DEEPFAKES E RESPONSABILIDADE PENAL: NOVOS DESAFIOS PARA O DIREITO PENAL DIGITAL

DEEPFAKES AND CRIMINAL LIABILITY: NEW CHALLENGES FOR DIGITAL CRIMINAL LAW

DEEPFAKES Y RESPONSABILIDAD PENAL: NUEVOS DESAFÍOS PARA EL DERECHO PENAL DIGITAL

Antônio Ferreira do Norte Filho[1], Maria Elvira Ferreira Vieira[2], Isaac Saymon de Souza Rodrigues[3]

DOI: 10.54899/dcs.v22i83.3563
Recibido: 13/10/2025 | Aceptado: 14/10/2025 | Publicación en línea: 27/10/2025.

RESUMO

O artigo analisa os desafios que a tecnologia das deepfakes impõe ao Direito Penal Digital contemporâneo. As deepfakes, produzidas por meio de algoritmos de deep learning, permitem a criação de imagens, áudios e vídeos falsos altamente realistas, com potencial para causar graves danos à honra, à imagem, à privacidade, à fé pública e à segurança da informação. A pesquisa discute a ausência de tipificação penal específica no ordenamento jurídico brasileiro, ressaltando a insuficiência dos tipos tradicionais diante da complexidade técnica e da autoria mediada por sistemas autônomos de inteligência artificial. Examina-se a necessidade de reinterpretação de conceitos clássicos, como dolo, culpa e autoria, bem como a ampliação da proteção penal dos bens jurídicos ameaçados pelas falsificações digitais. O estudo conclui pela urgência de atualização normativa e teórica do Direito Penal, pautada nos princípios da legalidade, proporcionalidade e intervenção mínima, a fim de equilibrar repressão, liberdade e inovação tecnológica no campo do Direito Penal brasileiro.

Palavras-chave: Deepfakes. Direito Penal Digital. Responsabilidade Penal. Inteligência Artificial. Bens Jurídicos.

ABSTRACT

The article analyzes the challenges posed by deepfake technology to contemporary Digital Criminal Law. Deepfakes, generated through deep learning algorithms, allow the creation of highly realistic fake images, audios, and videos capable of causing severe harm to legal goods such as honor, image, privacy, public trust, and information security. The research discusses the

[1] Doutor em Ciências do Ambiente e Sustentabilidade na Amazônia pela Universidade Federal do Amazonas (UFAM), Faculdade Santa Teresa, Manaus, Amazonas, Brasil. E-mail: nortefilho@gmail.com
Orcid: https://orcid.org/0000-0002-5946-3291
[2] Graduanda em Direito, Faculdade Santa Teresa, Manaus, Amazonas, Brasil. E-mail: mariaelvirafv16@gmail.com
Orcid: https://orcid.org/0009-0006-9283-3481
[3] Graduando em Direito, Faculdade Santa Teresa, Manaus, Amazonas, Brasil. E-mail: isaacsaymon11@gmail.com
Orcid: https://orcid.org/0009-0001-4268-6616

absence of specific criminal provisions in Brazilian law, emphasizing the insufficiency of traditional criminal types to address the technical complexity and algorithmic autonomy of artificial intelligence systems. It examines the need to reinterpret classical criminal concepts, such as intent, culpability, and authorship, and to expand the protection of threatened legal interests. The study concludes that the modernization of criminal law is urgent, guided by the principles of legality, proportionality, and minimal intervention, to balance repression, freedom, and technological innovation in the context of Brazilian Criminal Law.

Keywords: Deepfakes. Digital Criminal Law. Criminal Liability. Artificial Intelligence. Legal Heritage.

RESUMEN

El artículo analiza los desafíos que la tecnología de las deepfakes impone al Derecho Penal Digital contemporáneo. Las deepfakes, generadas mediante algoritmos de deep learning, permiten crear imágenes, audios y videos falsos de gran realismo, con potencial para causar graves daños a bienes jurídicos como el honor, la imagen, la privacidad, la fe pública y la seguridad de la información. La investigación aborda la falta de tipificación penal específica en el ordenamiento jurídico brasileño, destacando la insuficiencia de los tipos tradicionales frente a la complejidad técnica y la autoría mediada por sistemas autónomos de inteligencia artificial. Se examina la necesidad de reinterpretar conceptos clásicos como dolo, culpa y autoría, así como de ampliar la protección penal de los bienes jurídicos amenazados por las falsificaciones digitales. El estudio concluye que es urgente una actualización normativa y teórica del Derecho Penal, guiada por los principios de legalidad, proporcionalidad e intervención mínima, para equilibrar represión, libertad e innovación tecnológica en el contexto del Derecho Penal Brasileño.

Palabras clave: Deepfakes. Derecho Penal Digital. Responsabilidad Penal. Inteligencia Artificial. Bienes Jurídicos.

INTRODUÇÃO

A inteligência artificial (IA) e o deep learning transformaram radicalmente a forma como a sociedade produz, compartilha e consome informação. O que antes era restrito a ambientes de pesquisa hoje permeia o cotidiano digital, moldando comportamentos, decisões políticas e relações sociais. Entre as inovações, destaca-se a tecnologia das deepfakes, fusão de deep learning e fake, que utiliza algoritmos de redes neurais para reproduzir, com impressionante fidelidade, rostos, vozes e expressões humanas.

Essas falsificações sintéticas permitem criar vídeos, áudios e imagens de pessoas em situações que jamais ocorreram. Isso abre um vasto campo para aplicações legítimas, como na indústria cinematográfica, em acessibilidade e em reconstruções históricas e, perigosamente, para usos criminosos e antiéticos. Sua difusão está provocando graves consequências sociais, especialmente quando empregadas na disseminação de fake news, fraudes digitais, pornografia não consensual, manipulação de provas judiciais e campanhas de desinformação política.

Essas novas formas de manipulação da realidade digital desafiam a estrutura tradicional do Direito Penal, concebida para lidar com condutas humanas tangíveis, não com criações autônomas geradas por sistemas de IA.

A manipulação algorítmica de imagens e sons altera a própria percepção da verdade. O avanço tecnológico cria um "ambiente jurídico de constante adaptação", onde a ausência de respostas normativas pode ampliar a sensação de impunidade e fragilizar a proteção de bens jurídicos fundamentais (Pinheiro, 2021).

No território brasileiro, o ordenamento jurídico ainda não dispõe de tipificação penal específica para a criação e o uso ilícito de deepfakes. As condutas são enquadradas em dispositivos já existentes, como os crimes contra a honra, a falsidade ideológica e documental ou as normas previstas na Lei nº 12.737/2012, Lei Carolina Dieckmann e no Marco Civil da Internet, Lei nº 12.965/2014. Todavia, esses instrumentos mostram-se insuficientes diante da complexidade técnica e da difusão em massa proporcionada pelas plataformas digitais.

São necessárias propostas teóricas que reconheçam o poder normativo dos códigos de computadores gestados dentro do segredo da iniciativa privada. Essas teorias devem ser capazes de propor soluções para garantir que aquele primeiro vácuo (relativo à ausência de poder legitimado democraticamente no meio virtual) seja preenchido por normas constitucionais, protetoras dos Direitos Humanos (Neto; Morais, 2018).

A dificuldade em identificar a autoria, a fronteira entre dolo humano e automação algorítmica e a volatilidade das provas digitais revelam um vazio normativo que exige reflexão doutrinária e atualização legislativa.

A discussão sobre responsabilidade penal em ambientes mediados por Inteligência Artificial (IA) exige a redefinição de conceitos clássicos, como ação, culpa e autoria. A imputação penal pressupõe a capacidade de autodeterminação consciente e voluntária do agente, um requisito que se torna problemático quando parte da ação é executada por um sistema autônomo (Bitencourt, 2022).

Surge, assim, a questão central deste trabalho: como atribuir responsabilidade penal por condutas mediadas ou executadas por inteligência artificial, especialmente no caso das deepfakes?

O uso indevido dessas tecnologias ameaça diversos bens jurídicos tutelados como a honra, a imagem, a privacidade, a fé pública e a segurança da informação, uma vez que a produção e divulgação de deepfakes pode gerar danos irreparáveis à reputação de indivíduos e à credibilidade das instituições democráticas.

A manipulação de discursos de autoridades, por exemplo, tem potencial de abalar a confiança pública e desestabilizar processos eleitorais. Essa dimensão política e social amplia o papel do Direito Penal como instrumento de proteção não apenas individual, mas também da integridade informacional da sociedade digital.

Diante desse cenário, o presente artigo tem como objetivo analisar os desafios que as tecnologias de deepfake impõem ao Direito Penal Digital, discutindo os limites da responsabilidade penal diante da autonomia tecnológica e da ausência de previsão normativa específica, buscando compreender como o Direito Penal brasileiro pode reagir a essas novas formas de criminalidade informacional, sem comprometer o equilíbrio entre repressão e liberdade de expressão, entre proteção de bens jurídicos e incentivo à inovação tecnológica.

Assim, o artigo examinará as lacunas legislativas, os possíveis enquadramentos jurídicos e as implicações éticas e penais do uso de deepfakes, propondo caminhos interpretativos e legislativos para a consolidação de um Direito Penal Digital contemporâneo, capaz de responder às transformações tecnológicas que redefinem a própria noção de verdade e realidade no mundo jurídico.

A TECNOLOGIA DAS DEEPFAKES: ORIGEM, FUNCIONAMENTO E RISCOS SOCIAIS

O termo deepfake é a fusão de deep learning (aprendizado profundo) e fake (falso). Ele designa conteúdos audiovisuais manipulados digitalmente por redes neurais artificiais que replicam padrões visuais e sonoros com altíssimo realismo.

A inteligência artificial generativa, enquanto uma das expressões mais sofisticadas, comporta sistemas de aprendizado profundo, treinados com vastas bases de dados de rostos, vozes e expressões humanas para a geração de imagens ou vídeos sintéticos que imitam pessoas reais.

Tecnicamente o sistema das deepfakes opera com dois algoritmos: o gerador, que cria o conteúdo falso, e o discriminador, que tenta detectar essa falsificação e essa interação constante aperfeiçoa o gerador, permitindo que ele produza vídeos quase indistinguíveis da realidade, o que confere às deepfakes sua aparência autêntica e dificulta sua detecção, mesmo por especialistas.

Originalmente, essa tecnologia visava fins artísticos e educacionais, como a reconstrução digital de atores, a dublagem automatizada e o aprimoramento de softwares de acessibilidade. No entanto, a popularização das ferramentas de código aberto transformou as deepfakes em instrumentos de potencial delituoso.

A partir de 2017, vídeos falsos envolvendo celebridades proliferaram, dando origem à pornografia de vingança sintética, surgindo posteriormente, usos ainda mais complexos, como a criação de discursos falsos de líderes políticos e a falsificação de provas em investigações judiciais.

No campo jurídico, a manipulação por IA desafia a noção tradicional de autoria e dolo, o criador do conteúdo pode não ser o responsável por sua difusão e vice-versa.

Além disso, a automação de parte do processo de criação onde a máquina aprende a falsificar de forma autônoma levanta dúvidas sobre a imputação subjetiva da conduta, ponto central da responsabilização penal.

Assim, as deepfakes inauguram uma zona cinzenta entre a ação humana consciente e a produção algorítmica, exigindo a reinterpretação dos elementos do tipo penal e dos fundamentos da culpabilidade e o seu estudo das deepfakes transcende à tecnologia, alcançando uma dimensão ética e jurídica ao colocar em xeque a confiança social na imagem e no som como representações da verdade.

Se o Direito historicamente se apoiou em provas visuais e documentais para estabelecer a veracidade dos fatos, o surgimento das falsificações digitais obriga a repensar o próprio conceito de prova e de veracidade no ambiente virtual sendo essa reflexão a base para a compreensão dos desafios do enquadramento penal das condutas, tema que será aprofundado na próxima seção.

A TIPIFICAÇÃO PENAL DAS DEEPFAKES NO CONTEXTO DO DIREITO PENAL DIGITAL

O avanço tecnológico transformou profundamente as relações sociais, criando formas de interação e, consequentemente, novas modalidades de infração penal, bem como a emergência

do Direito Penal Digital resulta da necessidade de adaptar o sistema jurídico às condutas ilícitas praticadas em ambiente virtual. Nesse cenário, o meio digital atua como instrumento, cenário, ou até mesmo sujeito da ação. As deepfakes representam uma das manifestações mais complexas dessa criminalidade tecnológica.

O Direito Penal Digital é o ramo que aplica normas penais a condutas ilícitas praticadas por tecnologias da informação. Sua finalidade não é criar um Direito Penal, mas interpretar e adaptar os tipos existentes, respeitando os princípios da legalidade e da intervenção mínima (Greco, 2021).

Nesse sentido, as deepfakes desafiam esses limites, posto causarem danos graves a bens jurídicos essenciais sem que exista uma tipificação penal específica que abarque suas particularidades técnicas e subjetivas.

Atualmente, o ordenamento jurídico brasileiro não possui um tipo penal próprio para incriminar a criação ou disseminação de deepfakes, sendo as condutas delituosas enquadradas por analogia em tipos tradicionais, ou por normas do Marco Civil da Internet, Lei nº 12.965/2014, relativas à responsabilidade dos provedores e à remoção de conteúdo.

A prova digital impõe desafios inéditos de rastreabilidade e autenticidade, pois o conteúdo deepfake reproduz características biométricas de forma tão verossímil que pode escapar aos mecanismos periciais tradicionais.

Esse panorama coloca o Direito Penal diante de um dilema: Como distinguir a verdade dos fatos em meio a representações artificiais indistinguíveis da realidade? Essa questão afeta o princípio da presunção de veracidade das provas, pilar da persecução penal.

O Brasil ainda carece de legislação específica para manipulação digital lesiva. Tramitam no Congresso projetos relevantes, como o PL nº 2.338/2023, denominado Marco Legal da IA, que propõe regular a responsabilidade civil e penal por atos automatizados, mas nenhuma proposta foi sancionada.

Tal lacuna normativa reflete o descompasso entre a velocidade do avanço tecnológico e a resposta legislativa e essa omissão, contudo, deve ser tratada com cautela, posto criar tipos penais amplos e genéricos pode ameaçar o Princípio da Legalidade estrita, abrindo espaço para interpretações arbitrárias e a criminalização excessiva.

Torna-se fundamental buscar um equilíbrio entre tipicidade, proporcionalidade e política criminal para que o Direito Penal Digital proteja os bens jurídicos sem gerar insegurança.

Em síntese, o desafio da tipificação das deepfakes é conceitual e técnico, posto exigir

repensar categorias fundamentais do Direito Penal à luz da autonomia algorítmica e da desmaterialização das condutas.

A ausência de um tipo penal específico evidencia a urgência de um debate legislativo que considere a singularidade das falsificações digitais, evitando a analogia *in malam partem* ou a ampliação indevida de tipos existentes, pois somente por meio de regulação precisa e coerente será possível enfrentar o fenômeno dentro dos marcos do Estado de Direito.

RESPONSABILIDADE PENAL POR DEEPFAKES

A questão da responsabilidade penal nas deepfakes reside no núcleo problemático do Direito Penal Digital uma vez que o ordenamento jurídico tradicional pressupõe que o crime decorre de uma ação humana consciente. Contudo, a Inteligência Artificial (IA) insere uma mediação algorítmica onde parte do resultado é gerado de forma autônoma por sistemas de aprendizado de máquina. Diante disso, a indagação central é: Quem responde penalmente por um conteúdo ilícito criado ou disseminado por meio de IA generativa?

A responsabilidade penal tradicional baseia-se na autoria e culpabilidade, pois ser autor de um crime exige conduta voluntária, consciente e dirigida a um fim, com dolo ou culpa (Bitencourt, 2022).

No caso das deepfakes, a materialização do dano depende da vontade humana e da capacidade autônoma dos algoritmos de gerar, aperfeiçoar e disseminar o conteúdo sem intervenção direta do usuário.

Nesse contexto, identificam-se três possíveis agentes de responsabilidade: o criador do algoritmo/plataforma, o usuário que manipula o conteúdo e o divulgador, possuindo cada agente, graus distintos de imputação subjetiva: O criador que só seria responsabilizado se comprovada a consciência de que seu sistema seria usado para fins criminosos, configurando dolo eventual ou participação; o mero fornecimento da tecnologia, sem dolo de lesar, não configura crime; o usuário, ao empregar a ferramenta com o propósito de difamar, enganar ou prejudicar, realizando diretamente a conduta típica e dolosa, sendo o principal destinatário da sanção penal e o divulgador que ao compartilhar um conteúdo falso sem verificar a autenticidade que pode responder por crimes contra a honra ou divulgação de material ofensivo.

Seja de modo escondido, criptografado ou virtualmente, às claras, o delito informático ocorre em toda a virtualidade constantemente e deve ser compreendido como um todo para evitar

resultados danosos (Sydow, 2021).

A dificuldade reside na autoria mediata por sistemas autônomos, logo, o raciocínio reforça que o domínio da decisão, e não apenas o domínio da execução, é suficiente para caracterizar a autoria penal, podendo o usuário que utiliza uma deepfake consciente de seu potencial danoso ser responsabilizado a título de dolo eventual ou culpa consciente.

No plano internacional, o debate sobre responsabilidade é refletido no Regulamento da UE nº 2024/1689 (AI Act), o qual, embora não trate expressamente de responsabilidade penal, estabelece princípios de accountability e cadeia de responsabilidade (desenvolvedores, operadores, usuários), especialmente para sistemas de alto risco. Esse conceito dialoga com a doutrina da imputação objetiva, segundo a qual o agente responde pelo risco juridicamente desaprovado que introduziu, desde que o dano seja consequência desse risco.

No Brasil, o usuário que cria e dissemina a deepfake com dolo deve ser responsabilizado nos termos do artigo 18, inciso I, do Código Penal. Além disso, o grau de consciência da ilicitude não exclui a responsabilidade, conforme a lei penal.

Por fim, a responsabilidade penal nas deepfakes transcende a punição, posto ser fundamental para a preservação da dignidade humana e da confiança social na verdade, uma vez que a manipulação de identidade e imagem fere valores constitucionais como a honra e a privacidade, atuando o Direito Penal, como mecanismo de garantia da integridade moral e comunicativa da sociedade digital.

Portanto, a responsabilização penal exige uma releitura dos conceitos de autoria, dolo e culpabilidade para abranger condutas mediadas por sistemas autônomos sem violar o princípio da pessoalidade da pena. A intersecção entre técnica e direito requer uma resposta legislativa e doutrinária que harmonize inovação tecnológica e segurança jurídica.

RESULTADOS E DISCUSSÃO

O Direito Penal moderno se baseia na proteção de bens jurídicos enquanto valores essenciais cuja preservação justifica a intervenção estatal. No contexto das deepfakes, essa função protetiva é posta à prova, pois as condutas digitais ameaçam bens jurídicos de forma inédita.

As deepfakes violam múltiplos bens jurídicos simultaneamente devido à sua natureza massiva e multifacetada, sendo o primeiro deles é a honra e Imagem, como exemplo, um vídeo manipulado que atribui comportamentos inexistentes ou vexatórios agride diretamente a honra

objetiva e subjetiva e ainda que a conduta delituosa esteja tipificada nos crimes contra a honra previstos nos artigos 138 a 140 do Código Penal, a sua ocorrência amplifica o dano uma vez que o seu poder de convencimento torna a ofensa eterna, viral e de difícil reparação.

Outro bem jurídico atingido é o direito à imagem e privacidade, traduzido na possibilidade de reprodução digital do rosto, voz ou gestos sem consentimento em patente violação da individualidade e a autodeterminação informativa.

Nesse sentido, as deepfakes subvertem esse controle ao transformar elementos da identidade humana em matéria-prima algorítmica, com potencial lesivo à dignidade enquanto direito assegurado pelo artigo 5º, inciso X, da Constituição Federal de 1988 e pela Lei nº 13.709/2018, denominada Lei Geral de Proteção de Dados (LGPD).

Ademais, a fé Pública também é severamente afetada em virtude da falsificação de conteúdos com aparência de veracidade que ameaça a confiança coletiva na autenticidade das comunicações. A justiça só tem valor real quando sustentada pela integridade ética de todos os envolvidos (Pinto, 2020).

Deepfake que simula pronunciamentos de autoridades ou provas judiciais compromete a credibilidade das instituições e gera efeitos em decisões políticas e econômicas, representando, assim, uma nova forma de falsificação ideológica, que atinge sobremaneira a confiança social.

A segurança da informação e integridade democrática também são afetadas, em razão das deepfakes denotar ameaça à segurança da informação, propagando desinformação e corroendo a credibilidade das fontes.

No âmbito eleitoral, as deepfakes distorcem o debate público, influenciando a vontade popular, consequentemente colocando em risco o próprio funcionamento do regime democrático e esse conjunto de ameaças evidencia que as deepfakes geram um fenômeno de desordem informacional, com impacto transversal que se estende da vítima individual ao tecido social.

A inteligência artificial não precisa criar a ideia de cometer a ofensa específica, mas, para ser responsável criminalmente, precisa apenas cometer a ofensa específica com os elementos factuais dessa ofensa (Hallevy, 2020).

O Direito Penal deve, portanto, assumir uma função garantista e preventiva, focada na proteção proativa da dignidade humana e da verdade comunicativa no ambiente digital, devendo, contudo, a ampliação da tutela penal deve respeitar os princípios da proporcionalidade e da intervenção mínima.

Com efeito, o desafio reside em distinguir o uso socialmente aceitável da manipulação

dolosa, avaliando contexto, intenção e potencial de dano. Essa ponderação exige uma abordagem humanista e tecnológica, que reconheça a dignidade da pessoa sem recorrer ao punitivismo desmedido e arbitrário.

Assim, as deepfakes colocam em risco um amplo leque de bens jurídicos essenciais, devendo o Direito Penal Digital atuar de forma equilibrada, ou seja, nem omisso diante do dano, tampouco excessivo a ponto de restringir a liberdade de criação tecnológica.

METODOLOGIA

A metodologia referente à presente pesquisa contempla como objetivo fundamental descobrir respostas para problemas, mediante o emprego de procedimentos científicos (Gil, 1994), se apresentando, quanto à natureza o objetivo da contribuição com novos conhecimentos para a ciências, se traduzindo como uma pesquisa básica.

Quanto aos objetivos, visa proporcionar maior familiaridade com o problema, visando torná- lo mais explícito, classificando-se como pesquisa exploratória e descritiva. Quanto à abordagem, consiste numa pesquisa qualitativa posto buscar um aprofundamento da compreensão da relação do tema estudado, ou seja, o vínculo primordial entre o universo objetivo e a subjetividade do sujeito. A pesquisa qualitativa responde a questões muito particulares. Ela se preocupa, nas ciências sociais, com um nível de realidade que não pode ser quantificado, ou seja, ela trabalha com o universo de significados, motivos, aspirações, crenças, valores e atitudes, o que corresponde a um espaço mais profundo das relações, dos processos e dos fenômenos que não podem ser reduzidos à operacionalização de variáveis (Minayo, 2001).

Para a efetivação do presente estudo, foi realizada pesquisa bibliográfica, sendo, ao longo do estudo, apresentados entendimentos doutrinários e jurisprudenciais especializados na matéria, seguindo-se à necessária reflexão acerca do tema.

O objetivo deste artigo é contribuir para a reflexão acadêmica acerca da eficácia da normatividade referente às lacunas legislativas, os possíveis enquadramentos jurídicos e as implicações éticas e penais do uso de deepfakes no contexto do Direito Penal Digital contemporâneo. Busca-se sobretudo, proporcionar uma compreensão dos desafios enfrentados atualmente no contexto dos direitos fundamentais em face da tecnologia e da inovação no universo da modernidade.

CONCLUSÃO

O avanço das tecnologias de Inteligência Artificial (IA), com ênfase nas deepfakes, inaugura um novo capítulo para o Direito Penal, devido a essa tecnologia desafiar os pilares da justiça, sobretudo quanto à prova, à autoria e à verdade ao permitir a criação de imagens e vídeos falsificados com alto realismo.

Esse fenômeno não apenas amplia a ofensa a bens jurídicos tutelados como a honra e a privacidade, mas também abala a fé pública e a confiança social na autenticidade das informações.

A análise do presente estudo demonstrou que o ordenamento jurídico brasileiro carece de instrumentos normativos específicos para combater as deepfakes, restando insuficientes os tipos penais existentes diante da complexidade técnica, da autoria difusa e do alcance global das falsificações digitais, posto a ausência de tipificação própria gerar insegurança jurídica bem como dificultar a responsabilização penal, especialmente em contextos de processos automatizados de aprendizado de máquina.

Diante desse cenário, o Direito Penal Digital deve repensar suas categorias clássicas, adaptando conceitos como dolo, culpa e autoria às ações mediadas por algoritmos, devendo a responsabilidade penal por deepfakes ir além do agente direto, considerando toda a cadeia de produção e disseminação conforme o grau de controle e de previsibilidade de cada um. Essa abordagem encontra respaldo na doutrina da imputação objetiva e no conceito internacional de accountability chain da IA.

Mais que criar tipos penais, o desafio reside em equilibrar repressão e liberdade, garantindo-se a proteção do Direito Penal sem barrar a inovação tecnológica, uma vez que o combate às deepfakes deve ser integrado a uma política ampla de alfabetização digital, regulação ética da IA e fortalecimento da prova pericial eletrônica, visando preservar a dignidade humana e a integridade informacional.

Portanto, conclui-se que o fenômeno das deepfakes impõe ao Direito Penal brasileiro a necessidade urgente de atualização normativa e teórica, orientada pelos princípios da proporcionalidade e legalidade.

O Estado deve buscar um modelo de regulação que reconheça os riscos da IA sem ignorar seu potencial positivo, posto a fronteira entre a realidade e a simulação ser cada vez mais tênue,

e é nesse território híbrido que o Direito Penal contemporâneo deve atuar, assegurando que a verdade, a honra e a confiança pública não se percam nas sombras da digitalidade.

REFERÊNCIAS

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília: Diário Oficial da União, 5 out. 1988.

BRASIL. **Decreto-Lei nº 2848.** Código Penal. Brasília: Diário Oficial [da] República Federativa do Brasil, Brasília, 31 dezembro, 1940.

BRASIL. **Lei nº 12.737**, de 30 de novembro de 2012. Dispõe sobre a tipificação criminal de delitos informáticos; altera o Decreto-Lei nº 2.848, de 7 de dezembro de 1940 - Código Penal; e dá outras providências. Brasília: Diário Oficial [da] República Federativa do Brasil, 3 de dezembro de 2012.

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). Brasília: Diário Oficial [da] República Federativa do Brasil, 24 de abril de 2014.

BRASIL. **Lei nº 13.709**, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília: Diário Oficial [da] República Federativa do Brasil, 15 de agosto de 2018.

BRASIL. **Projeto de Lei nº 2.338/2023**. Dispõe sobre o desenvolvimento, o fomento e o uso ético e responsável da inteligência artificial com base na centralidade da pessoa humana. Brasília: Senado Federal, 17 de março de 2025.

BITENCOURT, C. R. **Tratado de Direito Penal: Parte Especial**. 24. ed. v. 2. São Paulo: Saraiva Educação, 2024.

GIL, A. C. **Administração de recursos humanos**. São Paulo: Atlas, 1994.

GRECO, R. **Curso de Direito Penal: Parte Especial (arts. 121–212)**. 21. ed. Niterói: Impetus, 2024.

HALLEVY, Gabriel. **Responsabilidade Penal da Inteligência Artificial**. São Paulo: Editora XYZ, 2020.

MINAYO, M. C. S. **Pesquisa social: Teoria, método e criatividade**. Petrópolis: Vozes, 2001.

NETO, E. J. M.; MORAIS J. L. B. **A fragilização do Estado-Nação na proteção dos Direitos Humanos violados pelas tecnologias da informação e comunicação**. Revista de direitos fundamentais e democracia, v. 23, n. 3, p. 231-257, set./dez. 2018.

PINHEIRO, P. P. Direito Digital. 7. ed. São Paulo: Saraiva Jur. 2021.

PINTO, H. A. **A utilização da inteligência artificial no processo de tomada de decisões: por uma necessária accountability**. Revista de Informação Legislativa: RIL, Brasília, v. 57, n. 225, p. 43-60, jan./mar. 2020.

SYDOW, S. T. **Curso de Direito Penal Informático: Partes Geral e Especial**. Salvador: Editora Jvspodium, 2021.

UNIÃO EUROPEIA. **Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho**, de 13 de junho de 2024, que cria regras harmonizadas em matéria de inteligência artificial e que altera os Regulamentos (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e as Diretivas 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (Regulamento da Inteligência Artificial). Bruxelas: Parlamento Europeu, 2024.