **Deepfakes**: the coming infocalypse. New York: Twelve, 2020.

SIRLIN, Nathaniel et al. Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. **HKS Misinformation Review**, 6 dez. 2021. Disponível em: https://misinforeview.hks.harvard.edu/article/digital-literacy-is-associated-with-more-discerning-accuracy-judgments-but-not-sharing-intentions/. Acesso em: 25 fev. 2024.

TANDOC, Edson C.; LIM, Zheng Wei; LING, Richard. Defining “Fake News”. **Digital Journalism**, v. 6, n. 2, p. 137-153, 30 ago. 2017. Disponível em: https://doi.org/10.1080/21670811.2017.1360143. Acesso em: 29 fev. 2024.

VAN RAEMDONCK, Nathalie; MEYER, Trisha. Why disinformation is here to stay. A socio-technical analysis of disinformation as a hybrid treat. In: LONARDO, L. (ed.). **Adressing Hybrid Threats: European Law and Policies**, 2024, p. 57-83.

VIVAS, Fernanda. Eleições 2024: campanha terá regras para Inteligência Artificial e não poderá usar deepfakes; veja o que foi aprovado pelo TSE. **G1**, 27 fev. 2024. Disponível em: https://g1.globo.com/politica/noticia/2024/02/27/eleicoes-2024-tse-comeca-julgamento-de-regras-para-a-votacao-de-outubro.ghtml. Acesso em: 28 fev. 2024.

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

WARDLE, Claire. Fake news. It's complicated. **First Draw News**, 16 fev. 2017. Disponível em: https://firstdraftnews.org/articles/fake-news-complicated/. Acesso em: 29 fev. 2024.

WARDLE, Claire; DERAKHSHAN, Hosseion. **Information disorder: toward an interdisciplinary framework for research and policymaking**. Strasbourg: Council of Europe, 2017.