

**RELATÓRIOS PRODUZIDOS
POR INTELIGÊNCIA
ARTIFICIAL GENERATIVA
COMO MEIO DE PROVA NO
PROCESSO PENAL:
ANÁLISE CRÍTICA À LUZ DO
HC 1.059.475/SP E DOS
LIMITES EPISTÊMICOS DA
PROVA DIGITAL**

REPORTS PRODUCED BY GENERATIVE ARTIFICIAL INTELLIGENCE AS
EVIDENCE IN CRIMINAL PROCEEDINGS: A CRITICAL ANALYSIS IN LIGHT
OF HC 1.059.475/SP AND THE EPISTEMIC LIMITS OF DIGITAL EVIDENCE

Ciências Sociais Aplicadas • 06/06/2026

REGISTRO DOI: [10.70773/revistatopicos/780637669](https://doi.org/10.70773/revistatopicos/780637669)

Gustavo Henrique de Andrade Cordeiro¹

Rodrigo Murad Vitoriano²

RESUMO

O artigo examina a admissibilidade e a valoração de relatórios elaborados por sistemas de inteligência artificial generativa no processo penal brasileiro, à luz das garantias fundamentais e dos limites epistemológicos da prova digital. O problema de pesquisa consiste em investigar se tais relatórios satisfazem os requisitos mínimos de justificação racional exigidos pela teoria da prova penal, sem afronta ao contraditório, à ampla defesa e ao devido processo legal. Como referencial analítico, adota-se o julgamento do Habeas Corpus n. 1.059.475/SP, do Superior Tribunal de Justiça, em articulação com a teoria da prova e com aportes críticos sobre opacidade algorítmica. A pesquisa emprega metodologia qualitativa, de caráter teórico-dogmático, baseada em exame jurisprudencial e revisão bibliográfica interdisciplinar. Argumenta-se que sistemas generativos, por se fundarem em modelos probabilísticos, apresentam limitações estruturais, como baixa explicabilidade, risco de geração de conteúdos inverídicos e reduzida auditabilidade, o que fragiliza sua confiabilidade probatória. Sustenta-se que tais relatórios não se ajustam plenamente às categorias de prova documental ou pericial, devendo ser tratados como elementos informativos sujeitos à validação técnica. Conclui-se que sua utilização exige parâmetros rigorosos de transparência, reprodutibilidade, controle independente e contraditório técnico.

Palavras-chave: inteligência artificial generativa; prova penal; epistemologia jurídica; devido processo legal; transparência algorítmica.

ABSTRACT

This article examines the admissibility and evidentiary value of reports produced by generative artificial intelligence systems in Brazilian criminal procedure, in light of fundamental guarantees and

the epistemological limits of digital evidence. The research problem consists of investigating whether such reports meet the minimum standards of rational justification required by the theory of criminal evidence, without violating the adversarial principle, full defense and due process of law. As an analytical framework, the study adopts the judgment of Habeas Corpus No. 1.059.475/SP, decided by the Superior Court of Justice, in articulation with evidence theory and critical approaches to algorithmic opacity. The research employs a qualitative, theoretical-dogmatic methodology, based on jurisprudential analysis and interdisciplinary literature review. It is argued that generative systems, relying on probabilistic models, present structural limitations, such as low explainability, the risk of generating inaccurate content and limited auditability, which undermine their evidentiary reliability. It is held that such reports do not fully fit the categories of documentary or expert evidence and should be treated as informational elements subject to technical validation. It concludes that their use requires strict standards of transparency, reproducibility, independent oversight and technical adversarial scrutiny.

Keywords: generative artificial intelligence; criminal evidence; legal epistemology; due process of law; algorithmic transparency.

1. INTRODUÇÃO

A incorporação de sistemas de inteligência artificial generativa ao cotidiano forense deixou de configurar hipótese especulativa para impor-se como questão concreta da teoria da prova. Capazes de sintetizar grandes volumes de informação, identificar padrões e produzir narrativas verossímeis, esses sistemas tensionam categorias que, por décadas, estruturaram o pensamento processual penal. O fenômeno não é trivial. Onde antes se discutia a licitude e a cadeia

de custódia de um vestígio, passa-se agora a indagar se um conteúdo gerado por máquina pode, em alguma medida, sustentar a reconstrução judicial dos fatos. A pergunta toca o núcleo do processo penal, espaço em que a exigência de racionalidade probatória assume densidade peculiar, justamente por estar em jogo a liberdade do acusado e a contenção do poder punitivo.

Diante desse cenário, emerge o problema que orienta a investigação: em que medida relatórios produzidos por sistemas de inteligência artificial generativa satisfazem os requisitos epistêmicos mínimos de justificação racional exigidos pela teoria da prova penal, especialmente quanto à verificabilidade das inferências, à controlabilidade intersubjetiva e à possibilidade de refutação pelas partes, sem violação ao contraditório, à ampla defesa e ao devido processo legal? Trata-se de questão que ultrapassa a mera admissibilidade formal, alcançando a compatibilidade desses instrumentos com o modelo de legitimação que sustenta a decisão penal.

O objetivo geral consiste em analisar criticamente as condições de admissibilidade e valoração de relatórios gerados por inteligência artificial generativa no processo penal brasileiro, a partir de uma epistemologia da prova. Para tanto, estabelecem-se os seguintes objetivos específicos: examinar os fundamentos da teoria da prova penal e seus limites epistemológicos; investigar os condicionamentos técnicos dos sistemas generativos que afetam sua confiabilidade probatória; analisar o precedente firmado no Habeas Corpus n. 1.059.475/SP como expressão das tensões normativas envolvidas; e contrastar o tratamento brasileiro com parâmetros do direito comparado, identificando salvaguardas aptas a preservar a racionalidade da decisão.

Quanto à metodologia, adota-se abordagem qualitativa, de caráter teórico-dogmático, combinada a um exame analítico-conceitual orientado pela epistemologia jurídica. O método de procedimento é monográfico, com recorte na análise de precedente judicial paradigmático, compreendido como expressão das tensões decorrentes da incorporação de tecnologias emergentes ao processo penal. As técnicas empregadas são a pesquisa bibliográfica e a documental, de natureza exploratória e explicativa, com consulta a legislação, doutrina nacional e estrangeira, atos normativos do Conselho Nacional de Justiça, artigos publicados em periódicos qualificados e jurisprudência nacional e comparada. Parte-se da premissa de que a adoção de novas tecnologias não pode ocorrer à revelia das estruturas que asseguram a legitimidade da decisão judicial.

O primeiro eixo de desenvolvimento reconstrói a teoria da prova penal e seus limites epistemológicos. Procura-se demonstrar que a validade da prova não se esgota na admissibilidade formal, mas depende de condições materiais de confiabilidade, de controle racional e de efetiva abertura ao contraditório. Nesse percurso, examinam-se a função da cadeia de custódia, os critérios de autenticidade e integridade da prova digital e a articulação entre racionalidade probatória e os standards de prova exigidos para a condenação penal.

O segundo eixo desloca a análise para o funcionamento técnico dos sistemas de inteligência artificial generativa. Busca-se compreender de que modo características como opacidade algorítmica, comportamento não determinístico, verossimilhança linguística e propensão a alucinações comprometem a reprodutibilidade e a explicabilidade dos resultados produzidos. A hipótese é a de que

existe uma incompatibilidade estrutural entre modelos probabilísticos opacos e o paradigma justificativo do processo penal garantista.

Os eixos seguintes voltam-se ao exame do Habeas Corpus n. 1.059.475/SP e à crítica da chamada fetichização tecnológica. Analisa-se a ratio decidendi do precedente, a controversa qualificação jurídica desses relatórios e o diálogo com o direito comparado, em especial o Regulamento europeu sobre inteligência artificial e a jurisprudência estrangeira. Em seguida, problematiza-se a tendência de atribuir confiabilidade excessiva a sistemas automatizados, discutindo a ilusão de objetividade algorítmica, os vieses estruturais e a necessária redefinição do ônus argumentativo no processo penal.

O enfrentamento dessas questões mostra-se essencial não apenas para a adequação tecnológica da Justiça criminal às demandas contemporâneas, mas sobretudo para a preservação dos direitos fundamentais em um contexto de crescente automação decisória. A análise que se propõe pretende contribuir para um debate informado sobre os contornos epistêmicos dessa transição, delimitando parâmetros mínimos para a eventual incorporação de sistemas generativos ao campo probatório e superando abordagens meramente instrumentais ou utilitaristas.

2. A TEORIA DA PROVA PENAL E SEUS LIMITES EPISTEMOLÓGICOS

2.1. A Prova Como Instrumento de Justificação Racional

A prova penal não se reduz a instrumento de reconstrução factual: constitui também mecanismo de legitimação da decisão judicial.

Nessa perspectiva, o processo penal deve ser compreendido como sistema de garantias voltado à contenção do poder punitivo estatal, o que implica um modelo probatório rigorosamente comprometido com os direitos fundamentais e com a limitação do arbítrio judicial (LOPES JR., 2023).

A atividade probatória possui dimensão essencialmente epistêmica, orientada à obtenção de conhecimento justificado sobre os fatos. Como sustenta Taruffo, a função da prova no processo não é meramente persuasiva, mas voltada à construção de um conhecimento racionalmente controlável, o que pressupõe critérios de racionalidade e de controle intersubjetivo (TARUFFO, 2014). A prova, assim, não pode ser tratada como mero artefato procedimental, mas como caminho para a reconstrução racional dos fatos juridicamente relevantes.

No mesmo horizonte, Ferrer Beltrán enfatiza que a valoração da prova deve guiar-se por padrões de justificação racional, afastando concepções subjetivistas ou meramente formais, de modo que a decisão judicial possa fundar-se em razões públicas e controláveis (FERRER BELTRÁN, 2007; 2017). A consequência é direta: princípios como o contraditório e a ampla defesa não operam apenas como garantias formais, mas como autênticas condições epistêmicas de validade da prova.

Daí decorre a relevância da tradição garantista. A partir de Ferrajoli, reforça-se a necessidade de submeter toda atividade probatória a limites estritos, como forma de contenção do poder punitivo e de preservação das garantias fundamentais. A validade da prova não se vincula apenas à admissibilidade formal, mas também à sua obtenção e valoração segundo regras que assegurem controle

racional, publicidade e possibilidade de refutação, evitando decisões fundadas em elementos opacos ou insuscetíveis de verificação intersubjetiva (FERRAJOLI, 2002).

A introdução de elementos cuja origem, funcionamento ou lógica interna não sejam plenamente compreensíveis compromete diretamente essa estrutura. Quando o percurso inferencial que conduz à formação do convencimento não pode ser reconstruído nem criticado pelas partes, a prova perde seu lastro epistêmico, ainda que conserve aparência técnica. É precisamente nesse ponto que os sistemas de inteligência artificial generativa colocam em xeque os pressupostos clássicos da teoria da prova, tema que se aprofunda nos eixos seguintes.

Convém observar que a exigência de controle racional não é incompatível com a convicção do julgador, mas a disciplina. O livre convencimento motivado, longe de autorizar a valoração arbitrária, impõe ao juiz o dever de explicitar as razões que sustentam suas conclusões, de modo que possam ser submetidas à crítica das partes e das instâncias revisoras. A motivação é, nesse sentido, o reverso necessário da liberdade na apreciação da prova. Qualquer elemento que escape a essa exigência de justificação introduz no processo uma zona de opacidade incompatível com o controle democrático da decisão (FERRER BELTRÁN, 2017).

2.2. Cadeia de Custódia, Autenticidade e Integridade da Prova Digital

O advento das novas tecnologias implica desafios complexos e multifacetados para o sistema de justiça, sobretudo quanto à produção, coleta e confiabilidade das provas digitais no processo

penal (VILELA; ASSIS; EDUARDO, 2025). A digitalização das interações sociais facilita a comunicação e o acesso à informação, mas também produz implicações significativas para a administração da justiça e para a proteção dos direitos individuais, na medida em que desloca para o ambiente eletrônico boa parte dos vestígios relevantes à persecução penal.

Nesse contexto, a noção de cadeia de custódia ganha centralidade. Lopes Jr. (2023) observa que ela exige o estabelecimento de procedimento regrado e formalizado, capaz de documentar toda a cronologia existencial da prova, de modo a permitir sua posterior validação em juízo e o exercício do controle epistêmico. Não se trata de presumir boa-fé ou má-fé, mas de definir objetivamente um procedimento que assegure a confiabilidade da prova, independentemente do elemento subjetivo do agente que a coleta.

A tensão com os sistemas generativos torna-se evidente nesse ponto. A cadeia de custódia pressupõe a possibilidade de documentar cada etapa do trajeto da prova, do vestígio inicial à valoração final. Um conteúdo gerado por inferência probabilística, contudo, não possui propriamente uma origem rastreável no sentido tradicional, pois não é coletado, mas produzido. Falta-lhe o vestígio anterior que a custódia se destina a preservar, o que desloca o problema da preservação da prova para a verificação de sua própria gênese, terreno em que os instrumentos clássicos de rastreabilidade revelam-se insuficientes.

Ao tratar da integridade e da autenticidade da prova digital, a doutrina contemporânea enfatiza o rigor na verificação dos elementos técnicos que sustentam sua confiabilidade. A atribuição de força probante a documentos eletrônicos depende da análise do

grau de segurança quanto à sua autenticidade, entendida como a possibilidade de identificação da autoria, e à sua integridade, relacionada à garantia de inalterabilidade do conteúdo (DIDIER JR.; BRAGA; OLIVEIRA, 2026). Sem esses atributos, o dado digital, por mais sofisticado, não se converte em prova confiável.

A prova eletrônica, contudo, ainda carrega incertezas relevantes, não sendo tarefa simples conciliar evolução tecnológica e legislação. Como adverte Filho (2025, p. 670):

No cenário atual, marcado por transformações tecnológicas intensas, a prova eletrônica desponta como uma das manifestações mais emblemáticas da reconfiguração dos instrumentos de produção e valoração da verdade no processo penal. No entanto, à medida que sua presença se consolida no cotidiano da persecução penal e da defesa técnica, torna-se igualmente evidente que sua utilização plena, segura e legítima está envolta em um nevoeiro de incertezas. Problematizar esse cenário significa explorar os pontos de tensão que derivam da frágil previsibilidade do futuro da prova eletrônica, com especial atenção às dificuldades que comprometem sua eficácia jurídica diante da rápida evolução tecnológica combinada à constante defasagem normativa.

Os desafios da prova digital já tensionam, portanto, as exigências de autenticidade, integridade e rastreabilidade. A confiabilidade

probatória não decorre apenas da existência do dado, mas das condições em que ele é produzido, preservado e interpretado. Quando tais condições não podem ser adequadamente verificadas, emerge o risco de enfraquecimento das garantias, com impactos diretos sobre o modelo garantista. Esse problema, como se verá, agrava-se de modo qualitativo nos sistemas generativos, que não se limitam a armazenar ou transmitir dados, mas geram conteúdos novos a partir de inferências probabilísticas.

2.3. Standards de Prova e o Contraditório Como Condição Epistêmica

Os limites epistemológicos da prova revelam que não basta que a informação exista ou seja tecnicamente sofisticada: é imprescindível que ela seja suscetível de compreensão, controle e contestação em bases racionais. A introdução de sistemas cuja lógica interna não seja plenamente acessível desafia esse pressuposto e impõe a revisão crítica dos critérios tradicionais de admissibilidade e valoração.

A discussão articula-se com a noção de standards de prova, em especial o critério do além de qualquer dúvida razoável, que exige não apenas suficiência quantitativa de evidências, mas qualidade justificativa das inferências produzidas (FERRER BELTRÁN, 2007). A introdução de elementos cuja confiabilidade não possa ser devidamente aferida compromete diretamente esse parâmetro, na medida em que impede a exclusão racional de hipóteses alternativas, fragilizando o juízo de certeza exigido para a condenação penal.

Essa exigência conecta-se à própria estrutura dialética do processo. O contraditório, entendido em sua dimensão substancial, não se satisfaz com a simples ciência formal do elemento apresentado: pressupõe a real possibilidade de influência e de refutação. Quando a parte adversa não consegue compreender como determinada conclusão foi alcançada, o contraditório esvazia-se, convertendo-se em ritual desprovido de eficácia epistêmica. A prova, então, deixa de ser controlável e passa a operar como autoridade não justificada.

Convém ressaltar que nenhum meio de prova é infalível, e que a tradição processual sempre conviveu com graus de incerteza. A questão não reside na busca de uma certeza absoluta, inatingível, mas na exigência de que o material probatório seja submetido a critérios públicos de controle. O problema dos sistemas generativos não é, assim, a presença de erro, comum a qualquer técnica, mas a impossibilidade estrutural de identificar, mensurar e corrigir esse erro por meio de procedimentos verificáveis.

Estabelecidas essas premissas, impõe-se examinar tecnologias ainda mais complexas e potencialmente disruptivas. Diferentemente de outras ferramentas digitais, os sistemas de inteligência artificial generativa produzem conteúdos inéditos com base em modelos probabilísticos, o que intensifica os desafios relacionados à transparência, à auditabilidade e à confiabilidade epistêmica. A esse exame dedica-se o eixo seguinte.

3. INTELIGÊNCIA ARTIFICIAL GENERATIVA E SEUS CONDICIONAMENTOS TÉCNICOS

3.1. Arquitetura e Funcionamento dos Modelos Generativos

A evolução recente da inteligência artificial provocou transformações profundas no modo como informações são produzidas, analisadas e utilizadas. No campo jurídico, essas mudanças ganham destaque diante da crescente adoção de sistemas capazes de automatizar tarefas intelectuais complexas. Entre essas inovações, a inteligência artificial generativa sobressai pela habilidade de produzir conteúdos inéditos com elevado grau de sofisticação linguística.

Como observam Coelho e Barbosa (2025), trata-se de um salto qualitativo no campo jurídico:

A IAg é o salto evolutivo no campo jurídico. Enquanto os modelos anteriores reconheciam padrões e executavam comandos pré-programados, os modelos generativos conseguem criar conteúdo inédito, como textos, argumentos, resumos e pareceres com fluidez e contextualização impressionantes. Essa capacidade abre novas possibilidades para o setor jurídico, desde a síntese automática de petições complexas até a elaboração de pareceres iniciais baseados em legislação e jurisprudência.

Apesar do potencial, é fundamental compreender que tais sistemas operam com base em padrões probabilísticos, e não a partir de um compromisso intrínseco com a veracidade factual. Os modelos baseados em arquiteturas de redes neurais profundas, em especial os grandes modelos de linguagem (large language models), identificam padrões estatísticos em vastos volumes de dados

textuais e preveem a sequência linguística mais provável. Sua função não consiste em compreender o conteúdo, mas em produzir combinações discursivas que podem, ou não, corresponder a fatos verificáveis.

Essa distinção é epistemologicamente decisiva. Como adverte Bostrom (2018), sistemas avançados de inteligência artificial não operam com consciência ou intencionalidade no sentido humano, produzindo resultados a partir de processos computacionais e inferências formais. Disso decorre que seus outputs não correspondem a conhecimento em sentido epistemológico, mas a construções derivadas de processamento algorítmico, o que impõe cautela redobrada quando se cogita de utilizá-los como suporte de uma imputação penal.

Tal característica diferencia esses sistemas dos instrumentos tradicionais de busca ou das bases de dados jurídicas: estes recuperam informações preexistentes, ao passo que aqueles geram combinações novas. A diferença, longe de ser apenas técnica, projeta-se sobre o regime probatório, pois um instrumento que não recupera, mas cria, exige critérios próprios de validação, incompatíveis com a simples transposição das regras pensadas para a prova documental.

Essa distinção também separa os sistemas generativos das ferramentas algorítmicas de avaliação de risco, como as discutidas na experiência estrangeira. Estas operam sobre variáveis predefinidas e produzem escores a partir de modelos atuariais relativamente estáveis; aqueles geram linguagem nova, de comportamento variável e sem estrutura previamente fixada. Se mesmo as ferramentas atuariais já suscitaram intensa controvérsia

quanto à transparência e ao contraditório, com maior razão a cautela se impõe diante de sistemas generativos, cuja imprevisibilidade é qualitativamente superior e cujos resultados não admitem replicação garantida.

3.2. Opacidade, Não Determinismo e o Problema das Alucinações

Os sistemas generativos apresentam comportamento não determinístico: uma mesma entrada pode gerar saídas distintas em diferentes execuções, sobretudo em modelos que incorporam mecanismos de amostragem probabilística. Essa característica compromete a reprodutibilidade dos resultados, elemento central para a validação científica e, por extensão, para a confiabilidade probatória, uma vez que impede a verificação consistente das inferências produzidas.

Some-se a isso o fenômeno das alucinações, isto é, a geração de informações falsas, porém plausíveis, sem correspondência com dados verificáveis. Aliada à ausência de explicabilidade e à impossibilidade de reconstrução do raciocínio, essa propensão compromete a confiabilidade desses sistemas no contexto probatório. A literatura técnica designa por opacidade a dificuldade de compreender o modo como tais sistemas chegam a seus resultados, distinguindo-a da opacidade intencional ou daquela ligada ao analfabetismo técnico (BURRELL, 2016).

No plano jurídico, Doneda (2019) já alertava que a crescente utilização de sistemas automatizados de decisão impõe desafios relevantes ao direito, especialmente quanto à transparência, à possibilidade de compreensão dos critérios utilizados e ao controle dessas decisões pelos indivíduos afetados. A opacidade, nesse

sentido, não é um defeito acidental, mas uma propriedade estrutural de determinados modelos, o que a torna particularmente refratária às exigências de controle do processo penal.

Importa distinguir dois planos com frequência confundidos. Uma coisa é a taxa de acerto de um sistema, que pode ser elevada em determinadas tarefas; outra, bem diversa, é a possibilidade de auditar, no caso concreto, como e por que aquele resultado específico foi produzido. Mesmo um modelo estatisticamente confiável em larga escala não oferece, por si, garantia quanto à correção de um output individual, justamente o que o processo penal demanda. A confiabilidade agregada não se converte automaticamente em confiabilidade do elemento singular trazido aos autos, e é este, e não aquela, que sustenta ou afasta uma imputação.

Um dos problemas mais sensíveis reside na verossimilhança linguística dos outputs gerados, ou seja, na capacidade de produzir respostas formalmente coerentes e semanticamente plausíveis, ainda que desprovidas de lastro fático. Tal aparência de racionalidade intensifica o risco de que conteúdos incorretos sejam percebidos como confiáveis pelos operadores do direito, sobretudo quando apresentados em formato técnico ou estruturado. A discussão desloca-se, então, da mera precisão informacional para o terreno da epistemologia da prova.

Diante dessas limitações, a utilização de inteligência artificial generativa no âmbito jurídico não pode ser reduzida a questão de eficiência operacional. Ela deve ser compreendida como problema de racionalidade epistêmica e de validade cognitiva dos elementos produzidos. A adoção acrítica desses sistemas no contexto

probatório arrisca substituir a justificação racional por uma autoridade algorítmica implícita, comprometendo a estrutura dialógica do processo e enfraquecendo o controle intersubjetivo que fundamenta a decisão judicial. É esse o pano de fundo do precedente analisado a seguir.

4. O HC 1.059.475/SP E OS LIMITES EPISTÊMICOS DA ADMISSIBILIDADE

4.1. O Caso e a Ratio Decidendi do Superior Tribunal de Justiça

No julgamento do Habeas Corpus n. 1.059.475/SP, a Quinta Turma do Superior Tribunal de Justiça enfrentou, de modo direto e inédito, a problemática da utilização de relatórios produzidos por inteligência artificial generativa no âmbito probatório. O caso teve origem em acusação de injúria racial ocorrida após partida de futebol em Mirassol, no Estado de São Paulo, e revelou o crescente recurso a ferramentas tecnológicas como suporte à atividade investigativa (BRASIL, 2026).

A controvérsia central foi peculiar. Um investigador de polícia utilizou ferramentas de inteligência artificial generativa para analisar o áudio de um vídeo e produziu relatório concluindo pela presença de expressão ofensiva. A perícia oficial do Instituto de Criminalística, todavia, mediante análise fonética e acústica, concluía em sentido oposto, apontando a ausência de elementos compatíveis com o termo imputado. Ainda assim, o relatório de inteligência artificial foi utilizado para fundamentar a denúncia, posteriormente recebida.

A própria trajetória processual do writ é elucidativa. O pedido liminar foi inicialmente indeferido, vindo a ordem a ser concedida de ofício apenas por ocasião do julgamento de mérito pela Quinta Turma,

após reavaliação da questão. Esse percurso reflete a novidade e a complexidade do tema, cuja apreciação pela Corte Superior, como reconhecido ao longo do processo, ainda não havia ocorrido em sua especificidade. A hesitação inicial evidencia que a admissibilidade de provas geradas por inteligência artificial não comporta resposta intuitiva, exigindo a construção de critérios dogmáticos próprios, distintos dos tradicionalmente mobilizados para a cadeia de custódia.

Ao analisar a admissibilidade da prova, o relator deslocou o eixo argumentativo da licitude e da cadeia de custódia para o terreno da confiabilidade epistêmica do elemento informativo. Destacou que o sistema jurídico exige não apenas licitude na obtenção, mas também confiabilidade na formação da prova e capacidade de fundamentar inferências racionais. Ressaltou, ainda, limitações técnicas próprias da inteligência artificial generativa, como a possibilidade de alucinações e a inadequação para a análise de áudio, e advertiu que a divergência em relação à perícia oficial deveria ser fundamentada em critérios técnico-científicos, o que não ocorrera.

Por unanimidade, o colegiado concluiu que o relatório não apresentava confiabilidade epistêmica mínima para ser utilizado como prova penal, determinando sua exclusão e a prolação de nova decisão sem sua consideração (BRASIL, 2026). A solução dialoga com a compreensão contemporânea da prova como instrumento de justificação racional: a valoração probatória deve ancorar-se em critérios de racionalidade e controlabilidade intersubjetiva, de modo que as inferências possam ser reconstruídas e criticadas pelas partes (FERRER BELTRÁN, 2017).

Sob a perspectiva epistemológica, a decisão demonstra preocupação relevante com os limites cognitivos das tecnologias emergentes. Como observa Taruffo (2014), a função da prova não é meramente persuasiva, mas voltada à construção de conhecimento racionalmente justificado. A aparência de precisão técnica, nesses casos, pode mascarar déficits significativos de confiabilidade, sobretudo quando se considera que informações revestidas de tecnicidade tendem a exercer efeito de ancoragem sobre a formação do convencimento judicial, ainda que carentes de lastro epistêmico adequado.

4.2. Qualificação Jurídica: Entre Prova Documental, Pericial e Elemento Informativo

A qualificação jurídica de relatórios produzidos por inteligência artificial é controversa. Eles não se encaixam perfeitamente como prova documental, dada a ausência de autoria humana direta, nem como prova pericial, em razão da falta de metodologia verificável nos moldes tradicionais. Essa zona cinzenta exige cautela, pois a indevida assimilação a categorias consolidadas tende a emprestar a tais relatórios uma confiabilidade que sua estrutura técnica não autoriza.

Seria legítimo, então, reconhecê-los como meios de prova autônomos? A resposta tende à negativa. Em consonância com a tradição garantista delineada por Ferrajoli (2002), a atividade probatória deve submeter-se a critérios rigorosos de verificabilidade e controle, sob pena de legitimar decisões baseadas em elementos opacos e insuscetíveis de contraditório efetivo. A ausência de reconstrução transparente do método empregado obsta o enquadramento desses relatórios entre as provas plenas.

Disso não decorre, porém, sua absoluta imprestabilidade. A solução mais consentânea com o sistema é admiti-los, quando muito, como elementos informativos, sempre condicionados a exigentes parâmetros de transparência, auditabilidade, reprodutibilidade e submissão ao contraditório técnico. Trata-se de reconhecer-lhes, no máximo, função auxiliar e subsidiária, jamais a aptidão para, isoladamente, sustentar um juízo condenatório.

Parte da doutrina sugere acomodar esses elementos na categoria das provas atípicas ou inominadas, à semelhança do que se propõe para a prova eletrônica em geral. A solução, contudo, é apenas parcialmente satisfatória. A atipicidade probatória, admitida pelo sistema, não dispensa o atendimento aos requisitos materiais de confiabilidade: prova atípica não é prova isenta de controle. Reconhecer que um relatório gerado por inteligência artificial não se enquadra nas categorias clássicas resolve o problema da forma, mas não o da substância, que continua a residir na ausência de método verificável e de fonte de inferência identificável.

O risco mais grave reside na assimilação indevida desses relatórios ao regime da prova pericial. A perícia extrai sua força da existência de método científico controlável, da habilitação técnica do perito e da possibilidade de contraditório por meio de assistentes técnicos e quesitos. Um output generativo não reúne esses atributos: não há método replicável, não há responsável tecnicamente identificável pela inferência, e a lógica subjacente permanece inacessível. Tratá-lo como se laudo fosse equivaleria a emprestar-lhe a autoridade epistêmica da perícia sem as condições que a justificam, subvertendo a própria razão de ser do regime pericial (FERRER BELTRÁN, 2007).

Dessa distinção decorre uma consequência prática relevante. Mesmo quando admitido como elemento informativo, o relatório não pode operar como prova de corroboração apta a reforçar, por si, a hipótese acusatória. A corroboração pressupõe independência epistêmica entre os elementos: dois indícios reforçam-se mutuamente quando suas fontes de erro não coincidem. Um conteúdo cuja confiabilidade não é aferível não satisfaz esse requisito, pois não se sabe se aquilo que aparenta confirmar a acusação não é, na verdade, mero artefato do próprio sistema. No caso analisado, foi precisamente o conflito com a perícia oficial que expôs essa fragilidade, revelando que a aparência de robustez técnica não correspondia a lastro fático verificável (BRASIL, 2026).

Essa orientação encontra respaldo no marco regulatório do uso de inteligência artificial pelo Poder Judiciário. A Resolução n. 615/2025 do Conselho Nacional de Justiça, que sucedeu e ampliou a Resolução n. 332/2020, condiciona a adoção de soluções de inteligência artificial a requisitos de transparência, explicabilidade, rastreabilidade, supervisão humana e classificação de risco (CONSELHO NACIONAL DE JUSTIÇA, 2025). Embora dirigida primariamente ao Judiciário, a normativa expressa um princípio extensível à esfera probatória: a tecnologia funciona como apoio, não como substituta da racionalidade humana.

A questão ganha contornos ainda mais delicados sob a perspectiva da teoria da decisão judicial. O uso de relatórios gerados por inteligência artificial pode influenciar, de forma indireta, o processo cognitivo do julgador, pois informações apresentadas com aparência técnica tendem a exercer efeito de autoridade. Preservar a exigência de fundamentação racional e de controle das inferências é, portanto,

condição para que a incorporação dessas tecnologias não se converta em deslocamento silencioso do eixo decisório.

4.3. O Diálogo com o Direito Comparado

No plano comparado, observa-se crescente preocupação com os limites do uso de sistemas automatizados em contextos decisórios que envolvem direitos fundamentais. A experiência europeia é particularmente relevante. O Regulamento (UE) 2024/1689, conhecido como AI Act, adota abordagem baseada em risco e classifica como de alto risco os sistemas utilizados na aplicação da lei e na administração da justiça, submetendo-os a exigentes requisitos de transparência, supervisão humana, rastreabilidade e avaliação contínua ao longo de seu ciclo de vida (UNIÃO EUROPEIA, 2024).

A jurisprudência estrangeira também tem enfrentado, ainda que de forma incipiente, os desafios da utilização de sistemas algorítmicos na tomada de decisões. No caso *State v. Loomis*, a Suprema Corte de Wisconsin admitiu o uso de ferramentas algorítmicas de avaliação de risco como elemento auxiliar, desde que não constituíssem fundamento exclusivo do juízo condenatório, ressaltando, contudo, suas limitações inerentes, sobretudo quanto à opacidade dos critérios e às restrições ao contraditório (WISCONSIN SUPREME COURT, 2016).

No âmbito europeu, a Corte Europeia de Direitos Humanos tem afirmado a centralidade de critérios como transparência, proporcionalidade e possibilidade de controle no uso de tecnologias intrusivas. No caso *López Ribalda and Others v. Spain*, a Grande Câmara reconheceu que mecanismos de vigilância devem observar rigoroso teste de proporcionalidade e garantias adequadas contra

abusos, parâmetros que, embora formulados fora do contexto específico da inteligência artificial generativa, são plenamente transponíveis à análise de tecnologias algorítmicas no processo penal (EUROPEAN COURT OF HUMAN RIGHTS, 2019).

Desse conjunto emerge, no direito comparado, a noção de devido processo algorítmico, que enfatiza a adaptação das garantias processuais tradicionais às especificidades dos sistemas automatizados. Tal concepção abrange não apenas a transparência e a possibilidade de contestação, mas também o direito à explicação e à revisão humana significativa, reforçando a ideia de que a incorporação de tecnologias deve vir acompanhada de mecanismos institucionais aptos a preservar a racionalidade e a legitimidade da decisão.

Esses precedentes, embora não diretamente transponíveis ao sistema brasileiro, convergem para a necessidade de uma abordagem cautelosa e criticamente orientada. A linha que os une é nítida: sistemas automatizados podem auxiliar, jamais substituir, o juízo humano controlável. O precedente do Superior Tribunal de Justiça insere o Brasil nesse movimento, ao recusar que a aparência de objetividade técnica dispense a confiabilidade epistêmica como condição de admissibilidade.

4.4. O Devido Processo Algorítmico e o Direito à Explicação

A convergência entre as exigências epistêmicas da prova e os parâmetros do direito comparado conduz à noção de devido processo algorítmico, formulação que busca adaptar as garantias processuais tradicionais às especificidades dos sistemas automatizados. A expressão remonta ao trabalho seminal de Citron

(2008), que demonstrou como sistemas de decisão automatizada comprometem as garantias clássicas do devido processo ao colapsar a distinção entre a individualização da decisão e a generalização normativa, sem observar as salvaguardas próprias de qualquer uma delas.

Citron (2008) identifica um fenômeno particularmente sensível ao processo penal: a tendência de presumir a infalibilidade do sistema informático, que esvazia o sentido das audiências e da revisão das decisões. Quando o operador do direito trata o resultado algorítmico como dado objetivo e definitivo, o controle humano converte-se em mera homologação formal, e a garantia de ser ouvido perde substância. Daí a proposta de um novo modelo de salvaguardas, voltado a preservar transparência, prestação de contas e participação diante da automação.

Em desenvolvimento posterior, Citron e Pasquale (2014) sustentam que previsões automatizadas capazes de afetar direitos individuais devem submeter-se a um regime de devido processo, que abrange o direito à notificação, à contestação dos dados e da lógica empregada e à revisão por autoridade humana competente. Transposta ao processo penal, essa exigência implica que nenhum elemento gerado por inteligência artificial pode subtrair-se ao escrutínio do contraditório sob o argumento da complexidade técnica ou do sigilo do modelo.

Dois componentes desse modelo merecem destaque. O primeiro é o direito à explicação, que assegura ao afetado o conhecimento dos critérios e da lógica que conduziram ao resultado, condição sem a qual o contraditório se torna ilusório. O segundo é a revisão humana significativa, que não se satisfaz com a presença formal de um

juiz julgador, mas exige efetiva capacidade de compreender, questionar e, quando necessário, afastar a conclusão automatizada. Ambos figuram, em alguma medida, no Regulamento europeu, que impõe transparência e supervisão humana aos sistemas de alto risco (UNIÃO EUROPEIA, 2024).

No ordenamento brasileiro, embora inexista marco legal específico sobre o uso de inteligência artificial na prova penal, os atos normativos do Conselho Nacional de Justiça já incorporam parte desses princípios. A Resolução n. 615/2025 consagra a explicabilidade, a rastreabilidade e a supervisão humana como diretrizes obrigatórias para soluções de inteligência artificial no Judiciário, ainda que voltadas primariamente ao uso institucional pela própria magistratura (CONSELHO NACIONAL DE JUSTIÇA, 2025). Persiste, contudo, uma lacuna sensível quanto ao emprego desses sistemas na fase investigativa, momento em que, não por acaso, o relatório do caso analisado foi produzido.

A noção de devido processo algorítmico não importa a recusa da tecnologia, mas a sua sujeição às mesmas exigências de justificação que se impõem a qualquer manifestação do poder punitivo. Onde a explicação não é possível e a revisão humana não é significativa, o uso do sistema afronta o devido processo legal, ainda que o resultado, por coincidência, se mostre correto. A legitimidade da decisão penal não decorre do acerto eventual, mas da possibilidade de controlar racionalmente o caminho que a ela conduziu, exigência que a opacidade algorítmica, por definição, tende a frustrar.

Há, ademais, um risco cognitivo específico que o devido processo algorítmico procura conjurar: a deferência automática ao resultado da máquina, fenômeno descrito como viés de automação. Quanto

mais sofisticada e fluente a apresentação do output, maior a tendência de aceitá-lo sem o devido questionamento, sobretudo sob pressão de tempo e de volume de trabalho. Por isso, a salvaguarda não pode resumir-se à previsão formal de revisão humana, devendo abranger condições materiais que a tornem efetiva, sob pena de a supervisão converter-se em simples chancela do que o sistema produziu (CITRON, 2008).

5. CRÍTICA: O RISCO DA FETICHIZAÇÃO TECNOLÓGICA

5.1. A Ilusão de Objetividade Algorítmica e os Vieses Estruturais

Um dos principais perigos na adoção de inteligência artificial no processo penal é a fetichização tecnológica, isto é, a atribuição de grau excessivo de confiabilidade a sistemas automatizados. Há uma tendência contemporânea de naturalização das tecnologias digitais, frequentemente tratadas como neutras quando, na realidade, estão inseridas em estruturas econômicas e informacionais marcadas por assimetrias de poder, opacidade e capacidade de modulação comportamental (ZUBOFF, 2020).

No contexto penal, isso pode resultar em decisões baseadas em evidências frágeis, porém revestidas de aparência técnica. Trata-se de forma sofisticada de deslocamento da responsabilidade decisória, que fragiliza o papel crítico do julgador. A era digital, nesse sentido, exige equilíbrio entre o reconhecimento de padrões e a cautela quanto a conclusões precipitadas, demandando diretrizes éticas, confiáveis e transparentes para a valoração racional da prova em suporte digital (CAVET, 2025).

A advertência não conduz ao tecnoceticismo. Como pondera Carraro (2023, p. 228), por mais sofisticada que seja, a inteligência artificial “é

uma ferramenta, e como todas as ferramentas, deve ser usada com cuidado, consideração e sem alarmismos apocalípticos”. O ponto não é recusar a tecnologia, mas situá-la corretamente: instrumento sujeito a controle, e não fonte autônoma de autoridade epistêmica.

A fetichização manifesta-se ainda na ilusão de objetividade algorítmica, a crença de que decisões mediadas por inteligência artificial seriam necessariamente mais neutras ou imparciais do que as humanas. Tal percepção ignora que algoritmos são construídos a partir de bases de dados historicamente situadas, frequentemente marcadas por vieses estruturais, o que pode levar à reprodução, ou ao agravamento, de desigualdades. Muitos sistemas operam como estruturas opacas, dificilmente auditáveis e potencialmente amplificadoras de discriminações sociais, sobretudo quando alimentados por dados enviesados (O'NEIL, 2020).

No mesmo sentido, Pasquale (2015) sustenta que vivemos em uma sociedade da caixa-preta, na qual decisões automatizadas se tomam com base em critérios inacessíveis ao escrutínio público, comprometendo a transparência e a accountability. Eubanks (2018) demonstra, por sua vez, que a automação de decisões estatais tende a impactar desproporcionalmente populações vulneráveis, reproduzindo padrões históricos de exclusão. Ao conferir aparência de cientificidade às decisões automatizadas, corre-se o risco de reduzir o espaço do contraditório e de enfraquecer o dever de fundamentação.

Convém detalhar o mecanismo pelo qual o viés se instala. Sistemas treinados com dados de persecução penal pretérita aprendem padrões que refletem não a criminalidade real, mas a atuação seletiva das agências de controle. Se determinados territórios e

grupos foram historicamente mais policiados, os registros indicarão mais ocorrências ali, e o sistema tenderá a reproduzir e a intensificar esse direcionamento. Forma-se um circuito de retroalimentação: a previsão orienta a vigilância, a vigilância gera novos registros, e os novos registros confirmam a previsão inicial, conferindo aparência empírica a uma profecia que se autorrealiza (O'NEIL, 2020).

A neutralidade aparente dos dados é outra fonte de equívoco. Ainda que variáveis sensíveis, como a raça, sejam formalmente excluídas do modelo, outras informações funcionam como suas representantes indiretas, a exemplo do endereço, da escolaridade ou do histórico de contatos. A discriminação opera, então, por procuração, de modo dissimulado e dificilmente detectável, o que a torna ainda mais perigosa do que a discriminação explícita, pois se reveste do prestígio da objetividade técnica (EUBANKS, 2018).

No campo probatório, esses riscos assumem feição particular. Um relatório gerado por inteligência artificial que incorpore, ainda que involuntariamente, correlações enviesadas pode apresentar como achado técnico aquilo que é, na origem, reprodução de um padrão discriminatório. A aparência de cientificidade dificulta a refutação, pois desloca o debate do plano dos fatos para o da autoridade do método, terreno em que a defesa frequentemente carece de meios técnicos para uma contestação efetiva. A assimetria de recursos entre acusação e defesa, já presente no processo penal, agrava-se diante da opacidade algorítmica.

Por essas razões, a alegada objetividade dos sistemas automatizados deve ser recebida com reserva metódica. Não se trata de negar utilidade à tecnologia, mas de recusar a transferência irrefletida de autoridade epistêmica a artefatos cuja gênese e cujo funcionamento

permanecem, em larga medida, inacessíveis ao controle das partes. A imparcialidade não é propriedade intrínseca do algoritmo, e sim resultado de condições institucionais de transparência e contestação que precisam ser ativamente asseguradas.

5.2. A Redefinição do Ônus Argumentativo e o In Dubio Pro Reo

Diante desse quadro, a utilização de inteligência artificial no processo penal não demanda apenas critérios técnicos de validação, mas sobretudo uma postura epistemológica crítica por parte dos operadores do direito. Cabe-lhes resistir à tendência de delegação acrítica da função decisória a sistemas automatizados, preservando o papel central do juiz como garantidor dos direitos fundamentais.

Impõe-se, nesse sentido, a redefinição do ônus argumentativo. A parte que pretenda valer-se de sistemas de inteligência artificial deve assumir o dever de demonstrar, de forma clara e tecnicamente fundamentada, a confiabilidade, a validade e a auditabilidade dos resultados apresentados. Não se trata de inverter abstratamente o ônus da prova, mas de reconhecer que quem introduz um instrumento opaco no processo deve arcar com o encargo de torná-lo controlável.

A ausência dessa demonstração deve implicar a desconsideração do elemento informativo, em observância ao princípio do in dubio pro reo e às exigências epistêmicas da prova penal. A dúvida sobre a confiabilidade técnica do resultado não pode reverter em prejuízo do acusado, sob pena de se inverter a lógica de proteção que estrutura o processo penal. Onde a confiabilidade não é demonstrável, a consequência adequada é a exclusão, e não a admissão precária.

Essa redistribuição do ônus encontra amparo no próprio dever constitucional de motivação das decisões judiciais. Se a fundamentação deve permitir a reconstrução racional do raciocínio decisório, então o elemento que a impede, por ser opaco, não pode integrar legitimamente a base da decisão. A exigência de demonstração de confiabilidade não é, portanto, um ônus processual acessório, mas decorrência direta da garantia de que toda condenação se assente em razões públicas e controláveis (BRASIL, 1988).

Essa exigência harmoniza-se com os parâmetros já consolidados no direito comparado e nos atos normativos do Conselho Nacional de Justiça. Transparência, explicabilidade, rastreabilidade, reprodutibilidade e supervisão humana significativa convertem-se, assim, em condições de admissibilidade, e não em meras recomendações de boas práticas. A prova que não as satisfaça permanece, quando muito, no campo dos indícios sujeitos a corroboração independente.

Por fim, convém reconhecer que o equilíbrio aqui defendido não é estático. À medida que evoluam técnicas de explicabilidade e de auditoria de modelos, é possível que parte das objeções atualmente intransponíveis venha a ser mitigada. Sem pretensão de esgotar o tema, sustenta-se que o ônus de demonstrar essa evolução recai sobre quem invoca a tecnologia, e que, até lá, a prudência epistêmica deve prevalecer sobre o entusiasmo tecnológico.

6. CONSIDERAÇÕES FINAIS

O presente estudo dedicou-se à análise da admissibilidade e da valoração de relatórios produzidos por sistemas de inteligência

artificial generativa no processo penal brasileiro, tema cuja atualidade se impõe diante da acelerada incorporação dessas tecnologias à prática forense. Ao longo do percurso, procurou-se demonstrar que a questão não se resolve no plano da mera inovação tecnológica, mas no terreno mais profundo da epistemologia da prova e das garantias que estruturam o processo penal democrático.

Buscando responder ao questionamento sobre se tais relatórios satisfazem os requisitos epistêmicos mínimos de justificação racional, estabeleceu-se como objetivo analisar criticamente suas condições de admissibilidade e valoração. Por meio de metodologia qualitativa e teórico-dogmática, articulada ao exame de precedente paradigmático e ao diálogo com o direito comparado, foi possível compreender que a confiabilidade probatória depende de condições que esses sistemas, em sua configuração atual, não asseguram de modo satisfatório.

No primeiro eixo, constatou-se que a validade da prova penal não se esgota na admissibilidade formal, mas depende de requisitos epistêmicos rigorosos, como a verificabilidade das inferências, a controlabilidade intersubjetiva e a efetiva abertura ao contraditório. Verificou-se que a cadeia de custódia, a autenticidade e a integridade operam como condições de confiabilidade da prova digital, e que o contraditório, em sua dimensão substancial, funciona como autêntica condição de racionalidade da decisão, e não como simples formalidade.

No segundo eixo, evidenciou-se que os sistemas generativos operam segundo modelos probabilísticos, sem compromisso intrínseco com a veracidade factual. Características como comportamento não determinístico, opacidade estrutural, verossimilhança linguística e

propensão a alucinações comprometem a reprodutibilidade e a explicabilidade dos resultados, fragilizando sua aptidão para fundamentar inferências racionalmente controláveis no campo probatório.

Os eixos finais demonstraram que o precedente firmado no Habeas Corpus n. 1.059.475/SP, ao deslocar a análise da licitude para a confiabilidade epistêmica, alinha o Brasil a um movimento mais amplo de cautela observado no direito comparado. Ao mesmo tempo, revelou-se o risco da fetichização tecnológica e da ilusão de objetividade algorítmica, que podem reduzir o espaço do contraditório e enfraquecer o dever de fundamentação, sobretudo em prejuízo de populações vulneráveis.

Diante do exposto, conclui-se que relatórios produzidos por inteligência artificial generativa não devem ser admitidos como prova autônoma no processo penal, sob pena de fragilização do contraditório, da ampla defesa e do devido processo legal. Quando eventualmente utilizados, devem ser compreendidos apenas como elementos informativos auxiliares, sempre condicionados a rigorosos mecanismos de validação técnica, auditabilidade, reprodutibilidade e efetivo controle contraditório. A dúvida sobre sua confiabilidade técnica há de resolver-se em favor do acusado, e não contra ele.

Reconhece-se que o debate comporta posicionamentos divergentes. De um lado, argumenta-se que a vedação rígida desses instrumentos poderia retardar ganhos de eficiência e desperdiçar o potencial analítico das novas tecnologias na elucidação de fatos complexos. De outro, sustenta-se que a prioridade conferida às garantias é inegociável, pois a aparência de precisão técnica não substitui a confiabilidade epistêmica e pode encobrir erros de difícil

correção. Ambas as perspectivas apresentam fundamentos legítimos e merecem consideração no desenho de políticas públicas e de eventuais marcos normativos sobre o tema.

Ao término desta investigação, entende-se que a incorporação dessas tecnologias é provavelmente inevitável, mas não pode prescindir de salvaguardas epistêmicas robustas, sob pena de deslocar o eixo de fundamentação da decisão para uma aparente autoridade tecnológica. Preservar o núcleo racional e controlável da decisão penal constitui não apenas exigência técnica, mas imperativo democrático. Sugere-se, para estudos futuros, o aprofundamento de critérios objetivos de auditabilidade e explicabilidade aplicáveis ao contexto forense, bem como a investigação empírica dos efeitos de ancoragem que tais relatórios produzem sobre o convencimento judicial. O desafio que se coloca é compatibilizar inovação e racionalidade jurídica, tarefa que exige vigilância permanente e debate qualificado.

REFERÊNCIAS BIBLIOGRÁFICAS

BOSTROM, Nick. *Superinteligência: caminhos, perigos, estratégias*. Tradução de Paulo Geiger. Rio de Janeiro: DarkSide Books, 2018.

BRASIL. *Constituição (1988). Constituição da República Federativa do Brasil*. Brasília, DF: Senado Federal, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 22 maio 2026.

BRASIL. Superior Tribunal de Justiça. *Habeas Corpus n. 1.059.475/SP (2025/0487202-0)*. Relator: Min. Reynaldo Soares da Fonseca. Quinta Turma. Julgado em: 7 abr. 2026. DJEN, Brasília, DF, 14 abr. 2026. Disponível em:

<https://www.stj.jus.br/sites/portalp/Paginas/Comunicacao/Noticias/2026/08042026-Quinta-Turma-rejeita-relatorio-produzido-por-IA-como-prova-em-acao-penal.aspx>. Acesso em: 22 maio 2026.

BURRELL, Jenna. How the machine “thinks”: understanding opacity in machine learning algorithms. *Big Data & Society*, London, v. 3, n. 1, p. 1-12, jan./jun. 2016. Disponível em: <https://doi.org/10.1177/2053951715622512>. Acesso em: 22 maio 2026.

CARRARO, Felipe. *Inteligência artificial e ChatGPT: da revolução dos modelos de IA generativa à engenharia de prompt*. 1. ed. [S. l.]: Casa do Código, 2023.

CAVET, Caroline Amadori. *Prova digital*. São Paulo: Revista dos Tribunais, 2025.

CITRON, Danielle Keats. Technological due process. *Washington University Law Review*, St. Louis, v. 85, n. 6, p. 1249-1313, 2008. Disponível em: https://openscholarship.wustl.edu/law_lawreview/vol85/iss6/2. Acesso em: 22 maio 2026.

CITRON, Danielle Keats; PASQUALE, Frank. The scored society: due process for automated predictions. *Washington Law Review*, Seattle, v. 89, n. 1, p. 1-33, 2014. Disponível em: <https://digitalcommons.law.uw.edu/wlr/vol89/iss1/2/>. Acesso em: 22 maio 2026.

COELHO, Alexandre Zavaglia; BARBOSA, Maria Juliana do P. *Como a IA generativa está moldando os serviços jurídicos*. São Paulo: Revista dos Tribunais, 2025. Disponível em:

proview.thomsonreuters.com/launchapp/title/rt/monografias/374797972/v1/page/RB-1.12. Acesso em: 22 maio 2026.

CONSELHO NACIONAL DE JUSTIÇA. *Resolução n. 332, de 21 de agosto de 2020*. Dispõe sobre a ética, a transparência e a governança na produção e no uso de inteligência artificial no Poder Judiciário. Brasília, DF: CNJ, 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/3429>. Acesso em: 22 maio 2026.

CONSELHO NACIONAL DE JUSTIÇA. *Resolução n. 615, de 11 de março de 2025*. Estabelece diretrizes para o desenvolvimento, a utilização e a governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário. Brasília, DF: CNJ, 2025. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/6001>. Acesso em: 22 maio 2026.

DIDIER JR., Fredie; BRAGA, Paula Sarno; OLIVEIRA, Rafael Alexandria de. *Curso de direito processual civil: teoria da prova, direito probatório, decisão, precedente, coisa julgada, processo estrutural e tutela provisória*. v. 2. 21. ed. Salvador: JusPodivm, 2026.

DONEDA, Danilo. *Da privacidade à proteção de dados pessoais: fundamentos da Lei Geral de Proteção de Dados*. 2. ed. São Paulo: Revista dos Tribunais, 2019.

EUBANKS, Virginia. *Automating inequality: how high-tech tools profile, police, and punish the poor*. New York: St. Martin's Press, 2018.

EUROPEAN COURT OF HUMAN RIGHTS. *López Ribalda and Others v. Spain* (Applications nos. 1874/13 and 8567/13). Grand Chamber. Judgment of 17 October 2019. Disponível em:

<https://hudoc.echr.coe.int/eng?i=001-197098>. Acesso em: 22 maio 2026.

FERRAJOLI, Luigi. *Direito e razão: teoria do garantismo penal*. 2. ed. Tradução de Ana Paula Zomer Sica et al. São Paulo: Revista dos Tribunais, 2002.

FERRER BELTRÁN, Jordi. *La valoración racional de la prueba*. Madrid: Marcial Pons, 2007.

FERRER BELTRÁN, Jordi. *Prova e verdade no direito*. Tradução de Vítor de Paula Ramos. São Paulo: Revista dos Tribunais, 2017.

FILHO, Sérgio Lúcio. *Prova eletrônica em matéria penal: requisitos e formalidades*. 1. ed. [S. l.]: Editora Dialética, 2025.

LOPES JR., Aury. *Direito processual penal*. 20. ed. São Paulo: Saraiva, 2023.

O'NEIL, Cathy. *Algoritmos de destruição em massa: como o Big Data aumenta a desigualdade e ameaça a democracia*. Tradução de Rafael Abraham. Santo André: Editora Rua do Sabão, 2020.

PASQUALE, Frank. *The black box society: the secret algorithms that control money and information*. Cambridge: Harvard University Press, 2015.

TARUFFO, Michele. *A prova*. Tradução de João Gabriel Couto. São Paulo: Marcial Pons, 2014.

UNIÃO EUROPEIA. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down*

harmonised rules on artificial intelligence (Artificial Intelligence Act). Bruxelas: União Europeia, 2024. Disponível em: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32024R1689>. Acesso em: 22 maio 2026.

VILELA, Matheus Dantas; ASSIS, Daniel Araújo de; EDUARDO, Pablo. Blockchain e processo penal: os reflexos dos avanços tecnológicos na confiabilidade da prova penal. *Virtuajus*, Belo Horizonte, v. 10, n. 18, p. 295-308, 2025. Disponível em: <https://periodicos.pucminas.br/virtuajus/article/view/33086/23748>. Acesso em: 22 maio 2026.

WISCONSIN SUPREME COURT. *State v. Loomis*, 371 Wis. 2d 235, 881 N.W.2d 749 (Wis. 2016). Decidido em 13 jul. 2016. Disponível em: <https://law.justia.com/cases/wisconsin/supreme-court/2016/2015ap000157-cr.html>. Acesso em: 22 maio 2026.

ZUBOFF, Shoshana. *A era do capitalismo de vigilância: a luta por um futuro humano na nova fronteira do poder*. Tradução de George Schlesinger. Rio de Janeiro: Intrínseca, 2020.

¹ Doutor em Direito pela Instituição Toledo de Ensino (ITE) e Mestre em Direito pelo Centro Universitário Eurípides de Marília (UNIVEM). Atua como Promotor de Justiça no Ministério Público do Estado de São Paulo, sendo Pró-Reitor Acadêmico e Professor Titular do UNIVEM. É Vice-Líder do Grupo de Pesquisa DiFuSo – Direitos Fundamentais Sociais e Coordenador Regional da Escola Superior do Ministério Público do Estado de São Paulo. Currículo Lattes: <http://lattes.cnpq.br/6687308419664444>

² Rodrigo Murad Vitoriano é Mestrando em Direito Digital pelo Centro Universitário Eurípides de Marília (UNIVEM). Pós-graduado em Direito Público pela Universidade Potiguar e pela Universidade Gama Filho, com ênfases em Direito Penal, Direito Constitucional, Direito Civil e Direito Processual Civil. Exerce o cargo de Procurador Jurídico do Poder Legislativo de Jales-SP. Integra o grupo de pesquisa NODICO, com atuação em direito digital e inteligência artificial. Currículo Lattes: <https://lattes.cnpq.br/2051996455632539>