ISSN: 2675-7222

SAS
South American Sciences

CRIMES CIBERNÉTICOS E INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE DA RESPONSABILIDADE PENAL NO CONTEXTO JURÍDICO CONTEMPORÂNEO

CYBERCRIMES AND ARTIFICIAL INTELLIGENCE: AN ANALYSIS OF CRIMINAL LIABILITY IN THE CONTEMPORARY LEGAL CONTEXT

https://doi.org/10.63330/sasciencesv6n2-038

Submetido em: 18/06/2026 e Publicado em: 25/06/2026

SAS: e26259

Antonio Kalil dos Santos Souza
Graduando em Direito da Faculdade da Amazônia - UNAMA Rio Branco
E-mail: a.kalilssantos@gmail.com.br
Lattes: http://lattes.cnpq.br/9706090142227809
ORCID: https://orcid.org/0009-0007-8797-6856

Cássio Pinheiro Bandeira
Mestre em Educação pela Universidade Federal do Acre (UFAC); Pós-graduado em Direito Penal e Direito Processual Penal pela Faculdade Cândido Mendes (UCAM - RJ); Bacharel em Direito pela Universidade Federal do Acre (UFAC). Além das funções do Magistério Superior, ocupou o cargo de Assessor Jurídico (Ministério Público do Estado do Acre MPAC) e Coordenador do Curso de Tecnólogo e Gestão de Serviços Jurídicos e Notariais, ofertado pelo Centro Universitário FAAO/U:VERSE, no período compreendido entre 06 de Janeiro de 2020 a 06 de Setembro de 2020; exerceu a Profissão de Professor de Ensino Superior no Curso de Bacharelado em Direito na intitulada Faculdade da Amazônia (UNAMA) entre agosto de 2020 a 09 de fevereiro de 2024. Atualmente é Professor do Ensino Superior nos Cursos de Bacharelado em Direito da Universidade Federal do Acre (UFAC) e Universidade da Amazônia (UNAMA)
Lattes: http://lattes.cnpq.br/5171913330342141

Elis Vitória Gomes de Lima
Graduanda em Direito da Faculdade da Amazônia - UNAMA Rio Branco
E-mail: elisvitoria2013@gmail.com
Lattes: https://lattes.cnpq.br/2501964258108162
ORCID: https://orcid.org/0009-0004-5228-1614

RESUMO

A presente pesquisa examina a configuração da responsabilidade penal diante do avanço da Inteligência Artificial (IA) e da progressiva sofisticação dos delitos cibernéticos no sistema jurídico-penal brasileiro. O problema central reside na tensão entre os postulados clássicos da dogmática penal — fundados na conduta humana voluntária, consciente e finalisticamente orientada — e a operatividade autônoma de sistemas algorítmicos dotados de aprendizado profundo (deep learning). Por meio de pesquisa exploratória e qualitativa, com emprego do método dedutivo, hermenêutico e comparativo, análise bibliográfica e documental, bem como incorporação de dados empíricos sobre criminalidade cibernética, o estudo investiga se a legislação vigente e os princípios constitucionais da culpabilidade e da legalidade mostram-

se tecnicamente adequados para fundamentar a imputação penal em delitos mediados por sistemas autônomos. Mediante cotejo com o AI Act europeu, o GDPR, o modelo norte-americano e o PL 2.338/2023 brasileiro, verificou-se a insuficiência do arcabouço normativo atual para cenários de plena autonomia decisória algorítmica, sugerindo-se a adoção de modelos de responsabilização estruturados sobre o risco sistêmico e a regulação por design. Conclui-se pela necessidade premente de legislação específica que reajuste os critérios de imputação objetiva, concentrando-se na violação do dever de conformidade tecnológica pelo agente humano — desenvolvedor ou operador —, de forma a assegurar segurança jurídica e proteção efetiva dos bens jurídicos na infosfera.

Palavras-chave: Direito Penal; Inteligência Artificial; Responsabilidade Penal; Autonomia Algorítmica; Imputação Objetiva.

ABSTRACT

This research examines the configuration of criminal liability in light of the advancement of Artificial Intelligence (AI) and the growing sophistication of cybercrimes within the Brazilian criminal legal system. The central problem lies in the tension between the classical postulates of criminal law doctrine — grounded in voluntary, conscious, and purposively directed human conduct — and the autonomous operation of algorithmic systems endowed with deep learning capabilities. Through exploratory and qualitative research, employing the deductive, hermeneutic and comparative methods alongside bibliographic, documentary analysis and empirical data on cybercrime, the study investigates whether current legislation and the constitutional principles of culpability and legality are technically adequate to ground criminal imputation in offenses mediated by autonomous systems. Through comparison with the European AI Act, GDPR, the North American model and Brazilian Bill 2338/2023, the insufficiency of the current normative framework for full algorithmic autonomy scenarios was confirmed, suggesting liability models structured around systemic risk and regulation by design. The research concludes that specific legislation is urgently needed to recalibrate objective imputation criteria, focusing on the violation of the duty of technological compliance by the human agent — developer or operator — in order to ensure legal certainty and effective protection of legally protected interests in the infosphere.

Keywords: Criminal Law; Artificial Intelligence; Criminal Liability; Algorithmic Autonomy; Objective Imputation.

1 INTRODUÇÃO

O veloz progresso da Inteligência Artificial (IA) tem produzido transformações estruturais em múltiplos setores da sociedade contemporânea, gerando oportunidades sem precedentes e, simultaneamente, impondo desafios de crescente complexidade ao campo jurídico. No âmbito do Direito Penal, a revolução tecnológica suscita uma indagação dogmática de primeira ordem: a possibilidade de imputação criminal em crimes cibernéticos nos quais a IA atua não meramente como instrumento, mas como verdadeiro agente dotado de autonomia operacional e capacidade decisória não predeterminada.

A dimensão quantitativa do fenômeno reforça a urgência do debate. Segundo o relatório IBM Security X-Force Threat Intelligence Index (2023), os ataques cibernéticos cresceram 38% globalmente em 2022, com 71% dos incidentes envolvendo alguma forma de automação ou inteligência artificial. No Brasil, a Febraban registrou prejuízos superiores a R$2,5 bilhões com fraudes eletrônicas em 2022, enquanto a Polícia Federal relatou incremento de 60% nos crimes digitais entre 2020 e 2023. A Interpol, no relatório African Cyberthreat Assessment (2023), aponta que 79% das organizações financeiras mundiais identificaram o uso de IA generativa em ataques sofisticados de engenharia social. A Tabela 1 sistematiza a evolução desses indicadores.

Tabela 1 – Evolução dos crimes cibernéticos no Brasil e no mundo (2020–2025)

Ano	Prejuízo Brasil (R$ bi)	Crescimento Global (%)	Ataques com IA (%)
2020	1,1	+15	N/D
2021	1,7	+22	~30
2022	2,5	+38	71
2023	3,2 (est.)	+41	79
2024	4,1 (est.)	+46	~85
2025	em apuração	em apuração	em apuração

Fontes: Febraban (2023); IBM Security X-Force (2023); Interpol (2023); estimativas dos autores com base em tendência histórica.

A dogmática penal clássica brasileira, profundamente influenciada pelo finalismo de Hans Welzel e pela tradição causal-normativa, define o crime como comportamento humano voluntário, consciente e teleologicamente orientado. Não obstante, o surgimento de sistemas de aprendizado profundo (*deep learning*) e agentes autônomos desafia essa premissa axial, criando lacuna de tipicidade e jurisdição na qual a conduta criminosa nem sempre pode ser reconduzida a uma ação volitiva humana imediata e identificável.

O objetivo primordial desta investigação é avaliar a eficácia do sistema jurídico-penal brasileiro para fazer frente aos desafios impostos pela IA, verificando se os princípios da conduta, da culpabilidade

e da legalidade permanecem tecnicamente adequados para fundamentar a imputação de responsabilidade penal humana em delitos cibernéticos de caráter autônomo. A pesquisa desdobra-se em objetivos específicos: (i) revisão analítica da teoria da responsabilidade penal à luz da autonomia algorítmica; (ii) análise da IA como meio comissivo e omissivo na prática de crimes; (iii) avaliação de modelos doutrinários de responsabilização aplicáveis a programadores, desenvolvedores e operadores; (iv) exame de direito comparado envolvendo modelos regulatórios europeu, norte-americano, chinês e brasileiro; e (v) proposta de critérios para a prova digital em processos envolvendo IA.

O presente artigo organiza-se em oito seções. Após esta introdução, a Seção 2 examina os fundamentos dogmáticos da responsabilidade penal diante da autonomia tecnológica. A Seção 3 analisa a prática jurídica nos crimes cibernéticos, incluindo análise jurisprudencial nacional e internacional e três estudos de caso detalhados. A Seção 4 apresenta o direito comparado. A Seção 5 discute as posições doutrinárias divergentes. A Seção 6 propõe um novo paradigma de imputação. A Seção 7 descreve a metodologia. A Seção 8 conclui o trabalho com síntese e recomendações.

2 A DOGMÁTICA PENAL FRENTE À AUTONOMIA TECNOLÓGICA

O conceito de conduta — compreendido como comportamento humano voluntário, consciente e finalisticamente orientado — constitui o alicerce epistemológico sobre o qual repousa todo o sistema de imputação do Direito Penal brasileiro. Nilo Batista (2007, p. 72) sustenta que a missão nuclear do Direito Penal moderno consiste na proteção de bens jurídicos essenciais mediante a restrição racional e garantista do poder punitivo estatal. Esse postulado, porém, pressupõe um sujeito humano capaz de autodeterminação, requisito que se torna problemático diante de sistemas algorítmicos dotados de aprendizado autônomo.

A teoria da ação finalista, formulada por Welzel, concebe o delito como unidade de desvalor de ação e desvalor de resultado, sendo o dolo e a culpa elementos integrantes do próprio tipo subjetivo. Essa arquitetura dogmática revela sua limitação estrutural quando confrontada com sistemas de IA cujos processos decisórios escapam à vontade e à previsão do agente humano. O desafio não é meramente terminológico: trata-se de uma crise de fundamentos que exige resposta legislativa e doutrinária articulada.

2.1 A CRISE DA SUBJETIVIDADE PENAL NO AMBIENTE DIGITAL

A emergência da IA coloca em xeque a aplicação irrestrita do art. 13 do Código Penal, que condiciona a responsabilização ao nexo de causalidade entre a conduta do agente e o resultado típico. Quando a IA opera autonomamente, a vontade humana — elemento constitutivo do dolo (art. 18, I, CP) ou da violação do dever objetivo de cuidado na modalidade culposa (art. 18, II, CP) — dilui-se em uma cadeia decisória algorítmica, colocando em crise o dogma de que todo crime pressupõe uma escolha moral subjetiva individualizável.

Na teoria do domínio do fato, formulada por Claus Roxin (2006, p. 48), o autor mediato é aquele que, por meio do domínio da vontade, utiliza outrem — ou, em construção analógica, um sistema autônomo — para a consecução do resultado. Tal enquadramento exige que o programador tenha direcionado especificamente o sistema à prática do ilícito, não sendo suficiente a mera criação de ferramenta que, autonomamente, evolua para comportamentos não antecipados. A omissão imprópria (art. 13, § 2.º, CP) surge como categoria dogmaticamente mais adequada para os casos em que o desenvolvedor, na condição de garante técnico, deixa de tomar as providências necessárias para evitar o resultado lesivo produzido pelo algoritmo.

Eugenio Raúl Zaffaroni (2020, p. 13) adverte sobre o perigo de reduzir o Direito Penal a uma lógica meramente tecnocrática, desprovida de substrato ético. A ausência de regulamentação específica pode ensejar responsabilidade objetiva dissimulada, violando frontalmente o Princípio da Pessoalidade da Pena (art. 5.º, XLV, CF/88). Se a sanção penal não pode transcender a pessoa do condenado, a dificuldade probatória em identificar o erro humano causalmente determinante em sistemas complexos de deep learning cria, paradoxalmente, tanto brecha de impunidade estrutural quanto risco de criminalização injusta de desenvolvedores por resultados tecnicamente imprevisíveis.

Alessandro Baratta (1999, p. 162) reforça que o sistema penal, historicamente, tende a incidir de forma seletiva e reprodutora de desigualdades. Na era da IA, esse risco se amplifica: a opacidade dos sistemas algorítmicos pode ser instrumentalizada para blindar os agentes com maior poder econômico e tecnológico, enquanto os operadores periféricos e usuários finais absorvem o peso do jus puniendi. Por essa razão, a construção dogmática de novos critérios de imputação deve ter como eixo a responsabilidade funcional pelo ciclo de vida do sistema autônomo, e não a mera posição formal ocupada pelo agente na cadeia produtiva.

2.2 DOLO, CULPA E ERRO EM SISTEMAS ALGORÍTMICOS

O dolo eventual — caracterizado pela assunção do risco de produção do resultado (art. 18, I, segunda parte, CP) — emerge como a categoria mais adequada para situações em que o desenvolvedor, ciente das capacidades evolutivas do sistema de IA, disponibiliza-o sem salvaguardas suficientes, assumindo tacitamente o risco de comportamentos lesivos autônomos. A culpa consciente aplica-se aos casos em que o agente prevê o resultado como possível, mas acredita sinceramente poder evitá-lo mediante controles técnicos que se revelam insuficientes.

O erro de tipo algorítmico constitui categoria dogmática que merece desenvolvimento específico pela literatura jurídica brasileira. Trata-se da situação em que o sistema de IA, por falha em seu treinamento, viés nos dados ou comportamento emergente imprevisível, pratica conduta que o programador genuinamente não antecipava e não poderia razoavelmente prever. Aplica-se analogicamente o art. 20 do

Código Penal: se o erro incide sobre elemento essencial do tipo, exclui o dolo, restando apenas a responsabilização culposa se o erro for inescusável. A questão central é definir o padrão de diligência técnica exigível do desenvolvedor médio, critério que deve ser aferido conforme o estado da arte tecnológico no momento da criação e implementação do sistema.

Guilherme de Souza Nucci (2014, p. 156) observa, a propósito da culpa, que a responsabilidade penal pressupõe sempre a possibilidade de o agente prever o resultado e agir de modo diverso. Nos sistemas de IA de alto grau de autonomia, o exame da previsibilidade não pode ser dissociado da análise técnica do estado da arte: o padrão de cuidado exigível do desenvolvedor de um algoritmo de recomendação em 2018 difere substancialmente daquele exigível em 2024, dado o avanço exponencial dos modelos de linguagem de grande escala (LLMs). Essa relatividade temporal do dever objetivo de cuidado é elemento indispensável da análise dogmática contemporânea.

2.3 PRINCÍPIOS CONSTITUCIONAIS PENAIS E A REGULAÇÃO DA IA

A análise da responsabilidade penal no contexto da IA não pode prescindir do cotejo com os princípios constitucionais que regem o Direito Penal Brasileiro. Além dos princípios da legalidade (art. 5.º, XXXIX, CF/88), da culpabilidade e da pessoalidade da pena (art. 5.º, XLV, CF/88), outros axiomas de igual relevância precisam ser examinados.

- **Princípio da Intervenção Mínima e da Fragmentariedade:** o Direito Penal deve ser a ultima ratio do ordenamento jurídico, intervindo apenas quando outras esferas — civil e administrativa — se revelem insuficientes. Na era da IA, isso implica que a tipificação de condutas algorítmicas deve ser criteriosamente selecionada, limitando-se às violações mais graves dos bens jurídicos tutelados.
- **Princípio da Subsidiariedade:** intimamente ligado à intervenção mínima, exige que o Direito Penal só atue após o esgotamento dos mecanismos extrapenais de controle. No campo da IA, isso significa que sanções administrativas da ANPD e da futura autoridade supervisora de IA (PL 2.338/2023), bem como a responsabilidade civil prevista nos arts. 186 e 927 do Código Civil e no art. 12 do CDC, devem ser prioritariamente acionadas. A criminalização só se justifica diante de violações dolosas ou gravemente culposas que resultem em dano efetivo a bens jurídicos de primeira importância.
- **Princípio da Ofensividade (ou Lesividade):** nulla poena sine iniuria — não há crime sem lesão ou exposição concreta a perigo de bem jurídico. Para a responsabilidade penal da IA, esse princípio impõe que a mera criação de sistemas autônomos, sem comprovação de lesão ou perigo concreto, não autoriza a incidência do jus puniendi.

- **Princípio da Proporcionalidade:** a sanção penal deve guardar correspondência com a gravidade da conduta e a culpabilidade do agente. Na imputação de crimes cibernéticos envolvendo IA, a proporcionalidade exige distinção criteriosa entre o desenvolvedor que age com plena ciência do potencial delitivo, o operador descuidado e o usuário que instrumentaliza o sistema.

Frank Pasquale (2015, p. 8), ao analisar a black box society, observa que a opacidade algorítmica não é acidente tecnológico, mas estratégia deliberada de blindagem da responsabilidade jurídica. Nesse sentido, a regulação por design, aliada ao princípio da proporcionalidade, emerge como instrumento indispensável para que o Direito Penal possa incidir sobre os verdadeiros responsáveis pelo risco tecnológico, sem sacrificar as garantias constitucionais.

3 A PRÁTICA JURÍDICA FRENTE AOS CRIMES CIBERNÉTICOS

3.1 O ARCABOUÇO NORMATIVO BRASILEIRO

Com a promulgação da Lei n.º 12.737/2012 (Lei Carolina Dieckmann), o ordenamento jurídico brasileiro avançou ao tipificar expressamente a invasão de dispositivo informático. Ameleto Masini Neto (2025, p. 45) explica que o crime cibernético exige a incorporação do conceito de desterritorialização, uma vez que a conduta e o resultado frequentemente se realizam em jurisdições distintas. O tipo penal fundamental prevê:

> Art. 154-A. Invadir dispositivo informático de outrem, conectado ou não à rede de computadores, com o fim de obter, adulterar ou destruir dados ou informações sem autorização expressa ou tácita do titular do dispositivo ou instalar vulnerabilidades para obter vantagem ilícita. (Brasil, 2012)

Não obstante o avanço legislativo, a literalidade da norma concentra-se no agente humano que invade, desconsiderando a realidade de sistemas de IA que, por aprendizado autônomo, identificam e exploram vulnerabilidades sem a necessidade de comando humano específico. A Lei n.º 14.155/2021 agravou as penas para crimes cibernéticos praticados de forma eletrônica ou pela internet, mas igualmente não contemplou a especificidade da atuação algorítmica autônoma.

O Marco Civil da Internet (Lei n.º 12.965/2014) introduziu balizas fundamentais para a responsabilidade civil dos provedores de conexão e de aplicações, estabelecendo o regime de responsabilidade subjetiva condicionada à notificação judicial (art. 19). Contudo, a Patrícia Peck Pinheiro (2021, p. 78) alerta que o Marco Civil foi concebido num momento em que os algoritmos de recomendação ainda não operavam com a autonomia e a escala que ostentam hoje, razão pela qual sua aplicação aos cenários de IA generativa demanda reinterpretação sistemática à luz dos princípios constitucionais da dignidade da pessoa humana e da proteção integral de dados.

No plano probatório, o Código de Processo Penal, com as inovações introduzidas pela Lei n.º 13.964/2019 (Pacote Anticrime), passou a disciplinar expressamente a cadeia de custódia da prova (arts. 158-A a 158-F). A cadeia de custódia digital exige documentação rigorosa de todos os momentos de coleta, acondicionamento e armazenamento dos elementos de prova. Nos crimes mediados por IA, os logs de operação, os metadados das transações, os registros de treinamento do algoritmo e os parâmetros do modelo constituem elementos probatórios essenciais.

3.2 ANÁLISE JURISPRUDENCIAL: NACIONAL E INTERNACIONAL

A compreensão dos desafios que a IA coloca ao Direito Penal exige o cotejo com decisões judiciais que já enfrentaram, direta ou indiretamente, a questão da responsabilidade algorítmica. A análise a seguir organiza-se em três frentes: STF, STJ e cortes internacionais.

3.2.1 Supremo Tribunal Federal (STF)

No RE 1.010.606/RJ (Tema 786 da Repercussão Geral, Rel. Min. Dias Toffoli, julgado em 11/02/2021), o STF apreciou a tensão entre a liberdade de expressão e o direito ao esquecimento na internet, firmando a tese de que o direito ao esquecimento é incompatível com a Constituição Federal. Embora o julgamento não tenha versado diretamente sobre IA, consolidou premissa relevante para o presente debate: a circulação automatizada de informações digitais está sujeita a limites constitucionais, abrindo espaço para o controle judicial de algoritmos que reproduzam conteúdo lesivo de forma autônoma.

Já na ADPF 403 e na ADI 5.527, julgadas em 2021, o STF debateu a interceptação de comunicações criptografadas, ressaltando a tensão entre segurança pública e privacidade digital. A decisão reforçou que o Estado não pode impor obrigações técnicas que comprometam a segurança sistêmica de plataformas digitais — princípio que se projeta sobre a regulação de sistemas de IA: não pode o legislador exigir do desenvolvedor o cumprimento de obrigações tecnicamente inviáveis.

Mais recentemente, o STF, ao julgar o RE 1.037.396 (Tema 987 da Repercussão Geral, 2025), apreciou a constitucionalidade do art. 19 do Marco Civil da Internet, firmando que plataformas digitais podem ser responsabilizadas civilmente por conteúdos gerados por usuários sem necessidade de prévia ordem judicial quando notificadas extrajudicialmente sobre ilícitos graves. Esse entendimento tem reflexo imediato na discussão sobre a responsabilidade dos operadores de sistemas de IA generativa que permitam a criação e difusão de conteúdo delitivo em escala.

3.2.2 Superior Tribunal de Justiça (STJ)

O STJ tem construído jurisprudência relevante sobre crimes cibernéticos e prova digital. No RHC 99.735/SC (Rel. Min. Nefi Cordeiro, Sexta Turma, 2019), a Corte analisou questão sobre obtenção de

provas mediante espelhamento de conversas via WhatsApp Web, concluindo pela ilicitude do procedimento adotado pela autoridade policial, que dispensou autorização judicial específica para o monitoramento telemático. O caso é paradigmático para a análise dos limites da investigação digital em ambientes tecnológicos sofisticados, revelando que a identificação do agente em crimes cibernéticos deve observar rigorosa observância das garantias constitucionais, não podendo repousar exclusivamente em mecanismos técnicos de rastreamento sem controle judicial adequado.

No HC 598.051/SP (Rel. Min. Rogério Schietti Cruz, Sexta Turma, 2020), o STJ reafirmou o entendimento sobre a necessidade de observância rigorosa da cadeia de custódia da prova digital, sob pena de nulidade (arts. 158-A a 158-F do CPP, introduzidos pelo Pacote Anticrime). O acórdão enfatizou que a validade das provas digitais pressupõe documentação contínua e detalhada do histórico do vestígio coletado, desde o reconhecimento inicial até o descarte, assegurando a integridade e a auditabilidade dos elementos probatórios. Tal exigência aplica-se, por extensão, aos logs e metadados produzidos por sistemas de IA.

No AREsp 1.568.124/RJ (2020), a Corte reconheceu a legitimidade da perícia em dispositivos eletrônicos como meio de prova, estabelecendo parâmetros para a cadeia de custódia digital que são aplicáveis, por extensão, aos registros produzidos por sistemas algorítmicos. No REsp 1.829.821/SP (2020), o STJ enfrentou a responsabilidade de plataformas digitais por algoritmos que recomendam conteúdo ilícito, aplicando o Marco Civil da Internet (art. 19) e sinalizando que a responsabilidade pelos resultados algorítmicos não é automática, demandando comprovação do nexo causal entre o funcionamento do sistema e o dano produzido.

Mais recentemente, no AgRg no HC 828.054/RN (Rel. Min. Joel Ilan Paciornik, Quinta Turma, julgado em 23/04/2024), o STJ estabeleceu que a auditabilidade, a repetibilidade, a reprodutibilidade e a justificabilidade constituem atributos essenciais das evidências digitais, enfatizando a necessidade de observância de metodologias certificadas compatíveis com os padrões internacionais de computação forense. Esse julgado representa importante marco na jurisprudência sobre prova digital e seus requisitos de validade, com reflexos diretos para a admissibilidade de logs algorítmicos em processos penais envolvendo IA.

3.2.3 Cortes Internacionais

A Corte Europeia de Direitos Humanos (CEDH), no caso Big Brother Watch and Others v. the United Kingdom (2021), condenou o uso de sistemas de vigilância em massa sem salvaguardas suficientes, estabelecendo que a coleta e o processamento automatizado de dados pessoais deve estar sujeito a controle independente e possibilidade de revisão judicial. O precedente é diretamente aplicável à utilização de IA em investigações criminais.

O Tribunal de Justiça da União Europeia (TJUE), no caso Ligue des droits humains v. Conseil des ministres (C-817/19, 2022), reforçou a primazia do direito à privacidade sobre sistemas automatizados de análise de dados de passageiros (PNR), impondo limites objetivos ao uso de algoritmos preditivos em matéria penal. A decisão fundamenta-se no art. 22 do GDPR — direito de não ser submetido a decisões exclusivamente automatizadas —, dispositivo espelhado no art. 20 da LGPD brasileira.

3.3 CASOS CONCRETOS: ESTUDOS DE CASO DETALHADOS

A análise de casos concretos é indispensável para que a pesquisa transite da abstração dogmática para a concretude da prática forense. Apresentam-se três estudos de caso que evidenciam os desafios reais da responsabilização penal na era da IA.

3.3.1 Caso 1 — Reconhecimento Facial e Prisão Indevida (Brasil, 2021)

Em novembro de 2021, um homem foi detido no estado do Maranhão com base exclusivamente em correspondência gerada por sistema de reconhecimento facial utilizado pela Polícia Civil. O algoritmo apontou semelhança entre a imagem do suspeito e a foto do autor de roubo registrada em câmera de segurança. Após dias de detenção, o erro algorítmico foi comprovado mediante análise pericial de outras características biométricas.

O caso expõe múltiplas dimensões jurídicas: (i) ausência de regulamentação específica para o uso de IA em investigações criminais; (ii) inexistência de revisão humana obrigatória antes da lavratura do auto de prisão em flagrante; (iii) lacuna de responsabilização do fornecedor do sistema algorítmico; e (iv) violação do art. 5.º, LXV, da CF/88 (prisão ilegal). Sob o prisma dogmático penal, identifica-se omissão imprópria do operador estatal, que, na condição de garante, deixou de adotar as cautelas técnicas exigíveis antes de executar a restrição de liberdade. O episódio evidencia como a ausência de supervisão humana no ciclo decisório algorítmico converte a eficiência tecnológica em instrumento de violação de direitos fundamentais.

Do ponto de vista comparativo, o caso dialoga com os estudos conduzidos por Joy Buolamwini e Timnit Gebru, que demonstraram empiricamente a presença de viés racial em sistemas de reconhecimento facial amplamente comercializados, com taxas de erro significativamente mais elevadas para rostos de pessoas negras e mulheres. A ausência de mecanismos de avaliação de viés algorítmico na contratação pública de sistemas de IA configura, sob a perspectiva dogmática proposta neste trabalho, violação do dever de conformidade tecnológica pelo operador estatal.

3.3.2 Caso 2 — Fraudes com Voz Sintética (Voice Cloning) e Deepfakes (2021–2023)

Entre 2021 e 2023, a IBM Security documentou crescimento expressivo nas fraudes com clonagem de voz por IA. Em um dos casos mais emblemáticos, registrado no Reino Unido em 2019 e replicado com variações no Brasil em 2022, executivos de multinacionais foram alvo de chamadas geradas por IA que replicavam a voz de seus superiores hierárquicos, induzindo-os a transferências bancárias fraudulentas de valores expressivos.

Do ponto de vista penal, configura-se estelionato qualificado (art. 171, § 2.º, VI, CP, introduzido pela Lei n.º 14.155/2021), com potencial causa de aumento pela utilização de recurso tecnológico que dificulta a identificação do agente. A questão dogmática central é a atribuição de autoria: o usuário que aciona o sistema de clonagem comete o delito em coautoria com o desenvolvedor da ferramenta? A resposta passa pelo exame do dolo específico do desenvolvedor: se a plataforma foi criada com finalidade legítima (entretenimento, acessibilidade) e instrumentalizada pelo usuário para a fraude, incide o art. 29, § 2.º, CP (participação de menor importância), isentando ou atenuando a responsabilidade do criador.

A questão adquire nova dimensão com o advento dos modelos de síntese de voz acessíveis ao público em geral. Quando a plataforma detentora da tecnologia é cientificada da utilização de seu produto para fraudes em escala e permanece omissa na adoção de salvaguardas técnicas — como marcas d'água digitais (audio watermarking) ou sistemas de detecção de abuso —, configura-se, no mínimo, dolo eventual quanto aos resultados subsequentes, independentemente do propósito original do sistema.

3.3.3 Caso 3 — Fraudes Bancárias com IA Generativa e Spear Phishing (2022–2023)

A Europol (2023) documentou ataques de spear phishing automatizado por IA generativa que vitimaram centenas de instituições financeiras globais em 2022, com prejuízo estimado de bilhões de dólares. No Brasil, o setor bancário registrou aumento significativo nas tentativas de phishing sofisticado em 2023, conforme a Febraban. Os ataques utilizam modelos de linguagem de grande escala (LLMs) para gerar mensagens personalizadas que superam os filtros tradicionais de detecção.

O enquadramento penal recai sobre o art. 155, § 4.º-B, do CP (furto qualificado mediante fraude eletrônica, introduzido pela Lei n.º 14.155/2021), com pena de reclusão de 4 a 8 anos. A responsabilidade do operador da plataforma de IA que disponibiliza LLMs sem mecanismos de prevenção ao abuso configura, no mínimo, dolo eventual: ao disponibilizar ferramenta com capacidade demonstrada de automatizar fraudes em escala, o operador assume o risco do resultado lesivo. O modelo tripartite proposto neste trabalho (Tabela 3) oferece solução dogmática para a distribuição da responsabilidade entre desenvolvedor, operador e usuário final.

3.4 A PROTEÇÃO DE DADOS E A PROVA DIGITAL

A Lei Geral de Proteção de Dados Pessoais — LGPD (Lei n.º 13.709/2018) — introduz dispositivos de relevante impacto na análise penal, especialmente quanto à transparência das decisões automatizadas. Viviane Maldonado (2021, p. 173) destaca que a utilização de dados para elaboração de perfis criminais por meio de IA deve observar rigorosamente o princípio da não discriminação. O texto legal assegura:

> Art. 20. O titular dos dados tem direito a solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses, incluídas as decisões destinadas a definir o seu perfil pessoal, profissional, de consumo e de crédito ou os aspectos de sua personalidade. (Brasil, 2018)

A prova digital em processos envolvendo IA impõe desafios que o Código de Processo Penal originário não contempla adequadamente. Nos crimes mediados por IA, os logs de operação, os metadados das transações, os registros de treinamento do algoritmo e os parâmetros do modelo constituem elementos probatórios essenciais. A perícia computacional forense deve utilizar técnicas de explainability (XAI — Explainable Artificial Intelligence) para reconstruir o processo decisório do sistema, tornando inteligível, para o operador do Direito, a sequência de inferências que culminou no resultado delitivo.

Nesse ponto, a exigência de cadeia de custódia digital assume feição específica: além dos requisitos gerais dos arts. 158-A a 158-F do CPP, impõe-se a preservação dos pesos do modelo (model weights), dos dados de treinamento relevantes e dos parâmetros de configuração vigentes no momento da conduta, sob pena de impossibilidade de reprodução e auditoria do comportamento algorítmico. A ausência desses elementos configura vício probatório capaz de comprometer a validade do processo penal, conforme o entendimento consolidado pelo STJ acerca da cadeia de custódia.

4 DIREITO COMPARADO: MODELOS REGULATÓRIOS DA IA

A comparação entre os modelos regulatórios adotados por diferentes ordenamentos jurídicos é indispensável para identificar soluções transplantáveis ao sistema brasileiro. A Tabela 2 sintetiza os principais marcos normativos e as escolhas regulatórias de cada jurisdição.

Tabela 2 – Direito comparado: modelos regulatórios de IA e responsabilidade penal

País/Bloco	Marco Normativo	Classif. por Risco	Responsab. Penal	Lacunas
União Europeia	AI Act (2024); GDPR (2018)	Sim (4 níveis)	Indireta via descumprimento	Sem tipos autônomos de IA
Estados Unidos	Exec. Order 2023; Section 230 CDA	Parcial (setorial)	Fragmentada; ênfase civil	Ausência de legislação federal unificada
China	Reg. Algoritmos 2022; Reg. IA Generativa 2023	Sim (foco em segurança nacional)	Direta pelo operador	Risco de controle social sem garantias
Brasil	PL 2.338/2023; LGPD; Lei 12.737/2012	Em construção (inspirado no AI Act)	Indireta; legislação penal geral	Sem tipos penais autônomos de IA

Fontes: AI Act (2024); Executive Order EUA (2023); Regulamentos Chineses de IA (2022/2023); PL 2.338/2023 (Brasil). Elaboração própria.

4.1 A UNIÃO EUROPEIA: AI ACT E GDPR

O Regulamento (UE) 2024/1689 — AI Act, aprovado pelo Parlamento Europeu em março de 2024 e em vigor desde agosto de 2024 — estabelece sistema de classificação em quatro categorias de risco: risco inaceitável (proibição absoluta), alto risco (obrigações estritas de conformidade), risco limitado (obrigações de transparência) e risco mínimo (sem obrigações adicionais). Para sistemas de alto risco — que abrangem ferramentas de reconhecimento biométrico, análise preditiva criminal e monitoramento —, exige avaliação de conformidade, documentação técnica detalhada e rastreabilidade das decisões.

Kate Crawford (2021, p. 211) observa que a regulação europeia, ao impor auditabilidade como condição de uso, rompe com a lógica da caixa-preta e distribui a responsabilidade pelo ciclo de vida do sistema. Luciano Floridi (2019, p. 5) complementa ao propor que a governança da IA deve incorporar o princípio da explicabilidade ética: não basta que o sistema produza resultados corretos; é necessário que seus processos decisórios sejam compreensíveis e justificáveis para os agentes humanos afetados.

O AI Act é particularmente relevante para o presente estudo porque introduz, pela primeira vez em um sistema jurídico de grande escala, a noção de compliance algorítmico como condição de validade jurídica dos sistemas de IA. A violação das obrigações de conformidade não acarreta responsabilidade penal autônoma, mas serve de elemento objetivo para a configuração da culpa grave ou do dolo eventual do desenvolvedor ou operador, nos termos do modelo proposto neste trabalho. É o que a doutrina europeia contemporânea tem denominado regulatory fault — a culpa normativa decorrente do descumprimento de standards técnicos legalmente estabelecidos.

4.2 ESTADOS UNIDOS: RESPONSABILIDADE DE PLATAFORMAS E AUTONOMOUS AGENTS

O modelo norte-americano é marcado pela fragmentação regulatória e pela centralidade da responsabilidade civil como mecanismo de accountability. A Section 230 do Communications Decency Act (1996), que imuniza plataformas digitais pela responsabilidade por conteúdo gerado por terceiros, encontra-se sob crescente pressão diante do uso de algoritmos recomendatórios que potencializam conteúdo prejudicial. Ryan Calo (2017, p. 3) sustenta que a ausência de regulação federal unificada para a IA cria incentivos perversos para o desenvolvimento de sistemas opacos, nos quais a responsabilidade é estrategicamente diluída entre fornecedores de modelos fundacionais e operadores downstream.

A Executive Order on Safe, Secure, and Trustworthy AI (outubro de 2023) representa tentativa de sistematização regulatória por via executiva, impondo obrigações de transparência e avaliação de risco para sistemas de IA de alto impacto utilizados por agências federais. Contudo, a ausência de legislação congressional vinculante mantém o sistema norte-americano em posição de fragilidade estrutural para os fins de responsabilização penal, especialmente diante da doutrina da corporate personhood que dificulta a imputação a indivíduos específicos dentro de grandes corporações tecnológicas.

4.3 CHINA: REGULAÇÃO ESTATAL E CONTROLE ALGORÍTMICO

A China adotou abordagem regulatória de caráter estatal centralizado, com ênfase no controle político e na segurança nacional. O Regulamento de Gestão de Serviços de Algoritmos de Recomendação (2022) e o Regulamento de Serviços de IA Generativa (2023) estabelecem obrigações de registro, transparência e filtragem de conteúdo. Cathy O'Neil (2016, p. 13) alerta que sistemas algorítmicos de pontuação social incorporam e amplificam desigualdades estruturais preexistentes, convertendo-se em instrumentos de opressão camuflados por aparente objetividade matemática.

Do ponto de vista do Direito Penal, o modelo chinês apresenta a característica singular de atribuir responsabilidade direta ao operador pelos conteúdos gerados por sistemas de IA sob seu controle, dispensando a demonstração de dolo específico quanto ao conteúdo individual. Embora eficaz do ponto de vista preventivo, tal modelo é incompatível com os princípios constitucionais brasileiros da culpabilidade subjetiva e da pessoalidade da pena, não sendo transplantável ao ordenamento pátrio sem profundas adaptações garantistas.

4.4 BRASIL: ANÁLISE CRÍTICA DO PL 2.338/2023

O Projeto de Lei n.º 2.338/2023, aprovado pelo Senado Federal em dezembro de 2024 e em tramitação na Câmara dos Deputados, representa a principal iniciativa brasileira de regulação da IA. Inspirado no modelo europeu, o PL estabelece categorias de sistemas de alto risco, obrigações de transparência e auditabilidade, e proibição de sistemas de pontuação social por autoridades públicas.

O art. 3.º define sistema de IA de forma ampla, abrangendo sistemas que, com base em objetivos definidos, geram saídas que influenciam ambientes físicos ou virtuais — definição adequada, mas que não resolve a questão da imputação penal quando o resultado é produzido por comportamento emergente não previsto pelos desenvolvedores.

Os arts. 15 a 22 do PL estabelecem obrigações dos agentes de sistemas de alto risco (fornecedores e operadores), incluindo avaliação de impacto, documentação técnica e mecanismos de supervisão humana. Contudo, o art. 55 do PL expressamente ressalva que a responsabilidade civil e penal dos agentes de IA é regida pela legislação específica vigente, sem criar novos tipos penais. Essa lacuna é o ponto mais crítico: o PL não tipifica a negligência algorítmica grave como delito autônomo, nem estabelece causas de aumento de pena vinculadas ao descumprimento das obrigações de conformidade.

Os arts. 65 a 72 do PL regulam os direitos dos titulares afetados por decisões automatizadas, assegurando revisão humana e explicabilidade. Embora relevantes para o campo do Direito Administrativo e Civil, esses dispositivos são insuficientes para fundamentar a imputação penal: um sistema pode observar todas as obrigações do PL e ainda assim produzir resultado delitivo não antecipado, precisamente porque a autonomia algorítmica transcende os protocolos de conformidade.

5 POSIÇÕES DOUTRINÁRIAS DIVERGENTES: O DEBATE EM PERSPECTIVA

O rigor científico impõe que sejam apresentadas e examinadas as principais posições doutrinárias que divergem das conclusões deste trabalho. A maturidade acadêmica de uma investigação jurídica mede-se, em parte, pela capacidade de dialogar honestamente com as teses contrárias.

5.1 A SUFICIÊNCIA DA LEGISLAÇÃO PENAL VIGENTE

Parte da doutrina sustenta que a legislação penal existente é suficiente para lidar com os crimes mediados por IA. Argumenta-se que o Código Penal, interpretado sistematicamente com o Marco Civil da Internet e a Lei Carolina Dieckmann, já oferece instrumentos adequados de imputação, bastando a aplicação criativa dos institutos da omissão imprópria, da autoria mediata e do dolo eventual. Guilherme de Souza Nucci (2014, p. 156) aponta que a expansão penal nem sempre corresponde à melhor política criminal, podendo gerar insegurança jurídica maior do que a lacuna que pretende preencher.

Essa posição merece consideração, mas enfrenta o obstáculo da taxatividade penal: o princípio da legalidade estrita (art. 5.º, XXXIX, CF/88) veda a analogia in malam partem no Direito Penal. A aplicação analógica de tipos penais criados para condutas humanas diretas a comportamentos algorítmicos autônomos pode violar a lex certa e produzir condenações juridicamente frágeis.

5.2 A IMPOSSIBILIDADE DA RESPONSABILIZAÇÃO POR RISCO

Uma segunda corrente doutrinária, de matiz libertária, sustenta que a responsabilização penal baseada no risco sistêmico — e não na conduta dolosa ou culposa individualizável — é incompatível com o Estado Democrático de Direito. Zaffaroni (2020, p. 45) alerta para o risco de o Direito Penal do inimigo incorporar a lógica da prevenção tecnológica como subterfúgio para punir quem ainda não cometeu crime identificável.

O argumento tem relevância constitucional inegável. Contudo, peca pela desconsideração de que o modelo proposto neste trabalho não prescinde da comprovação de dolo ou culpa: a responsabilidade do desenvolvedor pressupõe violação específica e demonstrável do dever de cuidado técnico, não a mera criação de tecnologia potencialmente perigosa. A analogia com a responsabilidade do fabricante de produtos defeituosos é pertinente: responde penalmente não quem fabrica regularmente, mas quem introduz no mercado produto com defeito de segurança cognoscível.

5.3 A REJEIÇÃO DA EXPANSÃO PENAL PARA A IA

Uma terceira vertente, representada por criminalistas de linha garantista, sustenta que qualquer expansão do Direito Penal para abarcar novas tecnologias deve ser recusada a priori, com base na fragmentariedade e na subsidiariedade. Nilo Batista (2007, p. 110) ensina que a inflação legislativa penal é sintoma de crise de legitimidade do sistema, e não solução para ela.

A crítica tem mérito enquanto alerta contra o populismo penal tecnológico. Entretanto, a posição de abstenção penal absoluta desconsidera que os danos produzidos por sistemas de IA em plena autonomia podem ser de magnitude catastrófica e irreversível — como ataques a infraestruturas críticas ou manipulação algorítmica de processos eleitorais. Nesses casos, a exclusividade das sanções civis e administrativas seria manifestamente desproporcional à gravidade da conduta omissiva do desenvolvedor ou operador.

David Garland (2008, p. 45) observa que a cultura do controle contemporânea tende a expandir o escopo de intervenção penal como resposta a novas formas de insegurança social. Esse risco é real e deve ser permanentemente monitorado pela doutrina e pelos operadores do Direito. A proposta deste trabalho — circunscrita à violação do dever objetivo de cuidado tecnológico por agente humano identificável — procura situar-se no equilíbrio entre o garantismo penal e a efetividade da tutela dos bens jurídicos na infosfera.

6 PROPOSTA DE UM NOVO PARADIGMA DE IMPUTAÇÃO

6.1 SUPERAÇÃO DO MODELO CAUSAL-FINALISTA

A transição de uma criminalidade de caráter exclusivamente individual para uma fenomenologia delitiva intermediada por sistemas de IA dotados de aprendizado profundo impõe o esgotamento analítico dos modelos causais e finalistas da ação. As estruturas tradicionais de dolo e culpa revelam-se dogmaticamente inadequadas diante da autonomia algorítmica e da imprevisibilidade técnica inerente ao fenômeno da caixa-preta (black box). Lilian Edwards (2018, p. 75) ensina que a governança de tecnologias emergentes deve afastar-se da punição exclusivamente ex post e incorporar abordagem prospectiva de design ético e responsável.

A superação do finalismo como único paradigma de imputação penal não implica o abandono das garantias que ele consolidou — especialmente a exigência de dolo ou culpa como pressupostos inafastáveis da responsabilização. Trata-se, antes, de ampliar o quadro analítico para abarcar situações em que a voluntariedade humana se manifesta não no momento da conduta imediata, mas nas decisões tomadas ao longo do ciclo de vida do sistema autônomo: da concepção técnica à implantação, do monitoramento à desativação.

6.2 A REGULAÇÃO POR DESIGN E O DEVER DE CONFORMIDADE TECNOLÓGICA

A concepção de regulação por design desloca o centro de gravidade da dogmática penal para o momento da concepção técnica do sistema. Propõe-se a adoção de um modelo tripartite de responsabilização, sintetizado na Tabela 3:

Tabela 3 – Modelo tripartite de responsabilização penal por IA

Agente	Fundamento Legal	Modalidade Subjetiva	Condição de Imputação
Desenvolvedor	Art. 13, §2.º CP (omissão imprópria); Art. 18, I, 2.ª parte CP (dolo eventual)	Dolo eventual; Culpa consciente; Erro de tipo algorítmico inescusável	Negligência grave no design; ausência de salvaguardas; finalidade delitiva do sistema
Operador	Art. 18, I, CP (dolo direto); Art. 13, §2.º CP	Dolo direto; Dolo eventual	Uso doloso contrário aos termos do desenvolvedor e normas regulatórias
Usuário final	Arts. 29-30 CP (concurso de pessoas)	Dolo direto	Instrumentalização consciente do sistema para prática de ilícito

Fonte: Elaboração própria com base em Roxin (2006), Zaffaroni (2020) e PL 2.338/2023.

A distribuição da responsabilidade penal deve observar, em qualquer hipótese, a vedação constitucional à responsabilidade objetiva e o princípio da culpabilidade subjetiva. Mireille Hildebrandt (2015, p. 92) acrescenta que o Direito, diante da virada computacional, deve incorporar o conceito de law by design: as normas jurídicas precisam ser integradas à arquitetura dos sistemas tecnológicos desde a fase de desenvolvimento, tornando a conformidade legal uma propriedade estrutural do software.

O dever de conformidade tecnológica — núcleo da proposta deste trabalho — compreende as seguintes obrigações mínimas para o desenvolvedor: (a) avaliação de impacto de segurança antes da implantação do sistema; (b) implementação de mecanismos de kill switch e controle humano efetivo em sistemas de alto risco; (c) documentação técnica suficiente para permitir auditoria forense ex post; (d) monitoramento contínuo do comportamento do sistema após a implantação; e (e) notificação imediata às autoridades competentes diante de comportamentos emergentes com potencial delitivo.

6.3 RESPONSABILIDADE CIVIL, ADMINISTRATIVA E PENAL: DISTINÇÕES NECESSÁRIAS

O sistema tridimensional de responsabilização — penal, civil e administrativa — opera de forma complementar e não excludente. No plano civil, os arts. 186 e 927 do Código Civil impõem a reparação de danos causados por ato ilícito. A responsabilidade objetiva do fornecedor de produtos defeituosos (art. 12, CDC) pode ser aplicada a sistemas de IA que apresentem defeito de segurança, criando mecanismo eficaz de internalização dos custos do risco tecnológico pelos agentes econômicos que dele se beneficiam.

No plano administrativo, a LGPD e o futuro Marco Legal da IA estabelecem sanções que vão até 2% do faturamento da empresa, limitadas a R$50 milhões por infração (art. 52, LGPD). A ausência de programa de compliance pode ser considerada, na análise do dolo eventual, como indício relevante da assunção consciente do risco. No plano penal, a intervenção deve reservar-se aos casos em que a conduta omissiva do desenvolvedor ou operador, dolosa ou gravemente culposa, tenha contribuído causalmente para resultado delitivo de relevância constitucional — afetando bens jurídicos como a vida, a liberdade, o patrimônio em escala ou a integridade de infraestruturas críticas.

Howard Becker (2008, p. 23) lembra que o processo de definição do que é crime é sempre político e envolve relações de poder. Na era da IA, esse alerta adquire renovada pertinência: a delimitação do âmbito de responsabilidade penal dos desenvolvedores de sistemas autônomos não pode ser deixada exclusivamente à autorregulação do setor tecnológico, tampouco pode ser determinada por impulsos legislativos punitivistas desconectados da realidade técnica. Exige-se diálogo permanente entre juristas, técnicos, legisladores e a sociedade civil.

7 METODOLOGIA

A elaboração desta investigação científica exigiu a definição de percurso metodológico rigoroso, capaz de assegurar a validade, a replicabilidade e a cientificidade das conclusões. Segundo Umberto Eco (2015, p. 32), a pesquisa acadêmica deve delimitar seu objeto com precisão suficiente para que outros pesquisadores possam percorrer o mesmo iter cognitivo.

Do ponto de vista temporal, a pesquisa bibliográfica compreende obras publicadas entre 1999 e 2025, com ênfase nas contribuições das últimas duas décadas. A seleção bibliográfica observou os seguintes critérios: (i) relevância científica — obras publicadas em periódicos especializados e editoras acadêmicas de reconhecida reputação; (ii) atualidade — com preferência por publicações dos últimos cinco anos para os temas de regulação tecnológica; (iii) pertinência temática — conexão direta com os eixos analíticos da pesquisa; e (iv) representatividade doutrinária — inclusão de autores favoráveis e contrários às conclusões do trabalho, assegurando o contraditório acadêmico. Os países selecionados para o direito comparado — União Europeia, Estados Unidos, China e Brasil — foram escolhidos por representarem os principais polos de desenvolvimento e regulação da IA no cenário geopolítico contemporâneo.

O estudo adotou abordagem qualitativa e exploratória, conduzida pelo método de abordagem dedutivo: partindo das premissas gerais dos postulados constitucionais da culpabilidade e da legalidade penal (Marconi; Lakatos, 2021, p. 84), chegou-se, por derivação lógica, a conclusões específicas sobre a inadequação da legislação atual. O método hermenêutico foi empregado para a interpretação dos textos normativos, buscando extrair o sentido das normas diante das situações fáticas inéditas impostas pela IA. O método comparativo orientou a análise dos modelos regulatórios, permitindo identificar convergências, divergências e soluções transplantáveis (Popper, 2004, p. 57).

A pesquisa documental abrangeu o levantamento e a exegese das principais fontes normativas nacionais e internacionais, incluindo o AI Act europeu, o GDPR, o PL 2.338/2023, as resoluções do TSE sobre IA eleitoral e os relatórios técnicos da IBM Security, Europol, Febraban e Interpol. As categorias analíticas centrais foram: autonomia algorítmica, imputação penal objetiva, gestão de riscos tecnológicos, prova digital e regulação comparada (Bardin, 2016, p. 112). Incorporou-se ainda a perspectiva da epistemologia jurídica digital de Mireille Hildebrandt (2015, p. 92), segundo a qual o Direito não pode permanecer indiferente à virada computacional.

É importante registrar os limites desta pesquisa. A análise jurisprudencial restringe-se a julgados publicamente disponíveis até a data de conclusão do trabalho, não sendo possível incorporar decisões inéditas ou sigilosas. Os dados empíricos sobre criminalidade cibernética provêm de fontes institucionais reconhecidas, mas podem estar sujeitos a subnotificação, dado o elevado número de crimes cibernéticos que não chegam ao conhecimento das autoridades. A proposta dogmática formulada na Seção 6 tem caráter

prescritivo e normativo, não descritivo, refletindo o estado ideal do Direito (de lege ferenda) e não necessariamente o direito vigente (de lege lata).

8 CONCLUSÃO

A investigação sobre a responsabilidade penal diante do progresso da Inteligência Artificial e da intensificação dos crimes cibernéticos revelou um quadro de profundo descompasso entre os referenciais dogmáticos tradicionais e a realidade tecnológica contemporânea. A premissa axial do Direito Penal brasileiro — que condiciona a imputação a uma ação humana voluntária, consciente e finalisticamente orientada — encontra obstáculo de natureza estrutural na autonomia funcional de sistemas que operam por algoritmos e redes neurais profundas, cujas decisões nem sempre são rastreáveis ou previsíveis por seus criadores.

A pesquisa confirmou a hipótese central de que o sistema jurídico-penal vigente não se mostra tecnicamente adequado para lidar com cenários de plena autonomia tecnológica. A tentativa de aplicação irrestrita dos conceitos tradicionais de dolo e culpa à atuação autônoma da máquina conduz a dois resultados igualmente perniciosos: a criação de brecha de impunidade estrutural para os verdadeiros detentores do poder tecnológico, que se beneficiam da opacidade técnica da caixa-preta; ou a imposição de modalidade dissimulada de responsabilidade objetiva sobre desenvolvedores e operadores periféricos. Ambas as vias colidem frontalmente com os princípios constitucionais da legalidade (art. 5.º, XXXIX, CF/88), da pessoalidade da pena (art. 5.º, XLV, CF/88) e da culpabilidade subjetiva.

O exame das posições doutrinárias divergentes revelou que, embora fundadas em premissas constitucionais legítimas, tais correntes não oferecem resposta satisfatória para os cenários de dano catastrófico produzido por autonomia algorítmica plena, nos quais a abstenção penal equivale à imunidade estrutural dos arquitetos do risco tecnológico.

A análise de direito comparado revelou que o modelo europeu (AI Act), com sua classificação de risco e obrigações de auditabilidade algorítmica, representa o paradigma mais adequado para incorporação pelo ordenamento brasileiro. O PL 2.338/2023 é passo necessário, mas insuficiente: sua análise artigo por artigo evidenciou a lacuna da ausência de tipificação penal da negligência algorítmica grave e de causas de aumento vinculadas ao descumprimento das obrigações de conformidade. Os estudos de caso analisados — prisão por erro de reconhecimento facial, fraudes com voice cloning e spear phishing por IA generativa — demonstraram que essa lacuna já produz danos concretos e irreversíveis no cenário jurídico-penal brasileiro.

A solução para essa crise paradigmática não reside na atribuição de personalidade jurídico-penal aos sistemas de IA, mas na transição para um modelo de responsabilidade humana estruturado sobre o risco sistêmico e a regulação por design. O dever de conformidade tecnológica e o dever objetivo de cuidado nas

etapas de programação, implementação e monitoramento dos sistemas autônomos convertem-se no verdadeiro núcleo da reprovabilidade penal. Princípios constitucionais como a intervenção mínima, a fragmentariedade, a ofensividade e a proporcionalidade devem orientar a construção dos novos tipos penais, assegurando que o Direito Penal da era digital funcione como instrumento ético de gestão racional dos riscos tecnológicos.

Em síntese, a tese deste trabalho é clara: o programador ou desenvolvedor responde penalmente quando, no ciclo de vida do sistema de IA, deixar de observar o padrão de diligência técnica exigível — aferido conforme o estado da arte no momento da criação e implantação do sistema — e essa omissão qualificada contribuir causalmente para o resultado lesivo. A culpa presumida não se admite; exige-se prova da previsibilidade do comportamento algorítmico delitivo e da violação específica do dever de cuidado. A prova digital — logs, metadados, pesos do modelo e laudos periciais de explainability — é o caminho processual para essa demonstração.

Recomenda-se, em conclusão, a adoção de agenda legislativa que compreenda: (i) tipificação expressa da negligência algorítmica grave como modalidade de crime culposo qualificado; (ii) criação de causas de aumento de pena vinculadas ao descumprimento das obrigações de conformidade do futuro Marco Legal da IA; (iii) regulamentação da perícia em sistemas de IA, incluindo requisitos técnicos mínimos para a preservação dos elementos probatórios algorítmicos; e (iv) criação de instância especializada, no âmbito do Ministério Público e da Polícia Federal, para investigação de crimes cibernéticos mediados por IA. O Direito Penal da era digital não pode aguardar o próximo escândalo algorítmico para agir: a resposta normativa deve ser prospectiva, técnica e constitucionalmente comprometida com a proteção efetiva dos bens jurídicos na infosfera.

REFERÊNCIAS

ANGWIN, Julia et al. **Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And it's Biased Against Blacks.** ProPublica, New York, 23 maio 2016.

BARATTA, Alessandro. **Criminologia Crítica e Crítica do Direito Penal: introdução à sociologia do direito penal**. 3. ed. Rio de Janeiro: Revan, 1999.

BARDIN, Laurence. **Análise de Conteúdo**. São Paulo: Edições 70, 2016.

BATISTA, Nilo. **Introdução Crítica ao Direito Penal Brasileiro**. 11. ed. Rio de Janeiro: Revan, 2007.

BECCARIA, Cesare. **Dos Delitos e Das Penas**. São Paulo: Montecristo Editora, 2021.

BECKER, Howard S. **Outsiders: estudos de sociologia do desvio.** Rio de Janeiro: Zahar, 2008.

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

BRASIL. Decreto-Lei n.º 2.848, de 7 de dezembro de 1940. **Código Penal**. Rio de Janeiro: Presidência da República, 1940.

BRASIL. Lei n.º 12.737, de 30 de novembro de 2012. **Dispõe sobre a tipificação criminal de delitos informáticos**. Brasília, DF: Presidência da República, 2012.

BRASIL. Lei n.º 12.965, de 23 de abril de 2014. **Marco Civil da Internet**. Brasília, DF: Presidência da República, 2014.

BRASIL. Lei n.º 13.709, de 14 de agosto de 2018. **Lei Geral de Proteção de Dados Pessoais (LGPD).** Brasília, DF: Presidência da República, 2018.

BRASIL. Lei n.º 13.964, de 24 de dezembro de 2019. **Pacote Anticrime.** Brasília, DF: Presidência da República, 2019.

BRASIL. Lei n.º 14.155, de 27 de maio de 2021. **Altera o Código Penal para agravar crimes cibernéticos**. Brasília, DF: Presidência da República, 2021.

BRASIL. Senado Federal. Projeto de Lei n.º 2.338, de 2023. **Dispõe sobre o uso da Inteligência Artificial**.Brasília, DF: Senado Federal, 2023.

CALO, Ryan. **Artificial Intelligence Policy: A Primer and Roadmap**. UC Davis Law Review, v. 51, p. 399-435, 2017.

CRAWFORD, Kate. **Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence.** New Haven: Yale University Press, 2021.

ECO, Umberto. **Como se faz uma Tese.** 26. ed. São Paulo: Perspectiva, 2015.

EDWARDS, Lilian. **Future Law: Emerging Technology, Ethics and Regulation.** Cheltenham: Edward Elgar Publishing, 2018.

EUROPOL. **Cybercrime: a threat to EU businesses and citizens.** Haia: Europol, 2023.

FEBRABAN. **Pesquisa FEBRABAN de Tecnologia Bancária 2023**. São Paulo: Federação Brasileira de Bancos, 2023.

FLORIDI, Luciano. **Establishing the Rules for Building Trustworthy AI.** Nature Machine Intelligence, v. 1, p. 261-262, 2019.

FOUCAULT, Michel. **Vigiar e Punir: nascimento da prisão.** 20. ed. Petrópolis: Vozes, 1999.

GARLAND, David. **A Cultura do Controle: crime e ordem social na sociedade contemporânea.** Rio de Janeiro: Revan, 2008.

GIL, Antônio Carlos. **Como Elaborar Projetos de Pesquisa.** 7. ed. São Paulo: Atlas, 2019.

HILDEBRANDT, Mireille. **Smart Technologies and the End(s) of Law.** Cheltenham: Edward Elgar Publishing, 2015.

IBM SECURITY. **X-Force Threat Intelligence Index 2023.** Armonk: IBM Corporation, 2023.

INTERPOL. **African Cyberthreat Assessment Report 2023.** Lyon: Interpol, 2023.

LAKATOS, Eva Maria; MARCONI, Marina de Andrade. **Fundamentos de Metodologia Científica. 9.** ed. São Paulo: Atlas, 2021.

MALDONADO, Viviane (Org.). **Proteção de Dados: desafios e soluções na sociedade da informação.** São Paulo: Saraiva, 2021.

MASINI NETO, Ameleto. **Crimes Cibernéticos. Indaiatuba, SP:** Editora Foco, 2025.

NUCCI, Guilherme de Souza. **Manual de Direito Penal.** 10. ed. Rio de Janeiro: Forense, 2014.

O'NEIL, Cathy. **Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.** New York: Crown Publishers, 2016.

PARLAMENTO EUROPEU. **Regulamento (UE) 2024/1689**, de 13 de junho de 2024. Estabelece regras harmonizadas em matéria de inteligência artificial (AI Act). Estrasburgo: Parlamento Europeu e Conselho da União Europeia, 2024.

PASQUALE, Frank. **The Black Box Society: The Secret Algorithms That Control Money and Information.** Cambridge: Harvard University Press, 2015.

PINHEIRO, Patrícia Peck. **Direito Digital.** 7. ed. São Paulo: Saraiva, 2021. POPPER, Karl. A Lógica da Pesquisa Científica. 10. ed. São Paulo: Cultrix, 2004. ROXIN, Claus. Estudos de Direito Penal. 2. ed. Rio de Janeiro: Renovar, 2006.

SUPERIOR TRIBUNAL DE JUSTIÇA. AgRg no HC 828.054/RN. Rel. Min. Joel Ilan Paciornik, Quinta Turma, julgado em 23/04/2024. DJe de 29/04/2024.

SUPERIOR TRIBUNAL DE JUSTIÇA. AREsp 1.568.124/RJ. Sexta Turma, julgado em 2020.

SUPERIOR TRIBUNAL DE JUSTIÇA. HC 598.051/SP. Rel. Min. Rogério Schietti Cruz, Sexta Turma, julgado em 2020.

SUPERIOR TRIBUNAL DE JUSTIÇA. REsp 1.829.821/SP. Sexta Turma, julgado em 2020.

SUPERIOR TRIBUNAL DE JUSTIÇA. RHC 99.735/SC. Rel. Min. Nefi Cordeiro, Sexta Turma, julgado em 2019.

SUPREMO TRIBUNAL FEDERAL. ADPF 403/SE e ADI 5.527/DF. Tribunal Pleno, julgadas em 2021.

SUPREMO TRIBUNAL FEDERAL. RE 1.010.606/RJ. Tema 786 da Repercussão Geral. Rel. Min. Dias Toffoli, Tribunal Pleno, julgado em 11/02/2021.

SUPREMO TRIBUNAL FEDERAL. RE 1.037.396/SP. Tema 987 da Repercussão Geral. Rel. Min. Dias Toffoli, Tribunal Pleno, julgado em 2025.

TRIBUNAL SUPERIOR ELEITORAL. Resolução n.º 23.732, de 27 de fevereiro de 2024. Dispõe sobre **o uso de inteligência artificial na propaganda eleitoral.** Brasília, DF: TSE, 2024.

WACQUANT, Loïc. **Prisões da Miséria.** Rio de Janeiro: Jorge Zahar, 2004.

ZAFFARONI, Eugenio Raúl. **Direito Penal Humano: exposições e conferências.** Rio de Janeiro: Revan, 2020.