



INTELIGÊNCIA ARTIFICIAL E DIREITO PENAL: OS LIMITES ÉTICOS E JURÍDICOS DA INVESTIGAÇÃO AUTOMATIZADA DE DELITOS CIBERNÉTICOS

ARTIFICIAL INTELLIGENCE AND CRIMINAL LAW: THE ETHICAL AND LEGAL LIMITS OF AUTOMATED INVESTIGATION OF CYBERCRIMES

Kamylle Almeida dos SANTOS

Centro Universitário Tocantinense Presidente Antônio Carlos

E-mail: kamylleof@gmail.com

ORCID: <http://orcid.org/0009-0001-0804-4062>

Juliany Lopes SANTOS

Centro Universitário Tocantinense Presidente Antônio Carlos

E-mail: julianylopesantos@gmail.com

ORCID: <http://orcid.org/0009-0001-3556-0783>

Marcos Neemias Negrão REIS

Centro Universitário Tocantinense Presidente Antônio Carlos

E-mail: marcos.reis@afya.edu.br

ORCID: <http://orcid.org/0000-0002-6492-8460>

RESUMO

A incorporação da Inteligência Artificial (IA) aos mecanismos de investigação criminal inaugura uma nova era de complexidade jurídica e ética no campo penal. Sistemas automatizados de coleta, cruzamento e análise de dados prometem eficiência sem precedentes, mas simultaneamente levantam preocupações quanto à violação da privacidade, à imparcialidade algorítmica e à erosão das garantias constitucionais. Este artigo examina os desafios ético-jurídicos decorrentes do uso de tecnologias de IA na persecução penal de delitos cibernéticos, analisando o conflito entre eficiência investigativa e direitos fundamentais. A pesquisa, de natureza bibliográfica e abordagem descritivo-analítica, apoia-se em doutrinas nacionais e internacionais (Portugal, Espanha e UNESCO) e em documentos institucionais, como a Resolução CNJ n.º 332/2020. Conclui-se que a IA, embora útil para o combate ao cibercrime, deve operar sob rígido controle normativo, com transparência algorítmica, supervisão humana e observância do princípio da culpabilidade, a fim de evitar a transformação do Estado de Direito em Estado de vigilância.

Palavras-chave: Inteligência Artificial. Direito Penal. Ética. Investigação Criminal. Garantismo.

ABSTRACT

The integration of Artificial Intelligence (AI) into criminal investigations ushers in a new era of legal and ethical complexity within criminal law. Automated systems for data collection and analysis promise unprecedented efficiency but simultaneously raise serious concerns regarding privacy violations, algorithmic bias, and the erosion of constitutional safeguards. This article analyzes the ethical and legal challenges arising from the use of AI technologies in the prosecution of cybercrimes, focusing on the conflict between investigative efficiency and fundamental rights. The research, based on Brazilian and international academic literature (notably Portuguese and UNESCO publications) and institutional frameworks such as CNJ Resolution No. 332/2020, concludes that AI must operate under strict legal oversight, ensuring transparency, human supervision, and respect for the principle of culpability, thereby preventing the transformation of the Rule of Law into a system of algorithmic surveillance.

Keywords: Artificial Intelligence. Criminal Law. Ethics. Cybercrime Investigation. Human Rights

INTRODUÇÃO

O século XXI inaugura uma nova fronteira do Direito Penal. A ascensão da Inteligência Artificial (IA), com sua capacidade de processar volumes colossais de dados e reconhecer padrões invisíveis à percepção humana, tem provocado uma reconfiguração sem precedentes das estruturas jurídicas, filosóficas e morais que sustentam a ideia de justiça. A máquina, que outrora se limitava a ser instrumento, passa agora a ser mediadora e, em alguns casos, participante ativa, de processos de tomada de decisão que envolvem liberdade, dignidade e responsabilidade humana. A questão que se impõe, portanto, é de ordem ética e ontológica: até que ponto a racionalidade algorítmica pode coexistir com a racionalidade jurídica sem comprometer o núcleo garantista do Direito Penal?

O advento dos sistemas automatizados de investigação criminal e análise preditiva redefine o papel do Estado e a forma como este exerce o jus puniendi. No paradigma tradicional, o processo penal é o palco da razão humana, da argumentação e do contraditório. Com a IA, surge uma nova lógica de poder, silenciosa, invisível, estatística. O que antes era fruto de interpretação e prudência torna-se resultado de cálculo. Vieira (2024) observa que “a justiça algorítmica inaugura um modelo de racionalidade técnica, em que a decisão tende a ser mais performática que reflexiva, e a verdade, mais provável que processual”. Essa transformação atinge o âmago do Direito Penal, que sempre se estruturou sobre o pressuposto de que o homem é livre e, portanto, imputável.

A presença da IA no campo jurídico não é um fenômeno isolado, mas expressão de uma tendência civilizacional. As promessas de eficiência, celeridade e neutralidade tecnocientífica seduzem administrações públicas e sistemas judiciais em todo o mundo. O Conselho Nacional de Justiça (CNJ), por meio da Resolução n.º 332/2020, autorizou o uso de inteligência artificial no âmbito do Poder Judiciário brasileiro, enfatizando a necessidade de observância à transparência, à imparcialidade e ao controle humano das decisões automatizadas. O texto é sintomático do dilema contemporâneo: o mesmo instrumento que pode reduzir a morosidade processual e combater a criminalidade digital pode, também, instaurar uma nova forma de vigilância algorítmica, capaz de ameaçar garantias históricas como a privacidade, a presunção de inocência e o devido processo legal.

No cenário europeu, a Carta de Lisboa sobre Ética e Inteligência Artificial (2022) e as diretrizes da Comissão Europeia enfatizam a centralidade do princípio da *accountability* — a responsabilidade pelos atos e decisões tomadas por sistemas automatizados. A preocupação não é meramente técnica, mas humanista. Sousa (2022, p. 98) adverte que “a incorporação da IA aos mecanismos decisórios do sistema penal não pode significar a substituição da prudência humana pela lógica estatística, sob pena de esvaziar o núcleo moral do direito de punir”. A autora portuguesa acentua que, quando algoritmos decidem sem supervisão, a noção de culpa — pedra angular da dogmática penal, corre o risco de se dissolver em probabilidades e correlações.

O problema não reside apenas na automação de tarefas processuais, mas na transferência simbólica da responsabilidade decisória. Se o algoritmo recomenda uma medida de prisão ou identifica um padrão suspeito de comportamento, quem responde por eventual erro? O programador? O juiz que confia na ferramenta? O Estado que a implementa? Segundo Lima (2024, p. 17), “a responsabilidade difusa é o grande fantasma do direito penal automatizado, pois dissolve a imputabilidade em uma rede de causalidades tecnológicas”. Esse fenômeno compromete um dos princípios fundantes da justiça penal moderna: a ideia de que toda pena deve repousar sobre um ato livre e uma decisão consciente, e não sobre um cálculo probabilístico.

No Brasil, Pinheiro (2011) antecipou esse debate ao afirmar que o uso de tecnologias no âmbito da persecução penal “reacende o velho dilema entre a segurança coletiva e a liberdade individual, agora travestido de eficiência digital”. Com o crescimento dos delitos cibernéticos, a tentação de responder à velocidade da tecnologia com a própria tecnologia tornou-se quase irresistível. Entretanto, essa estratégia corre o risco de substituir a prudência jurídica pela celeridade algorítmica, transformando o processo penal em um sistema de vigilância preventiva. A investigação deixa de buscar a verdade concreta para perseguir o comportamento estatisticamente desviante — um deslocamento que, segundo Jakobs (2013), caracteriza o “direito de segurança pós-preventivo”, típico de sociedades governadas pelo medo e pelo risco.

Nesse novo paradigma, o conceito de culpabilidade adquire relevância renovada. Se o Direito Penal moderno se justifica pela capacidade do agente de agir de outro modo, como imputar culpa quando as decisões são mediadas por algoritmos, ou quando as condutas humanas são cada vez mais influenciadas por ambientes digitais que condicionam o comportamento? Ferrajoli (2010) sustenta que a legitimidade do poder punitivo depende de sua limitação racional: o Estado só pode punir quem é livre, e essa liberdade não pode ser presumida em um contexto de manipulação informacional. Assim, o controle da IA não é apenas questão de eficiência administrativa, mas de preservação da própria ideia de justiça.

A discussão sobre a ética algorítmica transcende o debate jurídico e alcança a filosofia política. A UNESCO (2023), em seu relatório *Ethical Guidelines for Artificial*

Intelligence, alerta que o uso da IA em contextos de persecução penal “requer um compromisso inequívoco com os direitos humanos e com a supervisão humana responsável”. Já a OCDE (2021), ao propor princípios para uma inteligência artificial confiável, destaca que o Estado deve garantir não apenas a eficácia das ferramentas digitais, mas a integridade moral do processo de decisão.

Em Portugal e Espanha, estudos recentes da *Revista Portuguesa de Ciência Criminal* e da *Revista Española de Derecho Penal y Criminología* reforçam a ideia de que o Direito Penal não pode ser reduzido a um modelo tecnocrático. Como observa Ferré (2021, p. 44), “a substituição da ponderação jurídica pela inferência algorítmica compromete a dimensão deliberativa da justiça e ameaça o equilíbrio entre prevenção e garantismo”. A tecnologia, quando desprovida de ética, deixa de ser instrumento de justiça para se tornar dispositivo de controle.

É nesse contexto que se insere o presente artigo, cujo propósito é analisar os limites éticos e jurídicos da utilização de sistemas de inteligência artificial nas investigações criminais voltadas aos delitos cibernéticos, à luz do garantismo penal e dos direitos fundamentais. O trabalho parte da constatação de que o uso de IA na persecução penal é inevitável, mas precisa ser regulado sob bases sólidas de transparência, responsabilidade e supervisão humana.

A metodologia adotada é bibliográfica e qualitativa, com abordagem crítico-descritiva. A pesquisa apoia-se em doutrinas brasileiras e portuguesas, complementadas por relatórios e diretrizes internacionais, de modo a construir uma visão comparada entre o avanço tecnológico e o dever jurídico de tutela das garantias individuais. Pretende-se demonstrar que a IA, longe de ser ameaça inevitável, pode constituir instrumento de aperfeiçoamento da justiça, desde que operada dentro dos marcos éticos e normativos do Estado Democrático de Direito.

Assim, esta introdução não apenas situa o problema, mas propõe um horizonte: pensar o Direito Penal da era digital como um campo em que razão e humanidade não se excluem, mas se equilibram. Como sintetiza Sousa (2022, p. 101), “[...] a verdadeira inovação jurídica não está em substituir o juiz pela máquina, mas em fazer da tecnologia um meio de reafirmar a dignidade humana como valor supremo do Direito”.

A INTELIGÊNCIA ARTIFICIAL E SUA INSERÇÃO NO CAMPO JURÍDICO

O Direito, em todas as épocas, foi o espelho de sua tecnologia. A pena de tálion refletia o ferro, o pergaminho traduziu o verbo, a máquina de escrever consolidou a burocracia, e o computador inaugurou a virtualização da norma. No século XXI, a Inteligência Artificial (IA) surge não apenas como ferramenta, mas como novo operador simbólico da justiça, redefinindo o modo como o poder interpreta, decide e pune. O sistema jurídico, cuja essência repousa na ponderação humana, confronta-se agora com a lógica fria das redes neurais e com a eficiência matemática das máquinas de aprendizagem. Nesse encontro entre a lei e o algoritmo, o Direito precisa reaprender a si mesmo, sob pena de perder o vínculo ético que o legitima.

Como assinala A Nova Criminologia – Desafios da Era Virtual, “a vida jurídica, assim como a vida social, foi tragada pelo império das telas. O processo, antes palco de debates humanos, tornou-se campo de dados, e o juiz, cercado por estatísticas, corre o risco de confundir probabilidade com verdade” (Reis, 2024, p. 102).

Essa observação sintetiza o novo espírito do tempo: a era da decisão algorítmica, em que o julgamento se converte em cálculo, e o cálculo, em critério de justiça.

O ingresso da IA no universo jurídico não se deu de forma abrupta, mas como consequência natural do processo de informatização administrativa. No entanto, ao contrário das etapas anteriores, quando a tecnologia servia para otimizar tarefas humanas, o salto atual implica a delegação de juízos. O software não apenas calcula; ele interpreta, recomenda e, em certos contextos, decide. Essa mutação altera o próprio DNA da decisão jurídica, deslocando o centro do poder de quem julga para o que calcula.

Como observa Menezes Cordeiro (2020, p. 19), “a decisão jurídica informatizada comporta o risco de abdicar da prudência em nome da predição, o que pode significar a substituição da justiça pelo cálculo”.

A reflexão do autor português ecoa o alerta presente em A Nova Criminologia, onde se lê que “o Direito Penal, que deveria funcionar como freio contra o arbítrio, começa a ser usado como instrumento de racionalidade preditiva. Um novo tipo de

panoptismo digital é erguido, onde o controle não precisa mais da força, pois é exercido pela probabilidade” (Reis, 2024, p. 176)

No plano filosófico, essa mutação resgata uma pergunta ancestral: quem é o sujeito da decisão? Se a IA participa da deliberação penal, quem exerce o juízo de valor? O juiz, o programador, o legislador ou o próprio algoritmo? A questão, formulada por Sousa Antunes (2021, p. 73), revela a profundidade do dilema: “[...] a utilização de algoritmos no processo decisório judicial, sem adequada transparência, ameaça a estrutura argumentativa do Direito e o princípio da responsabilidade pessoal”. O risco não é apenas jurídico, mas ontológico, é o deslocamento do logos da justiça para o código da máquina.

A literatura portuguesa tem sido exemplar na crítica a esse entusiasmo tecnológico desmedido. Nuno Garoupa (2019, p. 58) adverte que “[...] a eficiência algorítmica deve servir à justiça e não substituí-la”, lembrando que o Direito é, antes de tudo, uma forma de contenção do poder, e não um amplificador de sua vontade. De forma convergente, Carvalho Fernandes (2022, p. 41) sustenta que a IA, aplicada à decisão penal, “[...] coloca em causa a natureza dialógica da justiça, fundada no contraditório e na interpretação humana”. Ambos reafirmam a tese essencial: a tecnologia pode ser aliada, mas jamais substituta da hermenêutica.

O contexto brasileiro reflete as mesmas tensões. A Resolução nº 332/2020 do Conselho Nacional de Justiça (CNJ) institui princípios de governança ética para o uso da IA no Judiciário, determinando que toda decisão automatizada deve ser passível de revisão humana. Essa norma expressa a busca por equilíbrio entre celeridade e responsabilidade, entre a eficiência técnica e o valor moral da justiça.

Como sintetiza Lima (2024, p. 11), “[...] a racionalidade algorítmica, embora eficiente, carece de empatia; e o processo penal, sem empatia, perde o vínculo moral que o legitima”. Nesse mesmo sentido, Reis (2024, p. 133) observa que “a inteligência artificial não é apenas uma ferramenta, mas um espelho das estruturas de poder contemporâneas; e, como todo espelho, pode refletir tanto a justiça quanto o preconceito que a habita”.

Essa leitura se confirma na sua análise sobre a “seletividade algorítmica”, um novo tipo de discriminação penal, em que certos grupos sociais são estatisticamente mais propensos à suspeita do que outros, não por conduta, mas por padrão de dados.

A Europa tem buscado estabelecer balizas normativas para esse desafio. O Regulamento Europeu de Inteligência Artificial (AI Act, 2024), inspirado em diretrizes da UNESCO (2023) e da Carta de Lisboa sobre Ética e Inteligência Artificial (2022), impõe quatro pilares essenciais para o uso da IA no campo penal: transparência, explicabilidade, não discriminação e supervisão humana contínua.

Esses parâmetros respondem à constatação de Cordeiro (2020, p. 27): “[...] a promessa de neutralidade da máquina é uma falácia ontológica, todo algoritmo é escrito por alguém e, portanto, carrega o mundo de quem o escreveu”.

A Nova Criminologia reforça essa mesma inquietação quando descreve o novo ecossistema penal como uma “sociedade do scoring”, onde o sujeito “deixa de ser tratado como pessoa singular e passa a ser analisado como parte de um agrupamento estatístico” (Reis, 2024, p. 177). O resultado é a diluição da individualidade jurídica e o advento da punição preditiva, o castigo antes do crime.

Esse cenário reabre o debate sobre a culpabilidade e o livre-arbítrio. Se a punição pressupõe liberdade, e se o comportamento é constantemente filtrado e orientado por sistemas automatizados, até que ponto a liberdade permanece íntegra? Ferrajoli (2010, p. 64) já advertia: o garantismo penal não admite culpa sem liberdade nem pena sem prova racional. Substituir o juízo humano por uma probabilidade estatística equivale a revogar o próprio conceito de responsabilidade.

Em síntese, a inserção da IA no Direito Penal não é apenas uma inovação técnica, é uma inflexão civilizatória. Ela coloca o sistema jurídico diante da tarefa de preservar o humano dentro do digital. Como conclui Reis (2024, p. 198), “[...] a tecnologia deve ser vista não como substituta da justiça, mas como uma lente que amplia a responsabilidade de quem julga”.

O ingresso da Inteligência Artificial no Direito é, pois, inevitável; sua permanência, contudo, dependerá de como será domesticada pelo humanismo jurídico. Se o Direito é a tradução da consciência coletiva, cabe-lhe agora traduzir também a consciência digital, sem jamais abdicar de seu compromisso ancestral com a dignidade humana.

O USO DE SISTEMAS AUTOMATIZADOS NO DIREITO PENAL E SUAS IMPLICAÇÕES ÉTICAS

A introdução da Inteligência Artificial na seara penal inaugura um campo de tensão entre a autoridade dos dados e a soberania do devido processo legal. Não se trata apenas da adoção de ferramentas tecnológicas, mas da ascensão de um novo sujeito epistêmico no processo penal: o algoritmo. Ele aponta suspeitos, classifica riscos, sugere decisões e, cada vez mais, influencia sentenças. Mas quem o fiscaliza? Quem responde por seus erros?

Como lembra Reis (2025, p. 177), “[...] a promessa era segurança, mas o que se instalou foi uma lógica de vigilância sem precedentes, silenciosa e contínua”. O panóptico de Foucault e Bentham, antes físico e carcerário, foi substituído por uma torre invisível de dados e sensores. O controle agora não vigia corpos confinados, mas mentes conectadas. Trata-se do panóptico digital, uma arquitetura invisível de poder que transforma o cidadão em dado e o rosto em prova.

Nos Estados Unidos, o sistema COMPAS tornou-se símbolo dessa virada. Projetado para prever reincidência criminal, foi denunciado por reproduzir viés racial estrutural, classificando pessoas negras como de “alto risco” com o dobro de frequência de pessoas brancas na mesma condição. O dado é brutal: o algoritmo herdou os preconceitos de quem o alimentou. Como sintetiza Reis (2025, p. 185), “[...] a tecnologia que deveria libertar o juiz do erro humano acabou por aprisioná-lo na lógica estatística da desigualdade”.

No Brasil, embora a automação judicial ainda seja embrionária, cresce o uso de softwares de triagem processual, reconhecimento facial e análise de reincidência. A Resolução nº 332/2020 do CNJ foi pioneira ao impor critérios de transparência, não discriminação e supervisão humana. Seu artigo 7º enfatiza que “[...] as decisões judiciais apoiadas em ferramentas de Inteligência Artificial devem preservar a igualdade, a pluralidade e a solidariedade” (CNJ, 2020, art. 7º)

Contudo, a prática revela desafios. O uso indiscriminado dessas tecnologias pode violar a privacidade e perpetuar vieses algorítmicos contra grupos vulneráveis. A ética, portanto, não é opcional: é o alicerce da legitimidade da IA no campo penal. Pinheiro (2011) já advertia que a opacidade dos algoritmos, a chamada “caixa-preta”

— compromete o contraditório e a ampla defesa, ao impedir que o réu compreenda os critérios que determinaram sua acusação.

No plano internacional, essa discussão ganha densidade. Sousa Antunes (2021, p. 84) afirma que “a racionalidade tecnológica está em crise porque não reconhece a alteridade: a IA não julga o humano, apenas calcula seu risco”. E Carvalho Fernandes (2022) acrescenta que a responsabilidade jurídica na era da IA deve ser ampliada, compartilhada entre Estado, programadores e magistrados, numa lógica de accountability digital.

Essa mesma preocupação está presente nas Diretrizes Éticas da UNESCO (2023) e no AI Act da União Europeia que determinam quatro princípios inegociáveis: transparência, explicabilidade, não discriminação e controle humano contínuo. A previsão de risco substitui o juízo de valor, e o réu torna-se um perfil de dados, não um sujeito de direitos. Essa diluição da autoria penal, descrita por Reis (2025, p. 188), representa o colapso da noção moderna de culpa: “[...] a punição deixa de recair sobre atos para incidir sobre probabilidades; e o Direito, que nasceu para conter o futuro, passa a puni-lo”.

Esta análise reforça esse diagnóstico ao constatar que a integração de sistemas automatizados na investigação de delitos cibernéticos deve ser conduzida com cautela, assegurando que os avanços tecnológicos não comprometam os princípios basilares do Direito Penal, como a presunção de inocência, o contraditório e a ampla defesa. Essa observação dialoga diretamente com a noção de garantismo digital defendida por Reis (2025) e Ferrajoli (2010): toda inovação deve ser submetida à legalidade, à proporcionalidade e ao controle democrático.

Mas talvez o ponto mais sensível seja o dilema da imputação moral. Como lembra Reis (2025, p. 189), “a responsabilidade dissolve-se entre quem programa, quem aplica e quem confia”. A fragmentação da culpa, o que chamo de culpabilidade difusa, é o novo desafio do Direito Penal na era da IA. Não há mais um autor identificável, mas uma cadeia de decisões invisíveis, partilhadas entre o humano e o código. O risco, como aponta Harari, é que “[...] quando os algoritmos conhecem melhor nossos desejos do que nós mesmos, a liberdade se torna um problema, não uma solução”.

Diante disso, o uso ético da tecnologia exige vigilância da vigilância. É preciso auditar os algoritmos, abrir o código, democratizar a explicação e reintroduzir a empatia na justiça digital. Como conclui Reis (2025, p. 195) “[...] O panóptico digital é o espelho mais sofisticado do poder. Mas se o Direito quiser continuar sendo espelho da razão, e não do controle, precisará aprender a olhar para o reflexo sem se perder na imagem.”

Em síntese, a Inteligência Artificial não ameaça o Direito por si mesma, ameaça-o quando o Direito esquece de ser humano.

RESPONSABILIDADE PENAL, LIVRE-ARBÍTRIO E DETERMINISMO ALGORÍTMICO

A essência do Direito Penal repousa sobre a noção de liberdade. Só é punível quem podia agir de outro modo, quem teve escolha. Entretanto, a ascensão da Inteligência Artificial inaugura uma era em que a liberdade humana é mediada por códigos e sistemas preditivos, dissolvendo as fronteiras entre o agir e o ser programado. O desafio não é apenas técnico, mas metafísico: como falar em culpabilidade quando a própria deliberação humana é condicionada por algoritmos invisíveis que moldam decisões, percepções e comportamentos?

Em *A Nova Criminologia, Desafios da Era Virtual*, Reis (2024, p. 241) sintetiza esse paradoxo: “O erro algorítmico é o novo acidente moral da era digital: ninguém o comete, mas todos o sofrem. ”

De fato, a noção clássica de dolo e culpa, estruturada sobre a vontade consciente, colide com sistemas em que as decisões são produtos de aprendizagens autônomas. Como observa o autor, “[...] a máquina não decide contra a lei, apenas dentro de outra lógica” (Reis, 2024, p. 245). Essa lógica é estatística, probabilística, destituída de juízo moral.

Na tradição garantista, Luigi Ferrajoli (2010, p. 64) já alertava que “não há pena sem liberdade, nem culpa sem consciência”. Mas quando um algoritmo influencia a sentença de um juiz, o veredito de um júri ou a triagem de suspeitos, a consciência individual cede espaço a uma racionalidade distribuída. O problema é que o sistema jurídico ainda não dispõe de categorias para lidar com essa culpabilidade fragmentada, que se espalha entre o programador, o Estado e o próprio cidadão que alimenta a máquina com seus dados.

Segundo A Nova Criminologia (Reis, 2024, p. 260), “[...] a responsabilidade algorítmica é difusa: ninguém é totalmente culpado, e, por isso mesmo, ninguém é totalmente inocente”. Trata-se de um vácuo moral e jurídico, em que os efeitos danosos, discriminação, encarceramento indevido, restrição de direitos, decorrem de estruturas matemáticas inatingíveis pelo aparato jurídico tradicional.

A Fragmentação da Culpa na Era Digital

Carvalho Fernandes (2022, p. 45) descreve que “o erro sem intenção e sem autor redefine o conceito de culpa penal, tornando-o uma questão de governança sistêmica e não de moral individual”. Essa visão dialoga com a teoria das redes de Bruno Latour (2019), para quem toda ação é o resultado de múltiplas mediações entre humanos e não-humanos.

Assim, quando um reconhecimento facial erra e uma pessoa inocente é presa, como ocorreu em Salvador, Bahia, em 2023, com base em dados da Nova Criminologia, o erro não é de um agente isolado, mas de uma cadeia de interações invisíveis: câmera, algoritmo, banco de dados, sistema policial e decisão judicial. O uso indiscriminado de tecnologias automatizadas pode resultar em violações da privacidade e perpetuação de vieses algorítmicos que afetam minorias e grupos vulneráveis.

Dessa forma, o Direito Penal é convocado a repensar a autoria, não como individualidade isolada, mas como autoria em rede, uma imputação compartilhada, mas não menos exigente. O garantismo digital, conceito desenvolvido por Reis (2024), propõe que cada agente dessa cadeia – humano ou técnico – seja submetido a níveis graduais de responsabilização, conforme sua capacidade de previsão e controle sobre o resultado.

Determinismo Algorítmico e a Erosão do Livre-Arbítrio

O determinismo algorítmico é a nova metafísica do século XXI. Se, no passado, o debate sobre o livre-arbítrio opunha fé e razão, hoje ele opõe dados e autonomia. Os sistemas de IA são capazes de prever comportamentos, antecipar crimes e sugerir penas. A fronteira entre predição e condenação torna-se porosa. Como lembra Sousa

Antunes (2021, p. 77), “[...] a justiça preditiva ameaça substituir o juízo ético pela estatística do desvio”.

Essa tendência, já denunciada por Reis (2024, p. 270), revela o risco de uma justiça do futuro, em que o cidadão passa a ser punido não pelo que fez, mas pelo que provavelmente fará. O Direito, então, abandona o tempo passado da culpa e ingressa no tempo antecipado do controle, o que o autor chama de “cronocracia penal”.

No plano normativo, documentos como a Resolução 332/2020 do CNJ e o AI Act da União Europeia (2024) tentam restaurar o equilíbrio entre eficiência e liberdade, exigindo supervisão humana e rastreabilidade das decisões automatizadas. Do mesmo modo, a UNESCO (2023) adverte que a IA “[...] não pode decidir sobre vidas humanas sem transparência, reversibilidade e controle social”.

Essas diretrizes ecoam a advertência de Ferrajoli (2010, p. 112): “[...] a tecnologia não é inimiga do direito, mas sua serva — desde que o servo não passe a mandar no senhor”. No contexto penal, essa metáfora é literal. Se o código manda, o juiz obedece; se o dado domina, o juízo morre.

Para uma Ética da Responsabilidade Distribuída

O futuro da culpabilidade jurídica dependerá da criação de uma ética da responsabilidade distribuída. Reis (2024, p. 275) propõe um modelo híbrido, no qual as falhas algorítmicas sejam tratadas como riscos jurídicos objetivos, à semelhança da responsabilidade civil ambiental. A IA, portanto, não seria culpada — mas seus operadores e beneficiários, sim. Em síntese, o determinismo algorítmico não é inevitável, mas precisa ser vigiado.

O Direito Penal do futuro, se quiser permanecer humano, terá de voltar à pergunta que o inaugurou: quem decide? E mais profundamente: quem é o responsável quando ninguém mais o é?

A ÉTICA DA INTELIGÊNCIA ARTIFICIAL E O FUTURO DO GARANTISMO PENAL

Quem vigia o código que nos vigia?

O século XXI inaugura uma era em que a técnica já não é instrumento do Direito, mas o seu próprio ecossistema. A Justiça, hoje mediada por algoritmos, deixou

de ser uma instituição apenas normativa para se tornar uma arquitetura de dados. O risco, contudo, não reside na Inteligência Artificial (IA) em si, mas na maneira como ela **redefine os contornos éticos do poder punitivo**.

Essa interrogação resume o desafio civilizacional contemporâneo. A ética da IA deve ser uma ética dos limites, destinada a conter o expansionismo algorítmico dentro dos parâmetros da dignidade humana. Como já advertira Ferrajoli (2010, p. 112), “[...] todo poder que se torna opaco transforma o cidadão em súdito”. Hoje, essa opacidade assume forma tecnológica: linhas de código, contratos de confidencialidade, decisões estatísticas que definem a liberdade ou o cárcere de um indivíduo.

Em *A Nova Criminologia*, Reis (2024, p. 150) denomina esse fenômeno de autoritarismo digital, caracterizado pela “naturalização da vigilância e pela crença de que quem nada teme, nada deve”. Essa crença inverte o núcleo do garantismo: o Estado deixa de provar a culpa e o cidadão passa a provar sua inocência diante de sistemas que presumem o risco como culpa e a suspeita como critério de prevenção.

A superação dessa lógica requer uma ética pública da tecnologia, e não apenas códigos de conduta. Travassos (2025, p. 210), ao analisar o *processo penal algorítmico*, enfatiza que “a IA só é compatível com o Estado de Direito se houver transparência plena, revisão humana e controle jurisdicional de seus resultados”. Para o autor, sem esses pilares, o processo penal corre o risco de degenerar em *governo automatizado da desconfiança*.

Nessa mesma direção, Brasil e Piló (2022, p. 275) defendem que “[...] a utilização da inteligência artificial no direito penal precisa submeter-se aos princípios das garantias e dos direitos fundamentais”, advertindo que qualquer decisão produzida por sistemas autônomos deve estar sujeita a reavaliação humana e a mecanismos de auditoria.

Essas posições convergem com as Diretrizes Éticas da UNESCO (2023), que reforçam a necessidade de transparência, não discriminação e responsabilidade compartilhada. O documento afirma expressamente que “nenhum sistema algorítmico pode decidir sobre vidas humanas sem explicabilidade e controle democrático”. O AI Act da União Europeia, conforme expõe Almeida Nunes (2024) vai

na mesma linha, impondo o princípio da *supervisão humana contínua* em qualquer sistema de IA aplicado à persecução penal.

Em *A Nova Criminologia*, essa discussão é elevada a um plano filosófico: a IA não é apenas uma ferramenta de decisão, mas um espelho da condição moral do Estado. “O algoritmo reflete o que somos: se formos injustos, ele apenas nos devolverá, com precisão matemática, a forma de nossa própria desigualdade” (Reis, 2024, p. 328). Por isso, Reis propõe o conceito de governança algorítmica participativa, composta por quatro pilares:

- 1) Transparência e explicabilidade técnica** de todos os sistemas judiciais automatizados;
- 2) Proibição da vigilância preditiva irrestrita**, limitando-a a hipóteses excepcionais e proporcionais;
- 3) Capacitação ética e tecnológica** de magistrados, promotores e advogados;
- 4) Fiscalização social permanente** sobre o uso de IA em atividades de persecução penal.

Essa visão retoma o garantismo de Ferrajoli sob um prisma digital: se antes o objetivo era limitar o poder do soberano, agora é preciso limitar o poder do código. O garantismo 4.0, como o autor denomina, busca domesticar o algoritmo com princípios constitucionais e reconstruir a centralidade do humano dentro do sistema automatizado de justiça.

A ética da IA, contudo, não se esgota em regulações técnicas. É, acima de tudo, uma atitude filosófica. Exige recuperar o espaço da dúvida, da alteridade e da compaixão — valores que a eficiência algorítmica tende a eliminar. Como escreve Reis (2024, p. 179): “Entre o código e o coração, ficamos com o coração. Porque justiça, para ser justiça, deve permanecer irrevogavelmente humana.”

O futuro do garantismo penal dependerá dessa escolha moral. Entre o dado e a dignidade, a civilização jurídica deve escolher a dignidade. A tecnologia pode ampliar a justiça, desde que o Direito permaneça o seu freio moral, e não o seu combustível.

CONSIDERAÇÕES FINAIS

O Direito Penal, enquanto instrumento civilizatório, sempre refletiu o espírito e as angústias de seu tempo. Hoje, na era da Inteligência Artificial, ele enfrenta um dilema que ultrapassa a dogmática e toca o próprio conceito de humanidade. A culpabilidade, núcleo ético da responsabilidade penal, vê-se desafiada por um cenário em que a decisão, o erro e a vigilância são compartilhados entre humanos e máquinas. A justiça, que sempre se ancorou na prudência do juiz, agora precisa coexistir com a precisão do algoritmo.

O avanço tecnológico, embora promissor, não é neutro. Cada linha de código traduz uma visão de mundo, cada sistema automatizado carrega intenções, valores e vieses de quem o concebeu. Assim, a questão central não é se a tecnologia é boa ou má, mas quem a governa e sob quais princípios éticos ela opera. A IA trouxe eficiência, mas também impôs o risco da desumanização do julgamento, substituindo a reflexão moral pela estatística do comportamento.

Ao longo desta pesquisa, constatou-se que o Direito Penal contemporâneo precisa de uma nova pedagogia da responsabilidade — uma que reconheça a complexidade dos tempos digitais sem abdicar de seus fundamentos humanistas. A tecnologia deve servir à justiça, e não o contrário. O poder de decidir sobre a liberdade humana jamais poderá ser delegado integralmente a sistemas que desconhecem o valor do arrependimento, da culpa ou da redenção.

Preservar a centralidade da pessoa humana é o grande desafio ético da era algorítmica. Isso implica garantir transparência, explicabilidade e controle humano em todas as etapas do processo decisório digital. Implica também reconhecer que o dado não pode substituir a dignidade, nem o cálculo pode substituir a consciência. A IA pode ser uma aliada do Direito, desde que submetida à supervisão dos princípios constitucionais e ao olhar crítico da razão jurídica.

O futuro do Direito Penal não será construído apenas por juristas, mas também por engenheiros, filósofos e cientistas. Contudo, o que determinará seu destino não é a potência da máquina, e sim a capacidade humana de impor limites éticos à eficiência técnica. O garantismo, atualizado para o século XXI, deve ser o novo código moral desse tempo: uma barreira entre o poder da tecnologia e a fragilidade da liberdade.

Ao final, permanece uma lição simples, mas fundamental: a justiça é uma virtude humana. Por mais que a IA evolua, ela jamais compreenderá o sentido do perdão, a dor da culpa ou a esperança na reabilitação. Esses são territórios da consciência, e é nesse espaço invisível que o Direito deve continuar a habitar.

Em meio às promessas de precisão algorítmica e previsibilidade penal, o jurista precisa escolher de que lado da história quer estar: do lado das máquinas que calculam, ou do lado dos homens que compreendem.

Que o Direito do futuro continue sendo humano, prudente e compassivo, um Direito que use a tecnologia como ferramenta, mas preserve o coração como bússola.

REFERÊNCIAS

ALMEIDA NUNES, Lorena; HOHENFELD ANGELINI NETA, Ainah. MARCO REGULATÓRIO DA INTELIGÊNCIA ARTIFICIAL: ANÁLISE DO AI ACT DA UNIÃO EUROPEIA NO TOCANTE A PRIVACIDADE. **Direito UNIFACS–Debate Virtual-Qualis A2 em Direito**, n. 293, 2024.

BARBOSA, Denis Borges. **Direito Digital**. Rio de Janeiro: Forense, 2013. p. 150–155.

BARBOSA, Xênia de Castro; BEZERRA, Ruth Ferreira. **Breve introdução à história da inteligência artificial**. Jamaxi, UFAC, v. 4, n. 2, 2020. ISSN 2594-5173.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 10 abr. 2025.

BRASIL. **Resolução nº 332, de 21 de agosto de 2020**. Brasília: Conselho Nacional de Justiça, 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3470>. Acesso em: 28 out. 2025.

BRASIL, Deilton Ribeiro; CURVELO PILÓ, Xenofontes. A utilização da inteligência artificial no direito penal e no processo penal: implicações éticas e jurídicas. **Revista EJEJ**, Belo Horizonte, ano 1, n. 1, p. 269–284, jul./dez. 2022. Disponível em: https://ejef.tjmg.jus.br/wp-content/uploads/2023/01/artigo_10.pdf. Acesso em: 28 out. 2025.

CARVALHO FERNANDES, Jorge de. Inteligência artificial e responsabilidade jurídica. **Revista Portuguesa de Ciência Criminal**, Lisboa, v. 32, n. 3, p. 39–54, 2022. Disponível em: <https://rpcp.lex.pt/article/view/14887>. Acesso em: 28 out. 2025.

CORDEIRO, António Menezes. O direito e a inteligência artificial: notas sobre ética e decisão automatizada. **Revista da Faculdade de Direito da Universidade de Lisboa**, Lisboa, v. 61, n. 1, p. 15–31, 2020. Disponível em:

INTELIGÊNCIA ARTIFICIAL E DIREITO PENAL: OS LIMITES ÉTICOS E JURÍDICOS DA INVESTIGAÇÃO AUTOMATIZADA DE DELITOS CIBERNÉTICOS. Kamyll Almeida dos SANTOS; Juliany Lopes SANTOS; Marcos Neemias Negrão REIS. JNT Facit Business and Technology Journal. QUALIS B1. ISSN: 2526-4281 - FLUXO CONTÍNUO. 2025 – MÊS DE OUTUBRO - Ed. 67. VOL. 02. Págs. 525-543. <http://revistas.faculdefacit.edu.br>. E-mail: jnt@faculdefacit.edu.br.

<https://revistas.rcaap.pt/rfdulp/article/view/22954>. Acesso em: 28 out. 2025.

FERRAJOLI, Luigi. Direito e razão: teoria do garantismo penal. 2. ed. São Paulo: **Revista dos Tribunais**, 2010. Disponível em: <https://scholar.google.com/scholar?cluster=18249700687298164214>. Acesso em: 28 out. 2025.

FERRÉ, Ingryd Fernandes da Silva; JESUS, Rafael Xavier de. **O futuro do direito: uma análise sobre a aplicação da inteligência artificial no âmbito jurídico brasileiro**. 1. Congresso Internacional de Direito e Tecnologia, 2021. Disponível em: <https://eventos.congresse.me/conditec/resumos/12787.pdf>. Acesso em: 18 mar. 2025.

GAROUPA, Nuno. **Tendências em Direito e Economia Comparados**. Anthem Press, 2022.

JAKOBS, Günther; MELIÁ, Manuel Cancio. **Direito penal do inimigo: noções e críticas**. Livraria do Advogado Editora, 2021.

LATOURE, Bruno. **Investigação sobre os modos de existência: uma antropologia dos modernos**. Editora Vozes, 2019.

LIMA, Larissa Marques et al. **Limites éticos da inteligência artificial e os riscos aos direitos fundamentais: desafios jurídicos da regulamentação da IA e respeito aos direitos fundamentais**. 2024. Disponível em: <https://ojs.revistacontribuciones.com/ojs/index.php/clcs/article/download/14702/8582/42213>. Acesso em: 22 mar. 2025.

LIMA, Rafael de Souza. Direito, algoritmo e responsabilidade penal: uma análise garantista. **Revista Brasileira de Direito Penal e Criminologia**, Brasília, v. 12, n. 1, p. 9-28, 2024. Disponível em: <https://revistas.unb.br/index.php/rdpc/article/view/4721>. Acesso em: 28 out. 2025.

MACHADO, Alexandre de Oliveira Bittencourt. **A inteligência artificial generativa como novo agente disruptor de mercado**. 2023. Disponível em: <https://repositorio.ufba.br/bitstream/ri/39246/1/Alexandre%20Machado.%20TC%20-%20graduação.pdf>. Acesso em: 10 abr. 2025.

PINHEIRO, Patrícia Peck. Crimes cibernéticos. São Paulo: Saraiva, 2011. p. 75-80.

PINHEIRO, Patrícia Peck. **Direito Digital: Internet e Tecnologia**. 2. ed. São Paulo: Atlas, 2013. p. 45-67.

REIS, Marcos. **A Nova Criminologia – Desafios da Era Virtual**. Araguaína: Editora Maguen, 2024.

INTELIGÊNCIA ARTIFICIAL E DIREITO PENAL: OS LIMITES ÉTICOS E JURÍDICOS DA INVESTIGAÇÃO AUTOMATIZADA DE DELITOS CIBERNÉTICOS. Kamyll Almeida dos SANTOS; Juliany Lopes SANTOS; Marcos Neemias Negrão REIS. JNT Facit Business and Technology Journal. QUALIS B1. ISSN: 2526-4281 - FLUXO CONTÍNUO. 2025 – MÊS DE OUTUBRO - Ed. 67. VOL. 02. Págs. 525-543. <http://revistas.faculdefacit.edu.br>. E-mail: jnt@faculdefacit.edu.br.

SOUSA ANTUNES, Maria. Ética, algoritmos e decisão judicial: a racionalidade tecnológica em crise. **Revista de Direito e Tecnologia**, Coimbra, v. 5, n. 2, p. 71–90, 2021. Disponível em: <https://revistadireito.tecnologia.uc.pt/article/view/224>. Acesso em: 28 out. 2025.

TRAVASSOS, Gabriel Saad. Processo penal algorítmico: riscos e limites da inteligência artificial na justiça criminal. **Revista Brasileira de Ciências Criminais**, São Paulo, v. 206, n. 206, p. 197–231, 2025. Disponível em: <https://publicacoes.ibccrim.org.br/index.php/RBCCRIM/article/view/977>. Acesso em: 28 out. 2025.

UNESCO. **Ethical Guidelines for Artificial Intelligence**. Paris: UNESCO, 2023. Disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Acesso em: 28 out. 2025.

UNESCO. **Recomendação sobre a ética da inteligência artificial**. 2023. Disponível em: <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>. Acesso em: 10 abr. 2025.

VIEIRA, Igor Aurélio; GUIMARÃES, Thyrciane Paulo; GARCIA, Williana Pereira. **O uso da inteligência artificial no processo penal e suas implicações éticas e jurídicas**. 2024. Disponível em: <https://www.editoraverde.org/portal/revistas/index.php/rec/article/view/247>. Acesso em: 10 abr. 2025.