

INTELIGÊNCIA ARTIFICIAL E A EPISTEMOLOGIA DO DIREITO PENAL: OS LIMITES DA SUPERVISÃO HUMANA DIANTE DAS ALUCINAÇÕES JUDICIAIS

ARTIFICIAL INTELLIGENCE AND THE EPISTEMOLOGY OF CRIMINAL LAW:
THE LIMITS OF HUMAN OVERSIGHT IN THE FACE OF JUDICIAL
HALLUCINATIONS

Ciências Humanas • 09/10/2026

REGISTRO DOI: [10.70773/revistatopicos/791516507](https://doi.org/10.70773/revistatopicos/791516507)

Elias Saraiva dos Santos Bisneto¹

Alex Silva Gonçalves²

RESUMO

Este artigo investiga se a supervisão humana exigida pela Resolução CNJ nº 615/2025 é epistemologicamente suficiente para prevenir falhas na fundamentação de decisões penais apoiadas em Inteligência Artificial (IA) generativa. Adota-se abordagem dedutivo-analítica, combinando pesquisa bibliográfica e documental com o exame do caso julgado pelo Órgão Especial do Tribunal de Justiça de São Paulo (TJ-SP) em julho de 2026, no qual se manteve o arquivamento de reclamação disciplinar contra magistrado que fundamentou decisão em precedentes jurisprudenciais inexistentes, gerados por IA. Os resultados indicam que a Resolução CNJ nº 615/2025, ao classificar a interpretação de fatos como crime entre as hipóteses de alto risco e exigir supervisão humana “efetiva, periódica e adequada” (art. 3º, VII), estabelece dever formal de revisão que não garante, por si só, verificação epistêmica substantiva do conteúdo gerado pela IA, como evidencia o caso examinado, em que a revisão humana existiu sem impedir a incorporação de jurisprudência fictícia à decisão. Conclui-se que a supervisão humana, tal como estruturada pela norma vigente, é necessária, mas insuficiente para preservar o contraditório e a ampla defesa em matéria penal, recomendando-se a adoção de protocolos de verificação obrigatória de citações e precedentes antes da prolação de decisões apoiadas em IA generativa.

Palavras-chave: inteligência artificial; supervisão humana; alucinação algorítmica.

ABSTRACT

This article investigates whether the human oversight required by CNJ Resolution No. 615/2025 is epistemologically sufficient to prevent failures in the grounds of criminal decisions supported by generative Artificial Intelligence (AI). The study adopts a deductive-

analytical approach, combining bibliographic and documentary research with the case decided by the Special Body of the São Paulo Court of Justice (TJ-SP) in July 2026, which upheld the dismissal of a disciplinary complaint against a judge who grounded a decision on nonexistent case-law precedents generated by AI. The results indicate that CNJ Resolution No. 615/2025, by classifying the interpretation of facts as crimes among its high-risk hypotheses and requiring “effective, periodic and adequate” human oversight (Art. 3, VII), establishes a formal duty of review that does not, by itself, guarantee substantive epistemic verification of AI-generated content, as shown by the examined case, in which human review took place without preventing fictitious case law from being incorporated into the decision. The study concludes that human oversight, as structured, is necessary but insufficient to preserve adversarial proceedings and the right to a full defense in criminal matters, and recommends adopting mandatory verification protocols for citations and precedents before issuing decisions supported by generative AI.

Keywords: artificial intelligence; human oversight; algorithmic hallucination.

1. INTRODUÇÃO

A Inteligência Artificial (IA) está se consolidando como uma das maiores inovações tecnológicas da atualidade, revolucionando setores que vão desde a economia e a educação até o Direito. No campo jurídico, a IA oferece promessas de transformação, incluindo a possibilidade de agilizar processos, reduzir custos e ampliar o acesso à justiça por meio da automação de tarefas, da análise de dados e da previsão de resultados processuais.

Com o uso de tecnologias como *Big Data*, Jurimetria e *Machine Learning*, advogados e magistrados conseguem organizar e interpretar um grande volume de informações, o que contribui para decisões mais eficientes e fundamentadas. Como aponta Rosa (2020), a IA no Direito é capaz de aumentar a produtividade e, ao mesmo tempo, melhorar a qualidade da justiça ao simplificar o acesso a dados relevantes e padrões jurisprudenciais.

Contudo, o uso da IA no sistema jurídico também apresenta uma série de desafios éticos e epistemológicos. Em um campo onde a interpretação contextual e a análise humana são essenciais para garantir uma aplicação justa e equitativa das normas, o uso de algoritmos para decisões automatizadas levanta questionamentos sobre a transparência, a explicabilidade e os riscos de viés algorítmico.

Esses riscos deixaram de ser hipotéticos. Em julho de 2026, o Órgão Especial do Tribunal de Justiça de São Paulo (TJ-SP) manteve o arquivamento de reclamação disciplinar apresentada pela Ordem dos Advogados do Brasil (OAB/SP) contra magistrado que, ao fundamentar uma decisão, incluiu precedentes jurisprudenciais inexistentes, gerados por ferramenta de IA generativa, fenômeno conhecido como "alucinação". A relatora, desembargadora Silvia Rocha, corregedora-geral da Justiça de São Paulo, reconheceu o "erro indesejável", mas concluiu que ele não comprometeu a apreciação dos fatos nem o resultado do julgamento, afastando a hipótese de falta funcional (Migalhas, 2026a). O episódio é emblemático porque a supervisão humana formal existiu: o magistrado assinou e submeteu a decisão, sem que isso impedisse a incorporação de jurisprudência fictícia à fundamentação.

Gomes, Vaz e Dias (2020) destacam que, embora a IA possa contribuir para o aprimoramento da justiça, seu uso deve ser pautado por princípios éticos rigorosos e submetido a supervisão humana para evitar que decisões automatizadas comprometam direitos fundamentais e a própria integridade do sistema jurídico. A implementação de IA no Judiciário exige, portanto, uma reflexão sobre os limites e as responsabilidades dessa tecnologia, para que seu uso esteja em conformidade com os princípios de justiça, equidade e transparência que orientam o Direito.

A Resolução CNJ nº 615/2025, em vigor desde julho de 2025, é hoje o marco normativo que disciplina o desenvolvimento, a governança, a auditoria e o uso de soluções de IA no Poder Judiciário brasileiro, tendo revogado a Resolução CNJ nº 332/2020. Entre seus princípios está a "supervisão humana efetiva, periódica e adequada no ciclo de vida da inteligência artificial, considerando o grau de risco envolvido" (art. 3º, VII), e seu Anexo de Classificação de Riscos qualifica como de alto risco a "averiguação, valoração, tipificação e a interpretação de fatos como sendo crimes" (item AR3), reconhecendo, portanto, que o uso de IA em matéria penal exige cautela redobrada. A LGPD (Lei nº 13.709/2018) complementa esse arcabouço ao regular o tratamento de dados pessoais e reforçar o direito à explicação e à revisão das decisões automatizadas.

Diante desse contexto, o presente artigo tem como pergunta central: a supervisão humana exigida pela Resolução CNJ nº 615/2025 é suficiente, do ponto de vista epistemológico, para prevenir falhas como a ocorrida no caso do TJ-SP na fundamentação de decisões penais apoiadas em IA generativa? A hipótese defendida é que essa supervisão, tal como estruturada pela norma vigente, um dever formal de revisão e de possibilidade de

modificação do conteúdo gerado, é necessária, mas não suficiente para garantir a verificação epistêmica substantiva exigida pelo contraditório e pela ampla defesa em matéria penal.

Para testar essa hipótese, o artigo articula a fundamentação teórica sobre epistemologia jurídica e aplicações da IA no Direito com a análise normativa da Resolução CNJ nº 615/2025 e com o exame do caso concreto do TJ-SP. A metodologia empregada é dedutivo-analítica, combinando pesquisa bibliográfica e documental com estudo de caso, a partir da análise de textos acadêmicos, da legislação e regulamentação aplicáveis e de reportagem jurídica especializada sobre o episódio examinado.

O estudo está dividido em três partes, além desta introdução. A primeira apresenta a fundamentação teórica, abordando o conceito de IA, a epistemologia jurídica e as aplicações práticas da IA no sistema jurídico. A segunda examina os limites éticos e epistemológicos da IA no sistema jurídico à luz da Resolução CNJ nº 615/2025, culminando na análise do caso do TJ-SP e no que ele revela sobre os limites da supervisão humana formal. Por fim, a conclusão retoma a pergunta de pesquisa e apresenta recomendações para o uso de IA generativa em decisões penais.

2. REFERENCIAL TEÓRICO: FUNDAMENTAÇÃO E APLICAÇÕES DA INTELIGÊNCIA ARTIFICIAL NO DIREITO

A aplicação da Inteligência Artificial (IA) no sistema jurídico tem potencial para otimizar processos, aumentar a eficiência e promover o acesso à justiça. No entanto, a sua introdução em um campo em que o conhecimento e a prática são fortemente baseados na interpretação humana levanta questões epistemológicas e éticas

que exigem uma análise aprofundada. Esta seção explora os conceitos fundamentais de IA, as particularidades da epistemologia jurídica e as principais aplicações práticas da IA no Direito, destacando tanto suas contribuições quanto os desafios ao sistema de justiça.

2.1. Conceito de Inteligência Artificial e Evolução no Contexto Jurídico

A Inteligência Artificial é amplamente definida como a capacidade de sistemas computacionais realizarem tarefas que exigem habilidades humanas, como reconhecimento de padrões, aprendizado e tomada de decisão.

No contexto jurídico, a IA é aplicada principalmente por meio de algoritmos de *Machine Learning* e análise de grandes volumes de dados, permitindo a automação de tarefas e a análise preditiva de decisões. Segundo Frazão e Mulholland (2020), a IA aplicada ao Direito envolve sistemas que processam informações complexas para identificar padrões e gerar variações baseadas em decisões anteriores, o que facilita o trabalho de advogados, juízes e demais operadores do Direito ao oferecer subsídios para a resolução de processos.

A evolução da IA no Direito se dá principalmente pela introdução de tecnologias como *Big Data* e Jurimetria, que processam grandes conjuntos de dados para apoiar decisões judiciais e administrativas. O uso dessas tecnologias permite uma visão panorâmica dos precedentes, o que facilita a uniformidade e a previsibilidade das decisões. Rosa (2020) observa que a IA oferece uma nova dimensão para o Direito, possibilitando o uso de algoritmos para o

reconhecimento de padrões em processos e na identificação de tendências que informam o sistema jurídico.

2.1.1. Diversidade de Perspectivas Sobre a Inteligência Artificial no Direito

A literatura brasileira sobre o tema não se limita às obras citadas acima. Tacca e Rocha (2018) já apontavam, ainda antes da consolidação regulatória do CNJ, que a incorporação da IA ao sistema do Direito produz reflexos que extrapolam a mera automação, afetando a própria forma como o sistema jurídico processa e valida o conhecimento. Na mesma linha, Oliveira Júnior e Leonel (2021) descrevem o estado da arte da IA aplicada ao Direito no Brasil como um campo em que avanços tecnológicos convivem com lacunas regulatórias e desafios de governança, diagnóstico que viria a ser parcialmente respondido, quatro anos depois, pela edição da Resolução CNJ nº 615/2025.

2.2. Epistemologia Jurídica e a Importância da Interpretação Humana

A epistemologia jurídica é o campo que estuda a natureza, os limites e os métodos de construção do conhecimento jurídico. No Direito, o conhecimento é construído a partir da interpretação das normas jurídicas, em que o aplicador da lei, geralmente o juiz, analisa a legislação e os princípios jurídicos considerando o contexto social e histórico do caso. Para Gomes, Vaz e Dias (2020), o processo de interpretação e aplicação da lei no Direito depende de uma subjetividade e de uma sensibilidade que vão além dos dados brutos, o que torna o campo jurídico uma área única em sua relação com o conhecimento.

Diferentemente da ciência e da tecnologia, que operam com dados empíricos e objetividade, o Direito trabalha com conceitos normativos e valores sociais, cuja aplicação varia conforme o caso. Segundo Nieva Fenoll (2018), o conhecimento jurídico não se limita à literalidade da lei, mas é enriquecido pelo julgamento humano, que considera as especificidades e os valores de cada situação. Nesse contexto, a epistemologia jurídica valoriza a interpretação humana, que possibilita uma análise mais completa e contextual das normas aplicadas, ao contrário da IA, que opera com uma lógica estatística e quantitativa.

Outro ponto fundamental da epistemologia jurídica é a explicabilidade das decisões, que exige que o juiz fundamente claramente os motivos que levaram a determinado veredito. Essa prática não apenas fortalece a transparência, mas também respeita o direito ao contraditório e à ampla defesa, permitindo que as partes compreendam e questionem as decisões.

Gomes, Vaz e Dias (2020) destacam que, para preservar a integridade da justiça, é necessário que as decisões sejam compreensíveis, o que coloca em questão a aplicação da IA no Direito, pois os algoritmos nem sempre oferecem uma explicação clara sobre como chegaram a uma conclusão específica.

Com o uso de IA, surgem desafios importantes, especialmente no que tange à "caixa-preta" dos algoritmos, situação em que o processo decisório interno de um algoritmo é opaco e difícil de entender. A LGPD e a Resolução CNJ nº 615/2025 reforçam a importância da transparência e da supervisão humana em decisões baseadas em IA, ao exigir explicabilidade e possibilidade de revisão dos resultados apresentados (art. 3º, II e VII). Isso se alinha com a

epistemologia jurídica, que exige que o julgamento por trás de uma decisão seja acessível e compreensível para todas as partes envolvidas.

2.3. Aplicações Práticas da IA no Sistema Jurídico

A aplicação de IA no Direito abrange diversas áreas, cada uma com benefícios específicos e desafios próprios. A seguir, discutem-se as principais aplicações práticas e como elas interagem com os critérios epistemológicos do Direito.

2.3.1. Automatização de Tarefas Repetitivas

A automação de tarefas repetitivas no campo jurídico representa uma das aplicações mais práticas e imediatas da Inteligência Artificial (IA). Por meio da automação, atividades burocráticas e processuais, como a organização de documentos, a triagem de processos, a classificação de documentos e a pesquisa de jurisprudência, podem ser realizadas com maior rapidez e precisão, permitindo que advogados e magistrados dediquem mais tempo a tarefas de maior complexidade e análise crítica. Como destaca Rosa (2020), a automação tem o potencial de transformar o sistema jurídico, reduzindo a sobrecarga de trabalho e otimizando a tramitação de processos, sem interrupções, o que também reduz o tempo de espera para as partes envolvidas (Frazão; Mulholland, 2020).

Outro aspecto relevante da automação de tarefas repetitivas é a possibilidade de minimizar erros humanos, especialmente em atividades de grande volume, como a análise e a categorização de milhares de documentos processuais. De acordo com Gomes, Vaz e Dias (2020), a automatização permite que os dados sejam

processados com maior precisão, o que reduz a incidência de erros que podem comprometer o andamento de um processo judicial, contribuindo para aumentar a confiabilidade das etapas processuais. Rosa (2020) ressalta ainda que essa otimização de recursos é particularmente relevante em um contexto de crescente demanda por serviços jurídicos e de recursos limitados.

2.3.2. Análise Preditiva e Jurisprudência

A análise preditiva no contexto jurídico é uma aplicação da Inteligência Artificial (IA) que utiliza algoritmos para prever desfechos de processos judiciais com base em dados históricos e padrões jurisprudenciais, permitindo que advogados e magistrados analisem probabilidades de êxito e identifiquem tendências que podem influenciar a tomada de decisões. Conforme aponta Rosa (2020), essa tecnologia possibilita maior uniformidade e previsibilidade nas decisões judiciais, auxiliando a formulação de estratégias processuais.

Além de auxiliar a prática advocatícia, a análise preditiva contribui para a gestão do sistema judiciário, permitindo que os tribunais antecipem demandas e otimizem a distribuição de recursos. Gomes, Vaz e Dias (2020) destacam que a IA aplicada à jurisprudência fornece uma visão ampla e integrada do sistema; alertam, contudo, que o uso de IA na análise preditiva requer controle rigoroso para evitar distorções e vieses que comprometam a imparcialidade das decisões, sob pena de o Judiciário tornar-se excessivamente dependente de decisões automatizadas, abrindo margem para uma "jurisprudência automatizada" que não reflete os valores e princípios fundamentais do Direito.

Por essa razão, a análise preditiva e o uso de IA na jurisprudência precisam ser implementados com supervisão humana constante, funcionando como apoio, e não como substituto, do julgamento humano. Rosa (2020) argumenta que, embora a IA possa auxiliar na uniformização das decisões, é essencial que os magistrados mantenham autonomia para interpretar o Direito conforme o contexto e as especificidades de cada caso.

2.3.3. Experiências Internacionais e Nacionais no Uso de IA no Direito

O uso da Inteligência Artificial (IA) no Direito vem ganhando destaque em diversos países. Nos Estados Unidos, a plataforma "ROSS Intelligence" utiliza processamento de linguagem natural para auxiliar advogados na pesquisa de jurisprudência; conforme destaca Rosa (2020), a ferramenta permite pesquisa mais ágil e precisa. Na Estônia, um sistema de IA resolve disputas judiciais de pequeno valor e de baixa complexidade, permitindo que os juízes concentrem-se em casos que demandam maior análise humana, modelo que Frazão e Mulholland (2020) apontam como um dos mais inovadores, embora também levante questionamentos sobre transparência e supervisão das decisões automatizadas.

No Brasil, o Conselho Nacional de Justiça (CNJ) lidera iniciativas como o projeto "Sinapses", que utiliza IA para identificar padrões em processos judiciais e reduzir demandas repetitivas, exemplo, segundo Gomes, Vaz e Dias (2020), de como a IA pode otimizar o trabalho dos tribunais. O sistema "Victor", do Supremo Tribunal Federal (STF), auxilia na triagem de recursos extraordinários, classificando e priorizando processos para que os ministros concentrem-se nos casos de maior relevância; para Rosa (2020), o

Victor foi fundamental para reduzir o acúmulo de processos no Supremo. Essas experiências, nacionais e internacionais, evidenciam o potencial da IA para o sistema jurídico, mas reforçam a necessidade de regulamentação e supervisão que evitem que a automação comprometa os princípios fundamentais do Direito, exigência que ganhou contornos mais precisos com a edição da Resolução CNJ nº 615/2025, examinada na seção seguinte.

2.4. Potencialidades e Limitações da IA no Direito

A IA no Direito possui potencial transformador, promovendo eficiência, acessibilidade e previsibilidade no sistema de justiça. A automação de tarefas repetitivas permite economia de tempo e de recursos, enquanto a análise preditiva oferece um nível de previsibilidade que contribui para a segurança jurídica. Frazão e Mulholland (2020) observam que o uso da IA pode tornar o sistema jurídico mais ágil, acessível e menos oneroso para o cidadão, facilitando o acesso à justiça.

No entanto, a IA enfrenta limitações e desafios importantes, especialmente quanto à transparência e à explicabilidade dos algoritmos. A opacidade algorítmica compromete o direito das partes ao contraditório, pois a explicabilidade é essencial para que as decisões judiciais sejam compreensíveis e contestáveis.

Além disso, o viés algorítmico é uma preocupação central, pois algoritmos treinados com dados históricos podem perpetuar injustiças passadas. Como enfatiza Rosa (2020), é fundamental que a IA no Direito seja empregada de forma responsável, com supervisão humana ativa, garantindo que os princípios de equidade e justiça sejam respeitados.

A epistemologia jurídica, ao valorizar a subjetividade e o julgamento humano, corresponde a um modelo de prática jurídica que a IA ainda não é capaz de reproduzir integralmente. O uso de IA deve, portanto, ser limitado ao papel de assistência, em que a tecnologia complementa, mas nunca substitui, o julgamento interpretativo e analítico do operador do Direito.

3. LIMITES ÉTICOS E EPISTEMOLÓGICOS DA IA NO SISTEMA JURÍDICO

A aplicação da Inteligência Artificial (IA) no sistema jurídico tem o potencial de revolucionar a prática jurídica, mas também impõe desafios éticos e epistemológicos que precisam ser enfrentados para garantir a integridade do Direito. A utilização de algoritmos em decisões judiciais envolve questões como a transparência, a necessidade de supervisão humana e o risco de viés algorítmico, que afetam diretamente a imparcialidade e a justiça no processo decisório. Esta seção examina esses limites, abordando a importância da explicabilidade dos algoritmos, a relevância da supervisão humana e os riscos de discriminação algorítmica.

3.1. Transparência e Explicabilidade dos Algoritmos

A transparência é um princípio fundamental para o uso ético e responsável da Inteligência Artificial (IA) no sistema jurídico, pois permite que as partes compreendam como uma decisão foi alcançada. A "opacidade algorítmica" (fenômeno em que os processos internos de um algoritmo são complexos demais para serem interpretados por não especialistas) constitui preocupação central no uso da IA no Direito.

Como aponta Rosa (2020), a opacidade dos algoritmos compromete o direito ao contraditório e à ampla defesa, na medida em que impede que as partes compreendam e, conseqüentemente, contestem as decisões automatizadas. Em um sistema jurídico que valoriza a transparência, a ausência de explicabilidade mina a legitimidade e a confiabilidade das decisões.

A Resolução CNJ nº 615/2025 estabelece, entre seus princípios, a explicabilidade, a contestabilidade e a auditabilidade das soluções de IA (art. 3º, II), de modo que o raciocínio por trás das decisões automatizadas seja acessível às partes envolvidas. Para Gomes, Vaz e Dias (2020), a explicabilidade é condição indispensável para a legitimidade das decisões automatizadas, pois assegura que as partes possam entender os critérios utilizados pelo algoritmo e, assim, exercer seu direito ao contraditório de maneira plena. Além disso, a transparência permite que as decisões automatizadas sejam revisadas e questionadas, promovendo uma aplicação mais justa e confiável da IA no Direito.

A Lei Geral de Proteção de Dados (LGPD) também contribui para a transparência no uso da IA ao garantir o direito dos cidadãos a obterem informações claras e acessíveis sobre decisões automatizadas que impactem seus direitos. Conforme estabelecido pela LGPD, os cidadãos têm o direito de entender como e por que uma decisão automatizada foi tomada e, mais importante, o direito de solicitar revisão humana dessas decisões. Segundo Rosa (2020), esse direito à revisão humana é essencial para assegurar que o processo decisório automatizado respeite os princípios de justiça e imparcialidade, pois permite a correção caso o algoritmo tenha cometido erro ou apresentado viés.

Por fim, a explicabilidade dos algoritmos é aspecto fundamental para o cumprimento do direito ao contraditório, pois permite que as partes afetadas compreendam o raciocínio do sistema e possam contestar a decisão com base em fundamentos claros. Sem transparência, os algoritmos de IA funcionam como caixas-pretas, nas quais os operadores do Direito não conseguem verificar se as decisões são justas e adequadas. Conforme pontuam Frazão e Mulholland (2020), a explicabilidade é componente central para que a IA seja usada de forma ética, de modo a garantir que o sistema de justiça mantenha sua integridade e a possibilitar que as partes exerçam seu direito de defesa.

3.2. Supervisão Humana e Responsabilidade nas Decisões Automatizadas

A supervisão humana é componente essencial para o uso seguro e ético da Inteligência Artificial (IA) no sistema jurídico. A Resolução CNJ nº 615/2025 exige que os sistemas computacionais do Judiciário "exijam a supervisão humana e permitam a modificação pelo magistrado competente de qualquer produto gerado pela inteligência artificial" (art. 34), vedando expressamente soluções "que não possibilitem a revisão humana dos resultados propostos" (art. 10, I) e reafirmando, quanto ao uso de modelos de linguagem, que o magistrado "permanecerá integralmente responsável pelas decisões tomadas e pelas informações nelas contidas" (art. 19, § 3º, II). A norma estrutura, assim, a supervisão humana como uma exigência formal e uma atribuição de responsabilidade, não, porém, como um procedimento de verificação obrigatória do conteúdo gerado, distinção que se revela decisiva à luz do caso examinado na seção 3.5.

Rosa (2020) defende que a presença humana na supervisão de decisões automáticas é essencial para que essas decisões respeitem os valores fundamentais do Direito. Em outras palavras, a supervisão humana oferece camada de proteção contra decisões automatizadas que possam desconsiderar nuances específicas e complexidades dos casos, essenciais para uma análise justa e contextualizada.

Além de garantir que a IA seja aplicada de forma ética, a supervisão humana é fundamental para a *accountability*, ou seja, para a responsabilização no sistema jurídico. Uma decisão judicial não pode ser atribuída exclusivamente ao algoritmo, pois isso comprometeria a possibilidade de identificação de erros e violações de direitos. Eubanks (2018) alerta para os perigos da "responsabilização algorítmica", fenômeno em que decisões errôneas são justificadas com base na tecnologia, eximindo o operador do Direito de sua responsabilidade; para a autora, é justamente a supervisão humana que evita essa desresponsabilização e garante que o uso da IA permaneça alinhado aos valores fundamentais da justiça.

A supervisão humana também permite que os operadores do Direito ajustem ou corrijam decisões sugeridas pela IA em casos em que o sistema automatizado apresente falhas, como viés algorítmico ou interpretações limitadas das normas jurídicas. Para Gomes, Vaz e Dias (2020), o controle humano é indispensável para corrigir potenciais erros dos algoritmos, pois apenas o ser humano é capaz de interpretar o contexto de forma ampla e de considerar fatores que vão além dos dados quantitativos. Em resumo, a supervisão humana atua como salvaguarda contra erros e injustiças que

possam decorrer do uso de IA, permitindo que as decisões finais levem em conta o julgamento e a ética humana.

Outro ponto importante é a necessidade de que a supervisão humana seja substancial, e não meramente formal. Não basta que um ser humano revise superficialmente as decisões automatizadas; é essencial que os operadores tenham compreensão detalhada do funcionamento dos algoritmos e da lógica por trás das decisões sugeridas pela IA. Frazão e Mulholland (2020) defendem que a supervisão humana deve ser ativa e baseada no conhecimento dos operadores sobre os processos algorítmicos, de modo que possam avaliar criticamente a qualidade e a justiça das decisões. Isso significa que os operadores do Direito devem estar capacitados para entender e interpretar os resultados fornecidos pela IA, assegurando que a tecnologia seja ferramenta complementar, e não limitação, ao processo de análise.

Por fim, a supervisão humana é crucial para preservar a integridade do sistema jurídico e garantir que o uso de IA não comprometa os direitos das partes. A Resolução CNJ nº 615/2025 elenca, entre seus princípios, "o devido processo legal, a ampla defesa e o contraditório" (art. 3º, V), a serem observados na governança das soluções de IA. Segundo Rosa (2020), a supervisão ativa e bem informada permite que o operador do Direito garanta a justiça das decisões, intervindo sempre que necessário para ajustar ou reavaliar os resultados. Dessa forma, a presença de supervisão humana não apenas promove justiça e imparcialidade, mas também fortalece a confiança da sociedade no sistema de justiça, assegurando que as decisões sejam fundamentadas e justas.

3.3. Viés Algorítmico e Discriminação

O viés algorítmico é um dos maiores desafios éticos da IA aplicada ao Direito. Conforme destacado por Rosa (2020), algoritmos treinados com dados históricos podem perpetuar e até amplificar preconceitos, comprometendo a justiça das decisões. Isso ocorre principalmente porque os dados usados para treinar esses algoritmos refletem desigualdades históricas, resultando em decisões enviesadas contra determinados grupos sociais.

A Resolução CNJ nº 615/2025 também aborda essa questão, determinando que os produtos gerados pela IA para suporte a decisões judiciais devem preservar a igualdade e a não discriminação, com adoção de medidas preventivas e corretivas sempre que verificado viés discriminatório (art. 8º). Gomes, Vaz e Dias (2020) argumentam que a supervisão humana e a auditoria são fundamentais para identificar e corrigir vieses, assegurando que a IA no Direito promova a justiça e a equidade. Dessa forma, o viés algorítmico pode ser mitigado por protocolos de auditoria e revisão que assegurem que os dados utilizados sejam representativos e equilibrados.

3.4. Impacto na Epistemologia Jurídica: A Perda de Subjetividade e de Contexto

A utilização da IA no Direito desafia a epistemologia jurídica tradicional, que depende da análise contextual e da subjetividade para a interpretação das normas jurídicas. Segundo Gomes, Vaz e Dias (2020), a IA, por se basear em padrões estatísticos, tende a desconsiderar as especificidades de cada caso, comprometendo a adaptação necessária do Direito às transformações sociais. A prática jurídica requer interpretação que os algoritmos, dada sua natureza matemática, não conseguem fornecer integralmente.

Rosa (2020) ressalta que o conhecimento jurídico é construído a partir da interpretação humana, que leva em conta o contexto e os valores sociais. Assim, a IA deve ser utilizada com cautela, respeitando o papel interpretativo do operador do Direito, para que as decisões continuem refletindo os princípios fundamentais do sistema jurídico. A Resolução CNJ nº 615/2025 também sustenta esse entendimento, ao vedar a utilização de IA generativa "como instrumento autônomo de tomada de decisões judiciais sem a devida orientação, interpretação, verificação e revisão por parte do magistrado" (art. 19, § 3º, II).

3.5. O Caso do TJ-SP e os Limites da Supervisão Humana Formal

O caso julgado pelo Órgão Especial do TJ-SP em julho de 2026 oferece um teste empírico para a distinção traçada na seção 3.2 entre supervisão humana como exigência formal e supervisão humana como verificação epistêmica substantiva. Segundo o noticiado por Migalhas (2026a), a reclamação disciplinar foi apresentada pela OAB/SP após magistrado fundamentar decisão em precedentes jurisprudenciais que, segundo a petição, não existiam, tendo sido gerados por ferramenta de IA generativa. Em sessão de 1º de julho de 2026, a corregedora-geral Silvia Rocha, relatora do caso, reconheceu "erro indesejável do magistrado na reprodução de alguns precedentes jurisprudenciais", mas concluiu que o equívoco "não foi intencional" e "não interferiu na análise dos fatos e no resultado do julgamento", contextualizando o episódio no "atual cenário de adaptação tecnológica do sistema de justiça". O Órgão Especial, em 8 de julho de 2026, negou provimento ao recurso e manteve o arquivamento.

O contraste com o tratamento dispensado a advogados pelo mesmo Tribunal, no mesmo período, é revelador. Em maio de 2026, a própria desembargadora Silvia Rocha, como relatora no Conselho Superior da Magistratura do TJ-SP, não conheceu recurso cuja petição continha dispositivo legal inexistente e trechos de obras doutrinárias jamais publicadas, indícios que o voto qualificou como uso abusivo e temerário de IA generativa sem a supervisão humana devida, determinando o encaminhamento do caso ao Tribunal de Ética e Disciplina da OAB/SP, no Processo nº 1001904-24.2024.8.26.0543 (Migalhas, 2026b).

Na mesma linha, a 34ª Câmara de Direito Privado do TJ-SP aplicou multa de 5% sobre o valor da causa, por litigância de má-fé, a advogado que reconheceu ter inserido, sem conferência, precedentes gerados por ferramenta de IA, no Processo nº 1106363-51.2024.8.26.0002, e a 11ª Câmara de Direito Privado impôs multa de ofício em hipótese semelhante, invocando precedentes do próprio Tribunal que reiteradamente qualificam a citação de jurisprudência inexistente como litigância de má-fé, no Processo nº 2057948-55.2026.8.26.0000 (Migalhas, 2026b).

A mesma relatora que, em relação ao magistrado, atribuiu o erro a uma fase de aprendizado institucional, qualificou episódio estruturalmente idêntico, quando praticado por advogado, como uso abusivo, com encaminhamento disciplinar. Isso sugere que a insuficiência identificada neste artigo não é apenas normativa, mas também de aplicação: o dever de supervisão humana recai, em tese, sobre todos os atores do processo por força dos arts. 3º, VII, e 19, § 3º, II, da Resolução CNJ nº 615/2025, mas sua exigibilidade prática tem sido desigual: mais rigorosa para advogados do que para o

magistrado, justamente o agente cuja decisão pode produzir, sozinha, efeitos definitivos sobre a esfera jurídica das partes.

O episódio é revelador precisamente porque a supervisão humana, no sentido formal exigido pelo art. 34 da Resolução CNJ nº 615/2025, esteve presente: o magistrado assinou a decisão e permaneceu, nos termos do art. 19, § 3º, II, da mesma Resolução, "integralmente responsável pelas decisões tomadas e pelas informações nelas contidas".

Essa responsabilidade formal, contudo, não impediu que jurisprudência inexistente fosse incorporada à fundamentação e chegasse a produzir efeitos até ser questionada pela parte prejudicada. A supervisão humana, tal como estruturada pela norma, garante que exista um responsável identificável pela decisão, mas não garante, por si só, que esse responsável tenha verificado a veracidade de cada citação e precedente gerado pela IA antes de assiná-la.

Essa lacuna ganha relevância acentuada em matéria penal. O Anexo de Classificação de Riscos da Resolução CNJ nº 615/2025 qualifica como de alto risco, no item AR3, a "averiguação, valoração, tipificação e a interpretação de fatos como sendo crimes", e no item AR2, a "aferição da adequação dos meios de prova e a sua valoração nos processos de jurisdição contenciosa". Se uma decisão penal viesse a se fundamentar em precedente jurisprudencial inexistente (hipótese estruturalmente idêntica à do caso do TJ-SP), o prejuízo recairia sobre o contraditório e a ampla defesa do próprio réu, que teria o direito de refutar uma tese jurídica construída sobre fundamento fático inexistente. A gravidade potencial nesses casos é

qualitativamente distinta daquela associada a decisões cíveis ou administrativas, dado o bem jurídico em jogo: a liberdade.

Esse achado sustenta a hipótese central deste artigo: a supervisão humana exigida pela Resolução CNJ nº 615/2025 é necessária, mas epistemologicamente insuficiente, para prevenir falhas como a examinada. A norma exige que exista revisão e que exista responsabilidade; não exige, em nenhum de seus dispositivos, um protocolo específico de checagem da veracidade de citações e precedentes antes da prolação da decisão. Preencher essa lacuna, por meio de rotinas obrigatórias de verificação (análogas às exigidas de advogados após episódios semelhantes registrados no Brasil, nos Estados Unidos e em Portugal desde 2023), é o passo que a epistemologia jurídica, na forma descrita na seção 2.2, recomenda para que a supervisão humana deixe de ser apenas formal e passe a ser substantiva.

3.6. Limites Éticos e a Proteção dos Direitos Fundamentais

A LGPD e a Resolução CNJ nº 615/2025 estabelecem diretrizes para o uso ético da IA no Judiciário brasileiro, garantindo a proteção dos direitos dos cidadãos e promovendo uma aplicação justa e transparente da tecnologia. A LGPD assegura o direito dos cidadãos de contestar decisões automatizadas e de solicitar revisão humana, enquanto a Resolução CNJ nº 615/2025 exige que as decisões judiciais baseadas em IA respeitem princípios de ética e transparência.

Conforme ressaltam Gomes, Vaz e Dias (2020), a LGPD e as regulamentações do CNJ são essenciais para assegurar que a IA no Direito promova a justiça e os direitos dos cidadãos, sem

comprometer os valores fundamentais do sistema jurídico. Essas regulamentações refletem o compromisso do Judiciário com o uso responsável da IA, de modo que a tecnologia complemente, e não substitua, o julgamento humano, preservando a dignidade e a integridade das decisões.

4. CONCLUSÃO

Este artigo partiu de uma pergunta específica: a supervisão humana exigida pela Resolução CNJ nº 615/2025 é suficiente, do ponto de vista epistemológico, para prevenir falhas na fundamentação de decisões penais apoiadas em IA generativa? A resposta, à luz do referencial teórico desenvolvido e do caso examinado, é negativa: a supervisão humana, tal como estruturada pela norma vigente, é necessária, mas não suficiente.

A Resolução CNJ nº 615/2025 representa um avanço normativo relevante em relação ao regime anterior, ao classificar a interpretação de fatos como crime entre as hipóteses de alto risco (Anexo, item AR3), ao exigir supervisão humana efetiva, periódica e adequada (art. 3º, VII) e ao impor que o magistrado permaneça integralmente responsável pelas decisões tomadas com apoio de IA (art. 19, § 3º, II). Esses dispositivos, no entanto, definem um dever formal de revisão e de responsabilização, não um procedimento de verificação obrigatória da veracidade de citações, precedentes e fatos gerados pela IA antes da prolação da decisão.

O caso julgado pelo Órgão Especial do TJ-SP em julho de 2026 evidencia exatamente essa lacuna. A supervisão humana formal existiu: o magistrado assinou e submeteu a decisão, permanecendo identificável e responsável nos termos da norma. Ainda assim,

jurisprudência inexistente integrou a fundamentação e somente foi identificada porque a parte prejudicada a questionou.

O órgão disciplinar, ao arquivar a reclamação sob o fundamento de que o Judiciário "ainda está aprendendo" a lidar com a tecnologia, tratou o episódio como falha de adaptação, não como evidência de uma insuficiência estrutural do modelo de supervisão vigente, tratamento que contrasta com o dispensado pelo mesmo Tribunal, na mesma relatoria, a advogados que incorreram em falha análoga e foram encaminhados à OAB ou multados por litigância de má-fé, como visto na seção 3.5. Essa assimetria reforça que a lacuna identificada neste artigo não é meramente teórica.

Essa distinção entre supervisão formal e verificação epistêmica substantiva conecta-se diretamente à epistemologia jurídica desenvolvida na seção 2.2: o conhecimento jurídico depende de uma subjetividade e de uma sensibilidade interpretativa que a IA, operando por lógica estatística, não reproduz. Uma revisão humana que se limita a assinar o produto da IA sem verificar suas premissas fáticas não exerce essa função interpretativa: apenas transfere a responsabilidade formal, sem transferir o controle epistêmico efetivo sobre o conteúdo da decisão.

As implicações são particularmente graves em matéria penal, onde o bem jurídico em jogo é a liberdade e onde o contraditório e a ampla defesa pressupõem que o réu possa refutar os fundamentos fáticos e jurídicos de uma acusação ou condenação. Uma decisão penal fundamentada em precedente jurisprudencial inexistente, hipótese estruturalmente idêntica à do caso examinado, comprometeria essas garantias de forma potencialmente

irreversível, sobretudo se não identificada antes do trânsito em julgado.

Recomenda-se, portanto, que tribunais e órgãos de correição adotem protocolos específicos de verificação obrigatória de citações e precedentes sempre que IA generativa for utilizada como apoio à fundamentação de decisões classificadas como de alto risco pela Resolução CNJ nº 615/2025, com atenção redobrada às hipóteses do item AR3 de seu anexo. Tal protocolo poderia incluir a exigência de que o magistrado ateste, expressamente, a conferência individual de cada precedente citado, nos mesmos moldes já recomendados à advocacia diante de episódios semelhantes registrados no Brasil, nos Estados Unidos e em Portugal desde 2023.

Por fim, este estudo não pretende esgotar o tema, mas contribuir para o debate sobre um ponto específico: a distância entre exigir supervisão humana e garantir verificação epistêmica. Estudos futuros poderão examinar se casos análogos ao do TJ-SP se repetem em outros tribunais brasileiros e se os protocolos de auditoria previstos na Resolução CNJ nº 615/2025 são, na prática, capazes de captar falhas desse tipo antes que produzam efeitos sobre os jurisdicionados.

REFERÊNCIAS BIBLIOGRÁFICAS

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018.** Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 15 set. 2026.

CONSELHO NACIONAL DE JUSTIÇA (Brasil). **Resolução nº 615, de 11 de março de 2025**. Estabelece diretrizes para o desenvolvimento, utilização e 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: 15 set. 2026.

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

FRAZÃO, Ana; MULHOLLAND, Caitlin (coord.). **Inteligência artificial e Direito**: ética, regulação e responsabilidade. 2. ed. São Paulo: Thomson Reuters Brasil, 2020.

GOMES, Eduardo Biacchi; VAZ, Andréa Arruda; DIAS, Sandra Mara de Oliveira. Limites éticos para o uso da inteligência artificial no sistema de justiça brasileiro, de acordo com a Lei 13.709 de 2018 (LGPD) e Resoluções 331 e 332 do Conselho Nacional de Justiça. **Revista Internacional Consinter de Direito**, v. 6, n. 11, 2020.

MIGALHAS. Juiz cita jurisprudência inexistente e TJ/SP não aplica punição: “estamos aprendendo”. **Migalhas Quentes**, 14 jul. 2026a. Disponível em: <https://www.migalhas.com.br/quentes/460324/juiz-cita-jurisprudencia-inexistente-e-corregedora-vota-contra-punicao>. Acesso em: 15 set. 2026.

MIGALHAS. Uso de IA: decisões expõem contraste entre punições a advogados e magistrados. **Migalhas Quentes**, 22 jul. 2026b. Disponível em: <https://www.migalhas.com.br/quentes/460846/ia-decisoes-expoem-contraste-entre-punicoes-a-advogados-e-magistrados>. Acesso em: 15 set. 2026.

NIEVA FENOLL, Jordi. **Inteligencia artificial y proceso judicial**. Madrid: Marcial Pons, 2018.

OLIVEIRA JÚNIOR, Ézio; LEONEL, Vilson. Limites e possibilidades da inteligência artificial aplicada ao Direito: estado da arte no Brasil. **Unisul de Fato e de Direito**, v. 11, n. 22, p. 45-56, jan./jun. 2021. Disponível em: https://portaldeperiodicos.animaeducacao.com.br/index.php/U_Fato_Direito/article/view/19939/13428. Acesso em: 9 nov. 2024.

ROSA, Alexandre Moraes da. **Inteligência artificial e Direito: ensinando um robô a julgar**. Salvador: JusPodivm, 2020.

TACCA, Adriano; ROCHA, Leonel Severo. Inteligência artificial: reflexos no sistema do direito. **NOMOS**: Revista do Programa de Pós-Graduação em Direito da UFC, Fortaleza, v. 38, n. 2, p. 53-68, jul./dez. 2018. Disponível em: <https://periodicos.ufc.br/nomos/article/view/20493>. Acesso em: 9 nov. 2024.

¹ Professor de Direito Penal no Centro Universitário Paraíso (Juazeiro do Norte, CE, Brasil). Mestrando em Direito Constitucional pela UNIFOR, Especialista em Direito Penal e Criminologia pela Universidade Regional do Cariri. Membro Consultor da Comissão Nacional de Acesso à Justiça do Conselho Federal da OAB. E-mail: [acesse o artigo original para visualizar o e-mail](#). ORCID: <https://orcid.org/0009-0004-6593-6528>

² Doutor em Direito pela Universidade Federal de Santa Catarina (UFSC), na área de concentração Direito, Política e Sociedade (2023). Mestre em Direito pela Universidade de Santa Cruz do Sul (UNISC),

na área de concentração Direitos Sociais e Políticas Públicas, linha de pesquisa Políticas Públicas de Inclusão Social (2017). Graduado em Direito pela Universidade Regional do Cariri (URCA, 2006). Especialista em Direito Ambiental e em Direito Penal Militar e Processual Penal Militar. Professor da graduação e da pós-graduação lato sensu do Centro Universitário Paraíso (UniFAP) e da pós-graduação lato sensu da Universidade Regional do Cariri (URCA). Advogado. E-mail: [acesse o artigo original para visualizar o e-mail](#). ORCID: <https://orcid.org/0000-0001-5491-2257>.