

IA generativa como meio de prova: considerações desde o HC 1.059.475

27/02/2026

O tema discutido no Habeas Corpus nº 1.059.475/SP, relatado pelo ministro Reynaldo Soares da Fonseca, no Superior Tribunal de Justiça, coloca em evidência um problema probatório de considerável relevância dogmática: a utilização de sistemas de inteligência artificial (IA) generativa como subsídio à atividade investigativa, em substituição a laudos periciais oficialmente produzidos.

O caso percorrido naquele *writ* versa sobre crime de injúria racial. A autoridade policial submeteu as imagens do fato investigado ao Instituto de Criminalística, que, mediante dois laudos tecnicamente fundados, concluiu pela impossibilidade de estabelecer qualquer correlação entre o material imagético e a agressão investigada. Insatisfeita com esse resultado, a autoridade policial determinou a submissão do mesmo material a sistemas de inteligência artificial generativa (Perplexity e Gemini), cujos *outputs* foram utilizados para se contrapor às conclusões periciais. Os *prompts* formulados na ocasião são no mínimo reveladores: (1) transcreva na íntegra o áudio do vídeo; (2) gere imagens que ilustrem o contexto; (3) explique como a imagem gerada se relaciona aos fatos.



O resultado dessa diligência policial — um artefato de inteligência artificial denominado de “Relatório Técnico” — que contrariava a perícia oficial, foi utilizado para oferecimento da denúncia, recebida pelo magistrado.

Diante da instauração do processo penal com base nesse elemento confessadamente produzido por meio de inteligência artificial generativa, a defesa impetrou Habeas Corpus ao Tribunal de Justiça de São Paulo, que convalidou o material e denegou a ordem, determinando o prosseguimento do feito na origem. Então, em novo HC, formulado perante o STJ, o *writ* não foi conhecido pelo relator que, posteriormente, em agravo regimental interposto pela defesa, reconsiderou a decisão e deferiu a liminar para suspender o processo penal a que responde o paciente. A impetração, de todo modo, aguarda julgamento pela 5ª Turma.

A impugnação defensiva fundou-se, primariamente, na violação à cadeia de custódia, instituto positivado pelos artigos 158-A a 158-F do Código de Processo Penal, com redação dada pela Lei nº 13.964/2019. É necessário, porém, precisar os contornos desse argumento, já que essa categoria processual padece de imprecisões, especialmente por parte do judiciário e da doutrina, que sofrem para separar a documentação que demonstra a cadeia de custódia da própria cadeia de custódia como uma série de procedimentos adotados para garantia da confiabilidade probatória.

Esse instituto, contudo, não se restringe às exigências de rastreabilidade e auditabilidade do material probatório; estas são consequências instrumentais de uma função mais ampla, voltada à preservação da integridade e autenticidade da prova. Nas certas palavras de Parodi, a cadeia de custódia da prova é o “conjunto de procedimentos a serem adotados para garantir a manutenção da integridade (e integralidade) de um vestígio, evidência ou prova” [1].

O problema do caso em tela, contudo, antecede logicamente qualquer discussão sobre cadeia de custódia: para que se cogite de rastreabilidade, é necessário que exista, antes de mais, uma prova a ser rastreada. Solicitar à inteligência artificial generativa que elabore transcrições e produza imagens contextuais vinculadas ao episódio investigado não gera elemento dotado de conteúdo probatório, muito menos *fiável*. Cuida-se de criação de conteúdo sintético, computacionalmente gerado, sem correspondência necessária com a realidade dos fatos. O problema da cadeia de custódia é, portanto, secundário diante de um vício mais grave e anterior: a inexistência de prova.

A prova digital, amplamente debatida na doutrina e na jurisprudência comparadas, ingressa no processo mediante extração e translação: dados armazenados em suportes digitais são convertidos em linguagem inteligível, com observância de metodologia que garanta a integridade do dado extraído. O fenômeno ora em análise não se confunde com essa categoria. Aproxima-se, antes, do que a literatura e a jurisprudência norteamericanas denominam de *automated generated evidence* [2]: elementos produzidos diretamente por sistemas computacionais autônomos, sem correspondência direta com dados preexistentes.

A distinção é relevante porquanto, no caso em exame, a inteligência artificial foi instada não a processar dados reais, mas a criar conteúdo imagético e interpretativo. Instaurou-se, assim, uma circularidade epistêmica inaceitável: o sistema produziu o dado e, em seguida, interpretou o dado que ele próprio havia produzido. E o pior: foi tratado como prova (!) pelas autoridades de primeiro e segundo graus da Justiça paulista. Prova não é qualquer elemento informacional que sirva para apurar a pretensa verdade sobre um fato. Menos ainda é o produto de uma operação fechada sustentada por algoritmos que não podem sequer ser explicados.

Confiabilidade como pressuposto de admissibilidade probatória: distinção entre prova ilícita e prova não confiável

A inadmissibilidade do material produzido pelos sistemas de inteligência artificial não decorre, em sua essência, da ilicitude na obtenção; o tal “relatório técnico” objeto do debate no Superior Tribunal de Justiça não foi obtido mediante violação a direitos ou garantias fundamentais. O



problema é de natureza distinta e, em certa medida, mais grave: trata-se de uma *não-prova*, de um elemento desprovido do atributo mínimo que condiciona a sua admissão no processo penal, a saber, a confiabilidade epistêmica. A analogia mais precisa é aquela que compara essa utilização da inteligência artificial à consulta a um cartomante para avaliação da culpabilidade do investigado, cujo resultado seria inadmissível não por ilicitude, mas por ausência de qualquer credencial epistemológica.

O mesmo raciocínio aplica-se, com exatidão lógica, ao caso em apreço, e é também pertinente a comparação com a questão das cartas psicografadas no processo penal. A cartomante ou a carta psicografada poderia coincidir com o fato “real”? Evidentemente que sim, o que não autoriza o seu uso. Golpes de sorte ou escolhas acidentais ou contingenciais não legitimam decisões, posto que não são justificáveis. O problema é que o Direito não pode se dar ao luxo de se contentar com decisões que poderiam até ser moralmente defensáveis, embora normativamente questionáveis. Do contrário o juiz Bridoye descrito por Rabelais deveria ser admitido como um modelo judicial, já que estatisticamente talvez até errasse menos do que alguns juízes no Brasil, ao decidir a sorte dos acusados lançando dados.

Acresça-se o fenômeno das alucinações dos sistemas de inteligência artificial generativa – amplamente documentado – que os torna propensos a gerar informações inverídicas com aparência de veracidade, circunstância que, por si só, já comprometeria fatalmente a confiabilidade de seus *outputs* em contextos de exigência de precisão factual.

A distinção conceitual entre prova *ilícita* e prova *não confiável* tem implicações dogmáticas (e práticas) relevantes: embora ambas produzam a mesma consequência de exclusão processual, seus fundamentos são distintos e não se podem confundir — a falta de confiabilidade é um problema de qualidade epistêmica; a ilicitude, de regularidade no modo de obtenção. Ademais, a ausência de transparência da polícia também é notável: por que usar tais softwares e não outros? Nenhum *prompt* foi usado em outros softwares? Ou somente estes trouxeram a resposta que se almejava?

O princípio da liberdade das provas — que remete ao utilitarismo de Jeremy Bentham e que não passa de um arroubo da própria noção irrefreada de livre conhecimento, este por sua vez um epifenômeno da íntima convicção — tal como historicamente concebido no processo penal brasileiro, admite o ingresso do maior volume possível de elementos informativos, com a única vedação das provas ilicitamente obtidas ou que confrontem valores sociais relevantes.

Essa orientação revela-se insuficiente — e anacrônica — para fazer face aos desafios probatórios postos pelas novas tecnologias que se instalaram e colonizaram a práxis do processo penal. Na doutrina crítica nacional, a propósito, destaca-se a acertada construção do professor Aury Lopes Jr., que identifica um princípio da especialidade probatória, segundo o qual cada meio de prova deve ser utilizado para a finalidade que lhe é própria, vedado o seu emprego para fins distintos [3].

Essa formulação aponta na direção correta: os maus usos da prova devem ser defesos, e o que se verificou no caso em análise configura, sem dúvida, um mau uso qualificado. O aspecto de maior gravidade, porém, reside na utilização dos relatórios da inteligência artificial

(sugestivamente denominados de “Relatórios Técnicos”) para suplantar as conclusões dos laudos periciais do Instituto de Criminalística. Embora o artigo 182 do Código de Processo Penal faculte ao juiz afastar as conclusões periciais, essa liberdade não é arbitrária: pressupõe fundamento técnico-científico idôneo.

O princípio que subjaz a essa exigência é simples e incontornável: somente a técnica refuta a técnica. Afastar laudos periciais oficiais com apoio em *outputs* de inteligência artificial generativa, desprovidos de qualquer metodologia verificável, representa violação ao princípio da especialidade técnica e ao dever constitucional de motivação racional das decisões judiciais, fulminando qualquer chance de que o processo se forme desde a concepção fazzalariana de *procedimento em contraditório* [4].

Por ora, algumas considerações sobre um tema em aberto

Embora o ministro relator do HC nº 1.059.475/SP tenha reconsiderado sua decisão inicial, de não conhecer da impetração, e deferido a medida liminar para suspender o andamento do processo penal na origem, o Superior Tribunal de Justiça ainda não decidiu o mérito das questões trazidas na impetração, que, quiçá, poderão perpassar os pontos acima percorridos.

De toda forma, o iminente enfrentamento desses temas evidencia a urgência de que doutrina, jurisprudência e legislador avancem na construção de um marco normativo para o uso da inteligência artificial no processo penal. Dentre as principais balizas de tal marco estão, sem sombra de dúvidas, a necessidade de se distinguir, com rigor conceitual, prova ilícita de prova não confiável, reconhecendo que a falta de confiabilidade epistêmica é também causa autônoma de inadmissibilidade.

O aporte de novas tecnologias ao sistema probatório é inevitável e, em muitos contextos, desejável. O que não se pode admitir é que essa incorporação se faça à margem dos princípios inegociáveis do contraditório, da ampla defesa e da integridade epistêmica do processo penal.

[1] PARODI, Lorenzo. Perícia defensiva em provas digitais no processo penal: origem, custódia, integralidade e integridade. São Paulo: Revista dos Tribunais, 2024, p. 68.

[2] Cf., nesse sentido, QUATTROCOLO, Serena. An introduction to ai and criminal justice in Europe. Revista Brasileira de Direito Processual Penal, Porto Alegre, vol. 5, n. 3, p. 1519-1554, set./dez. 2019.

[3] LOPES JR., Aury. Direito processual penal. 19. ed. São Paulo: SaraivaJur, 2022, p. 452-453.

[4] Cf., nesse sentido, DALMAGRO JR., João Carlos. Mineração de dados, pensamento computacional e persecução penal: a difícil coexistência entre a opacidade e o processo como procedimento em contraditório. In GIACOMOLLI, Nereu José; AQUINO, Pâmela de Araújo. Processo penal contemporâneo em debate: processo penal na era digital. Vol. 10. 1. ed. São Paulo: Tirant Lo Blanch, p. 120-137, 2025.



Fonte: <https://conjur.jumps.com.br/2026-fev-27/inteligencia-artificial-generativa-como-meio-de-prova-consideracoes-desde-o-hc-1-059-475-sp-do-stj/>