

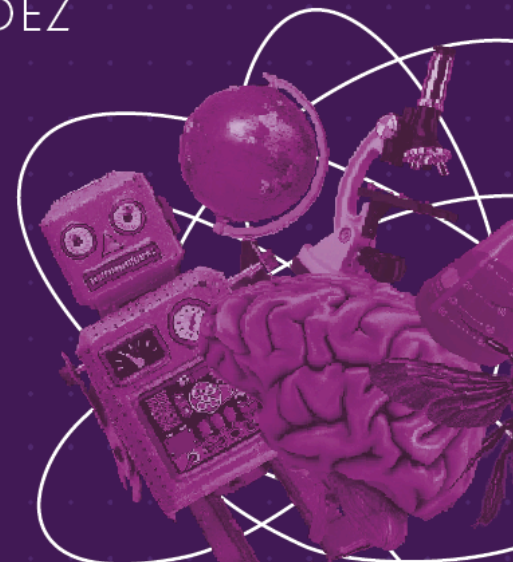
"AS CÂMERAS NUNCA MENTEM": UMA ANÁLISE DAS POLÍTICAS PÚBLICAS QUE PODEM SER UTILIZADAS PARA REGULAMENTAR O USO DAS DEEP FAKES DENTRO DO DIREITO ELEITORAL

Professor orientador: Alessandro Rodrigues da Costa

Alunas: Luiza Tanus Paixão Lopes e Debora Maria
Miura Nakayama

PROGRAMA DE
INICIAÇÃO CIENTÍFICA
PIC/CEUB

RELATÓRIOS DE PESQUISA
VOLUME 10 Nº 1- JAN/DEZ
2024



**CENTRO UNIVERSITÁRIO DE BRASÍLIA - CEUB
PROGRAMA DE INICIAÇÃO CIENTÍFICA**

**LUIZA TANUS PAIXÃO LOPES
DEBORA MARIA MIURA NAKAYAMA**

**"AS CÂMERAS NUNCA MENTEM": UMA ANÁLISE DAS POLÍTICAS
PÚBLICAS QUE PODEM SER UTILIZADAS PARA REGULAMENTAR O USO
DAS DEEP FAKES DENTRO DO DIREITO ELEITORAL**

Relatório final de pesquisa de Iniciação Científica apresentado à Assessoria de Pesquisa e Extensão.

Orientação: Alessandro Rodrigues da Costa

**BRASÍLIA
2025**

DEDICATÓRIA

“Amor, que a ninguém amado perdoa amar, levou-nos juntos ao labor de pensar.” (Verso livremente inspirado no Canto V da Divina Comédia, Dante Alighieri)

Dedicamos este trabalho aos nossos pais, cuja presença firme e silenciosa sustentou cada passo de nossa jornada acadêmica. A eles, que foram chão quando os caminhos tremiam, nossa mais profunda gratidão.

Ao nosso professor, ao nosso orientador que, como Virgílio à beira do Inferno, nos guiou com razão, paciência e coragem por entre as sombras da dúvida e da pesquisa.

E entre linhas, ideias e longas madrugadas, seguimos, como Paolo e Francesca, movidos por algo que também nos uniu mas, ao invés da paixão trágica, a busca comum pelo saber.

Por isso também dedicamos, sobretudo, às mentes curiosas que buscam compreender cada vez mais a relação entre tecnologia e direito, acreditando que o conhecimento é a ferramenta mais poderosa para proteger a democracia.

AGRADECIMENTOS

Agradecemos, primeiramente, aos nossos pais, que nos ouviram durante intermináveis horas sobre o andamento do trabalho e as barreiras que enfrentamos e superamos.

Também agradecemos, de forma especial, ao professor Alessandro Costa, nosso orientador, pela paciência, disponibilidade e pela orientação que guiou este projeto.

Por último, mas não menos importante, agradecemos aos nossos amigos, que nos deram *insights* e nos auxiliaram em tudo o que precisávamos.

RESUMO

A inteligência artificial (IA) está dominando o mundo tanto acadêmico quanto social. No Brasil, mais especificamente, nas eleições de 2024, a quantidade de notícias, fotos e vídeos fraudulentos foi considerável. Por esse motivo, a questão central deste trabalho é investigar os impactos das *fake news* e das *deep fakes*, no contexto do direito eleitoral. Contudo, rapidamente foi-se observado que esse é um problema mundial e nesse mesmo ano, metade da população mundial foi às urnas. Dessa forma, realizou-se uma pesquisa com abordagem qualitativa, visando revisões bibliográficas de artigos científicos, livros, legislações estrangeiras, especificamente a italiana e a da Nova Zelândia, legislação nacional, projetos de leis e reportagens que exemplificam os casos abordados. Por conseguinte, através dos estudos, foi possível concluir que as redes sociais são os maiores vetores de disseminação de informações, fotos e vídeos falsos, e mesmo com suas formas de prevenir postagens deste teor, essas publicações ainda chegam à população e tem influenciado significativamente a opinião pública, sobre um candidato ou um partido específico dentro do algoritmo recomendado, e portanto, atinge o debate democrático, que acaba contaminado de disparidades entre os concorrentes. Por fim, identificou-se que, apesar das lacunas no arcabouço normativo e da dificuldade em acompanhar a evolução das IAs, é possível fortalecer o enfrentamento dessas práticas por meio de: revisão e atualização das leis e projetos de lei existentes; fortalecimento de políticas públicas de educação midiática; implementação das medidas propostas, como agência reguladora, uso de inteligência artificial no monitoramento e combate à desinformação, parcerias com agências de checagem e aprimoramento de ferramentas colaborativas.

Palavras-chave: *deep fakes*; *fake news*; direito eleitoral; desinformação; democracia digital.

SUMÁRIO

1. Introdução	7
1.1 Contextualização da pesquisa	7
1.2 Objetivos	8
1.2.1 Objetivos gerais	9
1.2.2 Objetivos específicos	9
2 Fundamentação teórica	10
2.1 Definição	10
2.1.1 Definição de <i>fake news</i>	10
2.1.2 Diferenciação entre <i>fake news</i> e desinformação	12
2.1.3 Definição de <i>deep fakes</i>	13
2.2 Histórico e evolução	14
2.2.1 Evolução das <i>Fake News</i>	14
2.2.2 Evolução das <i>Deep Fakes</i>	16
2.2.3 Redes Sociais: Impacto nas <i>Fake News</i> e <i>Deep fakes</i>	17
2.3 Legislação e medidas de combate	18
2.3.1 Panorama Jurídico e Normativo	18
2.3.2 Campanhas Governamentais e Recursos para a População	18
2.3.3 Desafios Éticos e Legais no Contexto de IA	19
2.3.4 Desafios e Perspectivas	19
2.4 Comparações de legislação e medidas de combate com a Itália e Nova Zelândia	20
2.4.1 O Brasil	20
2.4.2 A Itália	21
2.4.3 A Nova Zelândia	22
2.4.4 Comparações	23
3 Método	24
4 Resultados e discussão	26
4.1 Impacto na sociedade	26
4.1.1 Efeitos na Política	26

4.1.2 Efeitos na Economia	27
4.1.3 Efeitos na Saúde Mental	28
4.2 Mecanismos de disseminação	31
4.2.1 Criação de <i>Fake News</i> e <i>Deep fakes</i>	31
4.2.2 Redes Sociais como Vetores de Disseminação	32
4.2.3 Aplicativos de Mensagens Privadas	35
4.2.4 Outras Plataformas Digitais	36
4.2.5 O Papel dos Algoritmos e da IA	37
4.3 Tecnologias de detecção	37
4.3.1 Detecção de <i>Fake News</i>	38
4.3.2 Detecção de <i>Deep Fakes</i>	39
4.4 Educação e conscientização	41
4.4.1 Incentivo governamental para influenciadores no combate às <i>fake news</i>	43
4.4.2 Unificação das emissoras para alertas sobre <i>fake news</i> durante eleições	44
4.4.3 Fortalecimento das Notas da Comunidade nas redes sociais	45
4.4.4 Análise de campanhas de conscientização bem-sucedidas	48
4.4.5 Educação básica como ferramenta de conscientização familiar	49
4.4.6 Campanha de conscientização sobre suicídio e desinformação	52
Considerações finais	57
Referências	60

1. Introdução

1.1 Contextualização da pesquisa

A inteligência artificial (IA), outrora vista como algo tão distante, quase saída de um filme como “Blade Runner 2049”, há algum tempo permeia nossa realidade. A IA evoluiu de tal maneira que possibilitou a criação das *deep fakes*, uma forma avançada desta tecnologia, que consiste em substituir faces e sons de certos indivíduos por outros em vídeos. Este tipo de tecnologia constitui um desafio no âmbito político, porque as *deep fakes* podem ser utilizadas para sabotar a integridade das eleições, como por exemplo, um vídeo falso mostrando um candidato específico recebendo propina (Chesney; Citron, 2018).

Para demonstrar que esses acontecimentos não estão apenas do outro lado do mundo, destaca-se notar um exemplo brasileiro que ocorreu durante as eleições presidenciais de 2022, quando um vídeo *deep fake*, que ganhou certa notoriedade, mostrava a apresentadora Renata Vasconcellos, do Jornal Nacional, supostamente anunciando uma pesquisa eleitoral com resultados manipulados, favorecendo o então candidato Jair Bolsonaro. A peça de desinformação, que alterava uma edição legítima do jornal exibida em 15 de agosto daquele ano, utilizava técnicas de IA, consistentes com as *deep fakes*, para simular a voz da jornalista e inserir dados falsos, conforme apurado em verificação publicada pelo UOL (Mutchnik, 2022).

Dessa forma, fica evidente como esse tipo de tecnologia já está afetando a democracia brasileira e a de diversos outros países. Na reportagem mencionada acima, ela também aborda como vários países também estão enfrentando os mesmos problemas. Contudo, já se passaram outros dois anos, a tecnologia evoluiu e pessoas comuns conseguiram aprender a dominar a IA e manipulá-la.

É comum acreditar que “câmeras nunca mentem” (Barnes; Barraclough, 2019), pois aceitar como verdadeiro o que é visto e ouvido em fotos e vídeos é costumeiro e não há motivo para questionar sua veracidade, contudo isso resulta uma agravada desinformação e distorção da percepção pública (Vaccari; Chadwick, 2020), tendo como consequência, impactos negativos profundos na democracia.

Em 2024, quase metade da população mundial foi às urnas (Brasil, 2024), e uma discussão surgiu como preocupação central entre os governos: a inteligência artificial.

Embora essa tecnologia ofereça inúmeras possibilidades de uso tanto para candidatos quanto para eleitores, uma de suas variantes tem gerado sérios desafios: às *deep fakes*. Com essa ferramenta, qualquer pessoa pode criar vídeos ou fotos falsos com um alto nível de realismo, levantando questões sobre como proteger as eleições para que as desinformações não causem a alteração de seus resultados.

Atualmente, este tema é amplamente discutido em áreas como Direito Eleitoral, Comunicação e Ciência da Computação. Pesquisadores como Chesney e Citron (2018), Vaccari e Chadwick (2020) e Robl Filho, Marrafon e Medón (2022) têm se dedicado a estudar tanto os aspectos técnicos dessas manipulações quanto seus impactos sociais, jurídicos e políticos.

Diante desse cenário, fica perceptível a relevância desta pesquisa em compreender como essas tecnologias afetam a democracia e de avaliar as medidas capazes de mitigar seus efeitos. Para isso, o trabalho realiza um estudo comparado entre as estratégias adotadas no Brasil, Itália e Nova Zelândia, investigando políticas públicas, regulamentações e ações de enfrentamento já em curso.

Portanto, em um mundo cada vez mais dominado pelas tecnologias, torna-se imperativo uma reflexão sobre os desafios jurídicos da regulamentação desse tipo de tecnologia para a preservação da integridade democrática brasileira.

1.2 Objetivos

1.2.1 Objetivos gerais

O objetivo Geral a ser alcançado pela pesquisa é analisar os impactos das *deep fakes* e *fake news* no contexto do Direito Eleitoral, com foco nas eleições de 2024, e avaliar as políticas públicas, legislações e estratégias de combate utilizadas no Brasil, Itália e Nova Zelândia. Almeja-se trazer propostas que possam auxiliar na redução do impacto causado pelas notícias, fotos e vídeos falsos.

1.2.2 Objetivos específicos

Com os objetivos Específicos, pretende-se:

1. Fazer uma revisão bibliográfica acerca do, Direito Eleitoral, o Marco Civil da *Internet* e projetos de lei relevantes para o caso em questão;

2. Contextualizar o surgimento e a evolução das *deep fakes* e *fake news* como fenômenos de desinformação digital;
3. Identificar e examinar casos relevantes de manipulação digital no cenário político-eleitoral brasileiro.
4. Comparar o conjunto normativo brasileiro com as legislações e iniciativas de combate adotadas na Itália e na Nova Zelândia;
5. Discutir os desafios trazidos pelas *deep fakes* e *fake news* e apresentar propostas de aprimoramento de políticas públicas e de mecanismos tecnológicos para prevenção e detecção de manipulações digitais.

2. Fundamentação teórica

2.1 Definição

Uma frase notória de Santo Agostinho é “*intellige ut credas, crede ut intelligas*” (“eu creio para compreender e compreendo para crer melhor” tradução nossa), destacando a relação entre fé e razão. Embora dita em um contexto religioso, essa frase pode ser reinterpretada à luz dos desafios contemporâneos relacionados à compreensão da verdade e da informação.

Tendo isso em mente, vale destacar que com avanços tecnológicos, surgiram diversos mecanismos e com eles, a desinformação. Dessa forma, é de fundamental importância conceituar os principais perpetuadores das manipulação de informações: as *fake news* e as *deep fakes*.

2.1.1 Definição de *fake news*

O direito brasileiro, com sua base em *civil law*, é extremamente rígido em relação ao que está escrito na Constituição, nos códigos e nas legislações, com o objetivo de evitar brechas. Por isso, é fundamental a conceituação precisa dos termos jurídicos, garantindo que haja apenas uma interpretação, evitando, assim, possíveis dúvidas ou confusões.

Dessa forma, o que melhor para descobrir o conceito de uma palavra do que em um dicionário?

O dicionário de Cambridge (2023) define *fake news* como “*false stories that appear to be news, spread on the internet or using other media, usually created to influence political views or as a joke.*”¹

Além disso, os autores Abreu e Adeodato (2020, p. 2), trouxeram um interessante ponto de vista:

Nesse sentido, pode-se identificar o caráter ambivalente da expressão “*fake news*” por ser empregado nestas duas direções: ao mesmo tempo em que se refere à desinformação, por meio de notícias e conteúdos cuja falsidade pode ser objetivamente constatada, ela também pode ser utilizada por sujeitos detentores do poder como recurso para desacreditar informações que, apesar de verídicas e sustentadas em evidências e argumentos, não sirvam aos seus interesses.

¹ Histórias falsas que parecem ser notícias, divulgadas na Internet ou usando outras mídias, geralmente criadas para influenciar opiniões políticas ou como piada. (tradução nossa).

As notícias falsas são utilizadas por diversos motivos, mas o principal é para atrair a atenção de pessoas seja qual for o motivo. O problema maior dessa situação é que, com a democratização do acesso a *internet* e a facilidade de criar um perfil em redes sociais, o volume de informações que são disseminadas aumentou bastante, podendo elas serem verdadeiras ou não. Tudo isso foi consumido vorazmente pela população, que além de ser atingida com diversas informações ao mesmo tempo, ainda recebiam elas com base em algoritmos, que são mecanismos que realizam uma triagem de assuntos e apenas mostram notícias com base em suas preferências pessoais e opiniões políticas. Tendo isso em vista, a justiça eleitoral determinou que iria combater os “fatos sabidamente inverídicos” (Abreu; Adeoato, 2020).

O resultado desse processo é ter uma população na qual existem *fake news* criadas especificamente para cada indivíduo e com isso, criando uma bolha de informação na qual a pessoa apenas receberá o que a IA acreditar que interesse-o, por isso diferenciar o verdadeiro do falso se tornou extremamente difícil. Dessa forma, Carvalho e Kanffer (2018, p. 1) afirmam que:

Com efeito, as fake news correspondem a uma espécie de “imprensa marrom” (ou yellow journalism), deliberadamente veiculando conteúdos falsos, sempre com a intenção de obter algum tipo de vantagem, seja financeira (mediante receitas oriundas de anúncios), política ou eleitoral.

As *fake news* tornaram-se tema de preocupação pública pelo seu efeito no campo político, uma vez que têm o poder de influenciar resultados eleitorais nas democracias do mundo ocidental (McNair, 2017).

Ter pensamento crítico é de suma importância para conseguir distinguir as notícias verdadeiras das falsas, contudo de acordo com o Indicador do Alfabetismo Funcional (Inaf, 2025), três em cada dez pessoas são considerados analfabetos funcionais, o que dificulta em manter a nossa democracia viva, pois eles não escolherão o candidato com as melhores propostas, mas, sim, aqueles que eles conseguirem se identificar com.

As *fake news* têm o poder de auxiliar na mudança de resultados eleitorais, porque pessoas sem pensamento crítico não saberão diferenciar o que é verdadeiro do falso e poderão mudar de opinião e até mesmo votar em um candidato diferente por causa de algo tão fútil quanto uma notícia espalhada pelo *whatsapp*.

Um exemplo desses acontecimentos foi o caso intitulado como “mamadeiras de piroca” de 2018, no qual foi-se publicado um vídeo nas redes sociais mostrando uma mamadeira com bico em formato de pênis e continha a informação de que esse tipo de objeto estava sendo distribuído nas escolas e creches de São Paulo por determinação do ex-prefeito Fernando Haddad, que naquela época era candidato à presidência da república. Diversas pessoas até hoje acreditam que isso aconteceu de fato, contudo foi apenas uma *fake news* que mostrava um brinquedo vendido em *sex shops* (Queiroga, 2021).

2.1.2 Diferenciação entre *fake news* e desinformação

Após definir as *fake news*, é necessário diferenciar as *fake news* de desinformação, pois embora ambos os termos se refiram à disseminação de informações falsas, é essencial compreender suas nuances.

Dessa forma, como apresentado anteriormente, as *fake news* são uma forma de desinformação intencional, pois são produzidas para induzir o público a acreditarem em algo que não seja verdadeiro, por outro lado, a desinformação é um conceito muito mais abrangente, já que inclui qualquer disseminação de conteúdo incorreto ou enganoso, porque pode ser por erro ou por má-fé (Abreu; Adeoato, 2020).

Um exemplo marcante de *fake news* nas eleições brasileiras ocorreu em 2018, quando circularam nas redes sociais mensagens falsas alegando que as urnas eletrônicas estavam programadas para transferir automaticamente os votos de um candidato para outro. Um caso específico afirmava que, ao votar no número de Jair Bolsonaro, o voto era registrado para Fernando Haddad. Essa notícia, completamente infundada, foi amplamente desmentida pelas autoridades eleitorais (Santos, 2018).

Ao contrário das *fake news*, a desinformação pode resultar de uma má interpretação ou da circulação de informações fora de contexto, sem que haja a intenção direta de manipular a quem se dirigiu a mensagem. Nesse sentido, a desinformação pode ser acidental, ocorrendo, por exemplo, quando uma pessoa ou entidade compartilha um dado equivocado sem ter consciência de sua falsidade, por isso é necessária uma análise mais cuidadosa, visto que sua intencionalidade é variável e nem sempre estará associada a condutas ilícitas.

Portanto, um exemplo no qual fica evidente que gerou desinformação foi durante as eleições brasileiras de 2022, na qual houve uma intensa circulação de vídeos antigos de entrevistas com o ex-presidente Luiz Inácio Lula da Silva. Nessas gravações, Lula mencionava revisões de leis e ações realizadas durante seus mandatos anteriores. No entanto, esses vídeos foram editados e compartilhados fora de contexto, criando a impressão de que ele estava propondo essas ideias como parte de seu plano de governo atual, durante a campanha eleitoral. Um exemplo disso foi a circulação de trechos de discursos em que Lula comentava sobre reformas econômicas e sociais já realizadas, como se fossem promessas futuras. Essa manipulação de informações gerou confusão e distorção sobre as verdadeiras propostas do candidato (Meireles, 2024).

In summa, enquanto as *fake news* envolvem a criação proposital de informações falsas para enganar indivíduos e utilizam de bolhas de conteúdos; a desinformação é um campo mais amplo de conteúdos incorretos, podendo ser compartilhada com ou sem intenção dolosa.

2.1.3 Definição de *deep fakes*

A inteligência artificial atingiu um patamar tão grandioso que a partir dela surgiram as *deep fakes*. Muitos ainda não têm um conceito definido para *deep fakes*, mas nosso entendimento a partir de diversos estudos que fizemos, definimos essa tecnologia como uma técnica avançada de manipulação digital aplicada a vídeos, áudios, textos ou imagens, comumente utilizada para produzir conteúdos falsificados que aparentam ser autênticos, alterando a realidade com o objetivo de enganar ou influenciar a percepção das pessoas.

De acordo com Chesney e Citron (2018):

Fueled by artificial intelligence, digital impersonation is on the rise. Machine-learning algorithms (often neural networks) combined with facial-mapping software enable the cheap and easy fabrication of content that hijacks one's identity—voice, face, body. Deepfake technology inserts individuals' faces into videos without their permission. The result is "believable videos of people doing and saying things they never did."²

² Impulsionada pela inteligência artificial, a falsificação de identidade digital está em alta. Algoritmos de aprendizado de máquina (geralmente redes neurais) combinados com software de mapeamento facial permitem a fabricação barata e fácil de conteúdo que sequestra a identidade de uma pessoa - voz, rosto, corpo. A tecnologia deep fake insere os rostos das pessoas em vídeos sem a permissão delas. O resultado são "vídeos críveis de pessoas fazendo e dizendo coisas que nunca fizeram". (tradução nossa).

Existem outros conceitos de *deep fake*, como: reconstruções digitais da imagem de uma pessoa, que passa a ser retratada de forma falsa, normalmente com o intuito de provocar desinformação, embora também esteja sendo amplamente utilizada para o humor e a sátira (Robl Filho; Marrafon; Medón, 2022).

A criatividade do povo brasileiro é inegável, e foi justamente essa característica que contribuiu para a rápida apropriação do uso de mecanismos de inteligência artificial como ferramenta para a disseminação de *fake news* em todo o país. Esse cenário é preocupante, sobretudo diante da importância do desenvolvimento do pensamento crítico, habilidade fundamental para distinguir informações verdadeiras das falsas. No entanto, considerando os dados do Indicador de Alfabetismo Funcional (Inaf, 2025), que em 2025 apontou que 29% da população brasileira era composta por analfabetos funcionais, é preciso reconhecer que muitos ainda não possuem as competências necessárias para essa análise crítica o que amplia significativamente os efeitos nocivos da desinformação.

Exemplificando as *deep fakes*, no contexto eleitoral brasileiro, é imprescindível não comentar sobre o que ocorreu em 2025. Nesse período, circulou um vídeo falso atribuído ao então presidente Lula, no qual ele supostamente admitia ter roubado recursos públicos. O G1 confirmou que se tratava de uma montagem, expondo a manipulação digital. Esse caso destacou a crescente sofisticação das *deep fakes* no Brasil, além da dificuldade que muitos eleitores enfrentam para distinguir entre vídeos legítimos e falsificações, agravando os riscos da desinformação no ambiente eleitoral (Fofano, 2025).

2.2 Histórico e evolução

2.2.1 Evolução das *Fake News*

O termo *fake news* refere-se a informações intencionalmente falsas que são apresentadas como se fossem verdadeiras, com o objetivo de manipular a opinião pública ou influenciar comportamentos. Embora o conceito de desinformação não seja novo, práticas de disseminação de boatos e manipulação de informações são antigas, esse termo ganhou proeminência e uma nova conotação no final do século XIX, sendo popularizado recentemente, a partir das eleições presidenciais nos Estados Unidos, em

2016 (Vieira, 2019), nesse contexto, a expressão foi impulsionada pelo debate político e pela cobertura midiática.

Essas práticas de desinformação desafiam o sistema democrático e a integridade das eleições ao explorar as redes sociais para manipular a opinião pública, por conseguinte, com o crescimento destas, as *fake news* se tornaram mais difíceis de controlar e monitorar, e sua propagação é praticamente instantânea. Isso levanta questões sobre a responsabilidade das plataformas digitais na moderação de conteúdo, bem como o papel do Estado e da legislação na contenção dessa prática.

No Brasil, o termo começou a ganhar destaque no período das eleições presidenciais de 2018 (Vieira, 2019). Foi amplamente utilizado pela mídia e pelos candidatos de diferentes espectros políticos, cada qual acusando o outro de disseminar informações falsas. Essa adoção da expressão *fake news* no cenário brasileiro refletiu uma preocupação crescente com os impactos da desinformação no processo eleitoral e sua capacidade de influenciar o voto e a opinião pública de maneira significativa.

Do ponto de vista jurídico, a evolução das *fake news* coloca desafios importantes para o Direito. De um lado, há uma necessidade urgente de atualização legislativa para lidar com essa nova realidade digital. Por outro, o direito à liberdade de expressão precisa ser cuidadosamente considerado, de modo a evitar censura indevida e preservar o debate democrático. A Legislação Brasileira, por exemplo, traz algumas previsões sobre a propaganda eleitoral, mas ainda não é completamente eficaz para lidar com o impacto massivo da desinformação digital.

Ademais, em pesquisa realizada pelo *Department of Information Management of the Chaoyang University of Technology* foi feita uma análise no âmbito eleitoral das eleições presidenciais dos Estados Unidos em 2016 para entender como seria possível diferenciar as *fake news* das notícias verdadeiras, e utilizando um sistema de inteligência artificial verificou-se que há 4 principais diferenças entre estas:

Firstly, fake news is based on social media and artificial fact, not based on fact. Secondly, fake news comes from a singular source, but real news is from multiple sources which usually cover both sides involved in the news. Thirdly, real news has both positive and negative viewpoints, but fake news has only one-sided information. Fourthly, identified and unidentified domains are websites or URLs that are used to publish or disseminate news or information that proves to be false or inaccurate. Additionally, credible news sources such as official news sites or media outlets can help identify areas associated with false or inaccurate news. In some cases, these news

sources often provide official statements about fake news or information circulating on social media or other sites.³

Entretanto, é especificado que mesmo que seja possível analisar esses dados e encontrar resultados que podem ajudar a população a discernir as informações que recebem, as *fake news* estão em constante evolução para se tornar cada vez mais convincentes, e um exemplo dessa constante evolução é a derivação das *fake news*, que surgiu recentemente e ficou conhecida pelos canais midiáticos como *deep fakes*.

2.2.2 Evolução das *Deep Fakes*

A tecnologia de *deep fake* é ainda mais recente e vem adicionando uma camada de complexidade. Com o uso da inteligência artificial, é possível editar vídeos manipulando o que está sendo declarado ou criar falsas atitudes de figuras públicas, como candidatos políticos. Esses vídeos são tão realistas que os espectadores muitas vezes têm dificuldade em identificar se são reais ou não. Assim, o uso de *deep fakes* pode comprometer ainda mais a confiança do eleitorado e criar um ambiente de insegurança e descrédito em relação ao processo eleitoral.

Em 2017, surgiu na plataforma de rede social, *Reddit*, essa derivação das *fake news*, conhecida como *deep fake* e diferente da manipulação de texto e imagens estáticas, o *deep fake* utiliza inteligência artificial para manipular vídeos, o que demonstra uma grande adversidade pela dificuldade que o público geral tem de discernir se de fato a informação é verdadeira ou não. Assim, essa tecnologia inicialmente foi usada de forma controversa, a qual, um usuário da plataforma publicou vídeos pornográficos nos quais sobrepunha o rosto de celebridades em cenas de atores de filmes adultos, sem o consentimento dos envolvidos (Vieira, 2019).

Esse tipo de manipulação evidencia uma nova e grave ameaça ao direito à privacidade, à honra e à integridade moral das pessoas. Além dos danos psicológicos e

³Em primeiro lugar, as notícias falsas baseiam-se nas redes sociais e em fatos artificiais, não em fatos reais. Em segundo lugar, as notícias falsas provêm de uma única fonte, mas as notícias reais provêm de várias fontes que geralmente cobrem ambos os lados envolvidos na notícia. Em terceiro lugar, as notícias reais têm pontos de vista positivos e negativos, mas as notícias falsas têm apenas informações unilaterais. Em quarto lugar, domínios identificados e não identificados são sites ou URLs usados para publicar ou divulgar notícias ou informações que se revelam falsas ou imprecisas. Além disso, fontes de notícias confiáveis, como sites oficiais de notícias ou meios de comunicação, podem ajudar a identificar áreas associadas a notícias falsas ou imprecisas. Em alguns casos, essas fontes de notícias costumam fornecer declarações oficiais sobre notícias falsas ou informações que circulam nas redes sociais ou em outros sites (tradução nossa).

sociais, os *deep fakes* colocam em risco a credibilidade de figuras públicas e a segurança de informações. No campo do Direito, esse fenômeno apresenta desafios inéditos, exigindo regulamentações específicas e uma compreensão aprofundada dos limites da liberdade digital, do direito à imagem e da proteção da dignidade humana diante das novas tecnologias.

2.2.3 Redes Sociais: Impacto nas *Fake News* e *Deep fakes*

É inegável que a desinformação e a propagação de notícias falsas possuem raízes antigas, com vestígios que remontam a períodos antes de Cristo. No Império Romano, por exemplo, era comum que o imperador realizasse anúncios estratégicos para encorajar figuras como Júlio César a enfraquecer o sistema republicano (Olan, 2024). Naquela época, não existiam mecanismos para verificar ou validar a autenticidade das informações divulgadas. Ademais, questionar a autoridade do imperador era considerado um ato de traição, o que podia levar à prisão. Por essa razão, era altamente improvável que alguém ousasse contestar a veracidade dessas informações.

Com o surgimento das redes sociais no final do século XX, a disseminação de desinformação foi amplificada, levando à popularização do termo *fake news*. O conceito de sociedade pode ser entendido como uma associação de indivíduos que compartilham valores culturais e éticos, operando sob um mesmo regime econômico e político, dentro de um território comum e submetidos a regras de convivência específicas.

E as redes sociais proporcionam um ambiente onde indivíduos podem compartilhar informações, seja com pessoas que fazem parte da mesma sociedade ou com aquelas que possuem valores culturais completamente diferentes. A *internet* possibilitou que as pessoas encontrassem outras que compartilhassem suas ideologias, independentemente de barreiras geográficas, ampliando o alcance da informação e permitindo que ideias circulassem entre sociedades com normas e valores diversos.

Nesse sentido, as redes sociais se tornaram um meio essencial para a propagação das *fake news*. Desse modo, a aceitação da desinformação pela sociedade torna-se um fator determinante para o controle ou intensificação de sua disseminação (Egelhofer, 2019). Em última análise, o impacto das redes sociais sobre as *fake news*

demonstra que a recepção e validação social são cruciais para o combate a esse fenômeno, dado o poder das redes de multiplicar rapidamente conteúdos, independentemente de sua veracidade.

Tendo em vista que as pessoas têm muito mais probabilidade de acreditar em histórias que favorecem seu candidato preferido, especialmente se tiverem redes de mídia social ideologicamente segregadas (Allcott, 2017), com isso percebe-se a importância dos mecanismos de disseminação.

2.3 Legislação e medidas de combate

2.3.1 Panorama Jurídico e Normativo

O Brasil tem adotado medidas para enfrentar a desinformação e mitigar os impactos das *fake news* e *deep fakes* sobre a democracia. A Resolução nº 23.610/2019 do Tribunal Superior Eleitoral (Brasil, 2019) é uma das principais iniciativas, proibindo o uso de *deep fakes* em propagandas eleitorais. A norma prevê multas entre R\$ 5.000,00 e R\$ 25.000,00 ou até o dobro do valor investido na disseminação de conteúdo, caso ultrapasse o limite estipulado.

Outras iniciativas legislativas incluem o Projeto de Lei nº 2630/2020 ("Lei das Fake News"), que busca regular a atuação de plataformas digitais e responsabilizar os envolvidos na disseminação de conteúdos falsos. Apesar desses avanços, há desafios para equilibrar a proteção contra desinformação com a garantia de liberdade de expressão.

2.3.2 Campanhas Governamentais e Recursos para a População

O Governo Federal lançou, em 2023, a campanha "Brasil contra Fake" (Brasil, 2023), com o objetivo de conscientizar a população sobre os perigos das *fake news* e fornecer ferramentas para identificar e evitar a disseminação de desinformação. A iniciativa também incentiva o uso de plataformas confiáveis de checagem de fatos, como "Agência Lupa", "Aos Fatos", "Projeto Comprova" e "Boatos.org".

Essas campanhas são complementadas por esforços de organizações não governamentais e de mídia para capacitar a sociedade em práticas de verificação de informações, contribuindo para o fortalecimento da democracia.

2.3.3 Desafios Éticos e Legais no Contexto de IA

A crescente sofisticação de tecnologias baseadas em inteligência artificial (IA), como *deep fakes* e personagens gerados por IA, adiciona complexidade ao enfrentamento da desinformação. Um ponto relevante é a determinação de responsabilidade quando um conteúdo gerado por IA causa danos, como difamação, violação de privacidade ou disseminação de informações falsas. Segundo estudo publicado na *Nature Machine Intelligence* (Pataranutaporn et al., 2021), a responsabilidade pode recair sobre diferentes agentes, incluindo o programador, a plataforma de disseminação ou até mesmo o usuário que interage com a IA. A falta de identificação clara de autoria nos conteúdos gerados por IA agrava o problema, dificultando a responsabilização legal e ética dos envolvidos.

2.3.4 Desafios e Perspectivas

Embora os esforços para combater a desinformação avancem, a tecnologia continua correndo em um ritmo mais acelerado que qualquer tentativa de controle. Monitorar plataformas de mensagem privadas, como *WhatsApp* e *Telegram*, ainda é um desafio real. E quando somamos isso à falta de uma alfabetização midiática decente, o resultado é um terreno fértil para *fake news* crescerem sem freio.

Por isso, atitudes simples como verificar a fonte, buscar mais de uma perspectiva e pensar duas vezes antes de apertar o botão de compartilhar deveriam ser o mínimo. Mas não dá pra jogar tudo nas costas do usuário. A discussão precisa ir além e encarar de frente o impacto da inteligência artificial na produção e na difusão dessas narrativas falsas. Já passou da hora de pensar em regulamentações que deem conta dessa realidade e, principalmente, de criar mecanismos sérios de responsabilização (Chesney; Citron, 2018).

Dessa forma, combater a desinformação no Brasil exige esforço coletivo: legislação atualizada, campanhas educativas, tecnologia afiada e uma população consciente. Só assim a gente protege o debate público e impede que a democracia seja engolida por um feed manipulado.

2.4 Comparações de legislação e medidas de combate com a Itália e Nova Zelândia

A desinformação, impulsionada por avanços tecnológicos e pela rápida disseminação em plataformas digitais, apresenta desafios significativos para democracias em todo o mundo. Brasil, Itália e Nova Zelândia adotaram estratégias distintas para mitigar os impactos de *fake news* e *deep fakes*, refletindo suas especificidades políticas, culturais e tecnológicas. Este artigo analisa as semelhanças e diferenças nas abordagens desses países, destacando suas legislações, iniciativas educativas e parcerias tecnológicas.

2.4.1 O Brasil

O Brasil tem implementado medidas para enfrentar a desinformação, particularmente em períodos eleitorais. No âmbito jurídico, a Resolução nº 23.610/2019 do Tribunal Superior Eleitoral (Brasil, 2019) proíbe o uso de *deep fakes* em campanhas eleitorais, prevendo penalidades financeiras rigorosas. Além disso, o Projeto de Lei nº 2630/2020, conhecido como "Lei das *Fake News*", busca regulamentar plataformas digitais e responsabilizar os disseminadores de conteúdos falsos.

As iniciativas educativas no país incluem campanhas públicas, como "Brasil Contra *Fake*" (Brasil, 2023), que visam conscientizar a população sobre os riscos da desinformação. No entanto, a alfabetização midiática ainda não é amplamente integrada ao sistema educacional, limitando o alcance das estratégias preventivas. Segundo a pesquisa "TIC Educação" realizada, em 2015, pelo Comitê Gestor da Internet no Brasil (CGI.br), apenas uma pequena parcela das escolas adota programas consistentes para ensinar os alunos a consumir informações de forma crítica (Comitê Gestor da Internet no Brasil, 2016).

No campo tecnológico, o uso massivo de plataformas como *WhatsApp* apresenta desafios significativos. A criptografia de ponta a ponta, que é essencial para proteger a privacidade dos usuários, também dificulta o monitoramento e a contenção de desinformação. Durante as eleições de 2018, por exemplo, houve relatos de disparos massivos de mensagens contendo *fake news* por meio do *WhatsApp* (Whatsapp [...], 2019). A partir disso, a plataforma implementou medidas como a

limitação no encaminhamento de mensagens e a identificação de mensagens encaminhadas.

Embora o Brasil tenha avançado em diversas frentes para enfrentar a desinformação, ainda há um longo caminho a percorrer.

2.4.2 A Itália

A Itália adota uma abordagem robusta, combinando legislação punitiva com estratégias educativas para combater a desinformação. A *Legge sulla Diffamazione*, prevista no artigo 595 do *Codice Penale* (Italia, 1930), estabelece penas que variam de multa a reclusão de até três anos, especialmente quando a ofensa ocorre por meio da imprensa, *internet* ou redes sociais. A legislação tipifica a conduta de ofender a reputação de alguém na ausência da vítima, diante de terceiros, e considera agravantes quando a difamação é realizada com atribuição de fato determinado ou mediante instrumento de divulgação massiva. Tais sanções visam proteger a dignidade, a honra e a imagem social do indivíduo, valores fundamentais no contexto democrático italiano.

No campo educacional, o projeto *Basta Bufale* surgiu em 2017, idealizado pela presidente da Câmara, Laura Boldrini, e pela ministra da Educação, Valeria Fedeli, com o apoio do MIUR, para formar professores e jovens no reconhecimento de *fake news*. A iniciativa incluiu um *decálogo anti-bufale*, distribuído a mais de 4,2 milhões de estudantes do ensino médio, que oferece ferramentas práticas para identificar e questionar notícias falsas (Italia, 2017). A ênfase na educação desde a base reflete o compromisso do país com a construção de uma cultura de responsabilidade digital, um passo essencial na defesa da qualidade do debate público desde a sala de aula.

Paralelamente, a Itália mantém parcerias ativas com empresas de tecnologia, como *Google* e *Facebook*, visando a identificação e remoção rápida de conteúdos falsos. Em 2018, o *Facebook*, em colaboração com organizações locais, lançou campanhas educativas direcionadas aos usuários italianos para conscientizar sobre a importância da verificação de fontes. Da mesma forma, o *Google* implementou ferramentas como o "*Google News Initiative*" para oferecer suporte aos jornalistas e combater a desinformação em vésperas de eventos importantes, como as eleições parlamentares de 2018 (Google, 2024).

A abordagem multifacetada da Itália reflete a compreensão de que o combate à desinformação exige a atuação integrada de diversos setores da sociedade: poder público, instituições educacionais, plataformas tecnológicas e cidadãos. Ao associar rigor normativo, formação educacional e cooperação com empresas digitais, o país sinaliza um modelo de enfrentamento à desinformação que alia repressão qualificada à promoção de uma cultura informativa crítica. No Brasil, embora avanços tenham sido alcançados, a ausência de integração plena entre os setores envolvidos evidencia a necessidade de um esforço coordenado e intersetorial mais efetivo para garantir um ecossistema informativo mais transparente e resiliente.

2.4.3 A Nova Zelândia

A Nova Zelândia adota uma abordagem preventiva e colaborativa no enfrentamento à desinformação no ambiente digital. O país conta com a legislação denominada *Harmful Digital Communications Act*, sancionada em 2015, que criminaliza a divulgação de informações prejudiciais, incluindo conteúdos difamatórios, ofensivos, abusivos ou que causem sofrimento emocional significativo. Tal norma amplia o escopo da regulação digital ao abordar não apenas a veracidade da informação, mas também os impactos psicológicos e sociais gerados por práticas comunicacionais nocivas (Nova Zelândia, 2015).

A legislação prevê que comunicações digitais, como mensagens em redes sociais, *e-mails*, comentários em *blogs* ou publicações em *sites*, não podem violar princípios como o respeito à privacidade, à honra, à imagem e à não incitação ao ódio. Caso seja constatada a infração a pelo menos um dos dez princípios legais estabelecidos, a conduta pode ser considerada ilícita e passível de sanção judicial ou penal (Nova Zelândia, 2015).

A resolução de conflitos inicia-se por meio da entidade *Netsafe*, órgão oficialmente responsável por receber denúncias, intermediar negociações entre as partes envolvidas e, se possível, alcançar soluções extrajudiciais, como a retirada do conteúdo, publicação de correções ou retratações. Nos casos em que não há resolução via *Netsafe*, a parte ofendida pode recorrer à Corte Distrital, que tem competência para determinar a remoção de conteúdo, impor obrigações de não repetição ou aplicar sanções civis (Nova Zelândia, 2015).

Em situações mais graves, sobretudo quando há indícios de intenção deliberada de causar sofrimento emocional, a atuação policial se torna necessária. A Polícia da Nova Zelândia é autorizada a investigar os fatos, preservar provas e instaurar procedimentos criminais, podendo inclusive levar à responsabilização penal do infrator. Entre as infrações penais previstas, destacam-se: a não observância de ordens judiciais, a publicação de conteúdo ofensivo com intenção de prejudicar e a divulgação não consentida de imagens íntimas (Nova Zelândia, 2015).

Além do aparato normativo, o país promove campanhas públicas voltadas ao incentivo do pensamento crítico e da checagem de informações, com o objetivo de formar uma cidadania digital consciente. Tais campanhas contam com apoio direto do governo e reforçam a necessidade de autorregulação por parte das plataformas tecnológicas. A Nova Zelândia, portanto, adota uma abordagem equilibrada, que busca conciliar a proteção contra os danos provocados pela desinformação com a garantia da liberdade de expressão e o estímulo à inovação tecnológica.

2.4.4 Comparações

Portanto, a análise das estratégias de Brasil, Itália e Nova Zelândia no combate à desinformação evidencia abordagens que combinam legislação, educação e parcerias tecnológicas, embora com ênfases distintas. O Brasil está focado na criação de um arcabouço normativo e no aumento da conscientização pública. A Itália prioriza a aplicação de sanções rigorosas e a formação educacional desde a base. Já a Nova Zelândia se destaca pela prevenção, promovendo autorregulação e campanhas voltadas para o fortalecimento do pensamento crítico.

Essas experiências destacam a necessidade de soluções contextuais e integradas, que combinem punição, prevenção e colaboração entre setores público e privado. Os três países oferecem modelos complementares que podem inspirar o aperfeiçoamento global das estratégias de combate à desinformação.

3. Método

A elaboração da parte teórica desta pesquisa segue uma abordagem qualitativa de revisão bibliográfica visando pesquisar sobre o tema, *fake news* e *deep fakes* dentro do contexto de direito eleitoral, em artigos científicos, livros, legislações nacionais e estrangeiras, projetos de lei e reportagens, para exemplificar as discussões.

A coleta do material bibliográfico foi realizada por meio de consultas a bases acadêmicas como Google Acadêmico (<https://scholar.google.com.br/?hl=pt>), Scielo (<https://www.scielo.br/>), e *research gate* (<https://www.researchgate.net/>).

Ademais, utilizou-se uma base extensa de referências bibliográficas brasileiras, principalmente no recorte jurisprudencial, pois foram estudadas, principalmente, notícias e resoluções do Tribunal Superior eleitoral (TSE), da Procuradoria Geral da República (PGR) e, especificamente, a decisão proferida pelo Supremo Tribunal Federal (STF) na ADI 4037, que fixou parâmetros para a responsabilização de plataformas por conteúdos de terceiros, interpretando o artigo 19 do Marco Civil da Internet e estabelecendo limites de atuação das empresas digitais.

Complementando esse panorama, as informações sobre as legislações estrangeiras fizeram parte desse estudo, mais especificamente, foram estudadas medidas preventivas e formas com as quais a Itália e a Nova Zelândia têm lidado com as *fake news* e *deep fakes*, todas essas informações foram retiradas de *sites* oficiais de governo de cada um desses países.

Portanto, para melhor compreensão da estrutura desta pesquisa, apresentam-se a seguir os principais elementos metodológicos:

a) Tipificação: a natureza desse trabalho é interdisciplinar e internacional, pois trabalha-se com um diálogo entre política, tecnologia e direito tanto brasileiro quanto estrangeiro, os escolhidos para serem avaliados, foram o italiano e o da Nova Zelândia. Dessa forma, possibilitando um estudo teleológico sobre métodos preventivos e tratamentos normativos das *deep fakes* e *fake news* para a manutenção do processo democrático brasileiro.

b) Caracterização do local de pesquisa: por se tratar de uma pesquisa em que o foco é inteligência artificial e notícias falsas veiculadas pela internet, a coleta de dados

foi feita pela internet, para que fosse possível acessar tanto conteúdos brasileiros como estrangeiros.

c) Objeto de estudo: o foco é o impacto das *deep fakes* e *fake news* nas eleições, especialmente nas de 2024 no Brasil e a análise e proposta de medidas legislativas e políticas de combate.

d) Delimitação e universo da amostra: a pesquisa feita através de leis, projetos de lei, publicações oficiais, artigos acadêmicos e reportagens.

e) Instrumento de coleta de dados: deu-se a partir de pesquisa bibliográfica e documental, reunindo leis, estudos acadêmicos, publicações governamentais e matérias jornalísticas de sítios disponíveis na *internet*.

f) Procedimentos metodológicos: a abordagem utilizada foi qualitativa, com análise comparativa e interpretação dos dados disponíveis em *sites* na *internet*. A pesquisa foi organizada em três etapas: definição e histórico do tema; análise dos impactos sociais, políticos e jurídicos; e estudo de medidas de combate no Brasil e em outros países.

4. Resultados e discussão

4.1 Impacto na sociedade

É imprescindível que a disseminação de *fake news* e *deep fakes* tenham impactos profundos e multifacetados na política, na economia, na saúde mental e na vida cotidiana dos brasileiros. Esses fenômenos têm potencial para desestabilizar a confiança nas instituições democráticas, alterar o curso de eleições e até agravar crises de saúde pública, como ocorreu durante a pandemia de COVID-19. Desse modo, apresenta-se alguns exemplos dos impactos que informações enganosas afetaram a sociedade brasileira.

4.1.1 Efeitos na Política

Nas eleições presidenciais de 2018 no Brasil, as *fake news* desempenharam um papel central. Diversas notícias falsas circulavam principalmente nas redes sociais, e como explicitado durante essas eleições as *fake news* começaram se espalhar rapidamente a ganhar uma grande visibilidade entre as notícias brasileiras. O Correio Braziliense realizou uma matéria em que explicita demonstra um estudo realizado pelo Instituto de Tecnologia de Massachusetts (MIT) apontou que as notícias falsas espalham 70% mais rápido que as verdadeiras, alcançando um número muito maior de indivíduos (Agência Estado, 2018).

E neste mesmo estudo foi comprovado que quando a notícia falsa é ligada a política o alastramento dessa informação é três vezes mais rápido que em demais assuntos. Sob essa conjuntura, demonstra-se a notícia publicada pela Câmara de Curitiba, a qual disserta que durante as eleições de 2018 do Brasil foram fabricadas cerca de seis *fake news* por dia, uma quantia exorbitante de informações falsas que incluíam desde boatos sobre candidatos até falsificações sobre propostas políticas, o que gerou desinformação significativa.

Um exemplo marcante foi o atentado sofrido por Bolsonaro em setembro de 2018, durante um evento de campanha em Minas Gerais. Na época, circularam rumores diversos sobre o caso, como a ideia de que o atentado teria sido orquestrado

pelo Partido dos Trabalhadores (PT), de Fernando Haddad, ou até mesmo que Bolsonaro teria planejado o ataque para se fazer passar por vítima. No entanto, nenhuma dessas alegações foi comprovada.

Além disso, ainda na matéria publicada pela Câmara de Curitiba, a pesquisa de Tatiana Dourado mostra que, no segundo turno das eleições de 2018, as *fake news* intensificaram ainda mais a polarização entre os candidatos, com o nome de Haddad aparecendo frequentemente em notícias falsas. Entre as mentiras mais espalhadas estavam as histórias sobre a urna eletrônica manipulada, o "Kit Gay" (suposto programa chamado Escola Sem Homofobia) e a falsa acusação de que uma jovem teria sido marcada por uma suástica por apoiadores de Bolsonaro (que, na verdade, se automutilou). Esse mar de desinformação afeta a opinião dos eleitores, podendo chegar a alterar o curso da eleição presidencial, pois como dito anteriormente o Indicador de Alfabetismo Funcional (Inaf, 2025) apontou que 29% dos brasileiros eram considerados analfabetos funcionais em 2018 que demonstra a incapacidade desses indivíduos de discernir notícias verdadeiras das notícias manipuladas.

4.1.2 Efeitos na Economia

A disseminação de *fake news* também envolve uma dinâmica econômica complexa, muitos dos atores envolvidos na propagação de desinformação têm interesses financeiros explícitos. E o Jornal da USP publicou uma reportagem na qual expõe que para monetizar *fake news* é necessário que o conteúdo seja de extrema relevância para o público, desse modo gerando muitos "*cliques*" na publicação feita (Alta [...], 2022).

E esse é um dos modelos de negócio mais comuns para ganhos com propagação de informação enganosa, os atores realizam a criação de *sites* e perfis em redes sociais destinados a espalhar notícias falsas para gerar cliques e, conseqüentemente, receitas publicitárias. Esses *sites*, geralmente com títulos sensacionalistas, atraem grande volume de tráfego de usuários que buscam informações que confirmem suas crenças ou que despertem emoções fortes, e o estudo do MIT citado acima afirma que a maioria das publicações falsas apresentam

notícias que geram medo, desgosto ou surpresa. Com isso, o modelo de monetização baseado em publicidade digital, como o *Google AdSense*, permite que esses *sites* lucrem com a audiência gerada pela desinformação.

Em contraposição a práticas prejudiciais, o Tribunal Superior Eleitoral (TSE) emitiu uma publicação por meio de suas redes sociais, alertando a população sobre os riscos legais associados à disseminação de informações falsas na internet. A medida visa enfatizar a relevância do combate à desinformação, esclarecendo que a propagação de conteúdos enganosos pode configurar crime, e reforçando a necessidade de conscientização da sociedade sobre o impacto negativo dessa prática. A disseminação de informações fraudulentas, muitas vezes camufladas para passarem despercebidas, pode ser um fator determinante para a manipulação da opinião pública, especialmente entre indivíduos mais vulneráveis à desinformação (Pílulas [...], 2022).

Outro ponto relevante no contexto da desinformação é o mercado de *deep fakes*, que se configura como uma indústria emergente e potencialmente nociva ao tecido social e à integridade das informações. Nesse cenário, o Ministério Público Federal (MPF) tem adotado medidas rigorosas, destacando que as plataformas de redes sociais estão proibidas de realizar impulsionamento de conteúdos falsos, como *fake news* e *deep fakes*. Além disso, tais plataformas não podem remunerar canais que promovam ou veiculem notícias falsas e manipulações digitais (Deepfake [...], 2024) .

4.1.3 Efeitos na Saúde Mental

As *fake news* se configuram atualmente como uma das maiores ameaças ao bem-estar social e psicológico, principalmente devido à crescente dependência da sociedade moderna em relação aos aparelhos celulares, estes dispositivos tornaram-se essenciais para a realização de atividades cotidianas, desde tarefas básicas até o trabalho remoto, sendo indispensáveis para o indivíduo.

Além disso, com apenas alguns *cliques*, qualquer pessoa pode criar uma conta e se inserir no vasto universo digital das redes sociais. Contudo, ao adentrar esse ambiente, o usuário está exposto a uma quantidade massiva de informações, muitas

das quais carecem de fundamentação científica e são veiculadas com o intuito explícito de manipular a opinião pública.

Essa propagação de *fake news* pode ser vista como um sério dano social, em um contexto onde os sentimentos e crenças pessoais têm mais peso do que os fatos objetivos, a manipulação da opinião pública torna-se uma prática recorrente, violando os princípios de boa-fé objetiva que devem nortear a convivência social.

A disseminação de notícias falsas cria um ambiente de desconfiança e polarização, onde a sociedade se vê dividida e fragmentada, principalmente quando os temas abordados envolvem questões políticas. Este fenômeno gera uma ruptura no tecido social e contribui para a diminuição da credibilidade dos meios jornalísticos tradicionais, ao mesmo tempo em que fomenta um ciclo vicioso de desinformação que acaba por reduzir a qualidade de vida social (Guimarães, 2019).

Outrossim, tem-se um impacto significativo sobre o exercício da cidadania, especialmente no contexto eleitoral, pois ao buscar informações sobre as propostas dos candidatos, o eleitor se vê constantemente exposto a conteúdos falsos ou distorcidos, que muitas vezes não estão relacionados às propostas políticas propriamente ditas, mas sim à vida pessoal dos candidatos. Esse fenômeno contribui para a desinformação da população, levando os cidadãos a tomarem decisões eleitorais baseadas em narrativas distorcidas ou mentirosas.

E esse volume excessivo de informações e a complexidade da tarefa de verificar a veracidade de cada uma delas geram um desgaste psicológico considerável nos eleitores. Diante disso, muitos optam por adotar uma postura passiva, preferindo acreditar nas informações que surgem em suas redes sociais, baseando-se em seu próprio senso comum e percepção subjetiva de veracidade, sem recorrer a fontes confiáveis ou checagem adequada dos fatos. Esse comportamento prejudica diretamente o exercício pleno da cidadania, uma vez que compromete a capacidade dos cidadãos de fazerem escolhas eleitorais informadas e responsáveis.

No cenário atual de proliferação de desinformação nas redes sociais, percebe-se a dificuldade na formação de uma opinião pública fundamentada em fatos

e argumentos racionais. Isso resulta em uma distorção do processo eleitoral, impactando negativamente a qualidade do voto e, conseqüentemente, no fortalecimento das instituições democráticas.

Ademais, o impacto das *fake news* na saúde mental dos indivíduos não pode ser subestimado, como apresentado anteriormente às notícias falsas normalmente geram emoções de medo, desgosto ou surpresa, e essas reações emocionais tornam-se gatilhos para o compartilhamento das notícias, o que amplia ainda mais sua propagação. O reconhecimento de uma notícia como falsa também pode gerar sentimentos de frustração e impotência nos usuários, que se veem frequentemente manipulados e sem poder para combater a desinformação que circula de forma desenfreada. Com isso, este ciclo de desinformação e frustração contribui diretamente para o desgaste psicológico dos indivíduos, afetando sua saúde mental e sua visão sobre o mundo (Abjaude, 2020).

No artigo supracitado (Abjaude, 2020), disserta-se que o uso das redes sociais por pesquisadores têm permitido uma comunicação mais eficiente e acessível sobre a ciência e assuntos atuais, traduzindo o conhecimento acadêmico para uma linguagem mais compreensível pelo público geral. Este tipo de divulgação científica contribui para a democratização do conhecimento, ajudando a combater a desinformação e oferecendo à sociedade informações de qualidade, baseadas em dados científicos.

Um exemplo significativo dessa abordagem é o Centro de Pesquisa em Assistência Farmacêutica e Farmácia Clínica da Universidade de São Paulo (CPAFF-USP), que, em 2017, criou uma rede de informação nas redes sociais para popularizar o conhecimento científico além das fronteiras do meio acadêmico (Abjaude, 2020). Através do uso de plataformas como *Instagram*, *Facebook* e *YouTube*, a instituição tem contribuído para a disseminação de informações confiáveis sobre saúde, medicamentos e outros temas relevantes, alcançando um público mais amplo e facilitando o acesso à informação de qualidade.

4.2 Mecanismos de disseminação

"Com a revolução digital, a informação tornou-se acessível em uma escala sem precedentes, possibilitando uma maior democratização do conhecimento, mas também desafios no controle da veracidade dos dados" (Kahn; Kellner, 2004).

Essa citação nos mostra como o mundo se transformou após a revolução digital, pois com a contribuição da globalização e do aumento exponencial do uso de redes sociais, a criação e disseminação de *fake news* e *deep fakes* também cresceu e com isso diversos problemas novos surgiram. Dessa forma, a compreensão dos mecanismos de disseminação dessas informações se torna essencial para que, nós, como sociedade, possamos enfrentar a crescente onda de desinformação, focando, especialmente, nos períodos eleitorais que virão, para garantir que a democracia permaneça justa.

4.2.1 Criação de *Fake News* e *Deep fakes*

Segundo Sun Tzu (2003), se você conhece o inimigo e conhece a si mesmo, por isso, não precisa temer o resultado de cem batalhas. Essa máxima reforça a importância do conhecimento para a estratégia e a vitória. Combater as *fake news* e as *deep fakes* funciona da mesma forma: antes de nos defendermos, precisamos entender como esses conteúdos são criados e disseminados, para que possamos criar estratégias para combatê-las. Assim como um soldado analisa as táticas e os movimentos do oponente, a sociedade precisa compreender as engrenagens por trás dessas falsas narrativas. Somente com esse entendimento poderemos construir defesas sólidas e proteger nossa democracia.

Ao analisar o trabalho de Vosoughi, Roy e Aral (2018), é abordado que o primeiro passo para a criação das *fake news* consiste em identificar temas sensíveis ou polêmicos, que geralmente incluem assuntos de grande relevância pública, como questões políticas, de saúde ou de segurança, e utilizá-los, pois como provocam fortes reações emocionais da população acabam tendo um maior potencial de viralização, já que provocam um engajamento expressivo, como por exemplo, diversos comentários na publicação, o que a faz ser mais impulsionada e conseqüentemente, chegando para mais pessoas.

Na sequência, os autores explicam que ocorre a manipulação e distorção de fatos, isso consiste em usar informações reais, mas removidas de seu contexto ou apresentá-las de forma manipulada. Um exemplo é o uso de citações fora de contexto ou a alteração de dados e estatísticas, criando uma falsa impressão sobre determinado acontecimento ou figura pública. Essa técnica confere às *fake news* uma aparência de legitimidade, dificultando a identificação da falsidade pelo público.

Outra etapa central citada pelos autores, no processo, é criar textos, vídeos e imagens que copiam o estilo e a formatação de fontes confiáveis, como por exemplo, *sites* de notícia, o que dá às *fake news* uma aparência de autenticidade.

A partir disso, foi-se popularizado o uso das *deep fakes*, pois elas surgiram através da inteligência artificial e permitiram a criação de vídeos e áudios que reproduzem fielmente as características visuais e vocais de figuras públicas, fortalecendo a ilusão de veracidade e a credibilidade do conteúdo fabricado. Esse avanço tecnológico, porém, traz um problema sério: dificulta a distinção entre o que é verdadeiro e o que é falso. Conforme apontam Vaccari e Chadwick (2020), a circulação desses vídeos políticos sintéticos aumenta a incerteza dos espectadores e abala a confiança nas notícias, aprofundando a crise da desinformação. Dessa forma, as *fake news* e as *deep fakes* ganham força e impacto significativo na sociedade, comprometendo o debate público e, em última instância, ameaçando a saúde dos processos democráticos.

4.2.2 Redes Sociais como Vetores de Disseminação

A disseminação nas redes sociais e em plataformas digitais é o um passo crucial para gerar uma ampla desinformação entre a população. Após a sua criação, as *fake news* e as *deep fakes* são amplamente distribuídas por meio de redes sociais, *blogs* e canais de notícias alternativos, que facilitam uma rápida propagação (Vaccari; Chadwick, 2020).

Dessa forma, a disseminação acelerada desses conteúdos ocorre, em grande parte, devido à estrutura e à lógica de funcionamento das redes sociais, como *Facebook*, *Twitter* (atualmente “X”) e *Instagram*, que favorecem o compartilhamento

em massa, intensificando a circulação de desinformação em larga escala. Além disso, para amplificar seu alcance, utilizam-se bots e algoritmos de recomendação, que aumentam o volume de interações, gerando uma falsa impressão de ampla aceitação e popularidade.

Ademais, outra maneira de tornar as *fake news* e as *deep fakes* mais verossímeis é a repetição da informação em múltiplos canais, o que gera o que se conhece como “ilusão de verdade”, esse é um fenômeno psicológico que ocorre quando a repetição de uma afirmação aumenta a percepção de sua veracidade, independentemente da sua precisão (Brashier; Eliseev; Marsh, 2020).

A exposição repetida ao mesmo conteúdo em diferentes fontes, mesmo que não confiáveis, reforça a impressão de veracidade, levando o público a considerar algo como verdadeiro apenas pela familiaridade e pela frequência de exposição (Hasher; Goldstein; Toppino, 1977 apud Pennycook; Cannon; Rand, 2018). Esse efeito é potencializado em ambientes digitais, onde informações são amplamente compartilhadas e reproduzidas em um curto espaço de tempo.

Para aprofundamento, vale a pena avaliar alguns dos motivos pelos quais as redes sociais se tornaram um dos principais disseminação de *fake news* e as *deep fakes*:

- **Viralidade:** O design das redes sociais, como por exemplo, *Twitter (X)*, facilita a disseminação em massa de conteúdo por meio de curtidas, compartilhamentos e comentários. Quando a plataforma observa que uma publicação está recebendo uma quantidade certa de atenção, ela será impulsionada pela plataforma, mesmo que seja uma *fake news* e assim atingindo milhões de pessoas em pouco tempo.

Pesquisas apontam que as *fake news* têm um poder de alcance muito maior do que as notícias verdadeiras. Um estudo publicado na revista *Science* em 2018 mostrou que conteúdos falsos se espalham com mais rapidez e atingem um público mais amplo, justamente porque as redes sociais priorizam o engajamento, seja por meio de curtidas, comentários ou compartilhamentos. Quando uma publicação viraliza, os

próprios algoritmos a impulsionam, ampliando ainda mais sua visibilidade, sem considerar se a informação é verdadeira ou não (Vosoughi; Roy; Aral, 2018).

- **Algoritmos de Recomendação:** Como disposto acima, as plataformas utilizam algoritmos específicos que são projetados para maximizar o engajamento, e conteúdos sensacionalistas ou polarizadores, como *fake news*, têm mais chance de serem promovidos. Tudo isso é bem explicado no trabalho de Recuero, Zago e Soares (2017, p. 6):

Muitos destes filtros estão baseados em processos de filtragem social (LERMAN, 2007), onde a própria rede dá visibilidade àquilo que considera importante, enquanto despreza aquilo que não o é. A filtragem social de informações não é um processo novo, uma vez que a circulação de informações é um dos principais pontos de estudo estrutural para a análise de redes sociais (vide BURT, 1992, por exemplo). Entretanto, em ferramentas de mídia social, essa atuação na difusão de informações atinge novos patamares (BAKSHY et al., 2012), inclusive porque não está baseada apenas na ação dos indivíduos, mas também, em algoritmos-curadores (TUFEKCI, 2015).

Aproveitando dessa lógica, criadores de conteúdo buscando aumentar suas visualizações, descobriram e se beneficiaram de um mecanismo conhecido como clickbait. Segundo Scott (2023), *clickbait headlines are teasers which are carefully designed to drive traffic to webpages by enticing users to click*.

- **Personalização de conteúdos:** Além disso, a segmentação baseada em dados de usuários permite que conteúdos sejam direcionados a públicos específicos, o que agrava o problema das "bolhas de informação". Isso significa que a desinformação pode ser personalizada para públicos distintos, reforçando suas crenças preexistentes e dificultando o contato com informações verdadeiras.

No livro *"The Filter Bubble"*, Eli Pariser (2011) argumenta que as bolhas de informação surgem quando os algoritmos das plataformas digitais personalizam o conteúdo de acordo com as preferências dos usuários. Além disso, ele afirma que, ao serem expostos apenas a opiniões e informações que reforçam suas crenças, os indivíduos podem se isolar em um ambiente que promove a polarização e a desinformação, ou seja, esse fenômeno pode levar a uma "realidade filtrada", onde as

peessoas se tornam menos informadas sobre questões sociais e políticas importantes, uma vez que não encontram visões opostas às suas.

4.2.3 Aplicativos de Mensagens Privadas

Os aplicativos de mensagens privadas, como *WhatsApp*, *Telegram* e *Signal*, desempenham um papel fundamental na disseminação de *fake news* e *deep fakes*. Esses aplicativos são especialmente desafiadores no combate à desinformação porque:

- **Criptografia de Ponta a Ponta:** A segurança oferecida pela criptografia de ponta a ponta nesses aplicativos significa que as mensagens enviadas entre os usuários são privadas, não podendo ser monitoradas por terceiros. Embora essa privacidade seja importante para garantir a segurança das comunicações, ela também impede que plataformas como *WhatsApp* impeçam a disseminação de conteúdos falsos. Como o próprio *site do WhatsApp (2025)* discorre:

A criptografia de ponta a ponta do WhatsApp protege suas conversas com outras pessoas no WhatsApp Messenger, garantindo que as mensagens e ligações fiquem apenas entre você e a pessoa com quem conversar. Ninguém mais pode ler ou ouvir suas conversas, nem mesmo o WhatsApp. Com a criptografia de ponta a ponta, as mensagens e ligações são protegidas com um cadeado exclusivo e somente você e a pessoa que recebe a mensagem têm acesso à chave especial para destrancá-lo e ler as mensagens. Não é necessário ativar configurações especiais para garantir a segurança das mensagens, pois todo esse processo ocorre automaticamente.

- **Grupos e Difusão Rápida:** Ademais, a criação de grupos com grande número de membros facilita a rápida disseminação de desinformação. Em aplicativos como o *WhatsApp*, uma única mensagem contendo *fake news* pode ser encaminhada repetidamente, atingindo um grande número de pessoas em pouco tempo. Durante o período eleitoral no Brasil, muitas *fake news* foram compartilhadas em grupos familiares e comunitários, onde a confiança nos remetentes fez com que as mensagens fossem recebidas sem questionamento.

Uma pesquisa realizada durante as eleições de 2018 identificou que grupos pró-Bolsonaro no *WhatsApp* foram usados para orquestrar a disseminação de notícias

falsas e ataques pessoais na internet, mostrando como essas plataformas podem ser exploradas para manipular informações em massa (Simões, 2018). Além disso, o Tribunal Regional Eleitoral da Bahia (TRE/BA, 2018) destacou que grupos no *WhatsApp* atuam como a principal porta de entrada para a propagação de notícias falsas e golpes, evidenciando o papel central desses espaços virtuais na disseminação de conteúdos enganosos.

Esses casos ilustram como a estrutura de grupos no *WhatsApp*, associada à dinâmica de confiança e proximidade entre os participantes, pode amplificar a disseminação de desinformação em períodos críticos, como campanhas eleitorais, tornando essencial o desenvolvimento de estratégias para conter esse problema.

- **Imunidade às Plataformas Públicas:** Diferente das redes sociais públicas, onde postagens podem ser denunciadas ou verificadas por serviços de checagem de fatos, as mensagens em aplicativos privados dificilmente passam por qualquer forma de controle externo. Isso facilita a propagação de conteúdos sem a mesma visibilidade pública.

Um exemplo notório desse fenômeno no Brasil foi a disseminação de mensagens em massa durante as eleições de 2018 via *WhatsApp*, como no caso das "mamadeiras de piroca", (Queiroga, 2021), que circulou amplamente em grupos familiares, gerando impacto significativo no debate público.

4.2.4 Outras Plataformas Digitais

Além das redes sociais e aplicativos de mensagens, outras plataformas digitais como *YouTube* e *TikTok* também contribuem para a disseminação de *fake news* e *deep fakes*. Vídeos falsificados podem se tornar virais nessas plataformas, atingindo milhões de visualizações antes de serem identificados como falsos. O formato audiovisual é especialmente perigoso, pois tende a ser mais envolvente e persuasivo do que o texto, aumentando o potencial de convencimento das *fake news* e *deep fakes* (Pennycook; Bear; Rand, 2021).

Estudos recentes mostram que, no caso específico de *deep fakes* em plataformas abertas como *YouTube* e *TikTok*, os principais alvos não são apenas figuras políticas de oposição ou personalidades vulneráveis, mas predominantemente homens brancos e, muitas vezes, líderes autoritários. Além disso, a maioria dos *deep fakes* encontrados nessas redes não tem apenas uma intenção política direta, mas também um caráter satírico ou simbólico, funcionando como uma “*profanation carnavalesca*” dos rostos do poder, em um uso cultural e humorístico mais do que estritamente ideológico (Öhman, 2025).

4.2.5 O Papel dos Algoritmos e da IA

A tecnologia que facilita a disseminação da informação também é responsável por amplificar o impacto das *fake news* e *deep fakes*. Isso acontece porque os algoritmos de inteligência artificial usados por plataformas como *Facebook* e *YouTube* priorizam conteúdos que geram alto engajamento. Ou seja, quanto mais curtidas, comentários e compartilhamentos, mais aquele conteúdo é impulsionado. E como as *fake news* sensacionalistas e as *deep fakes* mais chocantes costumam provocar reações fortes, elas têm muito mais chance de serem recomendadas e visualizadas. Segundo Vosoughi, Roy e Aral (2018), conteúdos falsos circulam com mais velocidade e alcance do que os verdadeiros justamente por despertarem emoções intensas como surpresa, medo ou indignação.

Além disso, ferramentas de geração automática de conteúdo e o uso de bots tornam o processo ainda mais rápido e massivo. Esses bots são programados para postar e compartilhar *fake news* em larga escala, criando a ilusão de popularidade e legitimidade. Isso influencia diretamente a percepção pública, porque dá a impressão de que aquele conteúdo é amplamente aceito ou verdadeiro, quando, na verdade, é parte de uma estratégia de manipulação e desinformação.

4.3 Tecnologias de detecção

Neste contexto, diversas ferramentas e métodos tecnológicos têm sido desenvolvidos para identificar e combater a desinformação. No entanto, a eficácia

dessas tecnologias ainda enfrenta diversos obstáculos, o que torna o seu uso um tema relevante para reflexão, principalmente no âmbito do direito.

4.3.1 Detecção de *Fake News*

No combate à disseminação de notícias falsas é possível distinguir dois grandes grupos de estratégias: aquelas que podem ser adotadas diretamente pela população e aquelas fundamentadas em tecnologias desenvolvidas por especialistas na área da informação e da computação.

No que tange às medidas acessíveis ao público em geral, destaca-se a análise de conteúdo e linguagem, a qual visa identificar padrões linguísticos, estilísticos e semânticos comumente presentes em *fake news*. Na reportagem da PBS, um telejornal estadunidense, destaca-se a análise de sentimento, uma vez que tais notícias, via de regra, recorrem a uma linguagem carregada de apelos emocionais com o intuito de manipular a percepção do leitor. Ademais, a verificação do estilo de escrita também se mostra relevante: inconsistências textuais, erros gramaticais em excesso e títulos sensacionalistas são elementos que podem indicar a inveracidade da informação (Torabi, 2019).

Outra medida relevante consiste na verificação da origem da informação, ou seja, investigar a procedência da notícia, o histórico do domínio eletrônico e os perfis que promoveram sua disseminação. Tais elementos podem revelar esquemas organizados de desinformação. Da mesma forma, a análise dos padrões de disseminação, considerando a rapidez e o formato com que determinada informação se espalha pelas redes sociais, pode ser indicativa de manipulação coordenada.

Paralelamente, do ponto de vista técnico, foram desenvolvidas ferramentas tecnológicas que visam auxiliar a população na identificação e no combate às *fake news*. No *website* estadunidense Hyperscience (2025), criado para rotular, anotar e estruturar automaticamente documentos complexos, gerando dados confiáveis e de alta qualidade para treinar modelos de IA com precisão específica para o negócio, destaca-se, entre essas ferramentas, o uso de técnicas de Processamento de Linguagem Natural, que envolvem o emprego de algoritmos de *machine learning* e

deep learning treinados para classificar textos como verídicos ou falsos, com base em grandes volumes de dados.

Outros mecanismos importantes incluem as plataformas de checagem de fatos, conduzidas por organizações especializadas em *fact-checking*, cuja finalidade é verificar a veracidade das alegações veiculadas nos meios de comunicação. Também se menciona o uso de bancos de dados de conhecimento, consistentes em bases compostas por informações previamente verificadas, que servem de parâmetro para validar novas alegações.

Em síntese, a atuação conjunta entre a população, por meio da adoção de uma postura crítica e investigativa, e as soluções tecnológicas desenvolvidas por especialistas é essencial para mitigar os impactos negativos causados pela propagação de notícias falsas, protegendo-se, assim, valores fundamentais como a liberdade de informação, a segurança jurídica e o regular funcionamento das instituições democráticas.

4.3.2 Detecção de *Deep Fakes*

O combate às *deep fakes* demanda uma abordagem integrada e coordenada, envolve quatro estratégias principais, a saber: legislação e regulamentação, políticas empresariais e ações voluntárias, educação e treinamento e, por fim, tecnologia *anti-deepfake* (Westerlund, 2019).

Em primeiro lugar, a legislação e regulamentação desempenham um papel fundamental, sendo necessário o desenvolvimento de normas específicas que responsabilizam a criação e disseminação de *deep fakes*, com a imposição de sanções adequadas. A criação de um marco legal robusto permitirá, assim, a punição dos responsáveis por tais práticas fraudulentas, protegendo os direitos da personalidade e a integridade das informações.

Em segundo lugar, as políticas empresariais e ações voluntárias devem ser adotadas pelas empresas de tecnologia, as quais têm a obrigação de implementar medidas preventivas. Tais medidas incluem a verificação de conteúdos e a colaboração

ativa com as autoridades competentes, a fim de mitigar a propagação da desinformação. A responsabilização das plataformas digitais pela disseminação de conteúdos manipulados é, portanto, uma questão que precisa ser abordada de forma clara, à luz da legislação vigente.

A educação e o treinamento também são essenciais no combate às *deep fakes*, especialmente por meio de programas educativos que promovam a conscientização pública sobre o impacto das manipulações digitais. Capacitar os cidadãos para identificar e prevenir a propagação de conteúdos fraudulentos contribui não só para a proteção individual, mas também para a preservação da ordem pública e da confiança nas informações disseminadas.

Por último, a tecnologia *anti-deepfake* representa uma ferramenta crucial para a detecção e bloqueio de conteúdos falsificados. O uso de tecnologias avançadas, como inteligência artificial e blockchain, pode ser empregado para autenticar mídias e identificar vídeos ou áudios manipulados. Estas tecnologias, aliadas a outros métodos de verificação, como as marcas d'água digitais e assinaturas digitais, podem garantir a autenticidade dos conteúdos, facilitando a detecção de manipulações posteriores.

Para sua detecção, diversas ferramentas tecnológicas têm sido desenvolvidas e aplicadas, como a análise de artefatos visuais e auditivos, que busca identificar falhas sutis nas imagens e no áudio. Apesar dos avanços tecnológicos, uma matéria do jornal estadunidense *New York Times* afirma que as *deep fakes* ainda apresentam inconsistências faciais, problemas de sincronização labial e artefatos visíveis no contorno do rosto e cabelo, que podem ser detectados tanto por algoritmos quanto por especialistas. Além disso, a análise de sinais fisiológicos sutis, como o pulso sanguíneo refletido na pele, também pode ser utilizada para identificar a manipulação de conteúdos (Thompson, 2024).

Outra abordagem importante para a detecção de *deep fakes* envolve o uso de modelos de *deep learning*, com destaque para as redes neurais convolucionais (CNNs) e as redes neurais recorrentes (RNNs). Tais modelos são amplamente utilizados para

analisar os frames dos vídeos e sequências de áudio, buscando padrões que possam indicar falsificação.

Ademais, no cenário digital atual, garantir a autenticidade do conteúdo é cada vez mais crucial, especialmente com o avanço rápido das tecnologias de IA e a crescente proliferação de imagens e vídeos manipulados para disseminar desinformação, com isso métodos tradicionais mostram-se insuficientes, tornando urgente a adoção de soluções técnicas robustas. Nesse contexto, destaca-se no *website* da castLabs (2024) a técnica de watermarking forense em quadro único, que incorpora marcas digitais em conteúdos autênticos para facilitar a identificação de manipulações e garantir a integridade e procedência da informação, oferecendo uma solução confiável para o futuro da mídia.

Em suma, o enfrentamento das *deep fakes* exige uma combinação de ações normativas, empresariais, educacionais e tecnológicas. O desenvolvimento de um arcabouço jurídico adequado, aliada ao uso de ferramentas avançadas de detecção, é imprescindível para proteger os direitos dos indivíduos e assegurar a veracidade da informação no ambiente digital.

4.4 Educação e conscientização

Tendo em vista tudo abordado anteriormente, é possível verificar que a disseminação de *deep fakes* e *fake news* constitui um dos desafios mais prementes para o Direito Eleitoral na contemporaneidade, afetando de maneira direta a lisura do processo democrático. A crescente sofisticação das tecnologias de manipulação de imagens e informações impõe novos desafios aos ordenamentos jurídicos, que precisam equilibrar a liberdade de expressão com a necessidade de preservar a integridade eleitoral e a confiança pública nas instituições. Neste contexto, este estudo buscou analisar os impactos das *deep fakes* e *fake news* dentro do contexto do direito eleitoral concomitantemente com o cenário político, bem como discutir estratégias eficazes para mitigar sua influência nociva.

No cenário atual, a regulação por meio da legislação é o caminho mais palpável, todavia de difícil fiscalização e execução. Nesse sentido, talvez a criação de uma

agência reguladora seria mais eficaz, mesmo diante das dificuldades existentes. Como alternativa, o apoio estatal às agências de *fact-checking* já consolidadas surge como forma de evitar danos políticos, democráticos e até sanitários decorrentes das *fake news* (Baldissera; Fortes, 2021).

A análise desenvolvida ao longo desta pesquisa evidenciou que a regulamentação do fenômeno é um desafio que exige uma abordagem multifacetada, combinando ações legislativas, tecnológicas e educativas. Em termos normativos, verificou-se a necessidade de um aprimoramento das legislações existentes, especialmente no que tange à responsabilização de indivíduos e entidades que deliberadamente criam e disseminam desinformação com o intuito de manipular o eleitorado. Além disso, é fundamental estabelecer parâmetros claros para a atuação das plataformas digitais, impondo-lhes deveres mais rigorosos no combate à propagação de conteúdos enganosos.

No âmbito tecnológico, observou-se a importância da implementação de ferramentas de verificação coletiva, como as "notas da comunidade" já empregadas em algumas redes sociais, como por exemplo, na plataforma X, antigo *Twitter*, que permite que usuários contribuam com contextos adicionais a publicações potencialmente enganosas. O diferencial está no fato de que a nota só é exibida publicamente quando pessoas com diferentes pontos de vista concordam com sua utilidade, promovendo um equilíbrio plural na avaliação da veracidade das informações (X, 2025). Além disso, o desenvolvimento de mecanismos automatizados de detecção de *deep fakes* mostra-se como um complemento estratégico, já que esses recursos podem ser aliados valiosos na identificação precoce de conteúdos manipulados, reduzindo seu impacto antes que alcancem grande disseminação. Paralelamente, o incentivo governamental a campanhas de conscientização conduzidas por influenciadores digitais pode ampliar o alcance das informações verídicas e promover maior engajamento social na luta contra a desinformação.

A educação midiática também desponta como um elemento essencial para a solução do problema. A introdução de disciplinas que estimulem o pensamento crítico e a verificação de fontes no ensino básico pode ter um efeito multiplicador na forma

como as novas gerações consomem e compartilham informações. Até mesmo o retorno das aulas de informática nas escolas pode contribuir e facilitar a didática dos professores sobre os perigos da internet, incluindo assuntos como *fake news* e *deep fakes*.

Ademais, campanhas educativas direcionadas a grupos mais vulneráveis à desinformação podem contribuir para a formação de uma sociedade mais preparada para enfrentar os desafios impostos pelas novas tecnologias de manipulação digital.

Dessa maneira, destacam-se as formas de combate às *fake news* e *deep fakes*:

4.4.1 Incentivo governamental para influenciadores no combate às *fake news*

Como já discutido anteriormente, as redes sociais têm um papel fundamental na disseminação de informações e influenciam diretamente o comportamento político e social da população (Vosoughi; Roy; Aral, 2018). Pensando nisso, propõe-se a criação de um programa governamental em parceria com influenciadores digitais. Essa parceria não se limitaria apenas aos criadores de conteúdo populares por vídeos de entretenimento, como os de “dancinhas” no TikTok, mas também incluiria influenciadores políticos, como Nikolas Ferreira e Erika Hilton. O objetivo seria que esses influenciadores atuem ativamente na produção de conteúdos que conscientizem a população sobre os riscos e impactos da desinformação.

Uma pesquisa de Woolley (2022) destaca que influenciadores políticos são usados em diversos países como estratégia para disseminar informações falsas, ampliando o alcance das mensagens e conferindo a elas uma aparência mais autêntica. Isso ocorre porque esses influenciadores atingem públicos específicos com narrativas locais e emocionais, utilizando conexões pessoais e comunitárias. Além disso, essa prática permite que campanhas políticas evitem a transparência, já que os pagamentos aos influenciadores muitas vezes ocorrem fora das plataformas digitais e as publicações nem sempre são identificadas como patrocinadas. Influenciadores também impulsionam candidatos, inclusive os com visões extremas, que dificilmente teriam visibilidade de outra forma.

Diante desse cenário, surge a seguinte pergunta: e se essa estratégia fosse utilizada para o benefício da população? Ou seja, em vez de disseminar desinformação, os influenciadores poderiam contribuir para reduzir a circulação de *fake news*?

Com base nessa ideia, propõe-se uma iniciativa que alinhe as estratégias de comunicação pública às novas formas de consumo de informação, nas quais os criadores de conteúdo têm papel fundamental na formação da opinião pública, principalmente entre os jovens. Um exemplo inicial seria a escolha de cinco influenciadores com grande engajamento para divulgar campanhas contra as *fake news*. Em contrapartida, eles poderiam receber benefícios fiscais, como isenção de impostos estaduais ou municipais, por exemplo Imposto sobre a Propriedade de Veículos Automotores (IPVA) ou Imposto Predial e Territorial Urbano (IPTU), durante um ano. Esses incentivos serviriam para valorizar social e economicamente o trabalho desses agentes digitais na defesa da integridade informacional.

Além disso, recomenda-se a criação de uma força-tarefa interinstitucional, formada por especialistas em comunicação digital, representantes de plataformas tecnológicas e órgãos públicos ligados à fiscalização e à educação. Essa equipe teria a responsabilidade de validar a veracidade dos conteúdos produzidos, assegurar que respeitem princípios éticos e legais, e promover estrategicamente os materiais com maior potencial positivo. Esse modelo colaborativo ajudaria a construir um ecossistema mais eficaz, democrático e tecnicamente preparado para combater a desinformação, reconhecendo também a legitimidade dos influenciadores como agentes importantes de transformação social no cenário atual da comunicação.

4.4.2 Unificação das emissoras para alertas sobre *fake news* durante eleições

Essa proposta consiste na formação de uma parceria entre as principais emissoras de televisão e rádio do país, com o objetivo de veicular diariamente campanhas educativas voltadas à conscientização da população sobre os riscos da desinformação, especialmente durante o período eleitoral. Trata-se da criação de uma frente unificada de comunicação, capaz de reforçar a importância da checagem de fatos e do consumo responsável de conteúdo midiático.

Nesse sentido, Porcello (2020, p. 1) afirma que o telejornalismo é um “eficiente antídoto contra mentiras, inverdades, falsidades, *fake news* e todo o tipo de manifestação inverídica divulgada para distorcer a realidade, enganar as pessoas e ameaçar a democracia”. A televisão brasileira, que cobre grande parte do território nacional, tem poder de alcance e influência sobre a população, tornando-se um meio privilegiado para campanhas de conscientização pública.

Ademais, uma pesquisa realizada por Silva Júnior e Carvalho (2019) evidenciou que a disseminação de *fake news* está diretamente relacionada a motivações pessoais e coletivas. Segundo os autores, os principais fatores que levam à propagação de notícias falsas são o interesse pessoal no conteúdo da mensagem, a familiaridade com o tema e o desejo de reforçar visões de mundo, muitas vezes desqualificando grupos com posicionamentos contrários. Em geral, essas informações são compartilhadas sem a devida verificação de sua veracidade.

Dessa forma, ao utilizar meios de comunicação tradicionais, com os quais a população já está habituada, para divulgar alertas educativos, torna-se possível reduzir a disseminação de *fake news* e estimular o senso crítico dos cidadãos. A veiculação desses conteúdos em horários de maior audiência (os chamados horários de pico) amplia significativamente o alcance da mensagem, promovendo maior familiaridade da população com as temáticas abordadas e fortalecendo a resistência coletiva à desinformação.

4.4.3 Fortalecimento das Notas da Comunidade nas redes sociais

Soluções inovadoras são necessárias para mitigar os danos causados pelas *deep fakes*, principalmente em períodos críticos, como as eleições. Uma proposta promissora consiste no fortalecimento das Notas da Comunidade, uma ferramenta colaborativa de verificação de conteúdo integrada diretamente às plataformas digitais, esta proposta explora o potencial das Notas da Comunidade como uma estratégia eficaz para reduzir a disseminação de *deep fakes* nas redes sociais durante o período eleitoral, considerando o papel das parcerias com agências de checagem de fatos e o uso de tecnologias avançadas, como inteligência artificial.

As Notas da Comunidade são uma ferramenta que permite aos usuários das redes sociais sinalizar conteúdos que consideram como desinformação, criando uma rede colaborativa de verificação. Essa abordagem tem o potencial de empoderar os próprios usuários na luta contra a desinformação, uma vez que, ao utilizar a ferramenta, os usuários podem identificar e questionar conteúdos potencialmente manipulados. No contexto das eleições, o fortalecimento do uso dessas notas poderia resultar na diminuição da propagação de *deep fakes*, pois os próprios eleitores estariam mais atentos e capacitados para verificar as informações consumidas e compartilhadas.

Ao analisar as notas da comunidade oferecidas pela rede social denominada X, anteriormente conhecida como Twitter, o artigo científico realizado por Nicolas Pröllochs, aferiu que estas tiveram um desempenho melhor do que de demais iniciativas, pois o Birdwatch difere em vários aspectos importantes: (i) ele está diretamente integrado à plataforma do Twitter, o que resulta em barreiras de entrada relativamente baixas para os verificadores de fatos; (ii) apresenta um sistema de avaliação relativamente sofisticado para promover verificações de fatos úteis; (iii) é uma abordagem puramente comunitária, o que pode reduzir o problema de muitos usuários desconfiarem dos verificadores de fatos profissionais; (iv) incentiva os usuários a fornecerem contexto e a vincular a fontes confiáveis, o que é conhecido por aumentar a credibilidade.

Com isso, para que as Notas da Comunidade se tornem uma ferramenta confiável e eficaz, é necessário investir em campanhas educativas que orientem os usuários sobre sua utilização. Tais campanhas podem ser realizadas pelas plataformas digitais e por organizações civis, com o objetivo de esclarecer o papel das notas na identificação de conteúdos falsificados e na promoção da transparência informacional. Além disso, é fundamental que as plataformas adotem mecanismos que incentivem a participação ativa dos usuários, como sistemas de recompensas ou visibilidade para aqueles que contribuem de forma significativa para a verificação de conteúdos.

A credibilidade das Notas da Comunidade pode ser aprimorada por meio da implementação de parcerias com agências de checagem de fatos, como PolitiFact,

Agência Lupa e outras organizações especializadas. Essas parcerias permitem que os conteúdos sinalizados pelos usuários sejam analisados e validados por profissionais, proporcionando uma camada adicional de verificação e conferindo maior confiabilidade às informações compartilhadas. A colaboração entre plataformas digitais e essas agências não só fortalece a integridade das notas, mas também aumenta a precisão e rapidez na resposta a conteúdos manipulados.

Além das abordagens colaborativas, o uso de tecnologias avançadas é essencial para a identificação de *deep fakes*. O emprego de inteligência artificial (IA), especialmente por meio de algoritmos de aprendizado de máquina, pode facilitar a detecção automatizada de vídeos e áudios manipulados. A IA pode ser utilizada para identificar características típicas de *deep fakes*, como inconsistências no movimento facial, falhas na sincronização labial, distorções no contorno do rosto e outros artefatos visuais e auditivos, como explicitado anteriormente. Quando combinado com as Notas da Comunidade, esse sistema pode aumentar significativamente a eficácia da verificação, permitindo que conteúdos manipulados sejam rapidamente sinalizados e removidos.

A implementação conjunta das Notas da Comunidade, das parcerias com agências de checagem e do uso de tecnologias de detecção pode ter um impacto significativo na redução da disseminação de *deep fakes* durante as eleições. Ao permitir que os cidadãos desempenhem um papel ativo na verificação de informações, essas estratégias contribuem para aumentar a conscientização e a vigilância pública em relação à desinformação. A combinação de esforço humano e ferramentas tecnológicas cria uma rede de proteção contra a manipulação de conteúdos, especialmente no contexto eleitoral, onde a precisão das informações é crucial para a tomada de decisões informadas pelos eleitores.

A implementação das Notas da Comunidade apresenta um potencial significativo para a verificação de fatos, mas ainda enfrenta desafios, especialmente quando se trata de conteúdo proveniente de contas de usuários influentes nas redes sociais. O artigo mencionado anteriormente aponta que, no caso do *Birdwatch*, uma parcela considerável das verificações de fatos é direcionada à *tweets* de políticos

influentes, tanto de um lado quanto do outro do espectro político nos Estados Unidos, sendo que, para esses *tweets*, a comunidade de usuários tende a reportar, de maneira predominante, desinformação. Isso sugere que existe um forte interesse por parte da população em desmascarar a desinformação política, evidenciando uma preocupação com o conhecimento verídico dos fatos, em oposição à propagação de *fake news* e *deep fakes*. (Pröllochs, 2022).

Essa análise levanta questões jurídicas e éticas relevantes, especialmente no que diz respeito à liberdade de expressão, à moderação de conteúdo nas plataformas digitais e à imparcialidade nas ferramentas de verificação de fatos. A dinâmica de verificação de desinformação, portanto, não apenas deve ser analisada sob o prisma tecnológico, mas também dentro dos parâmetros legais que garantem a equidade e a transparência no processo de identificação e remoção de conteúdos enganosos.

Em suma, o fortalecimento das Notas da Comunidade nas redes sociais, por meio de incentivos à participação ativa dos usuários, parcerias com agências de checagem e a integração de tecnologias avançadas, como inteligência artificial e *blockchain*, representa uma solução promissora para mitigar a disseminação de *deep fakes* durante as eleições. A efetividade dessa abordagem depende da colaboração entre plataformas digitais, organizações especializadas e a sociedade civil, criando um ecossistema de verificação de fatos mais transparente, ágil e confiável. À medida que a confiança pública nas informações digitais se deteriora, o fortalecimento dessas ferramentas é essencial para garantir que o processo eleitoral seja conduzido de maneira justa, informada e livre de manipulações.

4.4.4 Análise de campanhas de conscientização bem-sucedidas

Além dessas propostas, é necessário analisar campanhas de conscientização que obtiveram resultados positivos em outras áreas, para que seja possível o desenvolvimento de estratégias eficazes no combate às *fake news*. Dois exemplos notáveis são as campanhas antifumo e anti-buzina realizadas em Brasília, reconhecidas por terem conseguido ser eficazes.

A campanha antifumo, por exemplo, foi pioneira ao promover uma mudança cultural em ambientes fechados, tornando inaceitável o ato de fumar em locais públicos. Segundo informações do Distrito Federal (2023), essa transformação foi alcançada por meio de uma combinação de estratégias: forte apelo visual com mensagens diretas, apoio legislativo, fiscalização e, principalmente, ampla divulgação em diferentes meios de comunicação. Além disso, houve o envolvimento de figuras públicas e especialistas da saúde, que deram credibilidade à iniciativa e ampliaram seu alcance.

Já a campanha antibuzina foi uma iniciativa lançada em 1994 pelo Departamento de Trânsito do Distrito Federal (Detran-DF) com o objetivo de reduzir a poluição sonora causada pelo uso excessivo da buzina de veículos em Brasília. A campanha incluiu ações educativas, sinalização urbana com placas indicando "Cidade Antibuzina" e reforço na fiscalização para coibir o uso indevido da buzina. A mudança de comportamento não foi imposta de maneira punitiva, mas construída por meio de uma narrativa de empatia e convivência.

Ambas as campanhas têm em comum o uso de comunicação clara, repetição de mensagens consistentes e apelo à responsabilidade coletiva. Esses elementos podem ser adaptados para o enfrentamento à desinformação. Portanto, ao invés de apenas agir após a circulação das notícias falsas, é possível tomar atitudes preventivas.

A proposta, portanto, é aproveitar as lições dessas campanhas e aplicá-las ao contexto digital. Por exemplo, o uso de influenciadores, como já sugerido, poderia replicar o papel das figuras públicas nas campanhas antifumo, funcionando como vozes confiáveis que engajam o público com mensagens claras sobre como identificar e evitar *fake news*. A comunicação visual e a repetição de conteúdos curtos e impactantes, como ocorreu na campanha antibuzina, também pode ser utilizada em peças gráficas e vídeos educativos nas redes sociais.

4.4.5 Educação básica como ferramenta de conscientização familiar

A formação crítica das crianças no que tange à avaliação de informações e ao reconhecimento de desinformações, como *fake news* e *deep fakes*, é uma estratégia

pedagógica de relevante impacto social. Ao serem instruídas desde a educação básica, essas crianças tendem a levar tais conhecimentos para o ambiente familiar, funcionando como vetores de transformação social. Essa influência é potencializada pelo fato de estarem em fase de desenvolvimento cognitivo e, portanto, mais receptivas a novos aprendizados, além de geralmente apresentarem maior familiaridade com as Tecnologias Digitais de Informação e Comunicação (TDIC).

Nesse contexto, promover o letramento digital, compreendido como a capacidade de utilizar, compreender e interagir criticamente com as tecnologias digitais nos mais diversos contextos, torna-se imperativo. Tal abordagem vai além do simples uso de ferramentas tecnológicas, exigindo uma postura reflexiva, ética e cidadã diante do conteúdo veiculado no meio digital. O artigo de Maxmiller Ismael enfatiza que esse processo educativo deve ser iniciado já na educação básica. Contudo, uma barreira significativa identificada é a capacitação docente: enquanto os alunos já se inserem naturalmente no universo digital, os profissionais da educação muitas vezes encontram dificuldades para acompanhar as inovações tecnológicas, o que compromete a eficácia do letramento digital (Ismael, 2020).

A Base Nacional Comum Curricular (BNCC) reconhece o letramento digital como uma competência essencial à formação do estudante brasileiro, integrando-se de forma complementar aos conteúdos tradicionais da grade curricular. Dentre as propostas pedagógicas, destaca-se a utilização de imagens, sons e códigos comuns ao ambiente virtual como recursos didáticos, bem como a inserção de atividades de interpretação voltadas à educação infantil, promovendo uma leitura crítica e contextualizada do material disponível online.

Ademais, as instituições escolares desempenham papel central no desenvolvimento da autonomia intelectual e do senso crítico dos estudantes. Tal constatação é reforçada no artigo apresentado anteriormente, que cita as palavras de Cláudio Maierovitch, médico sanitário da Fiocruz Brasília, ao afirmar que “O celular se tornou um apêndice, faz parte do corpo humano. É uma mudança cultural mais profunda do que já notamos. É a sensação de carregar a verdade na mão o tempo todo. É a verdade portátil e instantânea.” Essa realidade sociotécnica evidencia que o

acesso à internet tornou-se uma extensão da vida cotidiana. Por conseguinte, torna-se essencial que os indivíduos, desde a infância, estejam preparados para discernir entre informação e desinformação, garantindo, assim, não apenas a proteção contra os efeitos nocivos da manipulação digital, mas também a construção de uma sociedade mais crítica, consciente e juridicamente informada.

Embora os conceitos de educação midiática, literacia ou alfabetização midiática sejam relativamente recentes no âmbito das políticas educacionais brasileiras, a ideia de uma educação crítica, tal como defendida por Paulo Freire, já possui um histórico consolidado na formação pedagógica nacional. A proposta freiriana, centrada na emancipação do sujeito por meio da reflexão e do diálogo, permanece atual e fundamental, especialmente quando transportada para o contexto das tecnologias da informação e comunicação (TDICs). Valores como empatia, persistência, representação, empoderamento e consciência crítica não apenas coexistem com o uso das TDICs, mas devem ser por elas potencializados.

No entanto, a conjuntura atual impõe desafios significativos, visto que os jovens brasileiros apresentam, simultaneamente, altos índices de conectividade e uma crescente exposição à desinformação, além de demonstrarem baixa confiança nas instituições políticas. Tal cenário revela uma crise de credibilidade e enfraquecimento do pacto democrático, o que exige intervenções formativas eficazes. Desse modo, reafirma-se a necessidade de inserção da educação crítica midiática como instrumento fundamental na formação dos estudantes, corroborando a hipótese de que jovens submetidos a esse tipo de formação estarão mais bem preparados para resistir à manipulação informacional, reconhecer discursos falaciosos e exercer a cidadania de maneira mais plena e consciente (Bernardi, 2021).

Sob o prisma jurídico, a efetivação dessa proposta encontra respaldo no artigo 205 da Constituição Federal, que estabelece a educação como direito de todos e dever do Estado e da família, sendo promovida e incentivada com a colaboração da sociedade, visando ao pleno desenvolvimento da pessoa, seu preparo para o exercício da cidadania e sua qualificação para o trabalho. Assim, a educação crítica midiática deve ser compreendida como uma expressão concreta desse mandamento

constitucional, voltada à formação de cidadãos capazes de atuar de forma ética, autônoma e participativa na vida pública.

Portanto, diante da crescente complexidade das relações digitais e do papel central que as plataformas tecnológicas desempenham na construção da realidade contemporânea, é imprescindível que o ordenamento jurídico e as políticas públicas educacionais caminhem em sintonia, promovendo uma educação que transcenda os limites tradicionais da sala de aula e prepare os jovens não apenas para o mercado de trabalho, mas também para o exercício pleno de seus direitos e deveres no ambiente democrático e digital.

4.4.6 Campanha de conscientização sobre suicídio e desinformação

A última estratégia proposta para combater a propagação de *deep fakes* nas eleições é a restrição severa à divulgação detalhada de casos manipulados, com o objetivo de evitar o efeito de contágio, ou seja, a amplificação de conteúdos falsificados ao serem destacados sem o devido contexto. A exposição excessiva de *deep fakes* pode reforçar a normalização da desinformação, levando o público a tratar esses conteúdos como legítimos, mesmo após a detecção de sua falsidade. Portanto, uma abordagem de contenção de danos se torna crucial para preservar a veracidade das informações.

Em paralelo, a saúde pública reconhece e estuda há décadas o efeito contágio no comportamento suicida. A exposição a relatos de suicídio, particularmente quando sensacionalistas ou glorificadores, pode, sob certas circunstâncias e em indivíduos vulneráveis, aumentar o risco de atos suicidas subsequentes. Campanhas de prevenção ao suicídio bem-sucedidas frequentemente incorporam estratégias rigorosas para mitigar este efeito, focando em comunicação responsável, promoção de resiliência e fortalecimento de redes de apoio.

Esta proposta propõe uma análise comparativa dos mecanismos de propagação subjacentes ao contágio suicida e à disseminação de *deep fakes*, com o objetivo de identificar lições e estratégias de mitigação do primeiro que possam ser adaptadas para combater o segundo no cenário eleitoral.

O contágio, em sua essência, refere-se à propagação de uma ideia, comportamento ou emoção através de uma população por meio de exposição social. No contexto do suicídio, o "efeito Werther", ou seja, o efeito imitativo que surge da aprendizagem social, na qual uma pessoa vulnerável se identifica com outra pessoa que morreu por suicídio e emula seu comportamento, descreve o aumento de atos suicidas após a publicidade de um caso. Este fenômeno é influenciado por fatores psicológicos como identificação, imitação e normalização de um comportamento anteriormente impensável, bem como por dinâmicas sociais que incluem a influência de pares e a cobertura midiática.

A disseminação de *deep fakes* em períodos eleitorais compartilha certas características com o contágio suicida no que diz respeito à sua propagação. Ambas dependem da transmissão de conteúdo (informacional no caso dos *deep fakes*, comportamental/ideacional no suicídio) dentro de redes sociais e comunidades. Similarmente, a vulnerabilidade individual (predisposição psicológica no suicídio, vieses cognitivos e falta de literacia midiática na desinformação) e o ambiente (cobertura midiática no suicídio, plataformas *online* e algoritmos na desinformação) desempenham papéis cruciais na amplificação do efeito.

O "contágio informacional" de *deep fakes* explora a tendência humana a acreditar em informações que confirmam suas crenças preexistentes (viés de confirmação) e a compartilhar conteúdo que evoca fortes respostas emocionais, como indignação ou medo. A velocidade e o alcance das plataformas digitais exacerbam esse potencial de contágio, permitindo que *deep fakes* se tornem virais antes que sua falsidade seja estabelecida e amplamente comunicada.

A abordagem utilizada em campanhas de prevenção ao suicídio se caracteriza por sua sensibilidade, considerando a natureza emocional e delicada do tema. Da mesma forma, o fenômeno das *deep fakes* e das *fake news*, apesar de ser relativamente novo, também possui um impacto significativo sobre a sociedade, exigindo, portanto, estratégias cuidadosamente elaboradas para mitigar seus efeitos. As lições aprendidas com campanhas de prevenção ao suicídio podem ser aplicadas no combate às *deep fakes* e à desinformação, uma vez que ambas as problemáticas

envolvem questões de saúde mental, manipulação emocional e disseminação de informações falsas.

Uma das principais estratégias utilizadas nas campanhas de prevenção ao suicídio é evitar a sensacionalização ou glorificação do ato. Relatos detalhados ou que romantizam o suicídio são expressamente desaconselhados, pois podem gerar um efeito de contágio, incentivando pessoas vulneráveis a seguir tais exemplos. De maneira análoga, a disseminação de *deep fakes*, que muitas vezes envolve conteúdos manipuladores e sensacionalistas, deve ser combatida com uma abordagem que evite amplificar seu impacto emocional ou validar suas premissas falsas. O foco deve ser, portanto, na análise factual, desmistificando o conteúdo e desafiando a narrativa distorcida apresentada pelas *deep fakes*.

Nas campanhas contra o suicídio, o fornecimento de informações claras e precisas sobre a complexidade do problema é essencial. A promoção da busca por ajuda profissional e o esclarecimento sobre os fatores que influenciam o comportamento suicida são elementos-chave para a prevenção. Analogamente, no combate às *deep fakes*, é crucial fornecer contexto sobre a origem e a natureza manipulada do conteúdo. A educação sobre as técnicas de criação de *deep fakes* e a importância de uma abordagem crítica ao consumir informações digitais deve ser uma prioridade, capacitando o público a identificar conteúdos sintéticos e a questionar a veracidade das informações antes de propagá-las.

Outro objetivo central é fortalecer a resiliência individual e comunitária, promovendo habilidades de enfrentamento e fomentando conexões sociais que funcionem como fatores de proteção. De maneira análoga, no contexto da desinformação, é essencial promover a literacia midiática e o pensamento crítico, incentivando os indivíduos a desenvolver a capacidade de verificar informações em fontes confiáveis e a adotar uma postura mais crítica diante do consumo de conteúdo digital. O fortalecimento da resiliência social, através de campanhas educativas sobre como distinguir notícias falsas e detectar *deep fakes*, pode ser uma ferramenta poderosa na construção de uma sociedade mais consciente e menos suscetível à manipulação digital.

A promoção de recursos de saúde mental, como números de emergência e serviços de apoio psicológico, é uma medida essencial na prevenção do suicídio. Da mesma forma, para combater as *deep fakes* e a desinformação, é necessário fortalecer as plataformas e os mecanismos para a denúncia de conteúdos manipulados. A criação de espaços digitais seguros e acessíveis, nos quais os usuários possam reportar casos de *deep fakes*, é fundamental para a proteção contra a disseminação de desinformação. Além disso, a promoção de fontes de informação verificada e de jornalismo de qualidade pode atuar como um antídoto eficaz contra os efeitos prejudiciais das *fake news* e *deep fakes*.

A comunicação sobre o suicídio deve ser cuidadosamente adaptada para diferentes públicos, respeitando as nuances culturais e emocionais de cada grupo, a fim de evitar impactos negativos. No combate às *deep fakes*, essa abordagem também é essencial, pois diferentes grupos demográficos podem ter padrões distintos de consumo de mídia e estar mais ou menos suscetíveis à desinformação. É necessário adaptar as mensagens e os canais de comunicação para atingir de forma eficaz as pessoas mais vulneráveis à manipulação digital, considerando seus vieses e comportamentos nas redes sociais. A comunicação responsável pode contribuir para a criação de uma cultura de verificação, diminuindo a propagação de *deep fakes* e incentivando o consumo consciente e informado de conteúdo digital.

Embora existam paralelos conceituais, a adaptação direta de estratégias de prevenção ao suicídio para o combate a *deep fakes* apresenta desafios. A natureza do dano difere (dano à vida no suicídio vs. dano à percepção e ao processo democrático nos *deep fakes*), assim como as motivações por trás da disseminação (sofrimento individual no suicídio vs. intenção de manipular no caso dos *deep fakes* eleitorais). Além disso, o ritmo da disseminação online de *deep fakes* pode ser significativamente mais rápido do que o do contágio suicida tradicional, exigindo respostas ágeis e escaláveis. A identificação e a atribuição da origem dos *deep fakes* também representam um desafio técnico considerável.

A epidemia de desinformação impulsionada por tecnologias como os *deep fakes* exige abordagens inovadoras e multifacetadas. Ao examinar as lições aprendidas com

décadas de pesquisa e prática em campanhas de prevenção ao suicídio, particularmente no que se refere à mitigação do efeito contágio, podemos obter insights valiosos. A compreensão dos mecanismos psicológicos e sociais que facilitam a propagação de conteúdos nocivos, sejam eles comportamentos suicidas ou *deep fakes* enganosos, é fundamental para o desenvolvimento de estratégias de comunicação mais eficazes, a promoção da resiliência informacional e o fortalecimento da infraestrutura democrática contra a manipulação em períodos eleitorais. A adaptação cuidadosa e baseada em evidências das melhores práticas em prevenção ao contágio oferece um caminho promissor para enfrentar o desafio crescente da desinformação sintética.

In summa, este estudo ressalta que o combate à desinformação eleitoral não deve ser uma responsabilidade exclusiva do Estado ou das plataformas digitais, mas sim um esforço coletivo que envolva a sociedade civil, a imprensa, os tribunais eleitorais, os atores políticos, mas, principalmente, o governo. A combinação de medidas regulatórias, tecnológicas e educativas pode representar um passo significativo na construção de um ambiente eleitoral mais transparente e confiável, garantindo a soberania do voto e a proteção dos princípios democráticos.

Considerações finais

Esta pesquisa teve como objetivo analisar os impactos das *deep fakes* e das *fake news* no processo eleitoral brasileiro, com foco nas eleições de 2024, à luz dos desafios jurídicos, sociais e políticos impostos por tecnologias emergentes de manipulação digital. Inicialmente, partiu-se da distinção técnica e jurídica entre esses fenômenos, evidenciando como evoluíram e se consolidaram como instrumentos cada vez mais sofisticados e perigosos para a democracia. Enquanto as *fake news* já comprometem a confiança coletiva na informação, as *deep fakes* aprofundam essa crise ao manipularem a realidade com elevado grau de verossimilhança.

No plano normativo, foram analisados dispositivos como a Resolução nº 23.610/2019 do Tribunal Superior Eleitoral e o Projeto de Lei nº 2.630/2020, além de experiências legislativas internacionais, como as da Itália e da Nova Zelândia. Apesar de alguns avanços, o ordenamento jurídico brasileiro ainda apresenta lacunas significativas quanto à aplicação eficaz dessas normas e à responsabilização dos agentes envolvidos.

Nesse contexto, destaca-se a recente decisão do Supremo Tribunal Federal, no julgamento da Ação Direta de Inconstitucionalidade nº 4.037, que declarou parcialmente inconstitucional o artigo 19 do Marco Civil da Internet. A referida decisão representa um marco importante ao abrir caminho para a responsabilização direta das plataformas digitais por conteúdos ilícitos, incluindo *deep fakes* e *fake news*, reforçando a necessidade de uma atuação mais firme do Poder Judiciário em defesa da integridade do processo democrático.

Com base na análise realizada, conclui-se que o enfrentamento desses desafios exige uma abordagem coordenada, multifacetada e interinstitucional. Embora medidas legislativas e tecnológicas sejam indispensáveis, sua efetividade depende diretamente da integração com ações de educação midiática, campanhas públicas de conscientização e maior engajamento social.

Entre as propostas discutidas, destacam-se a criação de uma agência reguladora, o fortalecimento de parcerias com agências de checagem, o incentivo ao uso de inteligência artificial e o aprimoramento de ferramentas colaborativas como as Notas da Comunidade. Tais medidas devem ser acompanhadas por investimentos

contínuos em letramento digital e pensamento crítico, com especial atenção à formação desde a educação básica.

Também se propôs a mobilização de influenciadores digitais como agentes de conscientização, a unificação de emissoras de rádio e televisão para veiculação de alertas em horários de pico e a replicação de estratégias exitosas de campanhas públicas anteriores, como as antifumo, antibuzina e de prevenção ao suicídio. Essas experiências revelaram a importância da comunicação clara, da repetição de mensagens consistentes e do apelo à responsabilidade coletiva, elementos essenciais no combate à desinformação.

Nesse cenário, o papel do Estado não deve se restringir à regulamentação, mas ampliar-se como promotor ativo da cidadania digital, garantindo políticas públicas educativas, incentivos à verificação de fatos e meios eficazes de denúncia. O combate à desinformação eleitoral, portanto, não pode ser tarefa exclusiva de instituições formais ou de plataformas digitais, mas um compromisso coletivo com a integridade democrática que envolva governos, sociedade civil, imprensa, educadores e o sistema de Justiça Eleitoral.

A proposta inicial desta pesquisa previa uma comparação mais profunda com ordenamentos jurídicos estrangeiros. Contudo, ao longo da investigação, observou-se que o Brasil já dispõe de iniciativas relevantes, embora ainda fragmentadas e de baixa efetividade. Por essa razão, optou-se por sistematizar tais medidas com vistas a potencializar sua aplicação prática e oferecer subsídios para o desenvolvimento de políticas públicas mais integradas e eficazes.

Conclui-se, portanto, que o enfrentamento das *deep fakes* e das *fake news* demanda uma ação proativa e articulada entre o Direito, a tecnologia e a educação. Somente com essa integração será possível construir um ecossistema informacional mais seguro, preservar a liberdade de expressão de forma responsável e proteger a autenticidade do voto como expressão genuína da vontade popular.

Por fim, é imprescindível reconhecer que os desafios não se encerram com este diagnóstico. Tecnologias como as *real-time deep fakes*, capazes de manipular imagens e sons ao vivo, já despontam como ameaças ainda mais sofisticadas. Diante disso, impõe-se ao Legislativo a necessidade de deixar de atuar de forma meramente reativa, assumindo postura proativa e preventiva, com respostas normativas à altura da

evolução tecnológica, cabe à eles acompanhar esse avanço com escuta ativa às necessidades da população, compromisso real com a preservação da democracia e segurança para a sociedade brasileira.

REFERÊNCIAS

- ABJAUDE, Samir Antonio Rodrigues et al. How do social media influence mental health?. **SMAD, Rev. Eletrônica Saúde Mental Álcool Drog.**, v. 16, n. 1, p. 1-3, 2020. DOI 10.11606/issn.1806-6976.smad.2020.0089. Disponível em: <https://revistas.usp.br/smad/article/view/166931/159489>. Acesso em: 12 ago. 2025.
- ABREU, Arthur Emanuel Leal; ADEODATO, João Maurício Leitão. Complexidades na conceituação jurídica de fake news. **Revista Em Tempo**, [S.l.], v. 19, n. 1, aug. 2020. DOI 10.26729/et.v19i1.3109. Disponível em: <https://revista.univem.edu.br/emtempo/article/view/3109>. Acesso em: 10 ago. 2025.
- AGÊNCIA ESTADO. 'Fake news' se espalham 70% mais rápido que notícias verdadeiras, diz MIT. **Correio Braziliense**, Brasília, 8 mar. 2018. Disponível em: https://www.correiobraziliense.com.br/app/noticia/tecnologia/2018/03/08/interna_tecnologia,664835/fake-news-se-espalham-70-mais-rapido-que-noticias-verdadeiras.shtml. Acesso em: 12 ago. 2025.
- ALTA lucratividade é o que mantém o mercado digital de fake news. **Jornal da USP**, São Paulo, 5 dez. 2022. Disponível em: <https://jornal.usp.br/radio-usp/alta-lucratividade-e-o-que-mantem-o-mercado-digital-de-fake-news/>. Acesso em: 12 ago. 2025.
- BAHIA. Tribunal Regional Eleitoral. **Grupos de WhatsApp são principais vetores para golpes na internet via fake news**. Salvador, 6 jun. 2018. Atualizado em: 12 ago. 2022. Disponível em: <https://www.tre-ba.jus.br/comunicacao/noticias/2018/Junho/grupos-de-whatsapp-sao-principais-vetores-para-golpes-na-internet-via-fake-news>. Acesso em: 12 ago. 2025.
- BALDISSERA, Wellington Antonio; FORTES, Vinícius Borges. Regulação das fake news: um dilema diante do direito à liberdade de expressão. **Direito e Desenvolvimento**, [S.l.], v. 12, n. 1, p. 18-36, 2021. DOI 10.26843/direitoedesenvolvimento.v12i1.957. Disponível em: <https://periodicos.unipe.edu.br/index.php/direitoedesenvolvimento/article/view/957>. Acesso em: 10 ago. 2025.
- BARNES, Curtis; BARRACLOUGH, Tom. **Perception inception: preparing for deepfakes and the synthetic media of tomorrow**. Auckland: Brainbox; Law Foundation of New Zealand, 2019. Disponível em: https://www.lawfoundation.org.nz/wp-content/uploads/2019/05/2018_ILP_5_RESEARCH-REPORT_Perception-Inception-Report-EMBARGOED-TILL-21-May-2019.pdf. Acesso em: 10 ago. 2025.
- BERNARDI, Ana Julia Bonzanini. **Educação crítica midiática: formação para cidadania de jovens no contexto de pós-verdade e fake news**. 2021. Tese (Doutorado em Ciência Política) – Instituto de Filosofia e Ciências Humanas, Universidade Federal do Rio

Grande do Sul, Porto Alegre, 2021. Disponível em:
<https://lume.ufrgs.br/handle/10183/234761>. Acesso em: 12 ago. 2025.

BRASHIER, N. M.; ELISEEV, E. D.; MARSH, E. J. An initial accuracy focus prevents illusory truth. **Cognition**, v. 194, p. 104054, 2020. Disponível em:
<https://pubmed.ncbi.nlm.nih.gov/31473395/>. Acesso em: 10 ago. 2025

BRASIL. Procuradoria-Geral da República. **Deepfake e inteligência artificial: saiba o que pode e o que é proibido nas campanhas eleitorais**. Brasília, 20 jun. 2024. Disponível em:
<https://www.mpf.mp.br/pgr/noticias-pgr2/2024/deepfake-e-inteligencia-artificial-saiba-o-que-pode-e-o-que-e-proibido-nas-campanhas-eleitorais>. Acesso em: 4 ago. 2025.

BRASIL. **Projeto de Lei nº 2630, de 2020**. Dispõe sobre a liberdade, responsabilidade e transparência na internet, institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet e dá outras providências. Senado Federal. Disponível em:
<https://www25.senado.leg.br/web/atividade/materias/-/materia/141944>. Acesso em: 10 ago. 2025.

BRASIL. Secretaria de Comunicação Social. **Governo Federal lança campanha Brasil contra Fake e reforça luta contra a desinformação**. 25 mar. 2023. Disponível em:
<https://www.gov.br/secom/pt-br/assuntos/noticias/2023/03/governo-federal-lanca-campanha-brasil-contrafake-e-reforca-luta-contraa-desinformacao>. Acesso em: 12 ago. 2025.

BRASIL. Supremo Tribunal Federal. **ADI 4037**. Decisão sobre o art. 19 do Marco Civil da Internet. Brasília, 26 jun. 2025. Disponível em:
<https://noticias.stf.jus.br/postsnoticias/stf-define-parametros-para-responsabilizacao-de-plataformas-por-conteudos-de-terceiros/>. Acesso em: 10 ago. 2025.

BRASIL. Tribunal Superior Eleitoral. **Eleições pelo mundo: quantos e quais países já foram às urnas em 2024?** 19 jul. 2024. Disponível em:
<https://www.tse.jus.br/comunicacao/noticias/2024/Julho/eleicoes-pelo-mundo-quantos-e-quais-paises-ja-foram-as-urnas-em-2024>. Acesso em: 13 ago. 2025.

BRASIL. Tribunal Superior Eleitoral. **Pílulas contra a desinformação: ganhar dinheiro com fake news é crime**. Brasília, 8 jun. 2022. Disponível em:
<https://www.tse.jus.br/comunicacao/noticias/2022/Junho/pilulas-contraa-desinformacao-ganhar-dinheiro-com-fake-news-e-crime>. Acesso em: 12 ago. 2025.

BRASIL. Tribunal Superior Eleitoral. **Resolução nº 23.610, de 18 de dezembro de 2019**. Dispõe sobre propaganda eleitoral, utilização e geração do horário gratuito e condutas ilícitas em campanha eleitoral. Disponível em:
<https://www.tse.jus.br/legislacao/compilada/res/2019/resolucao-no-23-610-de-18-de-dezembro-de-2019>. Acesso em: 12 ago. 2025.

CARVALHO, Gustavo Arthur Coelho Lobo de; KANFFER, Gustavo Guilherme Bezerra. O Tratamento Jurídico das Notícias Falsas (fake news). **Migalhas**, 2018. Disponível em: https://www.migalhas.com.br/arquivos/2020/6/61CC2C822E99EA_tratamento-juridico-noticias-f.pdf. Acesso em: 10 ago. 2025.

CHESNEY, Robert; CITRON, Danielle. Deep fakes: a looming crisis for national security, democracy and privacy? **Lawfare**, fev. 2018. Disponível em: <https://www.lawfaremedia.org/article/deepfakes-looming-crisis-national-security-democracy-and-privacy>. Acesso em: 10 ago. 2025.

COMITÊ GESTOR DA INTERNET NO BRASIL (CGI.br). **TIC Educação 2015**: pesquisa sobre o uso das tecnologias de informação e comunicação nas escolas brasileiras. São Paulo: Comitê Gestor da Internet no Brasil, 2016. Disponível em: https://www.cetic.br/media/docs/publicacoes/2/TIC_Edu_2015_LIVRO_ELETRONICO.pdf. Acesso em: 10 ago. 2025.

DISTRITO FEDERAL (Brasil). Secretaria de Estado de Saúde. **Tabagismo**. Secretaria de Saúde do Distrito Federal, 29 nov. 2023. Disponível em: <https://www.saude.df.gov.br/tabagismo>. Acesso em: 10 ago 2025.

EGELHOFER, Jana Laura; LECHERER, Sophie. Fake news as a two-dimensional phenomenon: a framework and research agenda. **Annals of the International Communication Association**, v. 43, n. 2, p. 97-116, 2019. DOI 10.1080/23808985.2019.1602782.

FOFANO, Jorge. É #FAKE vídeo em que Lula confessa roubo de recursos públicos: trecho mostra só parte de fala do presidente em evento no início de 2023 e muda sentido original. **G1**, 27 fev. 2025. Disponível em: <https://g1.globo.com/fato-ou-fake/noticia/2025/02/27/e-fake-video-em-que-lula-confessa-roubo-de-recursos-publicos.ghtml>. Acesso em: 14 ago. 2025.

GAMA, Sophia. Guerra de desinformação: as fake news nas eleições de 2018. **Câmara Municipal de Curitiba**, Curitiba, 15 jul. 2022. Disponível em: <https://www.curitiba.pr.leg.br/informacao/noticias/guerra-de-desinformacao-as-fake-news-nas-eleicoes-de-2018>. Acesso em: 10 ago. 2025.

GOOGLE. **Combater a desinformação online**: proteger a democracia com informação de qualidade. 2024. Disponível em: <https://portugal.googleblog.com/2024/03/combater-desinformacao-online-proteger.html>. Acesso em: 10 ago. 2025.

GUIMARÃES, Glayder Daywerth Pereira; SILVA, Michael César. Fake news à luz da responsabilidade civil digital: o surgimento de um novo dano social. **R. Jur. FA7**, Fortaleza, v. 16, n. 2, p. 99-114, jul./dez. 2019. Disponível em: <https://periodicos.uni7.edu.br/index.php/revistajuridica/article/view/940/775>. Acesso em: 10 ago. 2025.

HYNARD, James. Single-frame forensic watermarking for content authenticity. **castLabs**, 2 sep. 2024. Disponível em:
<https://castlabs.com/news/single-frame-watermarking-content-authenticity/>

INSTITUTO NACIONAL DE ALFABETIZAÇÃO FUNCIONAL (INAF). **INAF Brasil**: Alfabetização Funcional para Todos. 2025. Disponível em:
<https://alfabetismofuncional.org.br/>. Acesso em: 10 ago. 2025.

ISMAEL, Maxmiller Lima Lorangeira. **Letramento digital no combate à propagação das fake news no WhatsApp**. 2020. Artigo (Especialização em Estratégias Didáticas para Educação Básica, com uso das TIC) – Centro de Educação, Universidade Federal de Alagoas, Maceió, 2020. Disponível em:
<https://www.repositorio.ufal.br/bitstream/riufal/7491/1/Letramento%20digital%20no%20combate%20%C3%A0%20propaga%C3%A7%C3%A3o%20das%20fake%20news%20no%20whatsapp.pdf>. Acesso em: 10 ago. 2025.

ITALIA. **Codice Penale**: Regio Decreto 19 ottobre 1930, n. 1398 – Approvazione del testo definitivo del Codice Penale. Gazzetta Ufficiale della Repubblica Italiana, n. 251, 26 ott. 1930. Disponível em:
<https://www.gazzettaufficiale.it/sommario/codici/codicePenale>. Acesso em: 5 ago. 2025.

ITALIA. Ministero dell’Istruzione e del Merito. **Scuola, Boldrini e Fedeli presentano decalogo anti-bufale**: il progetto riguarderà 4,2 milioni di ragazzi. 31 ott. 2017. Disponível em:
<https://www.mim.gov.it/-/scuola-boldrini-e-fedeli-presentano-decalogo-anti-bufale-il-progetto-riguardera-4-2-milioni-di-ragazzi>. Acesso em: 12 ago. 2025.

MCNAIR, Brian. **Fake news**: falsehood, fabrication and fantasy in journalism. New York: Routledge, 2017. Disponível em:
https://www.researchgate.net/publication/323413083_Fake_News_Falsehood_Fabrication_and_Fantasy_in_Journalism. Acesso em: 10 ago. 2025.

MEIRELES, Gabriela. Postagem distorce declaração de Lula para parecer que ele confessa ter comprado votos em 2022. **Estadão**, São Paulo, 5 ago. 2024. Disponível em:
<https://www.estadao.com.br/estadao-verifica/lula-conta-como-comprou-todos-eleicao-2022/>. Acesso em: 12 ago. 2025.

MUTCHNIK, Letícia. É montagem vídeo que mostra pesquisa do Ipec no JN com Bolsonaro à frente. **UOL**, São Paulo, 19 ago. 2022. Disponível em:
<https://noticias.uol.com.br/confere/ultimas-noticias/2022/08/19/e-montagem-video-que-mostra-pesquisa-do-ipecc-no-jn-com-bolsonaro-a-frente.htm>. Acesso em: 10 ago. 2025.

NATURAL Language Processing (NLP): What is it and how is it used? **Hyperscience**, 2025. Disponível em:

<https://www.hyperscience.ai/resource/natural-language-processing/>. Acesso em: 12 ago. 2025.

NOVA ZELÂNDIA. **Harmful Digital Communications (HDC)**. Police New Zealand. 2025. Disponível em: <https://www.police.govt.nz/advice-services/cybercrime-and-internet/harmful-digital-communications-hdc>. Acesso em: 5 ago. 2025.

OGALE, Sanika; DESHPANDE, Rutuja. **The role of social media in the spread of misinformation and fake news**. 2024. Research Paper (Degree in Social Media Marketing, Digital Marketing) – Massachusetts Institute of Technology, World Peace University, Cambridge, 2024. Disponível em: https://www.researchgate.net/publication/379872133_The_Role_of_Social_Media_in_the_Spread_of_Misinformation_and_Fake_News. Acesso em: 10 ago. 2025

ÖHMAN, Carl. Who are the targets of deepfakes? evidence from flagged videos on TikTok and YouTube. **Journal of Digital Social Research**, v. 7, n. 2, p. 42-67, 2025. DOI 10.53420/jdsr.v7i2.55529. Disponível em: https://www.researchgate.net/publication/392682276_Who_Are_the_Targets_of_Deepfakes_Evidence_From_Flagged_Videos_on_TikTok_and_YouTube. Acesso em: 10 ago. 2025.

PARISER, Eli. **The filter bubble**. New York: The Penguin Press, 2011.

PATARANUTAPORN, Pat et al. If an AI-generated character misportrays or causes harm, who is to blame? **Nature Machine Intelligence**, v. 3, p. 1013-1022, 2021. Disponível em: <https://www.google.com/url?q=https://www.nature.com/articles/s42256-021-00417-9%23Sec18&sa=D&source=docs&ust=1737458672314383&usg=AOvVaw0s93weuCHz-uCtW59N1fzb>. Acesso em: 5 ago. 2025.

PENNYCOOK, G.; CANNON, T. D.; RAND, D. G. Prior exposure increases perceived accuracy of fake news. **Journal of Experimental Psychology: General**, v. 147, n. 12, p. 1865-1880, 2018. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC6279465/>. Acesso em: 10 ago. 2025

PENNYCOOK, Gordon; BEAR, Adam; RAND, David G. The (minimal) persuasive advantage of political video over text. **Proceedings of the National Academy of Sciences**, v. 118, n. 47, e2114388118, 2021. Disponível em: <https://www.pnas.org/doi/full/10.1073/pnas.2114388118>. Acesso em: 10 ago. 2025.

PORCELLO, Flávio A. C. A relevância do telejornalismo nos 70 anos da TV: como enfrentar as mentiras e fake news que abalam a democracia e ameaçam a liberdade de expressão no Brasil. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 43. 2020. **Anais [...]**. Disponível em: <https://lume.ufrgs.br/handle/10183/254625>. Acesso em: 12 ago. 2025.

PRÖLLOCHS, N. Community-Based Fact-Checking on Twitter's Birdwatch Platform. In: INTERNATIONAL AAAI CONFERENCE ON WEB AND SOCIAL MEDIA, [S. l.], **Proceedings** [...], v. 16, n. 1, p. 794-805, 2022. DOI 10.1609/icwsm.v16i1.19335. Disponível em: <https://ojs.aaai.org/index.php/ICWSM/article/view/19335>. Acesso em: 4 ago. 2025.

QUEIROGA, Louise. É #FAKE que PT distribuiu mamadeiras eróticas para crianças em creches pelo país. **G1**, 28 out. 2021. Disponível em: <https://g1.globo.com/fato-ou-fake/noticia/2021/10/28/e-fake-que-pt-distribuiu-mamadeiras-eroticas-para-criancas-em-creches-pelo-pais.ghtml>. Acesso em: 29 jun. 2025.

RECUERO, R. D. C.; ZAGO, G. D. S.; SOARES, F. B. Mídia social e filtros-bolha nas conversações políticas no Twitter. In: ENCONTRO ANUAL COMPÓS, 26., 2017, São Paulo, SP. **Anais** [...] São Paulo: Faculdade Cásper Líbero, 2017. Disponível em: <https://lume.ufrgs.br/handle/10183/166193>. Acesso em: 12 ago. 2025

ROBL FILHO, Ilton Norberto; MARRAFON, Marco Aurélio; MEDÓN, Filipe. 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**, Brasília, v. 13, n. 3, p. 32-47, out./dez. 2022. Disponível em: https://www.researchgate.net/publication/371905673_A_inteligencia_artificial_utilizada_para_desinformar_ou_a_servico_da_desinformacao_como_as_deepfakes_e_as_redes_automatizadas_abalam_o_mercado_livre_de_ideias_e_a_democracia_constitucional_e_deliberat. Acesso em: 12 ago. 2025.

SANTOS, Leona. Vídeo que mostra urna eletrônica 'preenchendo' voto em Haddad é falso: imagem foi impulsionada por Flávio Bolsonaro, candidato ao Senado pelo RJ e filho do presidente Jair Bolsonaro (PSL). **Veja**, 17 out. 2018. Disponível em: <https://veja.abril.com.br/coluna/me-engana-que-eu-posto/video-que-mostra-urna-preenchendo-voto-em-haddad-e-falso>. Acesso em: 16 jul. 2025.

SCOTT, Kate. "Deceptive" clickbait headlines: relevance, intentions, and lies. **Journal of Pragmatics**, v. 218, p. 71-82, 2023. Disponível em: <https://doi.org/10.1016/j.pragma.2023.10.004>. Acesso em: 12 ago. 2025.

SILVA JÚNIOR, Alvacir Siqueira da; CARVALHO, Célia Maria da Silva. A "onda" das fake news e a sua abordagem pelos diversos meios de comunicação. **Conexões: Revista de Relações Públicas e Comunicação Organizacional**, v. 2, n. 4, jul./dez. 2019.

SIMÕES, Mariana. Grupos pró-Bolsonaro no WhatsApp orquestram fake news e ataques pessoais na internet, diz pesquisa. **A Pública**, 22 out. 2018. Disponível em: <https://apublica.org/2018/10/grupos-pro-bolsonaro-no-whatsapp-orquestram-fake-news-e-ataques-pessoais-na-internet-diz-pesquisa/>. Acesso em: 12 ago. 2025.

SUN TZU. **A arte da guerra**. São Paulo: Martins Fontes, 2003.

THOMPSON, Stuart A. Elon Musk AI deepfake scam. **The New York Times**, 14 ago. 2024. Disponível em: <https://www.nytimes.com/interactive/2024/08/14/technology/elon-musk-ai-deepfake-scam.html>. Acesso em: 4 ago. 2025.

TORABI ASR, Fatemeh. The language gives it away: how an algorithm can help us detect fake news. **PBS NewsHour**, 25 ago. 2019. Disponível em: <https://www.pbs.org/newshour/science/the-language-gives-it-away-how-an-algorithm-can-help-us-detect-fake-news>. Acesso em: 4 ago. 2025.

VACCARI, Cristian; CHADWICK, Andrew. Deepfakes and disinformation: exploring the impact of synthetic political video on deception, uncertainty, and trust in news. **Social Media + Society**, v. 6, n. 1, 2020. DOI 10.1177/2056305120903408. Disponível em: https://www.researchgate.net/publication/339385335_Deepfakes_and_Disinformation_Exploring_the_Impact_of_Synthetic_Political_Video_on_Deception_Uncertainty_and_Trust_in_News. Acesso em: 12 ago. 2025.

VIEIRA, Ezequiel. Fake news: descentralização das informações e polarização política. **Observatório da Imprensa**, 9 abr. 2019. Disponível em: <https://www.observatoriodaimprensa.com.br/desinformacao/fake-news-descentralizacao-das-informacoes-e-polarizacao-politica/>. Acesso em: 12 ago. 2025.

VOSOUGHI, Soroush; ROY, Deb; ARAL, Sinan. The spread of true and false news online. **Science**, v. 359, n. 6380, p. 1146-1151, 2018. Disponível em: <https://news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308>. Acesso em: 21 jan. 2025.

WESTERLUND, Mika. The emergence of deepfake technology: a review. **Technology Innovation Management Review**, [S.l.], v. 9, n. 11, p. 3-14, nov. 2019. Disponível em: https://timreview.ca/sites/default/files/article_PDF/TIMReview_November2019%20-%20D%20-%20Final.pdf. Acesso em: 4 ago. 2025.

WHATSAPP diz que banuiu 400 mil contas na eleição de 2018 por disparos em massa. **Poder360**, 19 nov. 2019. Disponível em: <https://www.poder360.com.br/eleicoes/whatsapp-diz-que-baniu-400-mil-contas-na-eleicao-de-2018-por-disparos-em-massa/>. Acesso em: 12 ago. 2025.

WHATSAPP. **About end-to-end encryption**. WhatsApp Help Center, 2025. Disponível em: <https://faq.whatsapp.com/820124435853543/>. Acesso em: 16 jul. 2025.

WOOLLEY, S. C. Digital propaganda: the power of influencers. **Journal of Democracy**, v. 33, n. 3, p. 115-129, 2022. Disponível em: <https://dx.doi.org/10.1353/jod.2022.0027>. Acesso em: 16 maio 2025.

X. **Community notes**: sobre as notas da comunidade. 2025. Disponível em: <https://communitynotes.x.com/guide/pt/about/introduction>. Acesso em: 10 ago. 2022