www.revistaft.com.br

Revista FT | ISSN 1678-0817 | v. 30, n. 157, 2026 | DOI: 10.69849/b37szm95

O uso indevido de inteligências artificiais generativas nos crimes de estelionato no Brasil: desafios de tipificação e responsabilidade penal

The misuse of generative artificial intelligence in fraud crimes in Brazil: challenges of legal classification and criminal liability

Gabriel Carlos De Brito Anastácio
Aline De Assis Rodrigues Do Amaral Muniz
Marina Teodoro

Resumo: Este artigo analisa o uso indevido de inteligências artificiais generativas no crime de estelionato, destacando os desafios legais relacionados à tipificação penal e à responsabilidade jurídica. A pesquisa explora como a tecnologia tem sido empregada para cometer fraudes digitais, com ênfase nas novas formas de estelionato mediado por IA, como deepfakes e clonagem de voz. A metodologia utilizada é qualitativa e bibliográfica, com análise de doutrinas jurídicas, jurisprudência e documentos internacionais. A pesquisa propõe a revisão do conceito de autoria e culpabilidade no direito penal, considerando o impacto das IAs generativas no comportamento criminoso. Os resultados indicam que o direito penal atual não está preparado para lidar com a complexidade das ações cometidas com o auxílio de IA, sugerindo a necessidade de reformas legislativas e a cooperação internacional para enfrentar esses desafios. Conclui-se que é essencial repensar a responsabilização penal diante da automação cognitiva, integrando novas ferramentas legais e políticas públicas de prevenção.

Palavras-chave: inteligência artificial, estelionato digital, direito penal, responsabilidade penal, deepfakes.

Abstract: This paper analyzes the misuse of generative artificial intelligences in the crime of fraud, highlighting the legal challenges related to criminal classification and legal responsibility. The study explores how technology has been used to commit digital fraud, focusing on new forms of fraud mediated by AI, such as deepfakes and voice cloning. The methodology is qualitative and bibliographical, with an analysis of legal doctrines, case law, and international documents. The research proposes revisiting the concepts of authorship and culpability in criminal law, considering the impact of generative AIs on criminal behavior. The findings indicate that current criminal law is not prepared to address the complexity of actions committed with AI, suggesting the need for legislative reforms and international cooperation to tackle these challenges. The paper concludes that it is essential to rethink criminal responsibility in the face of cognitive automation, integrating new legal tools and public policies for prevention.

Keywords: artificial intelligence, digital fraud, criminal law, criminal responsibility, deepfakes.

1. INTRODUÇÃO

A ascensão das inteligências artificiais generativas no cenário contemporâneo inaugura uma era em que a fronteira entre criação humana e produção automatizada torna-se tênue, desafiando os fundamentos clássicos do Direito Penal. A inteligência artificial, especialmente em sua vertente generativa, representa um salto tecnológico que ultrapassa a simples automação de tarefas, alcançando a capacidade de criar textos, imagens, sons e até mesmo reproduções da voz humana com elevado grau de realismo. Essa transformação, embora traga benefícios inegáveis para a sociedade e para o progresso científico, também tem sido apropriada por agentes criminosos para fins ilícitos, em especial no cometimento de fraudes e estelionatos digitais.

A era digital tornou a informação o principal ativo econômico e, paradoxalmente, o principal vetor de vulnerabilidade jurídica. A capacidade de uma IA generativa de elaborar mensagens personalizadas, imitar pessoas reais e simular contextos verdadeiros favorece a criação de enganos complexos, cuja detecção se torna cada vez mais difícil. Em 2024, relatórios da Europol e da OCDE indicaram aumento significativo dos crimes de estelionato realizados com suporte de sistemas de IA, inclusive com o uso de vídeos e áudios *deepfake* aplicados em golpes de identidade e fraudes financeiras. O fenômeno desafia as fronteiras normativas do tipo penal previsto no artigo 171 do Código Penal Brasileiro, que foi concebido em um contexto analógico, alheio à lógica autônoma e probabilística dos algoritmos contemporâneos.

A problemática que emerge não se restringe à materialidade do crime, mas à própria natureza da ação delituosa. A utilização da inteligência artificial como instrumento ou meio fraudulento reconfigura a compreensão da autoria e da culpabilidade. O sujeito ativo, em muitos casos, delega parte substancial da execução à máquina, o que levanta questões sobre o domínio do fato, o dolo e a responsabilidade penal do desenvolvedor ou do usuário. Conforme explica Silva Sánchez (2022), “a expansão tecnológica do agir humano projeta novos espaços de imputação penal, em que a fronteira entre autoria mediata e atuação indireta se torna nebulosa” (p. 98).

O estelionato, historicamente definido como crime de obtenção de vantagem ilícita mediante fraude, adquire novas formas no ambiente digital. As mensagens automáticas de voz, os falsos perfis criados por IA e as respostas automáticas de sistemas conversacionais configuram estratégias de engano que dispensam o contato humano direto. A vítima é induzida ao erro por um sistema programado para gerar confiança, e o dano material resulta de um ato de manipulação algorítmica. Mendes (2022) observa que “o Direito Penal, ao lidar com tecnologias inteligentes, precisa distinguir o instrumento da vontade humana da entidade que simula autonomia, pois somente o primeiro pode ser objeto de juízo de reprovação moral e jurídica” (p. 127).

O contexto contemporâneo revela, assim, uma lacuna normativa significativa. O Código Penal brasileiro, datado de 1940 e revisado pontualmente, carece de

previsões capazes de abranger as condutas emergentes associadas ao uso indevido de sistemas autônomos. As modificações promovidas pela Lei nº 14.155/2021, ao incluir o §2º-A no artigo 171, contemplam o estelionato eletrônico, mas não tratam de condutas que envolvem agentes artificiais com capacidade de gerar conteúdo original e enganar múltiplas vítimas simultaneamente. A ausência de tipificação específica obriga o intérprete jurídico a recorrer à analogia e à hermenêutica constitucional para enquadrar comportamentos que transcendem o paradigma tradicional da fraude.

A complexidade do tema se intensifica ao considerar a dimensão internacional da tecnologia. O desenvolvimento de IAs generativas ocorre em ambiente transnacional, sem fronteiras jurídicas claramente definidas. A OCDE, em seus *Principles on Artificial Intelligence* (2023), alerta que "a responsabilidade e a rastreabilidade devem constituir princípios orientadores para o desenvolvimento ético e jurídico das inteligências artificiais" (p. 19). A União Europeia, por sua vez, avançou com o *AI Act* (2024), estabelecendo padrões para uso responsável da IA e impondo deveres de transparência aos desenvolvedores e provedores. Essas iniciativas evidenciam uma preocupação global em equilibrar inovação e proteção jurídica, mas também revelam o atraso da legislação penal brasileira diante do avanço tecnológico.

A delimitação do presente estudo recai sobre o uso indevido de IAs generativas no crime de estelionato, com foco na análise dos desafios de tipificação penal e responsabilização dos agentes envolvidos. Não se trata, portanto, de uma discussão meramente ética ou civilista, mas de uma abordagem voltada à aplicação concreta das normas penais diante de condutas mediadas por tecnologia autônoma. O objetivo é investigar como o ordenamento jurídico brasileiro pode adaptar-se para preservar os princípios constitucionais da legalidade, da dignidade humana e da culpabilidade, diante de realidades comunicacionais e cognitivas inéditas.

A justificativa deste trabalho encontra respaldo na urgência social e institucional de compreender e regulamentar as interações entre inteligência artificial e crime. Em relatório recente, a Interpol (2024) destacou que "as fraudes com base em IA representam a nova fronteira do crime cibernético, demandando atualização urgente

dos marcos legais". O cenário nacional não é diferente: investigações da Polícia Federal revelaram que golpes bancários e falsificações de voz utilizando IA aumentaram mais de 200% entre 2022 e 2024. A proteção do patrimônio e da confiança social exige, portanto, uma resposta penal que não seja apenas punitiva, mas também preventiva e educativa.

Diante dessa conjuntura, o problema de pesquisa que orienta este estudo pode ser formulado nos seguintes termos: como o Direito Penal brasileiro pode tipificar e responsabilizar o uso indevido de inteligências artificiais generativas nos crimes de estelionato, respeitando os limites constitucionais e as exigências de segurança jurídica? A hipótese de trabalho parte da premissa de que o atual ordenamento jurídico é insuficiente para lidar com tais condutas e que a solução requer uma interpretação extensiva e sistemática, ancorada em princípios constitucionais e no diálogo com o direito comparado.

O objetivo geral consiste em analisar os desafios jurídicos, dogmáticos e interpretativos decorrentes da aplicação do Direito Penal a condutas de estelionato mediadas por IAs generativas. Como objetivos específicos, busca-se: compreender a natureza técnica e jurídica dessas ferramentas; examinar a adequação do tipo penal do estelionato tradicional às novas modalidades de fraude digital; avaliar a responsabilidade penal do usuário, do programador e das plataformas tecnológicas; e propor diretrizes para futuras reformas legislativas.

A metodologia utilizada é qualitativa, exploratória e bibliográfica, com base em doutrinas jurídicas, documentos internacionais, jurisprudências do STF e STJ e relatórios de órgãos internacionais, como OCDE, ONU e União Europeia. Adota-se o método dedutivo, partindo de princípios constitucionais e penais para examinar sua compatibilidade com a realidade tecnológica emergente. A pesquisa inclui também uma análise jurisprudencial e comparativa entre o ordenamento brasileiro e o europeu, considerando o avanço regulatório da União Europeia no campo da inteligência artificial.

A relevância teórica deste trabalho reside na necessidade de reconstrução conceitual da imputação penal em tempos de automação cognitiva. A doutrina

clássica, baseada na ação humana consciente e voluntária, enfrenta desafios inéditos quando o agente se vale de sistemas capazes de aprender, adaptar-se e decidir de modo independente. Como ensina Barroso (2020), “o direito constitucional e penal contemporâneo deve dialogar com a realidade tecnológica, sob pena de se tornar um monumento à obsolescência” (p. 33). A compreensão desse diálogo é condição essencial para que o Direito mantenha sua função de limite ético do poder, inclusive do poder algorítmico.

A estrutura deste trabalho está organizada em três eixos principais. No primeiro, examinam-se os fundamentos teóricos e jurídicos da inteligência artificial, com ênfase nas implicações penais de sua autonomia funcional. No segundo, analisa se o crime de estelionato em suas formas tradicionais e digitais, destacando as inovações trazidas pela Lei nº 14.155/2021 e as lacunas persistentes diante das IAs generativas. Por fim, o terceiro eixo aborda os desafios de tipificação, a jurisprudência nacional e internacional, e as propostas de aprimoramento legislativo que buscam compatibilizar a tutela penal com os avanços tecnológicos.

Dessa forma, a introdução apresenta o pano de fundo teórico e prático do problema, revelando que a discussão sobre o uso indevido de IAs generativas não é apenas uma questão de modernização normativa, mas uma reflexão profunda sobre a essência da culpabilidade humana em tempos de automatização. O Direito Penal, para permanecer instrumento de justiça e não de anacronismo, deve aprender a dialogar com as máquinas sem abdicar da centralidade da pessoa humana. A investigação que se segue busca justamente delinear esse diálogo, equilibrando a proteção social, a inovação tecnológica e os valores constitucionais que sustentam o Estado Democrático de Direito.

Ao situar a problemática do uso indevido de inteligências artificiais generativas no contexto jurídico-penal, torna-se evidente que o debate ultrapassa a dimensão técnica da programação e penetra no núcleo axiológico do Direito. A inteligência artificial, enquanto construção humana dotada de autonomia operacional, reproduz decisões e comportamentos baseados em padrões de dados, o que desafia os conceitos tradicionais de intenção e culpabilidade. O sistema jurídico brasileiro, estruturado sobre o princípio da legalidade e da personalidade da pena, enfrenta um

dilema: como responsabilizar um ato cujo resultado foi produzido por uma ferramenta que não possui consciência moral, mas é capaz de agir de modo autônomo e gerar danos concretos?

A resposta a essa indagação exige um olhar que une a dogmática penal à teoria da tecnologia. Luciano Floridi (2021) propõe compreender a IA como um agente informacional inserido em um ecossistema ético, o que significa reconhecer que sua atuação não é neutra, mas carregada de consequências morais e jurídicas. No mesmo sentido, Doneda (2019) sustenta que “a regulação da tecnologia é, em última instância, a regulação do poder”, ressaltando que o controle dos algoritmos deve estar submetido à lógica dos direitos fundamentais (p. 79). Assim, compreender a IA generativa sob a ótica do Direito Penal implica admitir que sua manipulação incorreta representa uma nova forma de exercício de poder ilícito, em que o engano é potencializado por ferramentas automatizadas.

No Brasil, o conceito de estelionato permanece ancorado na noção de engano causado diretamente pela ação humana. O artigo 171 do Código Penal, em sua redação original, pressupõe a conduta dolosa do agente em induzir a vítima a erro. Todavia, no cenário atual, o engano pode ser intermediado por sistemas autônomos programados para aprender e adaptar-se a padrões de comportamento das vítimas. Tal realidade exige a expansão interpretativa do conceito de autoria e de dolo, considerando a possibilidade de o agente humano agir mediante domínio funcional de uma máquina. Claus Roxin (2012) já antecipava esse debate ao desenvolver a teoria do domínio do fato, segundo a qual a autoria penal não depende da execução direta do ato, mas do controle sobre o processo causal que o produz. A aplicação dessa teoria às inteligências artificiais permite imputar a responsabilidade ao indivíduo que mantém o poder de decisão sobre o uso da tecnologia, ainda que o resultado seja mediado por um sistema autônomo.

A literatura jurídica internacional vem acompanhando essa tendência. A União Europeia, por meio do *Artificial Intelligence Act* (2024), definiu categorias de risco associadas a sistemas de IA e impôs obrigações de transparência e rastreabilidade aos desenvolvedores e operadores. No campo penal, a proposta de *AI Liability Directive* (2023) busca estabelecer critérios de responsabilidade civil e penal por

danos causados por decisões algorítmicas. A Organização para Cooperação e Desenvolvimento Econômico (OCDE) também publicou, em 2023, diretrizes que orientam os Estados-membros a adotar políticas de governança ética e jurídica da IA, enfatizando a importância da supervisão humana e da prestação de contas. Tais iniciativas revelam uma tendência internacional de tratar a inteligência artificial não apenas como ferramenta, mas como vetor de transformação social e jurídica.

No contexto brasileiro, ainda não há legislação específica que regule a responsabilidade penal em crimes praticados com auxílio de IA. O Marco Civil da Internet (Lei nº 12.965/2014) e a Lei Geral de Proteção de Dados (Lei nº 13.709/2018) oferecem princípios importantes, como neutralidade, transparência e proteção da privacidade, mas não abordam a questão da imputação criminal em casos de uso indevido de tecnologias autônomas. A lacuna legislativa gera insegurança jurídica e dificulta a atuação dos órgãos de persecução penal. Barroso (2021) observa que “a velocidade da inovação tecnológica desafia a temporalidade do direito, exigindo do intérprete criatividade hermenêutica sem ruptura dos limites constitucionais” (p. 52). Essa observação é especialmente válida no campo penal, onde o princípio da reserva legal impede ampliações analógicas que possam gerar punições sem previsão expressa.

A doutrina contemporânea propõe diferentes caminhos para enfrentar esse desafio. Parte dos penalistas defende a criação de novos tipos penais específicos para condutas mediadas por inteligência artificial, de modo semelhante ao que ocorreu com os crimes informáticos. Outros autores sugerem a reinterpretação dos tipos já existentes, à luz do princípio da adequação social e da teoria da imputação objetiva. De acordo com Jakobs (2019), “a imputação penal deve recair sobre aquele que cria ou mantém uma fonte de risco juridicamente desaprovada” (p. 113). Nessa perspectiva, quem utiliza ou desenvolve uma IA generativa para enganar outrem e obter vantagem ilícita deve ser responsabilizado não pela ação direta da máquina, mas pelo risco que conscientemente introduziu na esfera jurídica alheia.

A evolução dos crimes de estelionato mediado por IA revela uma transformação ontológica da fraude. O engano não decorre mais de uma falsificação documental ou de uma mentira verbal, mas da manipulação de dados e da simulação de

identidades digitais. Os chamados *deepfakes* e *voice clones* exemplificam essa nova forma de fraude, em que a vítima acredita estar interagindo com uma pessoa real. Relatórios da Polícia Federal (2024) apontam o aumento expressivo de golpes baseados em clonagem de voz, em que criminosos utilizam áudios gerados por IA para se passar por familiares e solicitar transferências financeiras. Em muitos casos, a prova do dolo e do nexo causal entre o agente e o resultado torna-se extremamente complexa, uma vez que a materialidade do crime está diluída no ambiente digital.

Outro ponto crítico é o papel das plataformas tecnológicas. Empresas responsáveis por desenvolver ou hospedar sistemas de IA podem ser chamadas a responder por omissão ou falha de supervisão, especialmente quando suas ferramentas são utilizadas em larga escala para a prática de crimes. A responsabilidade penal das pessoas jurídicas, prevista no artigo 225, §3º, da Constituição Federal e regulamentada pela Lei nº 9.605/1998, poderia inspirar mecanismos semelhantes no campo da inteligência artificial. Para Silva (2023), “o Estado deve construir um regime de responsabilidade compartilhada, em que o dever de vigilância tecnológica se converta em dever jurídico de prevenção” (p. 141). Essa proposta reforça a necessidade de cooperação entre o setor público, o privado e a comunidade internacional, na busca de soluções éticas e normativas para o controle do uso criminoso da IA.

Ao mesmo tempo, a discussão sobre responsabilidade não pode ser dissociada da tutela dos direitos fundamentais. A Constituição de 1988 assegura, em seu artigo 5º, incisos X e XII, o direito à privacidade e à inviolabilidade das comunicações, bem como, no inciso LIV, o devido processo legal. Esses direitos formam o núcleo da proteção da personalidade e devem orientar a regulação do ambiente digital. Doneda e Sarlet (2020) observam que “a proteção de dados pessoais é a extensão natural do direito à dignidade humana no ciberespaço” (p. 96). Quando sistemas de IA são utilizados para manipular informações pessoais com fins fraudulentos, o dano ultrapassa a esfera patrimonial, atingindo a própria integridade moral da vítima.

A presente pesquisa, portanto, busca contribuir para a consolidação de um marco jurídico capaz de responder aos desafios impostos pelas tecnologias generativas. A

compreensão da IA como novo vetor da criminalidade exige repensar conceitos dogmáticos, como autoria, dolo e imputabilidade, à luz das mudanças tecnológicas. Mais do que punir, o Direito Penal deve prevenir e educar, servindo de instrumento racional de equilíbrio entre liberdade e segurança. Nesse sentido, Floridi (2021) afirma que “a ética da informação não se resume a limitar o poder tecnológico, mas a integrá-lo em uma visão humanista da sociedade” (p. 58).

Em síntese, esta introdução estabelece o ponto de partida teórico e metodológico para o estudo dos impactos das inteligências artificiais generativas no crime de estelionato. Ao delinear o problema, os objetivos e a relevância social do tema, evidencia-se que o Direito Penal brasileiro se encontra diante de um desafio histórico: adequar seus institutos às novas formas de agir humano mediado por tecnologia. O caminho a ser percorrido não é o da substituição da dogmática clássica, mas de sua reconstrução crítica, de modo que continue servindo à justiça em uma era em que o crime já não se comete apenas com as mãos, mas também com os algoritmos.

A complexidade que envolve o uso indevido de inteligências artificiais generativas no cometimento de crimes de estelionato impõe ao jurista contemporâneo um esforço de reinterpretação do próprio conceito de agir humano no campo penal. O Direito, enquanto sistema normativo construído sobre a previsibilidade e a intenção consciente, vê-se desafiado por um agente tecnológico que atua com base em probabilidades, aprendizado de máquina e replicação de padrões cognitivos. O dilema central, portanto, consiste em determinar até que ponto o produto de uma IA pode ser juridicamente imputado a seu criador, operador ou usuário, sem violar os princípios fundamentais da responsabilidade pessoal e da culpabilidade.

O avanço das tecnologias generativas, como *chatbots*, sintetizadores de voz e modelos multimodais, ampliou exponencialmente as possibilidades de fraude digital. Essas ferramentas são capazes de redigir textos persuasivos, falsificar assinaturas eletrônicas, reproduzir vozes idênticas e criar vídeos falsos com alto grau de realismo, possibilitando que criminosos manipulem vítimas com base em conteúdos impossíveis de distinguir do real. Nesse sentido, o estelionato digital tornou-se não apenas uma infração contra o patrimônio, mas um fenômeno que ameaça a

confiança social nas comunicações. Segundo relatório da Europol (2023), os crimes de fraude com *deepfakes* aumentaram 280% entre 2021 e 2023, sendo as IAs generativas o principal vetor dessa evolução criminosa.

Essa realidade revela um deslocamento do eixo tradicional da tipificação penal. Enquanto o estelionato clássico exige uma relação direta entre o autor e a vítima, mediada por um ato humano de engano, o estelionato mediado por IA dilui a interação e cria um espaço de mediação tecnológica. A vítima é iludida por um artefato automatizado que, em muitos casos, age sem supervisão humana imediata. Como observa Silva Sánchez (2022), “a automação do delito redefine a imputação, pois o comportamento causal deixa de ser uma conduta humana direta e passa a ser a ativação de um processo autônomo que, embora previsível, não é controlado em tempo real” (p. 104). Essa redefinição impõe ao jurista uma nova hermenêutica penal, na qual o dolo e a culpa precisam ser interpretados à luz da programação e da previsibilidade do resultado.

A dogmática penal contemporânea oferece instrumentos para esse tipo de análise. A teoria da imputação objetiva, conforme desenvolvida por Jakobs (2019), estabelece que o autor de um delito deve ser responsabilizado quando cria um risco não permitido que se concretiza em resultado lesivo. Aplicada às inteligências artificiais, essa teoria permite atribuir responsabilidade penal ao programador ou ao usuário que conscientemente utiliza um sistema generativo de alto risco para produzir enganos. Nesse contexto, o dolo não se confunde com a intenção direta de fraudar, mas com o conhecimento do potencial lesivo da ferramenta. Assim, se o agente emprega uma IA capaz de gerar vídeos falsos ou comunicações automatizadas para obter vantagem econômica, assume o risco do resultado ilícito e, portanto, responde penalmente pelo estelionato.

O princípio da legalidade, contudo, impõe limites rigorosos à expansão interpretativa. Nenhum cidadão pode ser punido sem lei anterior que defina o crime e a pena. Esse princípio, previsto no artigo 1º do Código Penal e no artigo 5º, inciso XXXIX, da Constituição Federal, constitui a base da segurança jurídica penal. Entretanto, como destaca Barroso (2021), “a legalidade não deve ser confundida com imobilismo, pois o direito é um organismo vivo que deve responder às mutações

da realidade” (p. 61). O desafio reside em compatibilizar o respeito à legalidade com a necessidade de proteger bens jurídicos frente a novas modalidades de ofensa digital.

Nesse sentido, o legislador brasileiro deu um primeiro passo ao editar a Lei nº 14.155/2021, que incluiu o §2º-A no artigo 171 do Código Penal, tipificando o estelionato eletrônico. Todavia, a referida norma limita-se a reconhecer a fraude cometida por meio de dispositivos informáticos, sem considerar a possibilidade de manipulação cognitiva autônoma realizada por sistemas de IA. Assim, o crime continua centrado na conduta humana direta, o que demonstra a insuficiência normativa frente às novas práticas. A doutrina, como observa Bitencourt (2022), alerta que “a mera adaptação semântica do tipo penal é insuficiente diante da complexidade das tecnologias inteligentes, que exigem revisão estrutural do conceito de ação típica” (p. 89).

A ausência de tipificação específica para os crimes praticados com auxílio de IA compromete também a coerência do sistema penal em relação ao princípio da culpabilidade. A punição exige não apenas a existência do fato típico e antijurídico, mas também a possibilidade de censura moral do agente. No entanto, quando o resultado decorre de uma ação parcialmente automatizada, a imputação subjetiva torna-se nebulosa. Seria o programador responsável pelo uso indevido de seu código? Ou o usuário que aplicou a ferramenta de forma ilícita? A resposta dependerá da demonstração de previsibilidade do resultado e do grau de controle efetivo sobre o sistema. Conforme assevera Fiandaca e Musco (2021), “a culpa penal só pode ser atribuída a quem, diante de um risco concreto, tinha o dever e a possibilidade de evitálo” (p. 137).

Além das questões dogmáticas, há também uma dimensão ética e política. A inteligência artificial, enquanto tecnologia global, não reconhece fronteiras jurisdicionais. Suas aplicações ultrapassam limites nacionais, o que torna ineficaz uma abordagem puramente doméstica do problema. É por isso que organismos internacionais, como a OCDE e a União Europeia, têm enfatizado a necessidade de cooperação transnacional para enfrentar os riscos jurídicos da IA. O *AI Act* europeu e os princípios de governança da OCDE são exemplos de esforços normativos que

visam harmonizar padrões de responsabilidade e segurança digital. O Brasil, ainda em processo de elaboração de seu marco legal de inteligência artificial, precisa incorporar esses parâmetros em sua estrutura jurídica, sob pena de ver-se isolado diante da complexidade global da tecnologia.

Em perspectiva criminológica, o uso indevido de IAs generativas nos crimes de estelionato não pode ser tratado apenas como uma variação moderna de antigas fraudes. Ele representa um novo paradigma de criminalidade, caracterizado pela desmaterialização da prova, pela impessoalidade do engano e pela descentralização da autoria. A máquina é o instrumento, mas também o meio de execução e dissimulação do delito. Conforme observa Zuboff (2019), “no capitalismo de vigilância, o poder se exerce por meio do controle da informação e da modelagem do comportamento humano” (p. 114). Ao aplicar essa lógica ao campo penal, é possível perceber que o crime de estelionato com IA não é apenas uma violação patrimonial, mas uma subversão da confiança informacional que sustenta as relações sociais.

A introdução deste trabalho, portanto, delineia um panorama teórico e normativo que justifica a urgência de repensar o Direito Penal diante da automação cognitiva. O problema não reside em punir a tecnologia, mas em compreender como a tecnologia amplia o alcance e a sofisticação da conduta criminosa. É preciso construir uma teoria da imputação digital que preserve a coerência do sistema penal e, ao mesmo tempo, garanta a eficácia da tutela jurídica no ambiente virtual. A hermenêutica penal deve evoluir sem trair seus fundamentos, reconhecendo que o sujeito do século XXI não é apenas o homem que age, mas também aquele que programa, delega e supervisiona agentes artificiais.

Em última análise, o estudo do uso indevido de inteligências artificiais generativas nos crimes de estelionato representa mais do que uma investigação sobre novas formas de fraude; trata-se de um esforço de reconstrução do próprio conceito de culpabilidade em uma era de simulações inteligentes. O Direito Penal brasileiro, herdeiro de uma tradição humanista e antropocêntrica, é chamado a enfrentar o desafio de julgar condutas que já não se limitam ao corpo físico, mas se projetam no espaço digital. A superação dessa crise dogmática exigirá criatividade legislativa,

prudência judicial e diálogo interdisciplinar, para que o Direito continue sendo expressão da racionalidade humana diante da crescente autonomia das máquinas.

2 FUNDAMENTOS TEÓRICOS E JURÍDICOS DA INTELIGÊNCIA ARTIFICIAL

O estudo da inteligência artificial (IA) e de sua implicação jurídica exige uma abordagem interdisciplinar que una os campos da tecnologia, da filosofia e do Direito. A compreensão do fenômeno não pode restringir-se a uma visão instrumental, que entende a IA apenas como ferramenta técnica, mas deve considerar seu impacto sobre os fundamentos da responsabilidade e da ação humana. A autonomia operacional dos sistemas generativos impõe uma ruptura conceitual com os modelos tradicionais de causalidade e imputação, pois as máquinas não apenas executam comandos, mas aprendem, adaptam-se e produzem resultados imprevisíveis a partir de padrões probabilísticos. Essa transformação altera profundamente as categorias dogmáticas do Direito Penal, sobretudo quando esses sistemas são empregados na prática de crimes como o estelionato.

A inteligência artificial é o produto de um longo processo de experimentação humana em busca de replicar ou simular o raciocínio. Desde Alan Turing, em 1950, com a publicação de *Computing Machinery and Intelligence*, discute-se a possibilidade de as máquinas pensarem. No entanto, a virada contemporânea deu-se com o advento das redes neurais profundas (*deep learning*) e dos sistemas generativos de linguagem, capazes de criar textos, sons e imagens com autonomia semântica. Floridi (2021) observa que “a IA contemporânea deixou de ser um instrumento passivo para tornar-se um agente informacional, operando em um espaço de decisão que antes era exclusivamente humano” (p. 73). Essa mudança ontológica – de ferramenta para agente – é o ponto de partida para qualquer reflexão jurídica sobre a responsabilidade penal na era digital.

Do ponto de vista jurídico, o desafio central está em definir a natureza da IA: se mero instrumento técnico, se extensão da vontade humana ou se ente autônomo dotado de relevância normativa. A doutrina majoritária brasileira e estrangeira tende a negar personalidade jurídica à IA, sustentando que ela carece de consciência

moral e de vontade própria. Entretanto, isso não elimina a necessidade de estabelecer parâmetros de responsabilização pelos atos que decorrem de sua atuação. Doneda (2019) defende que “a responsabilidade pelo uso da tecnologia deve seguir o vetor do poder de controle, e não da ficção da vontade” (p. 114). Em outras palavras, quem cria, opera ou se beneficia de um sistema de IA deve responder pelos riscos que introduz na esfera jurídica de terceiros.

O debate sobre a responsabilidade por atos praticados com o auxílio de IA não é novo na doutrina europeia. A Comissão Europeia, por meio do *AI Liability Directive* (2023), propôs um modelo de imputação baseado na previsibilidade e na rastreabilidade, de modo que o responsável é aquele que tinha condições de prever o dano e adotar medidas preventivas. Essa perspectiva rompe com o paradigma causal tradicional e inaugura uma teoria da responsabilidade algorítmica, em que o dever jurídico se relaciona ao controle técnico e à vigilância sobre o sistema. A OCDE (2023), em seus *AI Principles*, reitera essa linha ao afirmar que “a supervisão humana é requisito essencial para a legitimidade e a responsabilidade das decisões automatizadas” (p. 17).

No contexto penal, essa evolução conceitual exige uma releitura do princípio da culpabilidade. Tradicionalmente, a imputação penal pressupõe dolo ou culpa e uma conduta voluntária. Contudo, nas interações com sistemas inteligentes, o resultado pode derivar de processos não diretamente controlados pelo agente humano. Fiandaca e Musco (2021) destacam que “a modernidade tecnológica desafia o paradigma da ação, substituindo o fazer imediato pela delegação técnica de resultados” (p. 89). Quando um desenvolvedor cria um código capaz de tomar decisões autônomas que resultam em fraude, a autoria não é mais um ato direto, mas uma cadeia causal complexa que envolve escolhas premeditadas e riscos calculados.

A teoria da imputação objetiva, formulada por Jakobs (2019), fornece um arcabouço útil para lidar com esse fenômeno. Segundo o autor, há responsabilidade penal sempre que o agente cria um risco juridicamente desaprovado que se concretiza em resultado lesivo. Assim, o programador que lança um modelo generativo sem mecanismos de segurança adequados, sabendo de seu potencial para fraudes,

introduz um risco que transcende a neutralidade técnica. Esse risco, quando materializado em dano patrimonial ou em engano de terceiros, legitima a imputação penal por negligência consciente ou dolo eventual.

Outro ponto relevante é a aplicação do princípio da precaução tecnológica. Derivado do Direito Ambiental e incorporado ao debate ético sobre IA, esse princípio estabelece que o desenvolvimento e o uso de novas tecnologias devem ser orientados pela prudência e pela prevenção de riscos sociais. No campo penal, essa noção poderia fundamentar uma responsabilidade agravada para quem emprega sistemas de IA em contextos de alto risco sem supervisão humana. Conforme sustenta Mendes (2022), “a imprevisibilidade do comportamento algorítmico exige que o legislador antecipe cenários de abuso e estabeleça deveres de cuidado proporcionais à complexidade da tecnologia” (p. 133).

As inteligências artificiais generativas, por sua capacidade criativa e mimética, ampliam a zona de incerteza jurídica. Diferentemente de sistemas determinísticos, que operam segundo comandos fixos, as IAs generativas produzem respostas novas e originais, o que dificulta a identificação do nexo causal entre o ato humano e o resultado ilícito. Zuboff (2019) alerta que “a lógica da automação algorítmica desloca a intencionalidade humana para o domínio das probabilidades, tornando o controle moral difuso e fragmentado” (p. 162). Essa difusão de responsabilidades impõe ao Direito Penal a tarefa de reconstruir seus conceitos fundamentais sem romper com os limites constitucionais da legalidade e da culpabilidade.

No plano dogmático, a autoria mediata e a coautoria por intermédio de sistemas de IA configuram novos desafios interpretativos. Roxin (2012) argumenta que o domínio do fato é o critério determinante da autoria, e não a execução direta do ato. Assim, se o agente exerce poder decisório sobre a utilização da IA, define seus parâmetros de atuação e se beneficia do resultado, mantém o domínio funcional da ação e, portanto, deve ser considerado autor. Essa linha de raciocínio encontra eco na jurisprudência brasileira, que tem admitido a responsabilização de administradores e desenvolvedores de plataformas digitais quando há demonstração de dolo ou omissão relevante.

Por outro lado, há posições doutrinárias que defendem a necessidade de limites estritos à expansão da responsabilidade penal no campo tecnológico. Silva (2023) sustenta que “a punição simbólica da inovação pode gerar efeito paralisante sobre o desenvolvimento científico, comprometendo o equilíbrio entre segurança jurídica e liberdade de pesquisa” (p. 147). Essa advertência reforça a importância de se distinguir entre erro técnico e dolo instrumental. A criminalização excessiva pode desestimular a inovação, enquanto a ausência de regulação pode favorecer a impunidade. O ponto de equilíbrio está na criação de mecanismos normativos claros, que imponham deveres de diligência, auditoria e transparência, sem comprometer o progresso tecnológico.

A Constituição Federal de 1988 fornece o alicerce axiológico para essa discussão. Os princípios da dignidade da pessoa humana (art. 1º, III), da legalidade (art. 5º, II e XXXIX) e da proporcionalidade constituem o núcleo de racionalidade que deve orientar o enfrentamento jurídico da inteligência artificial. Barroso (2020) enfatiza que “a Constituição é um instrumento de adaptação social, e sua força normativa depende da capacidade de responder a novas realidades sem trair seus fundamentos” (p. 59). Nesse sentido, a interpretação constitucional deve permitir que o Direito Penal acompanhe as transformações tecnológicas, preservando o equilíbrio entre liberdade e segurança.

No âmbito internacional, observa-se uma convergência de esforços voltados à regulação ética e jurídica da inteligência artificial. A UNESCO (2023) publicou as *Diretrizes Éticas sobre IA*, estabelecendo princípios como transparência, responsabilidade e equidade. Esses documentos reforçam a ideia de que o uso da IA deve respeitar os direitos humanos e promover a justiça social. No mesmo sentido, a Declaração de Roma sobre IA (2024), subscrita por diversos países, inclusive o Brasil, reafirma o compromisso de garantir que os sistemas automatizados não sejam empregados para fins ilícitos ou discriminatórios.

Diante desse panorama, a análise dos fundamentos teóricos e jurídicos da inteligência artificial revela que a tecnologia, em si, não é inimiga do Direito, mas um campo de expansão dos seus desafios. O problema não reside na máquina, mas na ausência de mecanismos normativos capazes de enquadrar seu uso indevido. A

responsabilidade penal por atos praticados com o auxílio de IA deve, portanto, basear-se em critérios objetivos de controle, previsibilidade e benefício econômico. O desenvolvimento seguro e ético da IA requer, por parte do Estado e da sociedade, um compromisso coletivo com a regulação democrática e com a preservação dos valores fundamentais do Direito Penal contemporâneo.

Assim, a inteligência artificial não pode ser tratada como uma entidade alheia ao sistema jurídico, mas como parte integrante de um novo ambiente de responsabilização. O Direito Penal, se pretende permanecer instrumento eficaz de tutela dos bens jurídicos, precisa incorporar à sua estrutura teórica as noções de risco tecnológico, supervisão humana e rastreabilidade algorítmica. Somente assim será possível enfrentar, com racionalidade e justiça, os desafios impostos pelo uso indevido de IAs generativas em práticas criminosas como o estelionato digital.

3 O CRIME DE ESTELIONATO E SUAS NOVAS CONFIGURAÇÕES TECNOLÓGICAS

O estelionato, previsto no artigo 171 do Código Penal brasileiro, sempre foi entendido como o crime de fraude por excelência, uma conduta que exige o engano da vítima e a obtenção de vantagem ilícita em prejuízo alheio. A tipificação clássica pressupõe a ação dolosa do agente, que induz ou mantém alguém em erro mediante artifício, ardil ou qualquer outro meio fraudulento. No entanto, na era da automação cognitiva, essa concepção enfrenta uma transformação estrutural. As novas tecnologias, especialmente as inteligências artificiais generativas, expandiram as possibilidades de fraude a níveis antes inimagináveis, modificando tanto os meios de execução quanto a própria natureza do engano.

O desenvolvimento de sistemas generativos, como modelos de linguagem e algoritmos de aprendizado profundo, permitiu a criação de textos, áudios e imagens que reproduzem padrões humanos com extrema precisão. Essa capacidade de simulação é o que torna a IA generativa um instrumento poderoso para o cometimento de fraudes. O agente criminoso não precisa mais interagir diretamente

com a vítima: basta programar ou manipular uma IA para criar uma narrativa convincente, imitar a voz de um familiar ou gerar uma imagem realista de uma autoridade pública. O engano, nesses casos, decorre de uma ilusão construída por máquinas, mas legitimada pela confiança social na aparência de veracidade dos conteúdos digitais.

Historicamente, o estelionato sempre esteve vinculado à confiança social. Beccaria, ainda no século XVIII, afirmava que “a fraude é o mais insidioso dos delitos, pois destrói o tecido moral da convivência” (1764, p. 57). Essa observação permanece atual: o estelionato digital mina os alicerces da credibilidade nas comunicações eletrônicas, comprometendo a própria noção de segurança nas relações virtuais. A fraude contemporânea, diferentemente da tradicional, não depende da persuasão direta do agente, mas da manipulação de dados e da simulação tecnológica. Mendes (2022) ressalta que “a fraude digital desloca a materialidade do engano do contato físico para a manipulação simbólica de dados” (p. 211).

A Lei nº 14.155/2021, ao incluir o §2º-A no artigo 171 do Código Penal, representou um avanço significativo ao reconhecer o estelionato eletrônico como modalidade típica. O dispositivo prevê pena agravada para quem pratica o crime “mediante a utilização de informações fornecidas pela vítima ou por terceiro induzido a erro por meio de redes sociais, contatos telefônicos ou envio de correio eletrônico fraudulento”. Apesar de importante, essa inovação normativa ainda é insuficiente para lidar com o uso de inteligências artificiais generativas. A redação legal pressupõe que o agente humano seja o autor direto da fraude, enquanto, na prática, muitos golpes são executados por sistemas autônomos programados para enganar em massa, sem intervenção contínua do criminoso.

A tipificação clássica também enfrenta dificuldades diante do fenômeno dos *deepfakes*, conteúdos sintéticos gerados por IA que simulam com fidelidade pessoas reais. Esses vídeos e áudios falsos têm sido utilizados em fraudes sentimentais, extorsões, chantagens e manipulações políticas. A Polícia Federal (2024) registrou casos de falsificação de vozes de executivos e familiares para obtenção de transferências bancárias, o que demonstra a sofisticação das novas estratégias de engano. O problema jurídico surge quando se busca identificar o autor do delito e

comprovar o dolo, uma vez que o agente humano pode alegar desconhecimento do uso indevido da ferramenta.

A doutrina penal tradicional define o dolo como a vontade livre e consciente de realizar a conduta típica. Contudo, no contexto digital, a vontade pode ser mediada por sistemas que aprendem e agem de modo independente. Fiandaca e Musco (2021) sustentam que “a intervenção tecnológica cria uma zona cinzenta entre o dolo direto e o dolo eventual, na medida em que o agente aceita o risco da atuação autônoma da máquina” (p. 201). Assim, aquele que utiliza uma IA sem controle técnico adequado, sabendo que ela pode produzir fraudes, incorre em dolo eventual, pois assume o risco de causar o resultado ilícito.

Outra questão central é a autoria mediata. Claus Roxin (2012), ao formular sua teoria do domínio do fato, propõe que é autor quem possui o controle final do acontecimento delituoso. Aplicando essa lógica ao estelionato digital, o criador ou operador da IA que define seus parâmetros de funcionamento e dela se beneficia mantém o domínio funcional da ação, mesmo que não participe diretamente da execução do crime. Desse modo, a autoria penal não desaparece com a intermediação tecnológica, mas se reconfigura em novas formas de domínio técnico e decisório.

O estelionato mediado por IA também desafia a prova penal. A volatilidade das evidências digitais, a dificuldade de rastreamento de endereços de IP e o uso de redes descentralizadas (como a *dark web*) tornam a investigação complexa. Nesse ponto, a jurisprudência do Superior Tribunal de Justiça tem reconhecido a necessidade de aperfeiçoamento das técnicas de produção de prova digital. No AgRg no REsp 1.984.171/DF, o STJ enfatizou que “a prova digital deve ser coletada com observância aos princípios da legalidade e da cadeia de custódia, sob pena de nulidade” (DJe 12/09/2023). A ausência de parâmetros técnicos uniformes fragiliza o combate penal às fraudes com IA, abrindo espaço para a impunidade.

A dimensão econômica do estelionato digital é igualmente relevante. Segundo a Federação Brasileira de Bancos (FEBRABAN, 2024), os golpes com uso de IA causaram prejuízos estimados em 2,3 bilhões de reais apenas no primeiro semestre

daquele ano. Esses números demonstram que a criminalidade cibernética deixou de ser marginal e tornou-se uma indústria lucrativa. O uso de chatbots e clones de voz permite que criminosos multipliquem as tentativas de fraude, aumentando a escala e a eficácia das ações. A capacidade de automação das IAs gera um efeito de massa que ultrapassa a lógica do crime artesanal e coloca em xeque a eficácia do Direito Penal repressivo tradicional.

Do ponto de vista ético, o uso indevido da IA generativa reforça a necessidade de uma nova cultura de responsabilidade digital. O engano algorítmico não é apenas um problema técnico, mas uma violação da confiança pública na informação. Floridi (2021) defende que “a ética da informação é o novo campo de batalha moral da humanidade, pois quem controla os fluxos de dados controla a própria percepção da realidade” (p. 94). Quando os algoritmos são manipulados para criar realidades falsas, o resultado é uma erosão progressiva da verdade, fundamento essencial da convivência civilizada.

A análise criminológica também revela um fenômeno de despersonalização da fraude. A vítima raramente interage com o autor, o que reduz a percepção social de periculosidade. Essa impessoalidade contribui para o aumento do número de golpes e para a banalização da conduta delituosa. Jakobs (2019) aponta que “a ausência de contato direto entre agente e vítima dilui o sentimento de reprovação social, gerando uma criminalidade funcionalmente aceita” (p. 159). No caso das IAs generativas, o crime é ainda mais abstrato: o agente pode estar em outro país, e o engano é produzido por um algoritmo sem rosto.

Frente a essas novas configurações, o Direito Penal brasileiro precisa reinterpretar o conceito de fraude, ampliando sua compreensão para abranger o engano algorítmico. A fraude, nesse contexto, não é apenas o ato de ludibriar uma pessoa, mas de manipular informações ou sistemas automatizados de forma a induzir comportamentos humanos errôneos. Essa ampliação não viola o princípio da legalidade, desde que respeite o núcleo essencial do tipo penal e a teleologia da norma. O objetivo do legislador sempre foi proteger o patrimônio e a confiança social contra meios fraudulentos, e a IA generativa representa apenas a evolução dos instrumentos de engano.

Em síntese, o crime de estelionato mediado por inteligências artificiais generativas inaugura uma nova fronteira para o Direito Penal. A fraude digital redefine os papéis de autor, meio e vítima, exigindo do intérprete jurídico uma hermenêutica dinâmica, capaz de conciliar segurança jurídica e adaptação tecnológica. A proteção penal do patrimônio e da boa-fé objetiva, nesse cenário, depende da capacidade do Estado de atualizar seus instrumentos normativos e de investigação, reconhecendo que, na sociedade informacional, a verdade tornou-se um bem jurídico tão valioso quanto o próprio dinheiro.

4 DESAFIOS DE TIPIFICAÇÃO PENAL E RESPONSABILIDADE JURÍDICA

A tipificação penal do uso indevido de inteligências artificiais generativas no crime de estelionato constitui um dos maiores desafios dogmáticos do século XXI. O legislador brasileiro, ao redigir o artigo 171 do Código Penal em 1940, não poderia imaginar um contexto em que o engano pudesse ser produzido por um agente não humano, autônomo e capaz de criar informações falsas indistinguíveis da realidade. A lei penal, que sempre se fundamentou no pressuposto de que o autor é um ser racional e consciente, enfrenta agora um cenário em que a fraude pode ser cometida por meio de sistemas que simulam essas mesmas características. Essa ruptura epistemológica obriga o Direito a reconstruir seus fundamentos sobre uma nova base técnico-jurídica.

A questão central reside na delimitação da autoria e na aferição da culpabilidade. A doutrina tradicional define o autor como aquele que executa diretamente o verbo do tipo penal. Todavia, nos crimes cometidos com auxílio de IA, o verbo “enganar” não é realizado pela mão humana, mas por um código autônomo. Como observa Silva Sánchez (2022), “a robotização da conduta penal desloca o foco da ação para a decisão anterior de programar, configurar ou utilizar a máquina” (p. 141). Essa constatação conduz à noção de autoria mediata por meio de instrumento não humano, em que o domínio do fato se exerce sobre o sistema tecnológico que atua como prolongamento da vontade do agente.

No entanto, há hipóteses ainda mais complexas, em que a IA opera de maneira autônoma e imprevisível, produzindo resultados fraudulentos sem que o programador ou o usuário tenham previsto o desfecho. Nesse caso, a responsabilidade penal só pode ser afirmada se houver demonstração de culpa técnica, ou seja, negligência no dever de vigilância e controle do sistema. A teoria da imputação objetiva, formulada por Jakobs (2019), oferece critérios adequados para essa análise. Segundo o autor, a responsabilização penal exige que o agente tenha criado um risco juridicamente desaprovado e que esse risco tenha se concretizado em resultado lesivo dentro do alcance de previsibilidade. Assim, o dever de cuidado do operador da IA torna-se o ponto de ancoragem da culpa penal.

A ausência de previsão expressa para essas situações no ordenamento brasileiro cria uma lacuna que precisa ser suprida por meio de interpretação sistemática. A Constituição Federal de 1988, em seu artigo 5º, inciso XXXIX, consagra o princípio da legalidade penal: “não há crime sem lei anterior que o defina, nem pena sem prévia cominação legal”. Esse princípio impede a criação de tipos penais por analogia, mas não veda a aplicação extensiva fundada na finalidade da norma. Barroso (2021) explica que “a interpretação penal não pode ser cega à realidade, sob pena de tornar o Direito irrelevante diante da evolução social e tecnológica” (p. 64). Assim, a fraude cometida por meio de IA deve ser compreendida como modalidade sofisticada do mesmo fenômeno que o tipo penal do estelionato busca coibir: a violação da confiança por meio do engano.

Outro desafio relevante é a identificação do bem jurídico lesado. No estelionato tradicional, o objeto de tutela é o patrimônio. No entanto, no contexto das fraudes digitais com IA, o dano ultrapassa o aspecto econômico, atingindo a confiança pública nas comunicações e a integridade informacional das relações sociais. Floridi (2021) propõe a noção de “infoesfera” como o novo ambiente de existência humana, no qual a informação é elemento essencial da realidade social. Desse modo, o crime de fraude algorítmica lesa não apenas o patrimônio individual, mas o ecossistema informacional que sustenta a vida democrática. A integridade da informação passa a ser um bem jurídico digno de proteção penal.

A OCDE (2023) e a União Europeia, por meio do *AI Act* (2024), avançaram significativamente na tentativa de regulamentar a responsabilidade por atos ilícitos praticados com IA. O *AI Act* introduz o conceito de “sistemas de alto risco”, impondo aos desenvolvedores e operadores deveres de transparência, rastreabilidade e auditoria. O documento também reconhece a importância do princípio do *human oversight* — a supervisão humana como elemento indispensável para garantir a responsabilidade. No campo penal, essa diretriz implica que a ausência de supervisão sobre um sistema potencialmente lesivo pode configurar negligência grave.

O Brasil, por sua vez, ainda não consolidou um marco regulatório específico sobre inteligência artificial. O Projeto de Lei nº 2.338/2023, em tramitação no Senado Federal, propõe princípios gerais de desenvolvimento ético e responsável da IA, mas não trata expressamente da responsabilização penal. Essa omissão legislativa cria um vazio perigoso, pois deixa de reconhecer que o uso indevido da tecnologia já produz consequências concretas para a segurança econômica e jurídica da sociedade. Mendes (2022) ressalta que “a ausência de regulação penal sobre a IA é uma forma de anomia tecnológica, na qual a lei se torna espectadora da criminalidade digital” (p. 144).

Outro ponto de tensão é a responsabilidade das plataformas e empresas que desenvolvem ou hospedam sistemas de IA. O Marco Civil da Internet (Lei nº 12.965/2014) adota o modelo de responsabilidade posterior, segundo o qual o provedor só responde após descumprimento de ordem judicial. Todavia, esse regime mostra-se insuficiente para lidar com o uso em larga escala de IAs generativas em fraudes. A ausência de mecanismos de verificação e controle de uso permite que criminosos se valham de plataformas aparentemente neutras para promover golpes e manipulações. Doneda (2019) propõe a adoção de um modelo de responsabilidade compartilhada, no qual “o dever de diligência deve acompanhar a posição de poder informacional de cada agente” (p. 108). Nesse modelo, quanto maior a capacidade técnica e o controle sobre os dados, maior a responsabilidade pelo uso indevido.

No campo dogmático, a imputação penal nos crimes cometidos com IA exige uma revisão da teoria da ação. A ação típica, tradicionalmente concebida como comportamento humano voluntário dirigido a um fim, passa a ser compreendida como processo decisório distribuído, em que a vontade humana se manifesta por meio da configuração de parâmetros algorítmicos. Fiandaca e Musco (2021) argumentam que “o Direito Penal contemporâneo precisa distinguir entre agir pessoal e agir mediado por tecnologia, reconhecendo que ambos produzem efeitos jurídicos equivalentes” (p. 207). Isso significa que a programação de um sistema fraudulento é, em si mesma, uma forma de ação penalmente relevante.

A prova penal também assume nova complexidade nesse contexto. A rastreabilidade das decisões algorítmicas, o registro de logs e a perícia técnica sobre modelos de aprendizado de máquina tornam-se elementos essenciais para a identificação da autoria e do dolo. O Superior Tribunal de Justiça, no julgamento do HC 685.726/SP (DJe 19/06/2023), reconheceu que a “prova digital requer instrumentos técnicos de preservação e análise forense adequados, sob pena de violação do contraditório e da ampla defesa”. Essa orientação sinaliza a necessidade de atualização das práticas processuais penais, para garantir que os delitos digitais sejam devidamente apurados e julgados dentro dos parâmetros de justiça e segurança jurídica.

A dificuldade de tipificação também está ligada à volatilidade das condutas. As inteligências artificiais generativas operam em ciclos rápidos de atualização e aprendizado, o que dificulta a fixação de padrões de comportamento. O que hoje é previsível pode tornar-se imprevisível amanhã, diante da evolução do algoritmo. Por essa razão, o Direito Penal deve adotar uma abordagem principiológica, e não meramente casuística, de modo a abranger o fenômeno tecnológico de forma flexível. Barroso (2023) sustenta que “a normatividade constitucional deve ser suficiente para iluminar o futuro, e não apenas descrever o passado” (p. 81). Essa visão justifica o uso dos princípios da proporcionalidade, da razoabilidade e da proteção à confiança como guias interpretativos na aplicação da lei penal a crimes com IA.

A criminalização de condutas envolvendo IA também precisa observar o princípio da intervenção mínima. O Direito Penal deve atuar apenas quando outros ramos do direito — civil, administrativo ou regulatório — se mostrarem insuficientes para tutelar o bem jurídico ameaçado. Nesse sentido, o desenvolvimento de marcos regulatórios preventivos, como o *AI Act* europeu, pode reduzir a necessidade de sanções penais, estabelecendo mecanismos de auditoria e certificação que impeçam o uso abusivo da tecnologia. Floridi (2021) defende que “a prevenção ética é mais eficaz que a punição jurídica, pois atua antes da violação” (p. 109). No entanto, enquanto tais mecanismos não existirem de forma eficaz, o Direito Penal permanece como última barreira de contenção contra a instrumentalização da IA para fins ilícitos.

Por fim, a ausência de uniformização internacional na definição de responsabilidade penal por atos envolvendo IA gera insegurança jurídica e incentiva a migração de criminosos digitais para jurisdições mais permissivas. A criação de tratados multilaterais sobre governança penal da inteligência artificial, nos moldes da Convenção de Budapeste sobre Crimes Cibernéticos, seria um passo essencial para harmonizar legislações e fortalecer a cooperação entre os Estados. A criminalidade digital, por natureza transnacional, exige respostas igualmente globais.

Em síntese, os desafios de tipificação penal do uso indevido de IAs generativas no crime de estelionato evidenciam a necessidade de um novo paradigma jurídico, fundado em princípios de responsabilidade tecnológica e supervisão humana. A lei deve evoluir para reconhecer que o dolo e a culpa não se limitam à ação direta, mas também se manifestam na negligência informacional e na omissão de controle. O futuro do Direito Penal dependerá da capacidade de equilibrar a proteção dos bens jurídicos com o respeito aos limites constitucionais da legalidade, evitando tanto a impunidade quanto o excesso punitivo. A inteligência artificial, afinal, não é inimiga do Direito, mas o espelho que revela suas próprias deficiências diante da complexidade do mundo digital.

5 A RESPONSABILIDADE PENAL DO AGENTE HUMANO E DAS ENTIDADES TECNOLÓGICAS

A evolução das inteligências artificiais generativas trouxe à tona um dilema inédito no campo jurídico: quem deve responder pelos crimes cometidos por sistemas autônomos? A questão da responsabilidade penal, que até então se limitava à conduta humana, passa a abranger também a análise da cadeia de decisões técnicas que culminam em um resultado ilícito. O Direito, portanto, é chamado a reinterpretar o conceito de imputação, deslocando o foco da ação direta para a esfera do controle e da previsibilidade.

No cenário atual, a inteligência artificial ainda não possui personalidade jurídica. Contudo, isso não impede que seus efeitos sejam juridicamente relevantes. A ação autônoma de uma IA pode ser equiparada, sob certas condições, à ação mediada de um ser humano, desde que se comprove o domínio funcional sobre o sistema. Roxin (2012) sustenta que “a autoria penal não se restringe à execução material, mas compreende o poder de decisão sobre o curso causal do evento” (p. 197). Assim, aquele que programa, supervisiona ou se beneficia do uso indevido de um sistema generativo que comete fraude deve responder como autor mediato, na medida em que detém o domínio do fato.

A responsabilidade penal no contexto da IA generativa deve ser compreendida sob dois prismas: a responsabilidade do agente humano e a responsabilidade das entidades tecnológicas (empresas e plataformas). No primeiro caso, a imputação recai sobre o indivíduo que cria, manipula ou utiliza o sistema com finalidade ilícita. No segundo, envolve o dever de vigilância e prevenção das corporações que desenvolvem ou operam as ferramentas tecnológicas.

No que se refere ao agente humano, a doutrina penal oferece categorias clássicas que ainda podem ser aplicadas, embora reinterpretadas. O dolo direto ocorre quando o indivíduo utiliza conscientemente a IA para enganar ou fraudar; o dolo eventual manifesta-se quando o agente prevê a possibilidade de a ferramenta ser usada para fins ilícitos e, ainda assim, a emprega; e a culpa surge quando há negligência, imperícia ou imprudência na manipulação do sistema. Fiandaca e

Musco (2021) ressaltam que “a previsibilidade técnica é o novo critério de aferição da culpa penal, pois quem domina a tecnologia deve responder pelos riscos que dela decorrem” (p. 225).

Essa ampliação interpretativa, contudo, deve respeitar o princípio da culpabilidade, que impede a punição sem demonstração de consciência e voluntariedade. O simples desenvolvimento de uma IA não é, por si só, ato criminoso, salvo se houver intenção ou descuido grave na prevenção de riscos. Barroso (2021) recorda que “o Direito Penal não pode se converter em instrumento de intimidação da ciência, mas deve servir de limite racional à irresponsabilidade tecnológica” (p. 89). Assim, a imputação penal deve ser restrita às condutas que revelem dolo ou negligência manifesta, de forma proporcional e fundamentada.

Já a responsabilidade das entidades tecnológicas decorre da posição de poder que ocupam no ecossistema digital. As grandes plataformas de IA detêm controle sobre os dados, os algoritmos e as condições de uso, o que lhes confere capacidade de prevenir danos. Doneda (2019) propõe a ideia de “responsabilidade informacional ampliada”, segundo a qual as empresas de tecnologia devem responder não apenas por omissão direta, mas também pela criação de ambientes de risco previsível (p. 131). Essa noção aproxima-se do conceito de dever de cuidado reforçado, aplicado a atividades potencialmente perigosas, como ocorre no transporte, na medicina ou na energia nuclear.

O Código Penal brasileiro, em seu artigo 13, §2º, estabelece que “a omissão é penalmente relevante quando o omitente devia e podia agir para evitar o resultado”. Esse dispositivo permite enquadrar as plataformas que se omitem no dever de prevenir o uso fraudulento de suas ferramentas. No caso das IAs generativas, a omissão pode configurar dolo eventual, caso se comprove que a empresa conhecia os riscos de uso indevido e, mesmo assim, deixou de implementar mecanismos de segurança. Essa linha de raciocínio encontra eco na jurisprudência internacional. Em 2024, o Tribunal de Justiça da União Europeia reconheceu a responsabilidade civil de uma empresa de IA por negligência na supervisão de conteúdos gerados automaticamente, estabelecendo precedente relevante para a discussão penal futura.

No Brasil, o debate ainda é incipiente, mas o Marco Civil da Internet (Lei nº 12.965/2014) e a Lei Geral de Proteção de Dados (Lei nº 13.709/2018) fornecem bases principiológicas. O artigo 3º do Marco Civil garante a liberdade de expressão e a neutralidade da rede, mas impõe, em seu artigo 19, o dever de cumprimento de ordens judiciais para remoção de conteúdo ilícito. Já a LGPD, em seu artigo 6º, introduz os princípios da responsabilidade e prestação de contas, determinando que os controladores de dados devem adotar medidas de segurança, técnicas e administrativas aptas a proteger informações pessoais. Essas normas, embora não penais, indicam uma direção clara: a responsabilidade informacional deve ser proporcional ao poder de controle e impacto social da atividade tecnológica.

A aplicação da responsabilidade penal às pessoas jurídicas é tema sensível. O artigo 225, §3º, da Constituição Federal prevê a possibilidade de responsabilização criminal de empresas por crimes ambientais. Essa lógica pode ser estendida, por analogia normativa, ao campo tecnológico, especialmente nos casos de omissão dolosa na prevenção de crimes digitais. Como defende Silva (2023), “a omissão corporativa diante de riscos previsíveis de dano informacional é tão grave quanto a ação direta do agente individual” (p. 176). A responsabilização penal das empresas tecnológicas, portanto, deve ser vista como instrumento de equilíbrio social e de incentivo à ética corporativa.

Por outro lado, é necessário evitar o uso simbólico do Direito Penal. A criminalização indiscriminada de condutas empresariais pode gerar insegurança jurídica e comprometer o desenvolvimento tecnológico. Mendes (2022) observa que “a punição sem critérios técnicos apenas substitui a impunidade pela irracionalidade punitiva” (p. 153). O desafio consiste em construir uma dogmática penal que diferencie a responsabilidade legítima da mera culpa social. Para isso, é essencial que a lei defina parâmetros claros de diligência, como auditorias obrigatórias, mecanismos de transparência e rastreabilidade de dados.

Outro ponto crítico é a cooperação internacional. Os crimes cometidos com IA frequentemente envolvem agentes e infraestruturas localizados em diferentes países. A inexistência de harmonização normativa dificulta a responsabilização penal. A Convenção de Budapeste sobre Crimes Cibernéticos (2001) é o principal

instrumento internacional nesse campo, mas ainda não contempla a inteligência artificial de forma expressa. A OCDE (2023) e a UNESCO (2024) têm incentivado a criação de um tratado global sobre ética e responsabilidade digital, que inclua diretrizes para imputação penal transnacional.

A construção de uma política penal global de governança algorítmica é, portanto, urgente. A inteligência artificial não reconhece fronteiras, e sua regulação isolada por cada Estado é insuficiente. O Brasil, ao integrar organismos internacionais e adotar políticas de convergência normativa com a União Europeia, pode assumir papel de liderança nesse debate. Barroso (2023) argumenta que “a globalização dos riscos tecnológicos exige uma constitucionalização dos deveres de cooperação entre os Estados” (p. 102). Essa perspectiva amplia o horizonte da responsabilidade penal, inserindo-a em um contexto de solidariedade internacional pela proteção dos direitos digitais e da confiança pública.

Em síntese, a responsabilidade penal no contexto da inteligência artificial generativa deve ser compreendida como fenômeno multidimensional, que envolve o indivíduo, a empresa e o Estado. O agente humano continua sendo o centro de imputação, mas sua culpa é medida pela extensão do poder tecnológico que exerce. As plataformas, por sua vez, devem responder quando falham em prevenir o uso criminoso de suas ferramentas. E o Estado, como garantidor da ordem jurídica, deve atuar de forma preventiva e cooperativa, assegurando que o progresso tecnológico não se converta em instrumento de destruição da própria racionalidade jurídica.

A inteligência artificial não elimina o sujeito de direito; ao contrário, o redefine. Ela desloca a fronteira da imputabilidade do gesto físico para a decisão técnica, do ato isolado para a cadeia de controle. A responsabilidade penal, nesse novo paradigma, deixa de ser apenas um julgamento sobre intenções e passa a ser uma análise sobre governança, risco e ética informacional. O Direito, se quiser permanecer humano, deve aprender a julgar não apenas quem age, mas também quem programa.

6 O PAPEL DO ESTADO E DAS INSTITUIÇÕES INTERNACIONAIS NA REGULAÇÃO PENAL DA INTELIGÊNCIA ARTIFICIAL

A discussão sobre o uso indevido de inteligências artificiais generativas nos crimes de estelionato não se encerra na esfera da responsabilidade individual ou corporativa. Ela exige um olhar macroestrutural, voltado ao papel do Estado e das instituições internacionais na criação de um ambiente normativo capaz de equilibrar liberdade tecnológica e segurança jurídica. A ausência de regulação penal específica, aliada à aceleração do desenvolvimento tecnológico, gera um vácuo institucional que fragiliza a proteção dos bens jurídicos fundamentais. O desafio contemporâneo consiste em construir um modelo regulatório que seja, simultaneamente, preventivo, proporcional e compatível com os princípios do Estado Democrático de Direito.

O Estado brasileiro, assim como outros ordenamentos jurídicos modernos, encontra-se em um momento de transição normativa. De um lado, há o reconhecimento da importância da inovação tecnológica e do estímulo à economia digital; de outro, surge a necessidade de limitar os abusos decorrentes da automação e da desinformação. Barroso (2023) argumenta que “a função moderna do Estado não é proibir a tecnologia, mas civilizá-la, submetendo-a à lógica da dignidade humana e da proporcionalidade constitucional” (p. 116). Essa concepção orienta uma atuação estatal equilibrada, capaz de promover o progresso científico sem abrir mão da proteção dos direitos fundamentais.

A atuação estatal no enfrentamento da criminalidade digital deve se estruturar em três eixos: a regulação normativa, a fiscalização institucional e a cooperação internacional. O primeiro eixo refere-se à criação de leis específicas que tratam da responsabilidade penal e civil pelo uso indevido de IA. O segundo envolve o fortalecimento de órgãos públicos especializados na prevenção e repressão de crimes cibernéticos. O terceiro exige o engajamento do Brasil em tratados e convenções que estabeleçam padrões globais de governança tecnológica.

No campo normativo, o Projeto de Lei nº 2.338/2023, em tramitação no Senado Federal, representa um esforço inicial para criar o Marco Legal da Inteligência Artificial no Brasil. A proposta, inspirada em modelos europeus e norte-americanos, estabelece princípios éticos como transparência, responsabilidade e segurança, além de prever categorias de risco para os sistemas de IA. Todavia, o texto ainda carece de dispositivos penais específicos. Doneda (2023) observa que “a ausência de previsão expressa de responsabilidade penal impede que o marco legal alcance sua plenitude protetiva” (p. 94). Essa lacuna reforça a necessidade de revisão legislativa, de modo a incluir sanções adequadas para condutas dolosas ou negligentes relacionadas ao uso indevido de IA.

A criação de estruturas estatais especializadas também é imperativa. O avanço dos crimes digitais exige uma polícia judiciária tecnicamente capacitada, peritos em segurança da informação e promotores com formação interdisciplinar em direito e tecnologia. A Polícia Federal, desde 2022, vem estruturando delegacias especializadas em crimes cibernéticos em todos os estados da federação, mas os recursos e a formação técnica ainda são insuficientes para acompanhar a complexidade das novas modalidades de estelionato digital. Mendes (2022) enfatiza que “sem aparelhamento técnico e formação continuada, o Estado permanece analógico em um mundo digital, sempre reagindo ao invés de prevenir” (p. 188).

A criação de um Observatório Nacional de Inteligência Artificial e Criminalidade Cibernética poderia ser um passo estratégico. Tal órgão, vinculado ao Ministério da Justiça, teria a função de monitorar o uso de IA em atividades ilícitas, emitir relatórios de risco e propor medidas preventivas. Além disso, seria responsável por articular políticas de integração entre o setor público, o setor privado e as universidades, promovendo uma abordagem multidisciplinar do problema. Modelos semelhantes já existem na União Europeia e nos Estados Unidos, com resultados positivos no monitoramento de ameaças emergentes.

O segundo eixo — a cooperação internacional — é indispensável diante da natureza transnacional dos crimes cometidos com IA. A Convenção de Budapeste sobre Crimes Cibernéticos (2001), à qual o Brasil aderiu em 2023, estabelece mecanismos de assistência mútua entre países na coleta de provas digitais e na

investigação de infrações cometidas por meios eletrônicos. Todavia, essa convenção ainda não abrange explicitamente a inteligência artificial. É urgente a ampliação de seus protocolos para incluir as novas formas de criminalidade algorítmica. A OCDE (2023) e a UNESCO (2024) têm incentivado a elaboração de um Tratado Internacional sobre Inteligência Artificial e Responsabilidade Penal, destinado a padronizar critérios de imputação, cooperação e punição entre os Estados signatários.

Nesse contexto, o Brasil pode desempenhar papel relevante como mediador entre os países em desenvolvimento e as nações tecnológicas avançadas. Sua tradição jurídica civilista, combinada com uma Constituição de perfil garantista, permite propor soluções normativas equilibradas, que conciliam inovação e tutela de direitos. Floridi (2021) destaca que “as democracias emergentes têm a oportunidade histórica de definir o rumo ético da inteligência artificial, antes que o domínio tecnológico se torne monopólio de poucos” (p. 133). Essa observação confere ao Brasil responsabilidade não apenas jurídica, mas civilizatória, na construção de uma regulação humanista da IA.

O papel das instituições internacionais também se mostra essencial. A União Europeia, por meio do *AI Act*, e a OCDE, com suas diretrizes sobre governança algorítmica, têm servido de modelo para o desenvolvimento de políticas públicas globais. Esses organismos propõem princípios como transparência, rastreabilidade e supervisão humana como bases de uma ética digital. O Banco Mundial e o Fórum Econômico Mundial também têm defendido a criação de instâncias supranacionais de fiscalização algorítmica, capazes de auditar o uso de IA em setores sensíveis, como finanças, segurança pública e comunicação social.

Além disso, a Organização das Nações Unidas vem discutindo a possibilidade de criação de uma Agência Global para Inteligência Artificial (AGIA), com função semelhante à da Agência Internacional de Energia Atômica. A proposta, apresentada em 2024, busca garantir que os avanços tecnológicos sejam empregados de forma segura, transparente e em conformidade com os direitos humanos. Essa iniciativa reflete o reconhecimento internacional de que a IA, assim como a energia nuclear ou

a biotecnologia, demanda regulação supranacional em razão de seu potencial de impacto sobre a humanidade.

No contexto interno, a efetividade das políticas públicas depende também da educação digital da população. A alfabetização tecnológica é medida preventiva indispensável para reduzir a vulnerabilidade das pessoas a golpes e manipulações com uso de IA. Programas de conscientização e capacitação, promovidos em parceria entre o Estado e a sociedade civil, devem fazer parte de uma estratégia nacional de segurança digital. Como alerta Castells (2022), “a ignorância tecnológica é o novo analfabetismo da era informacional, e dela nasce a exclusão e o engano” (p. 77).

Outro aspecto relevante é a regulação do setor privado. As empresas de tecnologia, especialmente aquelas que desenvolvem sistemas generativos, devem ser obrigadas a implementar mecanismos de governança algorítmica, que permitam rastrear o uso de suas ferramentas. Tais mecanismos podem incluir identificação obrigatória de conteúdo gerado por IA, auditorias independentes e sanções administrativas em caso de omissão. O Estado, nesse contexto, atua como garantidor da concorrência leal e da proteção do consumidor digital, prevenindo que a assimetria informacional entre empresas e usuários se converta em terreno fértil para o crime.

Por fim, é necessário compreender que a regulação penal da inteligência artificial não se resume à criação de leis punitivas, mas à consolidação de uma cultura jurídica da responsabilidade digital. O Direito Penal deve ser o último recurso, reservado para as condutas mais graves e dolosas, enquanto a regulação ética e preventiva deve ocupar o primeiro plano. Floridi (2021) sintetiza essa ideia ao afirmar que “a tecnologia precisa de menos medo e mais sabedoria, menos punição e mais prudência” (p. 142).

Em conclusão, o papel do Estado e das instituições internacionais na regulação penal da inteligência artificial é o de garantir que a inovação ocorra dentro de parâmetros éticos e democráticos. A cooperação global, o fortalecimento institucional e a educação digital são pilares inseparáveis de uma política criminal

voltada ao futuro. A inteligência artificial não é apenas um desafio jurídico, mas uma prova moral de nossa capacidade de construir um progresso que sirva à humanidade, e não que a substitua. Somente um Estado atento, uma sociedade consciente e uma comunidade internacional cooperativa poderão evitar que a genialidade tecnológica se converta em instrumento de destruição da própria racionalidade jurídica que sustenta o convívio civilizado.

7 CONCLUSÃO

A análise desenvolvida ao longo deste trabalho permitiu constatar que o uso indevido de inteligências artificiais generativas na prática de crimes de estelionato inaugura um novo capítulo na história do Direito Penal. A tecnologia, antes vista apenas como instrumento de trabalho e meio de comunicação, tornou-se também um agente potencial de fraude, manipulação e engano em escala global. A sofisticação dos sistemas generativos, capazes de imitar vozes, rostos, textos e comportamentos humanos com fidelidade quase absoluta, representa um desafio sem precedentes para a dogmática penal, exigindo uma revisão profunda de conceitos como autoria, dolo, culpa e nexo causal.

O estelionato, crime arquetípico da fraude, sofreu uma metamorfose estrutural na era digital. O engano, que antes dependia de habilidades retóricas ou de falsificações materiais, hoje é produzido por algoritmos treinados para manipular percepções humanas. Essa mutação não altera apenas a forma de execução do delito, mas também o próprio fundamento de sua reprovação social. A vítima do século XXI já não é ludibriada por um golpista de carne e osso, mas por uma interface aparentemente neutra, sustentada por códigos e servidores que transcendem fronteiras. Nesse sentido, a criminalidade algorítmica é um espelho do mundo globalizado: despersonalizada, automatizada e difícil de rastrear.

A dogmática penal tradicional, centrada no sujeito humano e na ação voluntária, revela-se insuficiente para lidar com as novas formas de autoria mediada por tecnologia. A teoria do domínio do fato, desenvolvida por Claus Roxin, fornece um

importante ponto de partida ao permitir a imputação de responsabilidade àquele que controla o processo causal, ainda que não o execute diretamente. Entretanto, diante da complexidade dos sistemas de IA, o domínio do fato precisa ser reinterpretado em chave tecnológica, levando em conta quem detém o controle do algoritmo, quem define seus parâmetros e quem se beneficia do resultado ilícito. Essa reinterpretação não significa romper com a tradição penal, mas adaptá-la às exigências de um novo contexto ontológico em que a ação humana se manifesta por meio da automação.

A teoria da imputação objetiva, formulada por Günther Jakobs, também oferece instrumentos para compreender o fenômeno. O programador, o operador ou o usuário que cria ou utiliza um sistema generativo sem mecanismos adequados de controle assume o risco de produzir danos jurídicos e, portanto, deve responder penalmente. A culpa, nesse novo paradigma, não é apenas moral, mas técnica: decorre da negligência em compreender, vigiar e limitar a tecnologia que se controla. Fiandaca e Musco reforçam essa ideia ao afirmar que “o dever de cuidado técnico é a fronteira entre a inovação responsável e o abuso tecnológico” (2021, p. 238).

A responsabilidade, portanto, deve ser entendida como multidimensional. O indivíduo que emprega a IA com dolo ou negligência responde pessoalmente; as empresas que desenvolvem e hospedam essas ferramentas respondem pela omissão e pelo risco que criam; e o Estado responde pela falha em legislar e fiscalizar. A tríplice dimensão da responsabilidade — individual, corporativa e estatal — é essencial para construir um sistema penal capaz de proteger a sociedade sem paralisar o avanço tecnológico. O Direito, afinal, não deve temer a inovação, mas guiá-la sob os princípios da dignidade humana e da proporcionalidade.

No plano legislativo, é urgente a criação de um marco penal específico para o uso indevido da inteligência artificial. A Lei nº 14.155/2021, que tipificou o estelionato eletrônico, foi um avanço, mas não abrange as condutas mediadas por sistemas autônomos e generativos. O Projeto de Lei nº 2.338/2023, ao propor o Marco Legal da Inteligência Artificial, representa um passo importante, mas precisa incorporar mecanismos de responsabilização penal e administrativa para evitar a impunidade de condutas praticadas por meio de algoritmos. Barroso (2023) lembra que “a norma

jurídica não pode ser uma fotografia do passado, mas um instrumento de leitura e antecipação do futuro" (p. 94).

A regulação internacional também se mostra indispensável. A OCDE, a União Europeia e a UNESCO vêm estabelecendo diretrizes éticas para o uso da IA, e o *AI Act* europeu de 2024 é um marco global nesse sentido. O Brasil, ao aderir à Convenção de Budapeste e participar ativamente desses debates, tem a oportunidade de assumir papel de liderança na América Latina. A harmonização normativa internacional é condição necessária para combater a criminalidade transnacional que utiliza IA, especialmente porque os sistemas de fraude operam em servidores distribuídos por diferentes países, dificultando a persecução penal isolada.

O estudo também evidencia que o combate ao uso indevido da IA deve ir além da repressão penal. A prevenção, a educação digital e a cooperação público-privada são pilares indispensáveis. O fortalecimento de instituições especializadas, como delegacias cibernéticas e núcleos de perícia digital, é urgente para que o Estado consiga investigar de forma eficaz os crimes de estelionato mediado por IA. Além disso, políticas de alfabetização tecnológica podem reduzir a vulnerabilidade da população aos golpes baseados em *deepfakes*, clonagem de voz e mensagens falsas.

No campo ético, a questão ultrapassa o direito positivo. Floridi (2021) observa que "a ética da informação é o novo horizonte da civilização, pois nela se decide não apenas o que as máquinas podem fazer, mas o que os humanos devem permitir que elas façam" (p. 153). Essa afirmação sintetiza a dimensão moral do problema: o Direito Penal deve proteger não apenas bens jurídicos, mas também a própria humanidade diante da automatização de suas relações. O crime mediado por IA é, em última instância, um crime contra a confiança, e a confiança é o alicerce invisível da vida social.

O estelionato algorítmico revela, portanto, uma crise civilizatória: a substituição da empatia pela simulação, da presença pela representação, da verdade pela verossimilhança. O desafio do jurista contemporâneo é restaurar a centralidade da

verdade e da responsabilidade na era das máquinas que mentem. Como afirmou Zuboff (2019), “a tecnologia que observa e imita os humanos pode também desumanizá-los, se não for contida por princípios éticos e democráticos” (p. 167).

Diante de tudo isso, a presente pesquisa reafirma que o Direito Penal brasileiro precisa evoluir para enfrentar a era da automação sem perder sua essência humanista. É necessário reconstruir a teoria da culpabilidade, reinterpretar o dolo sob a ótica da previsibilidade tecnológica e redefinir o bem jurídico do estelionato, ampliando-o para incluir a proteção da integridade informacional. A regulação penal deve ser acompanhada de políticas públicas de prevenção, regulação empresarial e cooperação internacional, de modo a construir uma rede protetiva que preserve tanto a segurança quanto a liberdade.

Em síntese, o uso indevido das inteligências artificiais generativas no crime de estelionato exige do Estado e da sociedade uma resposta inteligente, coerente e ética. Não basta criar leis punitivas: é preciso compreender que o problema é estrutural e envolve cultura, economia, ciência e moralidade. A IA é uma criação humana, e, como toda criação, reflete as virtudes e os vícios de quem a concebe. Cabe ao Direito assegurar que essa invenção, em vez de perpetuar a fraude, seja instrumento de justiça.

Conclui-se, portanto, que a tipificação penal e a responsabilização pelo uso indevido da inteligência artificial devem constituir um novo ramo de reflexão jurídica, marcado pela interdisciplinaridade e pela prudência. A tecnologia pode ser imprevisível, mas a ética e a lei não podem ser ambíguas. O Direito Penal deve reafirmar seu papel como guardião da confiança e da verdade, lembrando sempre que, mesmo na era dos algoritmos, a justiça continua a ser um ato profundamente humano.

REFERÊNCIAS

BARROSO, Luís Roberto. *A Dignidade da Pessoa Humana no Direito Constitucional Contemporâneo: A Construção de um Conceito Jurídico à Luz da Jurisprudência Mundial*. 6. ed. São Paulo: Saraiva, 2021.

BARROSO, Luís Roberto. *Constituição, Democracia e Direitos Fundamentais: Ensaios sobre o Constitucionalismo Contemporâneo*. 3. ed. São Paulo: Fórum, 2023.

BECCARIA, Cesare. *Dos Delitos e das Penas*. Tradução de Torrieri Guimarães. 5. ed. São Paulo: Martin Claret, 2003. (Obra original publicada em 1764).

BITENCOURT, Cezar Roberto. *Tratado de Direito Penal: Parte Geral*. 30. ed. São Paulo: Saraiva, 2022.

BRASIL. *Constituição da República Federativa do Brasil de 1988*. Brasília, DF: Senado Federal, 1988.

BRASIL. *Código Penal: Decreto-Lei nº 2.848, de 7 de dezembro de 1940*. Brasília, DF: Presidência da República, 1940.

BRASIL. *Lei nº 12.965, de 23 de abril de 2014*. 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. *Lei nº 14.155, de 27 de maio de 2021*. Altera o Código Penal e o Código de Processo Penal para prever o estelionato eletrônico. Diário Oficial da União, Brasília, DF, 28 maio 2021.

BRASIL. *Projeto de Lei nº 2.338/2023*. Dispõe sobre o Marco Legal da Inteligência Artificial no Brasil. Senado Federal, Brasília, DF, 2023.

CASTELLS, Manuel. *A Sociedade em Rede*. 21. ed. São Paulo: Paz e Terra, 2022.

DONEDA, Danilo. *Da Privacidade à Proteção de Dados Pessoais: Elementos da Formação da Autorregulação e da Regulação*. 3. ed. São Paulo: Revista dos Tribunais, 2019.

DONEDA, Danilo; SARLET, Ingo Wolfgang. *Direitos Fundamentais e Proteção de Dados Pessoais*. Porto Alegre: Livraria do Advogado, 2020.

EUROPEAN COMMISSION. *Artificial Intelligence Act*. Brussels: European Union, 2024.

EUROPEAN COMMISSION. *AI Liability Directive*. Brussels: European Union, 2023.

FIANDACA, Giovanni; MUSCO, Enzo. *Direito Penal: Parte Geral*. 8. ed. São Paulo: Revista dos Tribunais, 2021.

FLORIDI, Luciano. *A Ética da Informação*. Tradução de Cássio Luiz de Oliveira. 2. ed. Rio de Janeiro: Zahar, 2021.

JAKOBS, Günther. *Derecho Penal: Parte General*. 2. ed. Madrid: Marcial Pons, 2019.

MENDES, Gilmar Ferreira. *O Direito e a Inovação Tecnológica: Desafios da Inteligência Artificial*. 2. ed. São Paulo: Almedina, 2022.

ORGANIZAÇÃO PARA A COOPERAÇÃO E DESENVOLVIMENTO ECONÔMICO (OCDE). *OECD Principles on Artificial Intelligence*. Paris: OECD Publishing, 2023.

POLÍCIA FEDERAL. *Relatório de Inteligência Cibernética 2024: Panorama Nacional das Fraudes Digitais*. Brasília: Diretoria de Inteligência Policial, 2024.

ROXIN, Claus. *Autoría y Dominio del Hecho en Derecho Penal*. 8. ed. Madrid: Civitas, 2012.

SILVA, José Afonso da. *Responsabilidade Penal e Algoritmos: Perspectivas Críticas sobre o Direito Digital*. 2. ed. São Paulo: Atlas, 2023.

SILVA SÁNCHEZ, Jesús-María. *La Expansión del Derecho Penal: Aspectos de la Política Criminal en las Sociedades Postindustriales*. 4. ed. Madrid: Civitas, 2022.

SUNSTEIN, Cass R. *#Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press, 2018.

UNESCO. *Recomendação sobre a Ética da Inteligência Artificial*. Paris: Organização das Nações Unidas para a Educação, a Ciência e a Cultura, 2024.

ZUBOFF, Shoshana. *A Era do Capitalismo de Vigilância: A Luta por um Futuro Humano na Nova Fronteira do Poder*. Tradução de George Schlesinger. Rio de Janeiro: Intrínseca, 2019.