www.revistaft.com.br

Revista FT | ISSN 1678-0817 | v. 30, n. 158, 2026 | DOI: 10.69849/j83tba67

A junção entre as deepfakes e deep web: como anonimidade e a infraestrutura da deep web podem potencializar a criação, disseminação e o uso malicioso das deepfakes.

The junction between deepfakes and the deep web: how anonymity and the deep web infrastructure can enhance the creation, dissemination and malicious use of deepfakes.

Ketllen Medly Carvalho Fontinele[1]
Orientadora: Profa. Rosana Reis de Melo Silva[2]

RESUMO

O presente artigo científico analisa os impactos jurídicos e os desafios práticos decorrentes da convergência entre a arquitetura de ocultação da *Deep Web* e a tecnologia de Inteligência Artificial generativa das *deepfakes*. Diante da popularização de mídias sintéticas altamente realistas e do manto de invisibilidade fornecido pelo protocolo de roteamento em camadas (*Onion Routing*), investiga-se como essa sinergia potencializa a prática de crimes cibernéticos e viola os direitos da personalidade, como a imagem, a honra e a privacidade. Sob a égide do método dedutivo e de uma pesquisa bibliográfica, documental e jurisprudencial, o estudo demonstra que a opacidade estrutural das redes anônimas gera um cenário de transvulnerabilidade para o usuário e mitiga a eficácia de diplomas normativos tradicionais, como o Marco Civil da Internet e o Código Penal, devido aos óbices de rastreabilidade de autoria e determinação de competência jurisdicional. Por fim, conclui-se que o enfrentamento

[1] Graduanda Ketllen Medly Carvalho Fontinele, do curso de Bacharelado em Direito, 10º período na Instituição de Ensino Superior Fametro (FAMETRO), Manaus, Amazonas - Brasil. E-mail: ketllenfontenele@gmail.com, ORCID: 0009-0001-8180-2237.

[2] Prof.ª Orientadora e Coordenadora do TCC II, no Centro Universitário FAMETRO: Prof.ª Esp. Rosana Reis de Melo Silva. Manaus, Amazonas, Brasil. E-mail: rosanareismello@gmail.com

dessa ameaça transfronteiriça exige não apenas a atualização dogmática e o fortalecimento da persecução penal através das diretrizes da Convenção de Budapeste, mas também a implementação de salvaguardas tecnológicas na origem das mídias e o investimento estatal em políticas públicas de educação e literacia digital.

Palavras-chave: Deepfakes. Deep Web. Direitos da Personalidade. Direito Digital. Transvulnerabilidade.

ABSTRACT

This paper analyzes the legal impacts and practical challenges arising from the convergence between the concealment architecture of the Deep Web and the generative Artificial Intelligence technology of deepfakes. In light of the popularization of highly realistic synthetic media and the cloak of invisibility provided by the layered routing protocol (Onion Routing), it investigates how this synergy enhances cybercrimes and violates personality rights, such as image, honor, and privacy. Under the deductive method and through bibliographical, documentary, and jurisprudential research, the study demonstrates that the structural opacity of anonymous networks creates a scenario of digital transvulnerability for users. Furthermore, it mitigates the effectiveness of traditional regulatory frameworks, such as the Brazilian Civil Rights Framework for the Internet and the Penal Code, due to obstacles in traceability and jurisdictional competence. Finally, the study concludes that tackling this cross-border threat requires not only dogmatic updates and the strengthening of criminal prosecution under the Budapest Convention guidelines, but also the implementation of technological safeguards in media origin and state investment in public policies for digital literacy.

Keywords: Deepfakes. Deep Web. Personality Rights. Digital Law. Transvulnerability.

1. INTRODUÇÃO

A virada tecnológica experimentada pela sociedade contemporânea na era da informação reconfigurou as fronteiras da comunicação, da segurança pública e da própria eficácia das normas jurídicas. Se, por um lado, o desenvolvimento de sistemas baseados em Inteligência Artificial (IA) generativa impulsionou a inovação e a automação de processos, por outro, inaugurou vetores inéditos e complexos de criminalidade cibernética. Entre essas inovações, destaca-se o fenômeno das *deepfakes* mídias sintéticas que utilizam algoritmos de

aprendizado profundo para replicar a imagem, a voz e as expressões humanas com um grau de verossimilhança quase indistinguível da realidade.

Paralelamente ao refinamento dessas ferramentas de manipulação, o ecossistema digital testemunha a consolidação da *Deep Web*, em especial da sua subcamada mais restrita, a *Dark Web*. Projetada originalmente para salvaguardar o anonimato e a privacidade legítima de dados institucionais, essa infraestrutura de roteamento em camadas e pulverização de rastros digitais passou a ser ativamente instrumentalizada por agentes maliciosos. A convergência entre a perfeição técnica das *deepfakes* e o manto de invisibilidade fornecido pela internet profunda potencializa de forma geométrica a criação, o armazenamento e a disseminação de fraudes, extorsões, campanhas de desinformação em massa e pornografia não consensual.

Diante desse panorama de opacidade técnica, delineia-se o problema central desta pesquisa: **De que maneira a anonimidade e a infraestrutura da Deep Web potencializam o uso malicioso das deepfakes e quais são os principais desafios enfrentados pelo ordenamento jurídico brasileiro para tutelar os direitos da personalidade violados por essa convergência?** A hipótese que se levanta sugere que a arquitetura descentralizada da rede mitiga a eficácia de mecanismos tradicionais de persecução penal e de responsabilização civil, colocando o cidadão comum em um estado de transvulnerabilidade digital.

A justificativa para a realização deste estudo reside na premente necessidade de analisar os limites dos diplomas normativos vigentes, tais como o Marco Civil da Internet (Lei nº 12.965/2014) e a Lei Geral de Proteção de Dados (Lei nº 13.709/2018), frente a uma ameaça que opera de forma transfronteiriça e descentralizada. Compreender esse fenômeno sob a ótica do Direito Digital e Penal é fundamental para propor soluções que conciliem o avanço tecnológico com a preservação da dignidade humana, da honra e da imagem do indivíduo.

Para cumprir tais objetivos, o presente artigo adota o método de abordagem dedutivo, amparado por uma pesquisa eminentemente bibliográfica e documental, com análise doutrinária e jurisprudencial. Estruturalmente, o trabalho divide-se em quatro seções de desenvolvimento, além desta introdução e das considerações finais. A primeira seção esquadrinha a infraestrutura técnica e as camadas da *Deep Web*. A segunda aborda o funcionamento das redes generativas adversariais (GANs) na criação das *deepfakes*. A terceira seção analisa os impactos jurídicos e os reflexos penais e civis no ordenamento brasileiro. Por fim, a quarta seção delineia recomendações, soluções técnicas e políticas públicas voltadas ao enfrentamento dessa ameaça cibernética.

2. A INFRAESTRUTURA DA DEEP WEB: CAMADAS, PROTOCOLOS E O ECOSSISTEMA DO ANONIMATO

2.1. A distinção metodológica entre surface web, deep web e dark web

Compreender a dinâmica dos crimes cibernéticos modernos, especialmente aqueles que envolvem a disseminação de conteúdos ilícitos de alta complexidade tecnológica, exige uma análise prévia e pormenorizada da própria estrutura e da divisão arquitetônica da rede mundial de computadores. A doutrina especializada costuma valer-se da clássica metáfora do iceberg para estratificar a internet em três dimensões perfeitamente distintas: a *Surface Web* (internet de superfície), a *Deep Web* (internet profunda) e a *Dark Web* (internet obscura), evidenciando que a maior parte dos dados trafegados permanece oculta do público geral (LONGHI; FALEIROS JÚNIOR; BORGES, 2020).

A *Surface Web* corresponde à camada visível, superficial e facilmente acessível do ambiente virtual, congregando todos os sítios eletrônicos, portais de notícias, plataformas de comércio eletrônico e redes sociais que são ativamente rastreados e indexados por motores de busca convencionais, tais como Google, Yahoo e Bing. Trata-se de uma porção estritamente superficial que, embora concentre o maior fluxo de usuários quotidianos e indexe bilhões de páginas públicas, representa apenas uma fração minoritária do volume total de dados trafegados globalmente na rede mundial.

Imediatamente abaixo dessa linha de superfície, encontra-se a *Deep Web*, a qual, diferente do que o senso comum e os meios de comunicação de massa frequentemente propagam, não consiste em um espaço intrinsecamente voltado à ilegalidade ou à marginalidade digital. Ao contrário, a internet profunda constitui uma infraestrutura vital e legítima para a manutenção da segurança, do sigilo e da privacidade das instituições contemporâneas, sendo definida tecnicamente como a porção da rede cujos conteúdos e páginas não são indexados pelos motores de busca tradicionais por razões de segurança e controle de acesso.

Nesta camada oculta residem bancos de dados corporativos, sistemas de *internet banking*, prontuários médicos hospitalares, arquivos acadêmicos restritos, intranets governamentais e caixas de correio eletrônico privado. O acesso a essas informações sensíveis é condicionado a barreiras de segurança rigorosas, tais como credenciais de autenticação, *firewalls*, redes privadas virtuais (VPNs) ou protocolos de criptografia específicos, visando resguardar o sigilo de dados institucionais e pessoais contra acessos não autorizados, em estrita observância aos ditames de governança digital e proteção de dados.

2.2. A dark web e o protocolo onion routing (tor)

O gargalo técnico onde os desafios à persecução penal se acentuam localiza-se em uma subcamada ainda mais restrita da internet profunda: a *Dark Web*. Esta dimensão exige a utilização de softwares, configurações e protocolos específicos para que a conexão seja estabelecida. É nesse ambiente que o ecossistema do anonimato é potencializado por redes sobrepostas (*overlay networks*), sendo a rede Tor (*The Onion Router*) a infraestrutura mais proeminente e utilizada para este fim.

O funcionamento da rede Tor baseia-se no princípio do roteamento em cebola (*Onion Routing*). Diferente de uma conexão convencional na internet de superfície, onde o pacote de dados do usuário viaja diretamente ao servidor de destino expondo o endereço de Protocolo de Internet (IP) de origem, o protocolo Tor fragmenta e criptografa a informação em múltiplas camadas de segurança. O tráfego é redirecionado de maneira aleatória por meio de uma série de nós (*relays*) operados por voluntários ao redor do globo.

Cada nó intermediário é capaz de descriptografar apenas a camada externa necessária para identificar a próxima etapa do circuito, desconhecendo por completo o conteúdo integral da mensagem e a identidade do emissor original. Somente o nó de saída (*exit node*) remove a última camada de criptografia e entrega o pacote ao destino final. Esse processo de triangulação e cifragem sucessiva resulta na pulverização dos rastros digitais, inviabilizando que provedores de conexão, plataformas ou órgãos de fiscalização estatal identifiquem a geolocalização ou a real identidade civil do internauta.

2.3. Os desafios jurisdicionais e a transvulnerabilidade do consumidor digital

A diluição dos rastros de conexão promovida pela arquitetura da *Dark Web* não engendra apenas um óbice de natureza estritamente técnica, mas projeta reflexos substanciais na eficácia do ordenamento jurídico pátrio. O primeiro grande desafio gravita em torno da determinação da competência jurisdicional e da aplicação da lei penal e civil no espaço. Dado que os pacotes de dados trafegam por múltiplos nós hospedados em diferentes países, a identificação da real localização do servidor que hospeda a aplicação ilícita torna-se, não raro, inviável pelos métodos tradicionais de cooperação jurídica internacional.

Sob a perspectiva do Direito do Consumidor e do Direito Digital, essa opacidade estrutural culmina no que a doutrina contemporânea qualifica como transvulnerabilidade. O cidadão que acessa ou é atraído para essas camadas profundas seja em busca de facilidades econômicas, seja

por desconhecimento dos limites técnicos da rede submete-se a um ambiente desprovido de salvaguardas institucionais básicas. Na *Dark Web*, os pilares da segurança da informação são subvertidos: a assimetria informativa entre o provedor do serviço anônimo e o usuário comum é absoluta.

Consequentemente, a eficácia de diplomas protetivos, como o Código de Defesa do Consumidor (CDC) e o Marco Civil da Internet (Lei nº 12.965/2014), resta severamente mitigada. Mecanismos cogentes de responsabilização civil, obrigações de transparência e o próprio dever de mitigar riscos de segurança tornam-se inócuos diante de agentes que operam sob o manto do anonimato tecnológico.

Assim, a infraestrutura que nasceu para resguardar a privacidade legítima contra regimes autoritários passa a ser instrumentalizada como um escudo de impunidade, pavimentando o caminho para a proliferação de fraudes e para a manipulação e disseminação de conteúdos digitais ilícitos de última geração.

3. A TECNOLOGIA DAS DEEPFAKES: CONCEITO, EVOLUÇÃO E REDES GENERATIVAS (GANs)

3.1. Conceito e evolução histórica da manipulação audiovisual

O avanço das tecnologias de informação e comunicação não apenas reformulou as interações sociais, mas também transformou radicalmente os mecanismos de produção e edição de conteúdos digitais. Nesse cenário, o termo *deepfake* resultante da aglutinação dos vocábulos *deep learning* (aprendizado profundo) e *fake* (falso) designa a técnica de síntese de imagens ou sons humanos baseada em Inteligência Artificial (IA) para criar mídias altamente realistas, capazes de simular a voz, as expressões faciais e os movimentos de indivíduos reais proferindo discursos ou realizando atos que, na realidade, jamais existiram.

Historicamente, a manipulação de registros audiovisuais não constitui um fenômeno estritamente contemporâneo; técnicas analógicas de montagem fotográfica e efeitos visuais cinematográficos são utilizadas desde o século XIX. Contudo, a virada de paradigma promovida pelas *deepfakes* reside na desnecessidade de intervenção humana minuciosa de um editor de vídeo avançado. O processo foi democratizado e automatizado por algoritmos computacionais que aprendem de forma autônoma a replicar padrões biométricos.

O fenômeno ganhou projeção global a partir do final de 2017, quando softwares de código aberto foram disseminados em fóruns de discussão na internet, permitindo que usuários

leigos fizessem a substituição de rostos (*face-swapping*) com extrema facilidade. O que inicialmente se manifestou como uma ferramenta de entretenimento ou sátira rapidamente transmutou-se em um vetor de severas violações a direitos da personalidade, uma vez que a tecnologia passou a ser direcionada para a criação de pornografia não consensual, falsificação de evidências e campanhas de desinformação em massa.

3.2. O funcionamento técnico das redes generativas adversariais (gans)

O núcleo tecnológico de inteligência artificial que confere às *deepfakes* modernas um grau de verossimilhança quase absoluto e frequentemente indistinguíveis a olho nu fundamenta-se em um revolucionário modelo de arquitetura de aprendizado de máquina (*machine learning*) introduzido no ano de 2014: às Redes Generativas Adversariais (*Generative Adversarial Networks* – GANs). Conforme sintetizado por Goodfellow et al. (2014, p. 2672), o modelo propõe um "framework onde dois modelos treinam simultaneamente: um modelo gerador, que captura a distribuição dos dados, e um modelo discriminador, que estima a probabilidade de uma amostra vir dos dados de treinamento".

O papel desempenhado pelo Gerador assemelha-se metricamente ao comportamento de um falsificador de obras de arte altamente especializado; o seu objetivo principal e exclusivo é criar amostras de dados sintéticos (como o mapeamento tridimensional e dinâmico do rosto de uma pessoa específica) a partir de um banco de imagens reais de treinamento, tentando aproximar-se ao máximo da realidade factual para enganar o seu oponente sistémico. Por sua vez, o Discriminador atua na função análoga à de um perito ou investigador criminal forense, cuja missão institucional é analisar minuciosamente o conteúdo produzido pelo Gerador e determinar, com base em probabilidades estatísticas, se aquela mídia específica é autêntica (proveniente do banco de dados real) ou artificial (sintetizada pelo algoritmo).

Essa disputa algorítmica ocorre em um ciclo contínuo, automatizado e autorregulado de milhares ou milhões de repetições sequenciais, conhecido no ecossistema da ciência de dados como processo de treinamento adversarial. À medida que o Discriminador aponta as falhas, ruídos e imperfeições da imagem ou áudio gerados, o Gerador absorve essa informação como um vetor de retroalimentação (*feedback loop*), ajustando imediatamente os seus parâmetros matemáticos internos para produzir uma nova versão refinada na iteração seguinte.

O processo repete-se sucessivamente até que o Gerador atinja tamanha sofisticação e precisão biométrica que o algoritmo do Discriminador se torna estatisticamente incapaz de diferenciar o dado real do dado sintetizado, encerrando o ciclo com uma falsificação perfeita.

Sob a ótica estrita do Direito Digital e da Ciência Forense, a eficácia desse autoaperfeiçoamento algorítmico engendra uma preocupante assimetria operacional, visto que a velocidade com que as ferramentas de inteligência artificial aprendem a forjar fraudes e suprimir vestígios digitais supera significativamente o desenvolvimento de métodos de perícia e *softwares* de detecção forense tradicionais. Essa realidade mitiga a eficácia do artigo 158 do Código de Processo Penal brasileiro, uma vez que a materialidade do crime cibernético se torna volátil e camuflada por uma perfeição matemática que desafia os limites da prova pericial contemporânea.

4. O USO MALICIOSO E OS IMPACTOS JURÍDICOS NO ORDENAMENTO BRASILEIRO

4.1. Reflexos nos direitos da personalidade: imagem, honra e privacidade

A disseminação e o uso malicioso de *deepfakes*, potencializados pelo anonimato estrutural das camadas profundas da internet, não representam apenas um desafio de policiamento técnico ou de engenharia de *software*, mas configuram uma agressão direta e violenta aos direitos da personalidade. Estes direitos encontram-se expressamente tutelados pelo artigo 5º, inciso X, da Constituição Federal de 1988, e pelos artigos 11 a 21 do Código Civil Brasileiro, constituindo cláusulas pétreas de proteção da dignidade humana. A tecnologia de substituição facial baseada em inteligência artificial e a síntese de voz de alta fidelidade atingem o núcleo imaterial da identidade do sujeito ao instrumentalizar e corporificar a sua identidade visual e fonética sem o devido consentimento.

No âmbito do direito à imagem, a violação opera-se em uma dupla dimensão jurídica indissociável: a imagem-atributo, que diz respeito à repercussão social, à boa fama e à reputação do sujeito perante a coletividade em que vive, e a imagem-retrato, que abarca a expressão física, os traços fisionômicos e as características biométricas intrínsecas do ser humano. Quando uma *deepfake* é dolosamente empregada para simular a participação de uma pessoa em atos degradantes, discursos espúrios ou em conteúdos pornográficos não consensuais fenômeno crescentemente denominado na literatura digital como *deepnude*, ocorre o esfacelamento imediato e devastador da sua honra em suas vertentes subjetiva e objetiva.

A infalibilidade aparente e o hiper-realismo da mídia sintética geram no espectador comum uma presunção automática de veracidade que torna qualquer desmentida jurídica ou nota de esclarecimento subsequente em grande parte ineficaz no imaginário social. Como leciona Silva (2023), o impacto na dignidade humana nesses cenários consolida um dano moral

permanente, devastador e de difícil reparação na esfera psíquica, familiar e profissional da vítima, exigindo do ordenamento uma postura que reconheça a gravidade da manipulação existencial sofrida pelo indivíduo.

4.2. Tipicidade penal e os desafios de enquadramento no código penal

Diante da ausência de um tipo penal específico, autônomo e abrangente no ordenamento jurídico brasileiro que criminalize genericamente a criação e a difusão de *deepfakes* maliciosas, o operador do Direito Penal é compelido a realizar um complexo e rigoroso esforço de subsunção interpretativa aos modelos proibitivos já existentes. Esse processo, contudo, esbarra frequentemente nas limitações intransponíveis decorrentes do princípio constitucional da legalidade estrita e da taxatividade penal, previstos no artigo 5º, inciso XXXIX, da Carta Magna, os quais vedam categoricamente a utilização da analogia *in malam partem* para punir condutas não expressamente tipificadas pelo legislador.

Tradicionalmente, as condutas que envolvem a difamação da reputação alheia, a criação de falsas declarações ou a imputação espúria de crimes por meio de mídias sintéticas hiper-realistas encontram abrigo nos clássicos crimes contra a honra tutelados pelo Estatuto Repressor, a saber: a Calúnia (art. 138), a Difamação (art. 139) e a Injúria (art. 140 do Código Penal). Quando a tecnologia é direcionada para forjar a nudez ou simular a participação de alguém em atos de conotação sexual com o intuito de constrangimento, humilhação ou vingança, a conduta é atraída pela moldura do artigo 218-C do Diploma Penal, introduzido pela Lei nº 13.718/2018, que pune criminalmente a divulgação, por qualquer meio, de cena de sexo, nudez ou pornografia sem autorização da vítima, incluindo de forma expressa em seu texto legal as montagens e os cortes audiovisuais.

Todavia, o grande entrave dogmático e prático na esfera criminal não reside unicamente na elasticidade dos tipos penais, mas sim na determinação inequívoca da autoria material e no isolamento do nexo de causalidade. A infraestrutura de ocultação e o roteamento em camadas que caracterizam a *Dark Web* impedem de forma quase absoluta que os órgãos de persecução penal, como a Polícia Civil e o Ministério Público, identifiquem o autor intelectual ou o desenvolvedor do código algorítmico que originou a fraude. Sem a identificação do endereço de IP real de origem, a autoria resta pulverizada no ciberespaço, impossibilitando a deflagração da ação penal por falta de justa causa.

Em decorrência dessa barreira tecnológica intransponível, o Estado acaba por conseguir responsabilizar apenas os replicadores secundários, isto é, os usuários comuns da internet de superfície que compartilham o vídeo falso em redes sociais, enquanto os verdadeiros mentores e operadores das plataformas criminosas profundas permanecem sob o manto da impunidade tecnológica. Essa desconexão entre a gravidade do fato e a capacidade de punição do Estado mitiga severamente a eficácia preventiva geral da sanção penal, demonstrando que as ferramentas tradicionais de investigação criminal são anacrônicas perante a volatilidade e o anonimato das redes sobrepostas.

4.3. A responsabilização civil e os limites normativos do marco civil da internet

No plano da responsabilidade civil extracontratual no ambiente virtual, o ecossistema das *deepfakes* tensiona severamente os limites normativos e os critérios de imputação estabelecidos pela Lei nº 12.965/2014, amplamente conhecida como o Marco Civil da Internet. O artigo 19 do referido diploma legal condiciona a responsabilidade civil subjetiva dos provedores de aplicações por danos decorrentes de conteúdos gerados por terceiros à prévia ordem judicial específica de indisponibilização, visando, em sua gênese, salvaguardar a liberdade de expressão e impedir a censura privada na rede de computadores.

Contudo, esse modelo regulatório clássico, estruturado para a internet de superfície (*Surface Web*), demonstra-se anacrônico, ineficaz e eivado de limitações práticas quando confrontado com a velocidade de replicação geométrica de uma mídia sintética manipulada. Sobre essa fragilidade regulatória, Patrícia Peck Pinheiro adverte que:

> A velocidade com que a informação circula na sociedade digital e o impacto gerado por conteúdos falsos ou manipulados demandam uma tutela jurídica muito mais ágil e preventiva, uma vez que os mecanismos tradicionais de responsabilização *a posteriori* muitas vezes não conseguem reverter o dano reputacional já consolidado. (PINHEIRO, 2021, p. 142).

A natureza descentralizada das redes sobrepostas impede que as ordens judiciais de remoção alcancem os verdadeiros hospedeiros do conteúdo ilícito, restando à vítima o ônus de tentar conter a propagação na internet de superfície, onde o dano à imagem já se consolidou de forma difusa. Essa assimetria regulatória faz com que o instituto da responsabilidade civil perca sua função precipuamente dissuasória e preventiva, transformando-se em um mecanismo meramente reativo e incapaz de estancar a progressão do dano existencial, o que impõe a

necessidade de se rediscutir os deveres de cuidado e a responsabilidade social das plataformas de compressão e compartilhamento de mídias digitais.

4.4. A proteção de dados biométricos sensíveis e a responsabilidade face à LGPD

Diante das lacunas deixadas pelo Marco Civil da Internet, a Lei Geral de Proteção de Dados Pessoais (Lei nº 13.709/2018 - LGPD) surge no cenário nacional como um potente microssistema de contenção e responsabilização civil e administrativa. Sob a ótica estrita do artigo 5º, inciso II, da LGPD, a imagem do rosto humano, as expressões faciais e os padrões de modulação fonética (voz) de um indivíduo não constituem meras informações acessórias, mas qualificam-se juridicamente como dados pessoais biométricos e, por conseguinte, dados sensíveis.

Essa categorização legal exige que qualquer operação de tratamento conceito amplo que abarca desde a coleta, o armazenamento, o compartilhamento e, fundamentalmente, a manipulação ou cruzamento de dados para o treinamento computacional de redes neurais generativas (GANs) esteja estritamente vinculada a uma base legal lícita, perfeitamente adequada às hipóteses taxativas do artigo 11 da referida lei, destacando-se o consentimento explícito, destacado e inequívoco do titular.

A criação de uma *deepfake* maliciosa sem autorização representa, portanto, um desvio total, absoluto e ilícito da finalidade do tratamento originário, violando frontalmente os princípios da segurança, da transparência, da não discriminação e da boa-fé objetiva insculpidos no artigo 6º da LGPD. Ademais, a responsabilidade civil regulada pela LGPD assume contornos de maior rigor ao estabelecer a responsabilidade solidária entre os agentes de tratamento (controladores e operadores) que descumprirem as obrigações de segurança da informação, permitindo incidentes como o vazamento de dados ou a raspagem não autorizada de dados (*data scraping*) em redes sociais.

Assim, a *accountability* jurídica deve ser direcionada não apenas contra o indivíduo que difunde a mídia sintética na internet, mas também contra as corporações tecnológicas e bancos de dados que negligenciam os deveres de custódia e proteção dos ativos biométricos da sociedade civil, permitindo que as identidades dos cidadãos sejam capturadas e instrumentalizadas como matéria-prima para fraudes algorítmicas no ciberespaço profundo.

5. RECOMENDAÇÕES E SOLUÇÕES NO ENFRENTAMENTO À AMEAÇA

5.1. Propostas de adaptação legislativa e políticas públicas

A superação dos vazios normativos que circundam o uso malicioso de *deepfakes* no ordenamento jurídico brasileiro exige uma postura legislativa ativa e responsiva. Embora o Código Penal ofereça respostas setoriais através dos crimes contra a honra, constata-se a premência da tipificação de condutas específicas que envolvam a criação e a difusão não autorizada de mídias sintéticas destinadas a induzir a erro ou causar prejuízos a terceiros. Projetos de lei em tramitação no Congresso Nacional buscam suprir essa lacuna ao prever causas de aumento de pena quando os crimes cibernéticos são perpetrados mediante o emprego de inteligência artificial generativa.

No âmbito das políticas públicas, a adesão formal do Brasil à Convenção de Budapeste sobre o Crime Cibernético (Decreto nº 11.129/2022) representa um marco institucional crucial. Contudo, torna-se imperioso converter as diretrizes internacionais em ações domésticas eficazes. Recomenda-se a estruturação e o aporte de recursos financeiros e tecnológicos perenes para as delegacias especializadas em repressão a crimes cibernéticos (DERCC) em nível estadual, capacitando as forças de segurança pública para atuarem na infiltração virtual em redes anônimas da *Dark Web* e na identificação de fluxos de criptoativos associados a organizações criminosas especializadas em extorsão e fraudes biométricas.

5.2. Desenvolvimento de ferramentas de detecção e combate (contribuição técnica)

O enfrentamento às ameaças cibernéticas não se exaure na coerção estatal-normativa; demanda, de igual modo, a mobilização de contramedidas tecnológicas. Uma vez que as redes GANs operam sob a lógica da evolução algorítmica contínua, o desenvolvimento de ferramentas de engenharia forense digital torna-se indispensável. Softwares de auditoria capazes de identificar anomalias imperceptíveis ao olho humano como inconsistências na taxa de piscada de olhos, padrões de reflexo de luz na retina ou variações espectrais na modulação fonética devem ser integrados à rotina das perícias criminais judiciais.

Ademais, desponta como solução arquitetural a implementação de protocolos de integridade de mídia na internet de superfície, tais como a criptografia de metadados e o uso de marcas d'água digitais (*watermarking*) indeléveis na origem do conteúdo. A criação de um ecossistema de "procedência da informação" permitiria que plataformas de redes sociais e

navegadores de internet identificassem automaticamente se um arquivo audiovisual sofreu modificações sintéticas substanciais antes de ser compartilhado, mitigando a progressão geométrica da desinformação originada nas camadas profundas da rede.

5.3. A importância da educação digital e da mídia-competência

Por fim, a última linha de defesa contra o uso nocivo da inteligência artificial reside no fortalecimento cognitivo do próprio usuário da rede. A vulnerabilidade digital, exaustivamente demonstrada ao longo deste estudo, é agravada pelo analfabetismo funcional digital, que impossibilita a identificação de conteúdo manipulados ou e-mails de *phishing* que servem de porta de entrada para golpes complexos.

Desenvolver a mídia-competência na sociedade civil significa qualificar o cidadão para exercer uma postura crítica perante o fluxo informacional. Políticas educacionais voltadas à literacia digital devem ser inseridas transversalmente nos currículos escolares e em campanhas de conscientização pública promovidas pelo Estado em parceria com a Autoridade Nacional de Proteção de Dados (ANPD).

Ao capacitar o indivíduo para compreender os riscos inerentes à exposição de seus dados biométricos e os mecanismos de funcionamento das fraudes em ambientes virtuais, obstaculiza-se a eficácia das engenhas sociais maliciosas, esvaziando o mercado de exploração ilícita que se alimenta do anonimato da *Dark Web*.

6. CONSIDERAÇÕES FINAIS

O presente estudo propôs-se a analisar os impactos jurídicos e os desafios práticos decorrentes da convergência entre a arquitetura de ocultação da *Deep Web* e a tecnologia de inteligência artificial generativa das *Deepfakes*. Ao longo da pesquisa, restou demonstrado que essa junção não opera apenas uma mudança quantitativa no volume de fraudes virtuais, mas uma alteração qualitativa na gravidade das lesões infligidas aos direitos da personalidade, em especial à imagem, à honra e à privacidade dos indivíduos.

No primeiro plano, a investigação sobre a infraestrutura da internet profunda evidenciou que os protocolos de segurança e o roteamento em camadas (como o sistema *Onion Routing* da rede Tor), embora legítimos em sua gênese protetiva, são ativamente instrumentalizados como escudos de impunidade transfronteiriça. Essa opacidade técnica gera um cenário de transvulnerabilidade para o usuário e impõe severas limitações à atividade persecutória do

Estado, uma vez que os critérios tradicionais de competência territorial e identificação de autoria civil e penal tornam-se inócuos diante do anonimato algorítmico.

No segundo plano, a análise do funcionamento das Redes Generativas Adversariais (GANs) revelou uma preocupante assimetria tecnológica. O autoaperfeiçoamento contínuo entre os sistemas geradores e discriminadores permite a criação de mídias sintéticas cujo realismo é praticamente indistinguível por métodos convencionais de observação, superando a velocidade de desenvolvimento dos mecanismos de detecção forense. Quando essas ferramentas são direcionadas para fins maliciosos nas subcamadas da rede, os danos aos direitos imateriais consolidam-se de forma permanente (*in re ipsa*), haja vista a ineficácia das retificações judiciais *a posteriori*.

Diante desse panorama, conclui-se que o ordenamento jurídico pátrio, malgrado o avanço representado pelo Marco Civil da Internet e pela Lei Geral de Proteção de Dados (LGPD), carece de atualizações normativas e instrumentais específicas. A tipificação penal pontual das condutas, o fortalecimento das delegacias especializadas por meio das diretrizes da Convenção de Budapeste e o fomento à infiltração policial virtual são passos urgentes na esfera repressiva.

Contudo, a resposta definitiva à ameaça não reside exclusivamente na sanção jurídica, mas sim na adoção de salvaguardas técnicas na origem das mídias (como marcas d'água indeléveis) e no investimento em políticas públicas de educação e literacia digital, capazes de reduzir a vulnerabilidade da sociedade civil na era da pós-verdade algorítmica.

7. REFERÊNCIAS

ALVES, Bruno Morais, et al. **Análise da responsabilização criminal dos criadores e propagadores de "deep fakes" no ordenamento jurídico brasileiro**. São Paulo: Editora Seven, 2024.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil**. Brasília, DF:
Presidente da República, 1988.

BRASIL. **Decreto-Lei nº 2.848, de 7 de dezembro de 1940**. Código Penal. Diário Oficial da União, Rio de Janeiro, 31 dez. 1940.

BRASIL. **Decreto nº 11.129, de 7 de julho de 2022**. Promulga a Convenção sobre o Crime Cibernético, adotada em Budapeste. Diário Oficial da União, Brasília, DF, 8 jul. 2022.

BRASIL. **Lei nº 12.965, de 23 de abril de 2014**. Estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil (Marco Civil da Internet). Diário Oficial da União, Brasília, DF, 24 abr. 2014.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Diário Oficial da União, Brasília, DF, 15 ago. 2018.

BRASIL. Superior Tribunal de Justiça. **Recurso Especial nº 1.774.956/SP**. Relator: Ministro Marco Aurélio Bellizze. Terceira Turma, julgado em 11 dez. 2018. Diário da Justiça Eletrônico, 17 dez. 2018.

GOODFELLOW, Ian et al. Generative Adversarial Nets. **Advances in Neural Information Processing Systems**, v. 27, p. 2672-2680, 2014.

LONGHI, João Victor Rozatti; FALEIROS JÚNIOR, José Luiz de Moura; BORGES, Gabriel Oliveira de Aguiar; REIS, Guilherme. **Fundamentos do Direito Digital**. Indaiatuba: Editora Foco, 2020.

PINHEIRO, Patrícia Peck. **Direito Digital**. 7. ed. São Paulo: Saraiva Educação, 2021.

RUBIM, B.; LUCIA, Karen. **As comunicações advindas do ciberespaço da Deepweb: uma análise sistêmica entre o risco e o direito no Brasil**. Londrina: Editora Thoth, 2024.

SILVA, Américo Luís Martins da. **O dano moral e a sua reparação civil**. 3. ed. Rio de Janeiro: Lumen Juris, 2023.

.