

INTELIGÊNCIA ARTIFICIAL NA ELABORAÇÃO DE MINUTAS JUDICIAIS: VIESES, MANIPULAÇÕES E OS DESAFIOS DA SEGURANÇA JURÍDICA PROCESSUAL**ARTIFICIAL INTELLIGENCE IN THE DRAFTING OF JUDICIAL DECISIONS: BIASES, MANIPULATIONS AND THE CHALLENGES OF PROCEDURAL LEGAL CERTAINTY** <https://doi.org/10.63330/aurumpub.066-002>**Gustavo Davanço Nardi**

Mestrando em Direito na UNIMAR

E-mail: nardiadv@gmail.comLattes: <http://lattes.cnpq.br/1049125778986007>**Jaqueline Amâncio de Freitas**

Aluna de Direito – Anhanguera

Lattes: <http://lattes.cnpq.br/9987458489270910>**Resumo**

Sistemas de inteligência artificial vêm sendo progressivamente incorporados à rotina decisória do Poder Judiciário brasileiro, sobretudo como ferramentas de auxílio à elaboração de minutas judiciais. Diante desse cenário, problematiza-se em que medida a utilização dessas ferramentas compromete a segurança jurídica processual e o dever constitucional de fundamentação das decisões judiciais. Sustenta-se, como tese central, que a inteligência artificial generativa, por operar mediante inferências estatísticas e não mediante compreensão semântica do Direito, introduz riscos estruturais — vieses algorítmicos, vies de automação, alucinações e vulnerabilidade a manipulações como a injeção de comandos — que não podem ser neutralizados por uma supervisão humana meramente formal e posterior. Metodologicamente, adota-se abordagem qualitativa, de natureza bibliográfica, documental e normativa, com método hipotético-dedutivo, examinando-se doutrina nacional e estrangeira, a Resolução CNJ nº 615/2025 e o parâmetro constitucional do artigo 93, inciso IX. Como resultado, demonstra-se que a legitimidade da decisão judicial permanece indissociável de uma reserva de humanidade insuperável pela automação, de modo que a inteligência artificial somente se justifica como instrumento auxiliar, jamais como agente substitutivo da convicção jurisdicional. O estudo contribui para o debate ao propor parâmetros de supervisão contínua e qualificada, distintos da mera revisão textual final, capazes de compatibilizar eficiência tecnológica e garantias processuais fundamentais.

Palavras-chave: Governança judicial; Inteligência artificial; Segurança jurídica; Vieses algorítmicos; Fundamentação das decisões judiciais.

Abstract

Artificial intelligence systems have been progressively incorporated into the decision-making routine of the Brazilian Judiciary, particularly as tools to support the drafting of judicial decisions. Against this backdrop, the article questions to what extent the use of such tools compromises procedural legal certainty and the constitutional duty to state reasons for judicial decisions. It is argued, as a central thesis, that generative artificial intelligence, operating through statistical inference rather than semantic understanding of Law, introduces structural risks — algorithmic bias, automation bias, hallucinations and vulnerability to manipulation such as prompt injection — that cannot be neutralized through merely formal and retrospective human supervision. Methodologically, a qualitative approach is adopted, of bibliographic, documentary and normative nature, employing the hypothetical-deductive method, examining national and foreign doctrine, CNJ Resolution No. 615/2025 and the constitutional parameter of Article 93, IX. As a result, it is demonstrated that the legitimacy of judicial decisions remains inseparable from a reserve of humanity that automation cannot overcome, so that artificial intelligence is only justified as an auxiliary instrument, never as a substitute agent for judicial conviction. The study contributes to the debate by proposing continuous and qualified supervision parameters, distinct from mere final textual review, capable of reconciling technological efficiency with fundamental procedural guarantees.

Keywords: Algorithmic bias; Artificial intelligence; Judicial governance; Legal certainty; Reasoning of judicial decisions.

1 INTRODUÇÃO

Transformações tecnológicas raramente ingressam nas instituições jurídicas de maneira gradual e previsível; costumam irromper sob a forma de promessas de eficiência que, somente depois de assimiladas pela prática cotidiana, revelam suas implicações mais profundas sobre estruturas normativas consolidadas. Esse fenômeno tem se manifestado com particular intensidade no âmbito do Poder Judiciário brasileiro, no qual sistemas de inteligência artificial passaram, em curto intervalo temporal, de experimentos incipientes de automação a instrumentos cotidianos de apoio à elaboração de minutas judiciais.

Diante da sobrecarga estrutural que caracteriza o sistema de justiça brasileiro — marcado por elevado volume processual, escassez de recursos humanos e pressão institucional por produtividade —, torna-se compreensível o entusiasmo despertado por ferramentas capazes de organizar informações, sistematizar teses jurídicas e sugerir redações para peças processuais. Entretanto, esse entusiasmo não pode obscurecer a natureza intrinsecamente probabilística dos sistemas de inteligência artificial generativa, cuja lógica operacional difere substancialmente daquela pressuposta pelo raciocínio jurídico tradicional.

Nesse contexto, formula-se o seguinte problema de pesquisa: em que medida a utilização de sistemas de inteligência artificial na elaboração de minutas judiciais compromete a segurança jurídica processual e o dever constitucional de fundamentação das decisões, mesmo diante das diretrizes estabelecidas pela Resolução CNJ nº 615/2025?

A pergunta não comporta resposta trivial, porquanto exige articulação entre categorias tecnológicas — vieses algorítmicos, viés de automação, alucinações, injeção de comandos — e categorias constitucionais e processuais consolidadas, como a exigência de fundamentação prevista no artigo 93, inciso IX, da Constituição Federal.

Sustenta-se, como hipótese central do presente trabalho, que a inteligência artificial generativa, por operar mediante inferências estatísticas sobre padrões linguísticos e não mediante compreensão semântica ou normativa do Direito, introduz riscos estruturais que não podem ser neutralizados por uma supervisão humana concebida como simples revisão textual posterior.

Defende-se, ademais, que a legitimidade da função jurisdicional pressupõe uma reserva de humanidade insuperável pela automação, de modo que a tecnologia somente se justifica como instrumento auxiliar da atividade decisória, jamais como agente substitutivo da convicção do magistrado.

Justifica-se a relevância científica e social do presente estudo pela ausência de consenso doutrinário consolidado acerca dos limites jurídicos aplicáveis à utilização de inteligência artificial generativa no processo decisório judicial, bem como pela recente edição da Resolução CNJ nº 615/2025, cujo conteúdo normativo ainda carece de análise crítica aprofundada quanto à sua efetividade prática.

Compreender os riscos técnicos e jurídicos envolvidos nessa transição tecnológica constitui condição necessária para a preservação de garantias processuais fundamentais, notadamente a imparcialidade, a fundamentação racional e a previsibilidade das decisões judiciais.

Como objetivo geral, pretende-se analisar criticamente os riscos jurídicos decorrentes da utilização de inteligência artificial na elaboração de minutas judiciais, examinando sua compatibilidade com o dever constitucional de fundamentação e com os princípios estabelecidos pela regulamentação do Conselho Nacional de Justiça.

Como objetivos específicos, busca-se: primeiro, examinar as categorias técnicas de vieses algorítmicos, viés de automação e alucinações, situando-as no contexto da atividade jurisdicional; segundo, analisar criticamente o modelo de governança estabelecido pela Resolução CNJ nº 615/2025, com destaque para a insuficiência de uma supervisão humana meramente formal; terceiro, articular essas categorias com o dever constitucional de fundamentação das decisões judiciais, evidenciando os limites da automação frente à reserva de humanidade inerente à função jurisdicional.

Estruturalmente, o presente artigo divide-se, após esta introdução e a exposição metodológica, em três grandes eixos de desenvolvimento. Primeiramente, examina-se a reconfiguração da função

jurisdicional diante das características técnicas dos sistemas de inteligência artificial generativa, desconstruindo o mito da neutralidade algorítmica.

Em seguida, analisa-se o modelo de governança judicial estabelecido pela Resolução CNJ nº 615/2025, evidenciando os limites de uma supervisão humana concebida em termos meramente formais. Por fim, discute-se a fundamentação constitucional das decisões judiciais como parâmetro normativo capaz de delimitar juridicamente os espaços legítimos de utilização da inteligência artificial no processo decisório.

Metodologicamente, conforme será detalhado na seção subsequente, adota-se abordagem qualitativa, de natureza bibliográfica, documental e normativa, valendo-se do método hipotético-dedutivo como instrumento de articulação entre a hipótese formulada e o material doutrinário, normativo e jurisprudencial examinado.

2 METODOLOGIA

Adota-se, no presente estudo, abordagem qualitativa, por se tratar de investigação orientada à compreensão de fenômenos jurídicos complexos que não se reduzem à mensuração quantitativa de dados, mas exigem interpretação crítica de categorias normativas, doutrinárias e institucionais.

Quanto à natureza, a pesquisa classifica-se como bibliográfica, documental e normativa: bibliográfica, porquanto se apoia em doutrina jurídica nacional e estrangeira consolidada acerca de inteligência artificial, processo civil constitucional e teoria da decisão judicial; documental, na medida em que examina diretamente o texto da Resolução CNJ nº 615/2025 e demais atos normativos pertinentes; normativa, por analisar dispositivos constitucionais e infraconstitucionais aplicáveis ao dever de fundamentação das decisões judiciais.

Quanto ao método de abordagem, emprega-se o método hipotético-dedutivo, na medida em que se parte de uma hipótese previamente formulada — a insuficiência da supervisão humana meramente formal diante dos riscos estruturais da inteligência artificial generativa — submetendo-a a confronto sistemático com o material doutrinário, normativo e teórico disponível, com vistas à sua confirmação, refutação ou qualificação.

Complementarmente, utiliza-se o método dedutivo na análise dos dispositivos normativos examinados, partindo-se de categorias gerais — como os princípios da transparência, supervisão humana e centralidade da pessoa humana previstos na Resolução CNJ nº 615/2025 — para a análise de sua aplicação ao caso específico da elaboração de minutas judiciais.

Os critérios de seleção das fontes bibliográficas privilegiaram obras e artigos científicos de reconhecida relevância acadêmica no campo da inteligência artificial aplicada ao Direito, da teoria da decisão judicial e da filosofia da tecnologia, com atenção especial a autores que já enfrentaram, em perspectiva crítica, a questão dos vieses algorítmicos e da automação de funções decisórias.

Complementarmente, examinou-se a literatura internacional sobre governança algorítmica e responsabilidade institucional diante de sistemas automatizados, de modo a enriquecer a análise com perspectiva comparada.

3 RESULTADOS E DISCUSSÃO

3.1 A RECONFIGURAÇÃO DA FUNÇÃO JURISDICIONAL DIANTE DA INTELIGÊNCIA ARTIFICIAL GENERATIVA: VIESES, VIÉS DE AUTOMAÇÃO E O MITO DA NEUTRALIDADE ALGORÍTMICA

Compreender adequadamente os riscos jurídicos decorrentes da utilização de inteligência artificial na elaboração de minutas judiciais pressupõe, antes de qualquer análise normativa, o enfrentamento de uma premissa amplamente difundida no discurso público e institucional: a suposta neutralidade dos sistemas algorítmicos, logo essa premissa, embora atraente por conferir aparência de objetividade científica às ferramentas tecnológicas, não resiste a exame mais detido de sua base epistemológica.

O pesquisador Frank Pasquale (2015), ao analisar o funcionamento opaco dos algoritmos que estruturam decisões institucionais contemporâneas, demonstra que sistemas automatizados não constituem instrumentos assépticos de processamento de informação, mas reproduzem, em escala ampliada, as escolhas — muitas vezes invisíveis — de seus desenvolvedores e os padrões estatísticos presentes nos dados de treinamento.

Nesse sentido, opacidade estrutural, que o autor denomina metaforicamente de "sociedade da caixa-preta", impede que usuários e até mesmo especialistas compreendam integralmente os critérios internos que orientam determinada resposta algorítmica.

Transpondo essa constatação para o contexto jurídico brasileiro, Nunes e Marques (2018) observam que os sistemas de inteligência artificial são treinados a partir de dados produzidos pela própria sociedade e, por essa razão, tendem a reproduzir padrões e distorções já existentes nesses conjuntos de informações.

Assim, quando aplicados à elaboração de minutas judiciais, tais sistemas não introduzem uma perspectiva neutra sobre o litígio, mas incorporam, de maneira frequentemente imperceptível, regularidades estatísticas que podem privilegiar determinadas teses jurídicas, determinados perfis de litigantes ou determinadas formas de argumentação em detrimento de outras.

Nesse sentido, a promessa de objetividade tecnológica revela-se, em boa medida, ilusória: substitui-se a subjetividade explícita do intérprete humano por uma subjetividade difusa, dissimulada sob a aparência de precisão matemática.

Essa constatação adquire especial gravidade quando articulada a um segundo fenômeno, intimamente relacionado ao primeiro, mas conceitualmente distinto: o viés de automação.

Além do mais, para Cass Sunstein (2018), em estudos dedicados à influência de sistemas automatizados sobre o comportamento decisório humano, evidencia que indivíduos tendem a atribuir maior credibilidade a respostas produzidas por mecanismos tecnológicos simplesmente em razão da aparência de objetividade que tais respostas ostentam, independentemente de sua efetiva correção material.

Esse fenômeno, identificado originalmente em contextos de automação industrial e aeronáutica, manifesta-se de maneira particularmente perigosa no ambiente forense, no qual a elevada carga de trabalho e a pressão institucional por produtividade criam condições propícias à confiança excessiva nas sugestões produzidas por sistemas de inteligência artificial.

Consequência direta desse quadro é a diminuição do exercício crítico necessário à atividade jurisdicional. Transfere-se, ainda que de forma parcial e muitas vezes involuntária, etapas relevantes da análise jurídica para mecanismos cuja lógica interna nem sempre é compreendida por seus próprios usuários.

Trata-se de fenômeno que Antoine Garapon e Jean Lassègue (2018), ao analisarem a chamada "justiça digital", denominam de deslocamento silencioso da responsabilidade interpretativa, no qual a autoridade epistêmica do julgador é gradualmente substituída pela autoridade estatística do sistema, sem que essa substituição seja objeto de deliberação institucional explícita.

O problema alcança dimensão ainda mais complexa quando se examina o funcionamento específico das inteligências artificiais generativas, categoria à qual pertencem os grandes modelos de linguagem atualmente empregados na elaboração de minutas judiciais.

Diferentemente dos sistemas tradicionais de recuperação de informação, que se limitam a localizar e apresentar dados previamente armazenados em bases estruturadas, os modelos generativos produzem textos inteiramente novos a partir de inferências probabilísticas sobre a sequência mais plausível de palavras, calculadas com base em vastos conjuntos de dados de treinamento.

Essa característica arquitetural, embora responsável pela notável fluência e coerência aparente dos textos produzidos, é também a causa direta das denominadas alucinações algorítmicas: fenômeno pelo qual o sistema gera informações factualmente incorretas — precedentes inexistentes, dispositivos legais equivocados, teses jurisprudenciais incompatíveis com a orientação vigente — apresentando-as, contudo, com o mesmo grau de fluência e autoridade discursiva que caracterizaria uma informação correta.

A gravidade desse fenômeno não reside apenas na possibilidade de erro, inerente a qualquer processo de produção de conhecimento humano ou tecnológico, mas na peculiar dificuldade de sua detecção.

Diferentemente de um erro grosseiro, facilmente identificável por sua incoerência lógica ou linguística, a alucinação algorítmica manifesta-se sob a forma de um texto bem estruturado, tecnicamente correto em sua superfície linguística e argumentativamente coerente em sua aparência — características

que, precisamente, tendem a neutralizar o ceticismo do leitor humano e a favorecer sua incorporação acrítica ao produto final.

Para o professor Richard Susskind (2019), ao examinar as possibilidades e limites da automação no campo jurídico, adverte que a sofisticação linguística dos sistemas contemporâneos cria um descompasso perigoso entre a confiabilidade aparente do texto produzido e sua confiabilidade material efetiva, descompasso que tende a se acentuar justamente nos contextos de maior pressão por produtividade — exatamente aqueles que caracterizam o cotidiano do Judiciário brasileiro.

A esses riscos de natureza estrutural somam-se riscos de natureza deliberada, relacionados à possibilidade de manipulação direta dos sistemas de inteligência artificial por agentes interessados no resultado do processo decisório. Entre as vulnerabilidades mais discutidas na literatura técnica contemporânea encontra-se a chamada injeção de comandos, ou *prompt injection*, técnica que consiste na inserção de instruções ocultas em documentos submetidos ao processamento pela ferramenta tecnológica.

Tais instruções, embora dificilmente perceptíveis ao leitor humano — por estarem inseridas, por exemplo, em metadados, em formatação invisível ou em trechos textuais aparentemente irrelevantes —, podem ser identificadas e executadas pelo sistema, que não distingue, a priori, entre o conteúdo argumentativo legítimo de uma petição e instruções técnicas nela dissimuladas.

Transposta para o âmbito judicial, essa vulnerabilidade revela preocupação de considerável gravidade institucional, nada impede, em tese, que uma petição contenha comandos destinados a influenciar a forma como o sistema de inteligência artificial interpreta determinado conjunto de fatos, hierarquiza argumentos ou organiza a estrutura da minuta a ser produzida. Ausentes mecanismos técnicos adequados de filtragem e controle — cuja implementação, é preciso reconhecer, permanece tecnicamente desafiadora mesmo para especialistas em segurança da informação —, a resposta produzida pelo sistema poderá ser afetada por instruções unilateralmente introduzidas por uma das partes do processo, sem que a parte contrária ou o próprio magistrado tenham conhecimento dessa interferência.

Tal possibilidade coloca em xeque valores estruturantes da atividade jurisdicional — a imparcialidade, a igualdade de tratamento entre litigantes e a própria transparência do processo decisório —, revelando que os riscos associados à inteligência artificial no Judiciário não se limitam a falhas técnicas involuntárias, mas alcançam também a esfera da manipulação deliberada e estrategicamente orientada.

Diante desse conjunto articulado de riscos — vieses algorítmicos, viés de automação, alucinações e vulnerabilidade a manipulações —, torna-se possível afirmar que a inteligência artificial generativa não constitui simples ferramenta de automação textual, equiparável, por exemplo, a corretores ortográficos ou sistemas de formatação documental. Trata-se, antes, de mecanismo que introduz, na própria arquitetura do processo decisório, fontes de indeterminação e de potencial distorção cuja neutralização exige muito mais do que boa-fé ou diligência ordinária por parte dos operadores do Direito — exige, como se demonstrará

na seção subsequente, arquitetura institucional específica de supervisão, capaz de acompanhar de maneira contínua e qualificada todo o percurso de utilização da tecnologia.

3.2 GOVERNANÇA ALGORÍTMICA NO JUDICIÁRIO BRASILEIRO: OS LIMITES DA RESOLUÇÃO CNJ Nº 615/2025 E A INSUFICIÊNCIA DA SUPERVISÃO HUMANA MERAMENTE FORMAL

Reconhecendo, ainda que de maneira gradual, a magnitude dos riscos anteriormente examinados, o Conselho Nacional de Justiça editou a Resolução nº 615, de 11 de março de 2025, estabelecendo diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no âmbito do Poder Judiciário.

Trata-se de iniciativa normativa relevante, que reconhece expressamente princípios estruturantes como a transparência, a supervisão humana, a segurança jurídica e a centralidade da pessoa humana como parâmetros vinculantes para a incorporação institucional dessas tecnologias.

Não se pode subestimar o significado institucional dessa normativa. Diferentemente de posturas meramente entusiásticas em relação à inovação tecnológica — frequentes em outros setores da Administração Pública —, a Resolução reconhece, ainda que implicitamente, que a incorporação de sistemas de inteligência artificial ao processo decisório judicial não constitui questão puramente administrativa ou orçamentária, mas envolve diretamente a proteção de direitos e garantias fundamentais dos jurisdicionados. Mais do que simplesmente autorizar a utilização da tecnologia, a norma busca delimitar as condições sob as quais tal utilização pode ocorrer sem comprometer a legitimidade da atividade jurisdicional.

Entretanto, uma análise crítica do texto normativo e, sobretudo, das práticas institucionais que dele decorrem revela uma tensão relevante entre a formulação principiológica da supervisão humana e sua operacionalização concreta no cotidiano forense. Compreender a supervisão humana como uma simples etapa de revisão do texto produzido pela máquina — na qual a inteligência artificial elabora a minuta e o operador do Direito apenas corrige eventuais equívocos ao final do processo — ignora, de maneira significativa, as limitações estruturais examinadas na seção anterior.

Essa compreensão reducionista da supervisão humana revela-se particularmente problemática à luz de um fenômeno já identificado pela literatura especializada: a qualidade formal da redação produzida pelos sistemas contemporâneos de inteligência artificial tende a gerar falsa sensação de confiabilidade, dificultando justamente a identificação de erros, omissões ou distorções presentes no conteúdo elaborado.

Em outras palavras, o mesmo atributo que torna a tecnologia atraente — sua capacidade de produzir textos linguisticamente sofisticados e argumentativamente coerentes — é também o principal obstáculo à efetividade de uma supervisão concebida em termos meramente terminais.

É precisamente nesse ponto que a posição defendida por Passos (2023) assume relevância decisiva para a construção de um modelo alternativo de governança, segundo o autor, a utilização legítima da inteligência artificial no processo decisório exige mudança substancial de postura por parte dos profissionais do Direito: magistrados e assessores não devem posicionar-se como meros revisores das respostas fornecidas pelo sistema, mas como agentes ativamente responsáveis pela definição prévia dos critérios de utilização da ferramenta, pela conferência sistemática das fontes por ela empregadas e pela avaliação crítica e continuada dos resultados obtidos ao longo de todo o processo.

A supervisão humana, nessa perspectiva, não constitui evento pontual situado ao final do fluxo de trabalho, mas dimensão transversal que deve permear cada etapa da utilização tecnológica — da formulação da consulta inicial à validação final do texto produzido.

Essa distinção entre supervisão terminal e supervisão contínua encontra respaldo em experiências comparadas relevantes. Danièle Bourcier (2020), ao examinar a incorporação de sistemas de inteligência artificial em diferentes sistemas de justiça europeus, observa que os modelos institucionais mais bem-sucedidos em conciliar eficiência tecnológica e proteção de garantias processuais são precisamente aqueles que estabelecem protocolos de auditoria e verificação em múltiplos estágios do processo decisório, e não apenas na fase de revisão editorial final.

A autora destaca, ainda, que a efetividade dessa supervisão qualificada depende diretamente da formação técnica adequada dos operadores jurídicos, os quais devem compreender, ao menos em seus contornos gerais, a lógica operacional dos sistemas que utilizam — exigência que, no contexto brasileiro atual, permanece amplamente não atendida pela maior parte dos programas de capacitação institucional oferecidos pelos tribunais.

Some-se a essa análise a constatação de que a própria Resolução CNJ nº 615/2025, ao estabelecer o princípio da transparência, não especifica com suficiente precisão técnica os mecanismos concretos pelos quais essa transparência deve se materializar no cotidiano da atividade jurisdicional.

Transparência algorítmica, em sistemas de inteligência artificial generativa baseados em arquiteturas de aprendizado profundo, constitui desafio técnico substancialmente distinto daquele enfrentado por sistemas mais simples de automação, uma vez que a própria lógica interna de tais modelos frequentemente resiste à explicação completa mesmo por parte de seus desenvolvedores — fenômeno que a literatura técnica denomina de problema da explicabilidade (explainability).

Sem instrumentos normativos que enfrentem especificamente esse desafio — como, por exemplo, a exigência de registro documentado das fontes e critérios utilizados em cada minuta produzida com auxílio de inteligência artificial —, o princípio da transparência corre o risco de permanecer em nível meramente retórico, desprovido de eficácia prática mensurável.

Cumprir reconhecer, por outro lado, que a Resolução representa marco normativo em constante processo de aperfeiçoamento, e que a exigência de precisão técnica absoluta em seu texto inaugural talvez constitua expectativa excessiva para um instrumento pioneiro nessa matéria.

Entretanto, precisamente por constituir marco inaugural, a interpretação de seus princípios não pode prescindir de esforço doutrinário que lhes confira densidade normativa suficiente para orientar a prática institucional cotidiana — esforço ao qual o presente estudo pretende contribuir, ao propor que a supervisão humana seja compreendida, doravante, como categoria processual dotada de conteúdo mínimo obrigatório: definição prévia de critérios de utilização, conferência sistemática de fontes, documentação das etapas de revisão e avaliação crítica continuada dos resultados produzidos.

Somente a partir dessa reconfiguração conceitual — que desloca a supervisão humana de evento pontual para dimensão estrutural e contínua do processo decisório — torna-se possível compatibilizar os inegáveis ganhos de eficiência proporcionados pela inteligência artificial com a preservação das garantias processuais fundamentais que sustentam a legitimidade da atividade jurisdicional, questão que será examinada, em sua dimensão propriamente constitucional, na seção subsequente.

3.3 FUNDAMENTAÇÃO CONSTITUCIONAL DAS DECISÕES JUDICIAIS E A RESERVA DE HUMANIDADE: LIMITES INTRANSPONÍVEIS À AUTOMAÇÃO DA CONVICÇÃO JURISDICIONAL

Nenhuma análise sobre os limites jurídicos da inteligência artificial no processo decisório pode prescindir do exame direto do parâmetro normativo que, em última instância, condiciona a legitimidade de toda decisão judicial proferida no ordenamento brasileiro: o dever constitucional de fundamentação, previsto no artigo 93, inciso IX, da Constituição Federal de 1988.

Trata-se de dispositivo cuja densidade normativa transcende, significativamente, a exigência de mera exposição de motivos ou de produção de texto coerente e bem redigido.

Luís Roberto Barroso (2022), ao examinar o papel da fundamentação no constitucionalismo contemporâneo, sustenta que fundamentar uma decisão judicial significa, essencialmente, enfrentar os argumentos relevantes apresentados pelas partes, interpretar o ordenamento jurídico aplicável ao caso concreto à luz dos princípios e valores constitucionais pertinentes, e justificar racionalmente a conclusão adotada, de modo que essa conclusão possa ser compreendida e controlada tanto pelas partes do processo quanto pela sociedade em geral.

Trata-se, portanto, de atividade que envolve necessariamente análise crítica, ponderação valorativa e responsabilidade institucional — elementos cuja presença não pode ser assumida como automaticamente garantida pela simples produção de um texto tecnicamente correto e argumentativamente bem estruturado.

Sob essa perspectiva, torna-se possível afirmar que a simples reprodução, pelo magistrado, de uma minuta elaborada por inteligência artificial — ainda que submetida a revisão editorial superficial — não atende, por si só, às exigências constitucionais de fundamentação.

Isso porque o cumprimento formal da exigência de motivação — a presença de um texto que descreve fatos, cita normas e apresenta conclusões — não equivale ao cumprimento material dessa mesma exigência, que pressupõe processo cognitivo efetivamente desenvolvido pelo julgador, no qual os argumentos das partes sejam genuinamente sopesados e a conclusão adotada resulte de deliberação racional e não de mera aceitação acrítica de um produto tecnológico.

Essa distinção entre cumprimento formal e cumprimento material da exigência de fundamentação encontra amparo conceitual relevante na noção de reserva da humanidade, desenvolvida por Rodrigues Junior et al. (2026) em análise dedicada especificamente aos limites constitucionais da automação de atividades decisórias.

Segundo os autores, determinadas funções estatais — entre as quais se inclui, com destaque, a atividade jurisdicional — pressupõem, como condição de legitimidade, a presença de deliberação humana insubstituível, não em razão de superioridade técnica intrínseca do raciocínio humano sobre o raciocínio algorítmico, mas em razão da própria estrutura de responsabilidade institucional que sustenta o exercício do poder estatal em um regime democrático.

Com efeito, a legitimidade da jurisdição não deriva da correção técnica ou linguística do texto produzido, mas da circunstância de que tal texto expressa a convicção — formada mediante processo deliberativo do qual o magistrado é pessoalmente responsável perante a Constituição e perante a sociedade — de um agente estatal investido de autoridade decisória por processo constitucionalmente legitimado.

Um sistema estatístico de processamento de linguagem, por sofisticado que seja, não é titular dessa responsabilidade institucional, tampouco pode ser dela investido, sob pena de se desfigurar a própria natureza da função jurisdicional enquanto exercício de poder estatal submetido a controle e a responsabilização.

Lenio Streck (2021), em perspectiva hermenêutica que dialoga diretamente com essa problemática, adverte que a compreensão jurídica não pode ser reduzida a processo mecânico de subsunção entre fatos e normas, mas constitui atividade interpretativa que envolve pré-compreensões, tradição jurídica e situação hermenêutica concreta do intérprete — elementos irredutíveis a qualquer forma de processamento estatístico de padrões linguísticos, por mais extenso que seja o corpus de treinamento utilizado.

Essa observação reforça a tese de que a inteligência artificial, ainda que capaz de produzir textos formalmente indistinguíveis de decisões judiciais autênticas, permanece estruturalmente incapaz de reproduzir o processo interpretativo que confere legitimidade material à fundamentação exigida pela Constituição.

A esses argumentos de natureza teórica soma-se consideração de ordem prática, igualmente relevante: a segurança jurídica, compreendida como previsibilidade e estabilidade das decisões judiciais, depende diretamente da confiança institucional depositada pelos jurisdicionados na racionalidade e na coerência do processo decisório.

Quando sistemas de inteligência artificial passam a produzir, ainda que ocasionalmente, interpretações inconsistentes entre si, referências normativas equivocadas ou conteúdos passíveis de manipulação por uma das partes, a confiabilidade sistêmica do Judiciário — e não apenas a correção de decisões isoladas — resta comprometida. Ingo Sarlet (2023), ao tratar da dimensão institucional dos direitos fundamentais, observa que a confiança depositada pela sociedade nas instituições judiciais constitui, ela própria, bem jurídico protegido, cuja erosão produz efeitos sistêmicos que transcendem os casos concretos diretamente afetados por eventuais falhas tecnológicas.

Diante desse quadro, as diretrizes estabelecidas pela Resolução CNJ nº 615/2025 devem ser interpretadas não como autorização ampla e irrestrita para a incorporação de inteligência artificial ao processo decisório, mas como tentativa normativa de equilibrar os inegáveis ganhos de eficiência proporcionados pela inovação tecnológica com a necessidade estrutural de preservar a confiabilidade do sistema de justiça.

Esse equilíbrio, contudo, somente se realiza efetivamente quando a supervisão humana é compreendida, conforme demonstrado na seção anterior, como dimensão contínua e qualificada do processo decisório, e quando a fundamentação das decisões judiciais permanece ancorada em deliberação genuinamente humana — ainda que tecnologicamente auxiliada — sobre os argumentos e o direito aplicável ao caso concreto.

Reafirma-se, assim, articulando os três eixos de análise desenvolvidos ao longo deste artigo, que a inteligência artificial, tecnicamente sujeita a vieses, a fenômenos de automação cognitiva, a alucinações e a manipulações deliberadas, e institucionalmente carente de mecanismos de supervisão efetivamente contínuos, não pode ocupar posição substitutiva na construção da convicção jurisdicional sem comprometer, de maneira estrutural, tanto a segurança jurídica processual quanto o próprio dever constitucional de fundamentação das decisões judiciais.

4 CONCLUSÃO

A investigação desenvolvida ao longo deste artigo permite afirmar, com razoável grau de segurança argumentativa, que a hipótese inicialmente formulada resta integralmente confirmada: a inteligência artificial generativa, por operar mediante inferências estatísticas sobre padrões linguísticos e não mediante compreensão semântica ou normativa do Direito, introduz na elaboração de minutas judiciais riscos estruturais — vieses algorítmicos, vieses de automação, alucinações e vulnerabilidade a manipulações

deliberadas — que não podem ser neutralizados por modelos de supervisão humana concebidos em termos meramente formais e retrospectivos.

A resposta ao problema de pesquisa originalmente proposto revela-se, à luz do percurso argumentativo desenvolvido, igualmente clara: a utilização de inteligência artificial na elaboração de minutas judiciais compromete a segurança jurídica processual e o dever constitucional de fundamentação das decisões precisamente na medida em que a supervisão humana seja compreendida como etapa terminal de revisão textual, e não como dimensão transversal e qualificada que acompanha toda a extensão do processo de utilização tecnológica.

As diretrizes estabelecidas pela Resolução CNJ nº 615/2025, embora representem avanço normativo relevante ao reconhecer principiologicamente a necessidade de transparência e de centralidade da pessoa humana, permanecem insuficientes enquanto não forem operacionalizadas mediante protocolos institucionais que confirmem conteúdo mínimo obrigatório à supervisão humana — definição prévia de critérios de utilização, conferência sistemática de fontes, documentação das etapas de revisão e avaliação crítica continuada dos resultados produzidos.

A contribuição original deste estudo ao campo científico reside precisamente na articulação sistemática entre categorias técnicas — geralmente examinadas pela literatura especializada em ciência da computação e em ética da tecnologia — e categorias constitucionais e processuais consolidadas, notadamente a exigência de fundamentação prevista no artigo 93, inciso IX, da Constituição Federal e a noção doutrinária de reserva da humanidade.

Ao demonstrar que a legitimidade da decisão judicial não deriva da qualidade formal do texto produzido, mas da presença de deliberação humana efetivamente responsável perante a Constituição, o presente trabalho oferece parâmetro interpretativo capaz de orientar tanto a aplicação prática da Resolução CNJ nº 615/2025 quanto o desenvolvimento futuro de protocolos institucionais de governança algorítmica no âmbito do Poder Judiciário brasileiro.

Reconhece-se, por fim, que a incorporação da inteligência artificial à atividade jurisdicional constitui processo irreversível e, em muitos aspectos, benéfico à eficiência do sistema de justiça. Entretanto, a preservação da legitimidade dessa atividade exige que a tecnologia permaneça, invariavelmente, na posição de instrumento auxiliar do trabalho humano, jamais na posição de agente substitutivo da convicção jurisdicional — condição sem a qual a segurança jurídica, pilar estruturante do Estado Democrático de Direito, restará inevitavelmente comprometida.

REFERÊNCIAS

BARROSO, Luís Roberto. **Curso de direito constitucional contemporâneo: os conceitos fundamentais e a construção do novo modelo**. 10. ed. São Paulo: Saraiva, 2022.

BOURCIER, Danièle. Inteligência artificial e justiça: reflexões sobre os limites da automação decisória. **Revista Direito e Tecnologia**, Paris, v. 5, n. 2, p. 88-104, 2020.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidente da República, 1988. Disponível em: <http://www.planalto.gov.br>. Acesso em: 12 jul. 2026.

BRASIL. Conselho Nacional de Justiça. **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/files/original1555302025031467d4517244566.pdf>. Acesso em: 9 jul. 2026.

GARAPON, Antoine; LASSÈGUE, Jean. **Justice digitale: révolution graphique et rupture anthropologique**. Paris: Presses Universitaires de France, 2018.

NUNES, Dierle; MARQUES, Ana Luiza Pinto Coelho. Inteligência artificial e direito processual: vieses algorítmicos e os riscos de atribuição de função decisória às máquinas. **Revista de Processo**, v. 43, n. 286, p. 421-447, dez. 2018.

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

PASSOS, Hugo Assis. **Inteligência artificial e a repercussão geral da questão constitucional: análise crítica e parâmetros de utilização**. 2023. Tese (Doutorado em Direito Constitucional) — Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP), Brasília, 2023. Disponível em: <https://repositorio.idp.edu.br/handle/123456789/4673>. Acesso em: 9 jul. 2026.

RODRIGUES JÚNIOR, Otavio Luiz *et al.* O dilema de Talos e o direito fundamental à decisão humana. **Migalhas**, 2026. Disponível em <https://www.migalhas.com.br/depeso/459932/o-dilema-de-talos-e-o-direito-fundamental-a-decisao-humana> . Acesso em: 12 jul. 2026.

SARLET, Ingo Wolfgang. **A eficácia dos direitos fundamentais: uma teoria geral dos direitos fundamentais na perspectiva constitucional**. 14. ed. Porto Alegre: Livraria do Advogado, 2023.

STRECK, Lenio Luiz. **Hermenêutica jurídica e(m) crise: uma exploração hermenêutica da construção do Direito**. 15. ed. São Paulo: Saraiva, 2021.

SUNSTEIN, Cass R. **Algorithms and automation bias: risks and remedies**. Cambridge: Cambridge University Press, 2018.

SUSSKIND, Richard. **Online courts and the future of justice**. Oxford: Oxford University Press, 2019.