Revista Ibero-Americana de Humanidades, Ciências e Educação — REASE

OPEN ACCESS

doi.org/10.51891/rease.v12i11.31903

O USO DA INTELIGÊNCIA ARTIFICIAL EM DECISÕES AUTOMATIZADAS NO PODER JUDICIÁRIO BRASILEIRO

THE USE OF ARTIFICIAL INTELLIGENCE IN AUTOMATED DECISIONS IN THE BRAZILIAN JUDICIARY

EL USO DE LA INTELIGENCIA ARTIFICIAL EN LAS DECISIONES AUTOMATIZADAS DEL PODER JUDICIAL BRASILEÑO

João Antonio Cirqueira de Almeida[1]
José Carlos de Freitas[2]

RESUMO: A incorporação da inteligência artificial ao Poder Judiciário brasileiro suscita questões que ultrapassam a eficiência operacional, especialmente quanto à influência dos sistemas na construção das decisões judiciais, à reprodução de vieses e à preservação das garantias processuais. Este artigo tem como objetivo analisar os limites éticos, interpretativos e institucionais da utilização dessas tecnologias na atividade jurisdicional, com ênfase na explicabilidade dos sistemas, na responsabilidade humana e no dever de fundamentação das decisões. Trata-se de pesquisa qualitativa, bibliográfica e documental, orientada pela análise de conteúdo de Bardin L (2016). A análise foi organizada em quatro eixos temáticos construídos durante a leitura das fontes: racionalidade automatizada, vieses algorítmicos, segurança e integridade processual e responsabilidade humana na fundamentação judicial. A discussão evidencia que a legitimidade do emprego da inteligência artificial depende da avaliação crítica de seus resultados, de mecanismos de segurança e controle institucional e da manutenção da responsabilidade humana. As tentativas de manipulação de sistemas identificadas em instituições judiciais brasileiras em 2026 reforçam essas preocupações, embora não permitam generalizar vulnerabilidades para todas as ferramentas utilizadas pelos tribunais. Conclui-se que a incorporação dessas tecnologias exige supervisão efetiva e observância das garantias constitucionais e processuais.

Palavras-chave: Inteligência artificial. Poder Judiciário. Decisão judicial. Fundamentação das decisões. Responsabilidade humana.

ABSTRACT: The incorporation of artificial intelligence into the Brazilian Judiciary raises issues that extend beyond operational efficiency, particularly regarding the influence of these systems on judicial decision-making, the reproduction of algorithmic biases, and the preservation of procedural safeguards. This article aims to analyze the ethical, interpretive, and institutional limits of using these technologies in judicial proceedings, focusing on system explainability, human responsibility, and the duty to provide reasoned judicial decisions. This is a qualitative, bibliographic, and documentary study guided by Bardin L (2016) content analysis approach. The analysis was organized into four thematic areas developed during the examination of the sources: automated rationality, algorithmic biases, procedural security and integrity, and human responsibility in judicial reasoning. The discussion indicates that the legitimacy of artificial intelligence in judicial practice depends on critical assessment of its outputs, effective security mechanisms, institutional oversight, and the preservation of human responsibility. Attempts to manipulate artificial intelligence systems identified within Brazilian judicial institutions in 2026 reinforce these concerns, although they do not justify generalizing vulnerabilities to all systems used by the courts. The study concludes that incorporating these technologies requires effective human supervision and compliance with constitutional and procedural safeguards.

Keywords: Artificial intelligence. Judiciary. Judicial decision-making. Reasoned judicial decisions. Human responsibility.

[1] Discente do curso de Direito na Universidade de Gurupi – UnirG.
[2] Professor de Filosofia do Direito da Universidade de Gurupi-UnirG. Licenciado em Filosofia, Mestre e Doutor em Letras.

RESUMEN: La incorporación de la inteligencia artificial al Poder Judicial brasileño plantea cuestiones que van más allá de la eficiencia operativa, especialmente en relación con la influencia de estos sistemas en la construcción de las decisiones judiciales, la reproducción de sesgos algorítmicos y la preservación de las garantías procesales. Este artículo tiene como objetivo analizar los límites éticos, interpretativos e institucionales del uso de estas tecnologías en la actividad jurisdiccional, con énfasis en la explicabilidad de los sistemas, la responsabilidad humana y el deber de fundamentación de las decisiones. Se trata de una investigación cualitativa, bibliográfica y documental, orientada por el análisis de contenido de Bardin L (2016). El análisis se organizó en cuatro ejes temáticos construidos durante la lectura de las fuentes: racionalidad automatizada, sesgos algorítmicos, seguridad e integridad procesal y responsabilidad humana en la fundamentación judicial. La discusión indica que la legitimidad del uso de la inteligencia artificial depende de la evaluación crítica de sus resultados, de mecanismos de seguridad y control institucional y del mantenimiento de la responsabilidad humana. Los intentos de manipulación de sistemas identificados en instituciones judiciales brasileñas en 2026 refuerzan estas preocupaciones, aunque no permiten generalizar las vulnerabilidades a todas las herramientas utilizadas por los tribunales. Se concluye que la incorporación de estas tecnologías exige una supervisión humana efectiva y el cumplimiento de las garantías constitucionales y procesales.

Palabras clave: Inteligencia artificial. Poder Judicial. Decisión judicial. Fundamentación de las decisiones. Responsabilidad humana.

1 INTRODUÇÃO

A incorporação da inteligência artificial (IA) ao Poder Judiciário brasileiro suscita questões que ultrapassam a modernização dos serviços judiciais. Embora essas tecnologias ofereçam recursos para organizar informações, identificar padrões e auxiliar no tratamento de grandes volumes processuais, sua utilização também interfere nas condições em que o Direito é interpretado e aplicado. A discussão torna-se especialmente relevante quando ferramentas computacionais deixam de atuar apenas na gestão dos processos e passam a influenciar a construção de argumentos ou o conteúdo das decisões. Nesse cenário, a busca por eficiência precisa ser examinada em relação às exigências de fundamentação, controle e legitimidade que caracterizam a atividade jurisdicional.

A distinção entre apoio tecnológico e participação no processo decisório é fundamental para compreender a extensão desse problema. Sistemas empregados na classificação de processos ou na identificação de demandas semelhantes desempenham funções diferentes daqueles utilizados para sugerir argumentos, selecionar precedentes ou elaborar minutas de decisões. Mesmo quando a decisão final permanece formalmente sob responsabilidade de um magistrado, os resultados produzidos por essas ferramentas podem orientar a análise do caso e influenciar quais informações recebem maior atenção. Por isso, a presença de supervisão humana, isoladamente, não esclarece em que medida o julgamento foi efetivamente examinado de forma crítica nem como eventuais erros do sistema podem ser identificados.

Toledo C e Pessoa DA (2023) examinam criticamente o uso da inteligência artificial na tomada de decisão judicial brasileira e discutem a opacidade dos algoritmos, os vieses e as dificuldades de compatibilizar o processamento automatizado de dados com a aplicação do Direito. Essas questões revelam uma diferença importante entre reconhecer regularidades e justificar uma conclusão jurídica. A identificação de semelhanças entre processos pode auxiliar o trabalho judicial, mas não determina, por si só, que casos aparentemente semelhantes devam receber a mesma solução. A relevância de um precedente depende de sua relação com os fatos examinados, dos fundamentos que sustentam sua aplicação e das possíveis distinções juridicamente significativas. Quando padrões identificados computacionalmente são tratados como respostas suficientes, existe o risco de que particularidades do caso concreto e questões interpretativas relevantes percam espaço na análise.

A fundamentação judicial constitui, portanto, um dos principais eixos para avaliar os limites dessas tecnologias. A motivação de uma decisão permite compreender o percurso argumentativo adotado, contestar seus fundamentos e verificar sua compatibilidade com o ordenamento jurídico. Siqueira DP, et al. (2022) analisam os limites da utilização da inteligência artificial na jurisdição a partir do dever de fundamentação analítica e das exigências do devido processo legal. Os autores discutem, nesse contexto, a relação entre explicabilidade e possibilidade de utilização de ferramentas tecnológicas na tomada de decisão. Nunes Pecego D e Lobato Collet Janny Teixeira R (2024), por sua vez, investigam a passagem da opacidade à explicabilidade das decisões judiciais, destacando os problemas éticos associados aos vieses, ao uso de dados pessoais e à dificuldade de compreender os resultados produzidos por determinados sistemas.

A explicabilidade, contudo, não se confunde com a simples apresentação de uma justificativa em linguagem compreensível. Uma ferramenta pode produzir uma explicação aparentemente coerente sem que isso permita verificar adequadamente os dados considerados, os critérios utilizados ou a influência de seus resultados sobre a decisão. Do ponto de vista jurídico, interessa compreender como as informações fornecidas pelo sistema foram avaliadas, quais razões sustentam a conclusão adotada e se a fundamentação responde aos argumentos relevantes apresentados no processo. Essa distinção permite deslocar a discussão da transparência entendida apenas como divulgação de informações técnicas para a possibilidade efetiva de examinar e contestar o raciocínio que fundamenta uma decisão.

A questão também envolve a própria compreensão da interpretação jurídica. Em Dworkin R (2007), a integridade do Direito está relacionada à interpretação das práticas jurídicas à luz de princípios que permitam compreender o ordenamento de maneira coerente. Streck LL (1999), no campo da hermenêutica jurídica, problematiza a interpretação como atividade que não se reduz à aplicação mecânica de textos normativos. Essas perspectivas oferecem fundamentos para investigar os limites de uma concepção de decisão baseada predominantemente na identificação de padrões anteriores. A contribuição desses autores para o debate sobre inteligência artificial não consiste em fornecer respostas diretas sobre tecnologias que não constituem o objeto central de suas obras, mas em oferecer referenciais para examinar o que está envolvido na interpretação e na justificação de uma decisão jurídica. Sob esse enfoque, a questão relevante é avaliar como a utilização de sistemas computacionais se relaciona com a análise das particularidades do caso, a coerência jurídica e a responsabilidade pelo julgamento.

No plano institucional, esses problemas exigem critérios de governança que permitam acompanhar o desenvolvimento e a utilização das ferramentas. A Resolução nº 615/2025 do Conselho Nacional de Justiça estabelece diretrizes para o desenvolvimento, o uso e a governança de soluções de inteligência artificial no Poder Judiciário, tendo sido posteriormente alterada pela Resolução nº 674/2026 (CONSELHO NACIONAL DE JUSTIÇA, 2025; CONSELHO NACIONAL DE JUSTIÇA, 2026c). A análise desse marco normativo permite situar o debate sobre transparência, supervisão e responsabilização no contexto regulatório brasileiro. A existência de diretrizes, entretanto, não encerra a discussão sobre sua aplicação prática: permanece necessário examinar como os parâmetros institucionais podem responder aos riscos associados à opacidade, à reprodução de vieses e à influência dos sistemas sobre a fundamentação judicial.

Embora a literatura brasileira já examine separadamente aspectos como vieses algorítmicos, explicabilidade e fundamentação das decisões, a articulação entre esses problemas permite investigar uma questão mais ampla: em quais condições o auxílio computacional é compatível com as exigências interpretativas e argumentativas da jurisdição? Essa articulação constitui o eixo da presente pesquisa. Em vez de pressupor que a inteligência artificial necessariamente compromete a legitimidade judicial ou que seus ganhos de eficiência justificam sua adoção, busca-se examinar os critérios pelos quais suas aplicações podem ser avaliadas, considerando as diferenças entre tarefas de apoio e atividades que interferem diretamente na justificação da decisão.

Diante disso, o artigo procura responder à seguinte questão: quais são os limites éticos, interpretativos e institucionais da utilização da inteligência artificial no processo decisório do Poder Judiciário brasileiro, especialmente quanto à fundamentação das decisões, à explicabilidade dos sistemas e à preservação da responsabilidade humana? Seu objetivo geral é analisar esses limites a partir da literatura jurídica e filosófica e dos documentos institucionais pertinentes. Para tanto, desenvolve-se uma pesquisa qualitativa, de natureza bibliográfica e documental, orientada pela análise crítica das fontes selecionadas. A investigação busca relacionar os problemas de interpretação, fundamentação e governança algorítmica, de modo a contribuir para uma avaliação fundamentada das condições em que a inteligência artificial pode auxiliar a atividade jurisdicional sem afastar as exigências de justificação e controle das decisões

2 MÉTODOS

A pesquisa caracteriza-se como qualitativa, de natureza bibliográfica e documental, com finalidade exploratória e abordagem analítico-interpretativa. O estudo tem como objetivo examinar os limites éticos, interpretativos e institucionais da utilização da inteligência artificial no processo decisório do Poder Judiciário brasileiro, com ênfase na fundamentação das decisões, na explicabilidade dos sistemas e na preservação da responsabilidade humana.

O levantamento bibliográfico priorizou publicações compreendidas entre 2010 e 2026, sem excluir obras anteriores consideradas relevantes para a fundamentação teórica da pesquisa. As fontes foram localizadas em portais de periódicos científicos e repositórios acadêmicos. Foram examinados trabalhos relacionados à aplicação da inteligência artificial à atividade jurisdicional, aos vieses algorítmicos, à explicabilidade dos sistemas, à hermenêutica jurídica e ao dever de fundamentação das decisões. Entre as obras utilizadas estão os estudos de Toledo C e Pessoa DA (2023), Siqueira DP, et al. (2022), Selbst AD e Barocas S (2018), Streck LL e Jung L (2024), Nunes Pecego D e Lobato Collet Janny Teixeira R (2024) e Cardoso OV (2026). A seleção das referências considerou sua pertinência para o problema de pesquisa e sua contribuição para a compreensão das dimensões jurídicas, éticas e institucionais envolvidas no tema.

A pesquisa documental compreendeu a análise de normas jurídicas e materiais institucionais relacionados à utilização da inteligência artificial no Judiciário brasileiro. Foram considerados a Constituição Federal de 1988, o Código de Processo Civil, as Resoluções nº 615/2025 e nº 674/2026 do Conselho Nacional de Justiça e comunicados de órgãos judiciais

relativos a tentativas de manipulação de sistemas de inteligência artificial em 2026. Esses documentos foram examinados em articulação com a literatura, de modo a relacionar as discussões teóricas às exigências normativas e aos episódios concretos divulgados pelas instituições do sistema de justiça.

A organização e a interpretação do material foram orientadas pela abordagem de análise de conteúdo de Bardin L (2016). A leitura das fontes permitiu identificar questões recorrentes e estabelecer relações entre os argumentos teóricos, as disposições normativas e os acontecimentos institucionais examinados. Os eixos temáticos foram construídos ao longo desse processo, a partir dos problemas identificados no material consultado, e não definidos previamente como categorias fechadas.

A análise foi estruturada em quatro eixos. O primeiro aborda a incorporação da inteligência artificial à atividade jurisdicional e os limites da racionalidade automatizada. O segundo examina os vieses algorítmicos e os limites do conhecimento produzido pelos sistemas. O terceiro discute as tentativas de manipulação por meio de prompt injection, considerando suas implicações para a segurança e a integridade processual. O quarto analisa a responsabilidade humana e o dever de fundamentação das decisões judiciais assistidas por inteligência artificial. Essa organização permite abordar aspectos distintos do problema de pesquisa e, ao mesmo tempo, estabelecer conexões entre os riscos tecnológicos, as exigências jurídicas e as condições institucionais necessárias à utilização dessas ferramentas.

Por se tratar de uma pesquisa bibliográfica e documental, não foram realizados testes técnicos dos sistemas empregados pelos tribunais, entrevistas com profissionais do Judiciário ou levantamentos empíricos destinados a mensurar a frequência de erros e vieses. Desse modo, os resultados correspondem à interpretação crítica das fontes consultadas e não permitem determinar a prevalência dos problemas identificados nem generalizar os episódios documentados para todos os sistemas utilizados pelo Poder Judiciário brasileiro.

3 RESULTADOS E DISCUSSÕES

A discussão sobre inteligência artificial no Poder Judiciário brasileiro exige uma distinção conceitual que nem sempre recebe a devida atenção na descrição das iniciativas institucionais: automatizar uma tarefa processual não equivale a utilizar um sistema capaz de produzir inferências, e nenhuma dessas operações corresponde, por si só, ao exercício da jurisdição. Essa diferenciação é decisiva para avaliar o alcance das tecnologias adotadas pelos

tribunais, pois a complexidade da atividade judicial não se mede apenas pelo volume de informações processadas, mas pela relação entre os fatos examinados, as normas aplicáveis e as razões que justificam a solução de um conflito.

Toledo C e Pessoa DA (2023) identificam justamente a imprecisão na classificação de ferramentas como sistemas de inteligência artificial ou mecanismos de automação como um dos obstáculos à compreensão do fenômeno no Brasil. Ao recuperarem levantamentos sobre iniciativas tecnológicas nos tribunais, os autores mostram que os números divulgados variavam conforme os critérios empregados para classificar os projetos. Essa constatação tem uma consequência para a análise do tema: a expansão quantitativa das iniciativas institucionais não pode ser tomada, sem exame adicional, como medida da expansão da inteligência artificial propriamente dita, muito menos como evidência de sua participação efetiva na produção de decisões judiciais. Os autores também apontam dificuldades para obter informações claras sobre o funcionamento e o uso concreto desses sistemas, o que limita a avaliação externa de seus efeitos (TOLEDO C e PESSOA DA, 2023).

O problema da classificação, portanto, antecede a avaliação dos benefícios e riscos. Se uma ferramenta organiza documentos ou identifica processos com características semelhantes, sua contribuição imediata está relacionada ao tratamento da informação. Quando um sistema recomenda precedentes, sugere fundamentos ou produz uma minuta, sua atuação pode alcançar a estrutura argumentativa da decisão. As duas situações não são equivalentes, ainda que ambas sejam apresentadas institucionalmente como inovação tecnológica. Na segunda, é necessário investigar como o conteúdo gerado se relaciona com os elementos específicos do processo, quais critérios orientam sua produção e de que maneira o julgador verifica a pertinência das sugestões. A distinção não estabelece uma fronteira rígida entre tarefas administrativas e atividade jurisdicional; ela oferece um critério para analisar o grau de influência que a tecnologia pode exercer sobre o raciocínio jurídico.

Essa questão adquire maior relevância quando confrontada com a análise de Siqueira DP, et al. (2022). Ao examinarem os limites da inteligência artificial na tomada de decisão judicial, os autores partem do devido processo legal e do dever de fundamentação analítica, relacionando a utilização desses sistemas às exigências previstas no artigo 489, § 1º, do Código de Processo Civil e no artigo 315, § 2º, do Código de Processo Penal. A hipótese defendida no estudo restringe a possibilidade de utilização decisória da inteligência artificial, em princípio, a casos repetitivos e desprovidos de complexidade, desde que a explicabilidade seja compatível

com o dever de fundamentação. Trata-se de uma posição teórica delimitada, e não de uma conclusão empírica de que sistemas automatizados já sejam capazes de decidir validamente toda essa categoria de processos (SIQUEIRA DP, et al., 2022).

O argumento desses autores permite aprofundar a diferença entre produzir uma resposta e justificar juridicamente uma decisão. Um sistema pode identificar regularidades em um conjunto de processos e gerar uma conclusão provável a partir delas. A validade de uma decisão judicial, contudo, depende de sua relação com o ordenamento jurídico, com as particularidades relevantes do caso e com os argumentos que precisam ser enfrentados. A regularidade estatística de uma solução anterior não demonstra, por si, que ela seja juridicamente adequada para um novo processo. Casos aparentemente semelhantes podem apresentar diferenças factuais ou jurídicas que alteram o resultado, enquanto decisões reiteradas podem reproduzir interpretações questionáveis. Por essa razão, a possibilidade técnica de repetir padrões decisórios não elimina a necessidade de examinar sua pertinência normativa.

A preocupação de Toledo e Pessoa com a incompatibilidade entre o processamento de dados e a aplicação do Direito aprofunda esse ponto. Os autores questionam a capacidade de sistemas de inteligência artificial realizarem juízos de valor e destacam os riscos de opacidade e de reprodução de vieses cognitivos (TOLEDO C e PESSOA DA, 2023). A questão não deve ser reduzida à afirmação genérica de que máquinas são incapazes de compreender qualquer conteúdo jurídico. Sistemas contemporâneos podem identificar relações linguísticas, recuperar informações, comparar textos e produzir respostas sofisticadas. O problema jurídico está em determinar o que essas capacidades demonstram e o que permanece sem demonstração. A fluência de uma resposta não comprova que o sistema tenha avaliado a relevância normativa dos elementos do caso, reconhecido adequadamente um conflito entre princípios ou assumido responsabilidade pelas consequências da conclusão apresentada.

Essa diferença também ajuda a compreender a crítica hermenêutica formulada por Streck LL e Jung L (2024). Ao discutirem a relação entre hermenêutica e inteligência artificial, os autores questionam os pressupostos teóricos que sustentam a transposição de modelos técnicos, cognitivos e comportamentais para a compreensão do fenômeno jurídico. Em contraposição a uma visão que aproxima a decisão de um procedimento de processamento de informações, defendem a preservação da dimensão linguística, interpretativa e normativa do Direito (STRECK LL e JUNG L, 2024). Essa perspectiva não exige negar a utilidade de

ferramentas computacionais; ela questiona a premissa de que o desempenho técnico seja suficiente para explicar a atividade de interpretar e decidir juridicamente.

A contribuição da hermenêutica está, assim, em deslocar a discussão da capacidade de gerar resultados para as condições em que uma interpretação pode ser juridicamente sustentada. A decisão não se legitima apenas porque coincide com padrões observados em julgamentos anteriores ou porque apresenta uma estrutura textual semelhante à de decisões reconhecidas como adequadas. É necessário examinar a relação entre o texto normativo, o contexto do caso, a tradição jurídica e os argumentos que sustentam a conclusão. Quando a inteligência artificial participa da elaboração de uma decisão, essa exigência permanece. A intervenção tecnológica precisa ser examinada a partir das exigências da própria atividade jurisdicional, e não utilizada como fundamento para redefini-las implicitamente.

É nesse sentido que a responsabilidade humana não pode ser compreendida como uma formalidade satisfeita pela assinatura do magistrado ao final do processo. A presença de um julgador pode coexistir com diferentes graus de dependência em relação ao conteúdo produzido pelo sistema. Se as sugestões algorítmicas forem incorporadas sem exame crítico, a decisão poderá conservar uma aparência de autoria humana enquanto parte relevante de seu percurso argumentativo permanece pouco examinada. Essa possibilidade constitui uma questão analítica, não uma descrição comprovada de todos os usos judiciais da inteligência artificial. Sua importância decorre da necessidade de verificar se a supervisão humana é efetiva, isto é, se o responsável pela decisão dispõe de condições para compreender, questionar e rejeitar os resultados oferecidos pela ferramenta.

A regulamentação do Conselho Nacional de Justiça reconhece a relevância institucional desses problemas. A Resolução CNJ nº 615/2025, em sua redação compilada após a Resolução nº 674/2026, estabelece diretrizes para o desenvolvimento, a utilização e a governança de soluções de inteligência artificial no Judiciário, contemplando preocupações relacionadas à transparência, à não discriminação, ao devido processo e à fundamentação da atividade jurisdicional (CONSELHO NACIONAL DE JUSTIÇA, 2025; CONSELHO NACIONAL DE JUSTIÇA, 2026c). O marco regulatório fornece parâmetros para avaliar as condições de utilização dessas ferramentas, mas sua existência não basta para comprovar que os mecanismos de controle sejam efetivos em todas as aplicações. Essa avaliação depende das características de cada sistema, da documentação disponível, das formas de supervisão e das possibilidades de revisão de seus resultados.

A análise deste primeiro eixo indica, portanto, que a questão central não consiste em estabelecer uma oposição abstrata entre inteligência artificial e atividade judicial. O ponto decisivo é identificar como cada aplicação interfere no percurso que liga as informações do processo à interpretação jurídica e à decisão fundamentada. Quanto mais a ferramenta ultrapassa a organização de dados e passa a sugerir argumentos ou formular conclusões, mais importante se torna examinar a origem de seus resultados, os critérios que os orientam e as condições para sua contestação. Essa delimitação conduz ao próximo problema da pesquisa: compreender de que maneira a interpretação jurídica impõe limites ao emprego de padrões computacionais e por que a explicabilidade de um sistema, embora necessária em determinados contextos, não resolve sozinha a questão da legitimidade da decisão.

Um dos problemas mais relevantes da utilização de inteligência artificial em contextos decisórios reside na maneira como os sistemas estabelecem relações entre informações e transformam essas relações em resultados. Em modelos de aprendizado de máquina, a identificação de padrões depende dos dados disponíveis, das escolhas feitas durante o desenvolvimento e dos objetivos definidos para o sistema. Essas condições influenciam quais características recebem maior peso e quais relações passam a ser consideradas relevantes. No âmbito judicial, essa questão merece atenção porque uma associação estatisticamente significativa não constitui, por si só, uma razão juridicamente válida para decidir um caso.

Toledo C e Pessoa DA (2023) abordam esse problema ao relacionarem o processamento algorítmico de dados à reprodução e à intensificação de vieses cognitivos. A preocupação dos autores envolve a possibilidade de que sistemas treinados com informações provenientes de práticas sociais e institucionais existentes reproduzam padrões problemáticos presentes nesses registros. Quando o histórico utilizado contém desigualdades ou decisões orientadas por pressupostos discriminatórios, o modelo pode aprender associações que preservam esses padrões, ainda que a discriminação não apareça expressamente como critério de decisão (TOLEDO C e PESSOA DA, 2023).

A dificuldade está em que um viés pode se manifestar sem que o sistema utilize diretamente uma característica protegida. Variáveis aparentemente neutras podem funcionar como indicadores indiretos de condições sociais, econômicas ou territoriais. Em um contexto judicial, por exemplo, um modelo que utilize dados históricos para identificar padrões de litigância pode atribuir relevância a características associadas ao acesso desigual à Justiça. A hipótese ilustra um mecanismo possível, não um resultado empírico demonstrado para os

sistemas dos tribunais brasileiros. Para verificar se esse tipo de problema ocorre em uma aplicação concreta, seria necessário conhecer as variáveis utilizadas, a composição dos dados, o objetivo do modelo e os efeitos produzidos sobre os grupos atingidos.

A análise de Selbst AD e Barocas S (2018) permite aprofundar a discussão para além da identificação de vieses. Em The Intuitive Appeal of Explainable Machines, os autores distinguem dois problemas que costumam ser tratados conjuntamente. Um modelo pode ser inscrutável, quando é difícil compreender como funciona, e pode ser não intuitivo, quando as relações que utiliza não correspondem às expectativas humanas sobre quais fatores deveriam importar. Essas propriedades estão relacionadas, mas não são equivalentes, e demandam respostas diferentes. Tornar um modelo compreensível exige apresentar de maneira inteligível as regras ou relações que orientam seus resultados. Justificar essas relações requer examinar por que foram adotadas e se há razões defensáveis para utilizá-las (SELBST AD e BAROCAS S, 2018).

Essa distinção modifica a maneira de avaliar a transparência algorítmica. Suponha-se que um sistema utilizado para auxiliar a análise de processos consiga indicar quais variáveis contribuíram para uma classificação. A informação pode esclarecer o funcionamento do modelo, mas permanece uma pergunta anterior à sua aceitação: por que essas variáveis devem ser consideradas relevantes para a tarefa judicial em questão? A explicação técnica pode revelar que uma associação foi utilizada sem demonstrar que ela seja pertinente à avaliação jurídica. A transparência, portanto, não elimina a necessidade de examinar os critérios que orientaram o desenvolvimento do sistema.

Selbst AD e Barocas S (2018) chamam atenção justamente para a insuficiência de explicações que descrevem o modelo sem examinar o processo pelo qual suas regras foram estabelecidas. Os autores observam que determinadas técnicas de interpretabilidade podem indicar quais fatores influenciaram um resultado ou aproximar um modelo complexo por meio de uma representação mais simples. Esses recursos ajudam a compreender o comportamento do sistema, mas não resolvem necessariamente o problema normativo. Uma relação pode ser descrita com clareza e continuar sem justificativa adequada. Para avaliar sua legitimidade, é preciso considerar as decisões tomadas durante o desenvolvimento, as alternativas consideradas, as restrições impostas e os fundamentos para a escolha dos critérios utilizados (SELBST AD e BAROCAS S, 2018).

Essa contribuição é especialmente pertinente ao Direito porque a avaliação de um resultado judicial não se encerra na identificação dos fatores que contribuíram para sua produção. A questão envolve a pertinência desses fatores à luz das normas aplicáveis e dos direitos em discussão. Um modelo pode apresentar regularidades consistentes nos dados e, ainda assim, apoiar-se em relações inadequadas para a finalidade pretendida. A precisão estatística responde a uma pergunta sobre o desempenho do sistema segundo determinado critério; ela não determina, isoladamente, se esse critério é juridicamente apropriado.

A crítica de Streck LL e Jung L (2024) permite situar esse problema em uma discussão mais ampla sobre os pressupostos da racionalidade jurídica. Os autores questionam a perspectiva técnico-naturalista que acompanha certas propostas de aplicação da inteligência artificial ao Direito e defendem a retomada da hermenêutica filosófica como forma de preservar a dimensão linguística, interpretativa e normativa do fenômeno jurídico. A contribuição dessa abordagem não consiste em oferecer uma técnica de auditoria de modelos, mas em questionar a ideia de que a explicação de um procedimento técnico seja suficiente para compreender ou legitimar a interpretação jurídica (STRECK LL e JUNG L, 2024).

As três perspectivas analisadas cumprem funções diferentes. Toledo e Pessoa permitem investigar como vieses e opacidade podem surgir na utilização de sistemas de inteligência artificial. Selbst e Barocas demonstram que a compreensão técnica do modelo não equivale à justificação das relações que ele emprega. Streck e Jung questionam os pressupostos filosóficos de uma concepção de Direito que atribua centralidade excessiva à racionalidade técnica. A aproximação entre esses estudos sustenta uma questão específica para o presente artigo: a legitimidade dos critérios algorítmicos depende de uma avaliação que ultrapasse a descrição de seu funcionamento e examine sua adequação à finalidade jurídica para a qual são utilizados.

Essa conclusão não implica que toda associação identificada por um sistema seja inválida ou que o emprego de modelos estatísticos seja necessariamente incompatível com a atividade judicial. Ela indica que a avaliação precisa ser contextualizada. É necessário verificar quais dados foram utilizados, quais variáveis influenciaram o resultado, que objetivo orientou o treinamento e se os critérios empregados são adequados à finalidade pretendida. Quando o sistema participa de uma atividade que pode afetar direitos, essa avaliação deve considerar ainda a possibilidade de identificar efeitos discriminatórios e de contestar resultados inadequados.

O problema dos vieses algorítmicos, portanto, não se limita à hipótese de uma máquina reproduzir preconceitos humanos. Ele envolve as condições pelas quais determinadas

informações são selecionadas, transformadas em padrões e utilizadas para orientar decisões. A explicabilidade pode tornar parte desse percurso mais compreensível, mas não substitui a avaliação crítica dos critérios adotados. É essa diferença entre compreender como um sistema produz um resultado e justificar por que esse resultado deve ser considerado aceitável que conduz ao próximo eixo, dedicado às exigências de fundamentação judicial e às condições para o controle das decisões influenciadas por inteligência artificial.

A utilização de inteligência artificial na análise de documentos judiciais introduz uma preocupação que não se limita à precisão dos resultados produzidos pelos sistemas. É necessário considerar a possibilidade de que os próprios documentos submetidos à análise contenham instruções destinadas a interferir no comportamento das ferramentas. Essa questão ganhou relevância institucional no Brasil em 2026, quando tribunais passaram a divulgar ocorrências de prompt injection em peças processuais. Os episódios oferecem um campo concreto para discutir a segurança dos sistemas de IA sem pressupor que todos apresentem as mesmas limitações ou que as tentativas identificadas tenham sido bem-sucedidas.

O prompt injection é uma técnica de ataque que busca alterar o comportamento esperado de um sistema de inteligência artificial por meio da inserção de instruções que entram em conflito com aquelas que deveriam orientar seu funcionamento. Greshake K et al. (2023) investigam essa ameaça no contexto de aplicações integradas a grandes modelos de linguagem e distinguem, em particular, os ataques indiretos, nos quais os comandos maliciosos são inseridos em conteúdos que o sistema posteriormente recupera ou processa. Os autores demonstram que a fronteira entre dados e instruções pode se tornar imprecisa nessas aplicações, criando oportunidades para interferências na funcionalidade dos sistemas e na utilização de ferramentas externas. O estudo oferece uma base técnica para compreender por que um documento aparentemente destinado apenas à leitura pode também servir como veículo para uma tentativa de manipulação (GRESHAKE K, et al., 2023).

No contexto processual, essa técnica pode envolver instruções dissimuladas por recursos de formatação que dificultam sua identificação durante a leitura convencional, como texto branco sobre fundo branco ou espaçamento entre caracteres. A intenção é fazer com que o sistema interprete o conteúdo como um comando a seguir, em vez de tratá-lo exclusivamente como parte do documento submetido à análise. Em maio de 2026, o Superior Tribunal de Justiça (STJ) informou ter identificado petições com esse tipo de comando e anunciou providências para investigar as ocorrências e registrar os fatos nos autos. O tribunal também descreveu

mecanismos de proteção empregados pelo sistema STJ Logos para impedir que instruções externas maliciosas fossem executadas (SUPERIOR TRIBUNAL DE JUSTIÇA, 2026c).

Os episódios divulgados pelo tribunal nos meses seguintes ajudam a dimensionar o problema. Em setembro de 2026, a Sexta Turma do STJ identificou um comando oculto em um recurso especial que solicitava à inteligência artificial que acolhesse os argumentos apresentados pela defesa. Na mesma ocasião, o colegiado determinou a comunicação dos fatos à Ordem dos Advogados do Brasil (OAB) e ao Ministério Público Federal (MPF). Em outro processo, a Segunda Turma aplicou multa por litigância de má-fé após identificar instruções ocultas em recurso apresentado por um procurador municipal, determinando também a comunicação do caso a outras autoridades. Esses episódios foram divulgados pelo próprio tribunal, que relatou as providências adotadas em cada processo (SUPERIOR TRIBUNAL DE JUSTIÇA, 2026a; SUPERIOR TRIBUNAL DE JUSTIÇA, 2026b).

Essas ocorrências devem ser interpretadas com precisão. A identificação de comandos maliciosos não demonstra, por si só, que os sistemas tenham obedecido às instruções ou que o resultado de um julgamento tenha sido alterado. Em comunicado de maio, o STJ afirmou que o sistema STJ Logos dispunha de camadas de segurança destinadas a impedir a execução de comandos externos maliciosos, incluindo mecanismos de separação entre instruções e dados, delimitação do contexto de uso e filtragem dos resultados gerados. Portanto, a existência das tentativas não equivale à comprovação de comprometimento do sistema. Ainda assim, a necessidade de identificar e neutralizar essas instruções evidencia que a segurança da IA precisa ser considerada desde o desenvolvimento até a utilização das ferramentas (SUPERIOR TRIBUNAL DE JUSTIÇA, 2026c).

A distinção é importante porque o problema envolve a relação entre o documento analisado e as instruções que orientam o sistema. Uma petição é apresentada para expor fatos, argumentos e pedidos da parte. Se um sistema utilizado para auxiliar a análise processual não conseguir distinguir adequadamente o conteúdo documental de uma instrução inserida para alterar seu comportamento, abre-se a possibilidade de interferência indevida no tratamento das informações. O risco não depende necessariamente de o sistema produzir uma decisão final de forma autônoma. Mesmo quando a ferramenta é utilizada para classificar documentos, elaborar sínteses ou sugerir encaminhamentos, uma saída influenciada por instruções impróprias pode afetar as etapas posteriores do trabalho judicial.

Esse cenário exige diferenciar a segurança técnica da regularidade processual. A primeira envolve medidas destinadas a prevenir, identificar e neutralizar tentativas de manipulação. A segunda diz respeito à preservação das condições legítimas de participação das partes e à integridade dos atos praticados no processo. Uma tentativa de induzir um sistema a favorecer determinada parte pode comprometer essa integridade mesmo quando o mecanismo de proteção impede que a instrução seja executada. A conduta e o resultado precisam ser avaliados separadamente, pois a existência de uma tentativa não comprova que houve influência efetiva sobre a análise, enquanto a ausência de alteração do resultado não torna irrelevante a necessidade de apuração.

A resposta institucional também merece atenção. Em maio de 2026, o Comitê Nacional de Inteligência Artificial no Judiciário aprovou medidas voltadas à prevenção de tentativas de manipulação, e o Conselho Nacional de Justiça (CNJ) divulgou orientações destinadas a reduzir os riscos associados à inserção de comandos ocultos em documentos processuais. O enfrentamento dessa questão, contudo, não se resume à adoção de filtros técnicos. Também envolve procedimentos de registro de incidentes, revisão das ferramentas, capacitação dos profissionais e definição de responsabilidades quando forem identificadas tentativas de interferência (CONSELHO NACIONAL DE JUSTIÇA, 2026a; CONSELHO NACIONAL DE JUSTIÇA, 2026b).

A partir desses episódios, é possível formular uma questão relevante para a governança judicial: como demonstrar que uma ferramenta mantém a separação entre as instruções legítimas de funcionamento e o conteúdo externo que deve apenas analisar? A resposta depende das características de cada sistema e das medidas de segurança efetivamente implementadas. Não seria adequado concluir, com base nos casos divulgados, que todos os sistemas de inteligência artificial utilizados pelos tribunais são vulneráveis. Tampouco seria justificável presumir que a presença de mecanismos de proteção elimina qualquer possibilidade de interferência. A avaliação precisa apoiar-se em documentação técnica, testes de segurança, registros de incidentes e informações sobre a forma como os sistemas são utilizados.

Os casos brasileiros de 2026 acrescentam uma dimensão específica ao debate sobre a inteligência artificial no Judiciário. Os riscos não decorrem exclusivamente de possíveis erros do modelo ou da reprodução de padrões discriminatórios. Eles também podem surgir da interação entre o sistema, os documentos processuais e as estratégias adotadas por usuários que procuram influenciar seu funcionamento. Nesse contexto, a confiança institucional não deve

repousar apenas na existência declarada de salvaguardas, mas na capacidade de prevenir interferências, detectar ocorrências, investigar os fatos e corrigir eventuais falhas.

Os episódios divulgados não permitem generalizar a existência de vulnerabilidades em todas as ferramentas judiciais, mas justificam tratar a segurança como parte constitutiva da governança da inteligência artificial. A questão deixa de ser apenas se um sistema consegue produzir resultados úteis e passa a incluir as condições sob as quais esses resultados são obtidos e protegidos contra interferências indevidas. Essa perspectiva complementa a discussão anterior sobre vieses algorítmicos e prepara o exame das responsabilidades institucionais envolvidas no uso dessas tecnologias.

A incorporação da inteligência artificial à atividade jurisdicional suscita questões sobre os limites de sua utilização e a preservação das garantias que legitimam o exercício da jurisdição. A busca por maior eficiência no tratamento dos processos deve respeitar as exigências constitucionais e processuais que orientam a atuação dos magistrados. Nesse sentido, o artigo 93, inciso IX, da Constituição Federal determina que todas as decisões judiciais sejam fundamentadas, enquanto o artigo 489, § 1º, do Código de Processo Civil estabelece critérios para a adequação dessa fundamentação. Tais exigências permanecem aplicáveis quando ferramentas tecnológicas auxiliam a análise dos autos ou a elaboração de decisões (BRASIL, 1988, art. 93, IX; BRASIL, 2015, art. 489, § 1º).

O problema se torna mais complexo quando se considera que a fundamentação judicial não se resume à apresentação de argumentos formalmente organizados ou à indicação de dispositivos legais pertinentes. Fundamentar significa demonstrar o percurso jurídico que sustenta a conclusão adotada, permitindo que as partes compreendam as razões da decisão e questionem seus fundamentos pelos meios processuais cabíveis. Essa exigência impõe limites à utilização de sistemas capazes de produzir textos convincentes, pois a coerência da redação não assegura, por si só, a correção da interpretação jurídica nem a consideração adequada das particularidades do processo.

Siqueira DP, et al. (2022) analisam essa questão a partir do dever de fundamentação analítica e dos limites da substituição da atuação humana por algoritmos. Os autores relacionam o emprego da inteligência artificial ao devido processo legal e à necessidade de estabelecer parâmetros para sua utilização na atividade decisória. A contribuição dessa perspectiva consiste em deslocar a discussão da capacidade técnica dos sistemas para as condições jurídicas sob as quais seus resultados podem ser incorporados à jurisdição. A produção automatizada de uma

resposta não resolve o problema da legitimidade da decisão, uma vez que a atividade jurisdicional exige interpretação, justificação e respeito às garantias processuais (SIQUEIRA DP, et al., 2022).

Essa distinção ganha importância diante da possibilidade de que ferramentas generativas produzam fundamentações aparentemente completas, mas insuficientes para responder às questões efetivamente discutidas no processo. Cardoso OV (2026), ao examinar a relação entre inteligência artificial e fundamentação judicial, chama a atenção para os riscos associados à padronização de modelos decisórios e à reutilização automatizada de argumentos. A preocupação não reside na padronização em si, que pode ser legítima em controvérsias semelhantes, mas na perda de correspondência entre os fundamentos utilizados e as circunstâncias concretas de cada caso. Uma decisão pode empregar terminologia jurídica adequada e citar precedentes pertinentes sem demonstrar que examinou os argumentos capazes de influenciar seu resultado (CARDOSO OV, 2026).

É nesse ponto que a supervisão humana precisa ser compreendida de maneira mais rigorosa. A presença de um magistrado no fluxo de trabalho não constitui, isoladamente, garantia de controle efetivo. Se a revisão se limita à leitura superficial de um texto aparentemente consistente, informações incorretas, referências inexistentes ou interpretações inadequadas podem ser incorporadas à decisão sem exame suficiente. O controle humano exige condições para verificar a procedência das informações, confrontar os argumentos gerados com os autos e avaliar se a conclusão encontra sustentação no ordenamento jurídico. A intervenção humana adquire relevância jurídica quando envolve análise crítica e possibilidade real de rejeitar o resultado sugerido pela ferramenta.

A questão também envolve a possibilidade de contestação. Para exercer o contraditório de maneira efetiva, as partes precisam conhecer os fundamentos que sustentam a decisão e dispor de condições para questioná-los. Quando a utilização de uma ferramenta interfere na seleção de informações, na síntese dos argumentos ou na elaboração dos fundamentos, torna-se relevante avaliar se essa participação comprometeu a compreensão das razões decisórias. Isso não significa que todo uso de inteligência artificial deva ser necessariamente exposto da mesma forma, independentemente de sua finalidade. Significa que a utilização da tecnologia não pode inviabilizar o controle processual sobre os fundamentos que efetivamente sustentam o pronunciamento judicial.

No plano institucional, a Resolução CNJ nº 615/2025, com as alterações posteriores da Resolução nº 674/2026, estabelece diretrizes para o desenvolvimento, a utilização e a governança de soluções de inteligência artificial no Poder Judiciário. O ato normativo contempla preocupações relacionadas à transparência, à supervisão, aos direitos fundamentais, ao devido processo e à fundamentação da atividade jurisdicional. Sua importância está em reconhecer que a incorporação dessas ferramentas requer condições institucionais de acompanhamento e controle, em vez de depender exclusivamente da confiança depositada no desempenho declarado dos sistemas (CONSELHO NACIONAL DE JUSTIÇA, 2025; CONSELHO NACIONAL DE JUSTIÇA, 2026c).

A responsabilização precisa ser examinada à luz dessa estrutura. O desenvolvimento e a manutenção das ferramentas envolvem questões técnicas e institucionais, enquanto a elaboração e a prolação das decisões permanecem submetidas às obrigações jurídicas próprias da atividade jurisdicional. Não seria adequado atribuir automaticamente qualquer falha ao magistrado, desconsiderando as condições técnicas e organizacionais em que a ferramenta foi disponibilizada. Tampouco seria aceitável invocar a complexidade algorítmica para afastar o dever de examinar a legalidade e a suficiência dos fundamentos utilizados. A identificação de responsabilidades depende das circunstâncias concretas, das atribuições de cada agente e das normas aplicáveis.

A legitimidade do uso judicial da inteligência artificial depende, portanto, de critérios que ultrapassam a rapidez da resposta ou a qualidade formal do texto produzido. É necessário preservar a relação entre os fatos do processo, os argumentos das partes, a interpretação jurídica e a conclusão adotada. Quando essa relação não pode ser verificada, a aparência de eficiência corre o risco de ocultar fragilidades na justificação da decisão. A tecnologia pode auxiliar a atividade jurisdicional, mas seus resultados precisam permanecer sujeitos ao exame crítico exigido pelo direito.

4 CONSIDERAÇÕES FINAIS

A análise desenvolvida neste estudo permite compreender que a incorporação da inteligência artificial ao Poder Judiciário brasileiro envolve questões que ultrapassam a eficiência operacional. Embora essas ferramentas possam auxiliar o tratamento de informações e determinadas atividades processuais, sua utilização em contextos relacionados à decisão judicial exige cautela diante dos riscos de reprodução de vieses, de manipulação dos sistemas e

de enfraquecimento da fundamentação jurídica. O problema central, portanto, não reside na presença da tecnologia em si, mas nas condições sob as quais ela é desenvolvida, utilizada e submetida a controle.

Em resposta ao problema de pesquisa, conclui-se que os limites éticos, interpretativos e institucionais da inteligência artificial no Judiciário estão associados à preservação das garantias processuais, à possibilidade de examinar criticamente os resultados produzidos pelos sistemas e à manutenção da responsabilidade humana sobre a atividade jurisdicional. A fundamentação das decisões não pode ser reduzida à produção de textos formalmente coerentes, assim como a existência de mecanismos técnicos de segurança não dispensa a investigação de possíveis interferências. A legitimidade do uso dessas ferramentas depende da articulação entre salvaguardas tecnológicas, controle institucional e observância das exigências jurídicas que orientam a atuação judicial.

Os casos de tentativas de manipulação identificados em 2026 evidenciam a atualidade dessas preocupações, embora não autorizem generalizações sobre a vulnerabilidade de todos os sistemas empregados pelos tribunais. Em conjunto com as discussões sobre vieses algorítmicos e fundamentação judicial, esses episódios reforçam a necessidade de avaliar a inteligência artificial a partir de seus efeitos e de suas condições concretas de utilização. A inovação tecnológica somente se mostra compatível com a legitimidade da jurisdição quando permanece subordinada aos direitos fundamentais e às exigências do devido processo legal.

Por se tratar de uma pesquisa bibliográfica e documental, o estudo apresenta limitações quanto à avaliação empírica do funcionamento das ferramentas de inteligência artificial utilizadas pelos tribunais brasileiros. A análise de literatura e de documentos institucionais permite identificar questões jurídicas relevantes, mas não é suficiente para determinar a frequência de falhas, mensurar a incidência de vieses ou estabelecer em que medida os sistemas efetivamente influenciam o conteúdo das decisões judiciais.

Diante dessas limitações, são necessárias novas pesquisas que investiguem empiricamente a utilização dessas ferramentas, especialmente quanto à transparência dos sistemas, à identificação de vieses, à resistência a tentativas de manipulação e à efetividade da supervisão humana. Também se mostra relevante examinar como magistrados e demais profissionais do sistema de justiça compreendem os resultados gerados por essas tecnologias e quais procedimentos adotam para verificar sua confiabilidade. Estudos comparativos entre

tribunais e avaliações longitudinais poderão contribuir para identificar mudanças nas práticas institucionais e verificar a efetividade das medidas de governança adotadas.

Por fim, a expansão da inteligência artificial no Judiciário exige acompanhamento científico e revisão contínua das práticas institucionais, uma vez que as ferramentas, os riscos e os marcos regulatórios permanecem em transformação. O aprofundamento dessas investigações é indispensável para que a adoção tecnológica seja orientada por evidências e critérios jurídicos verificáveis, evitando que a busca por produtividade comprometa a fundamentação das decisões, a transparência da atividade jurisdicional e a proteção dos direitos fundamentais.

REFERÊNCIAS

BARDIN L. Análise de conteúdo. Tradução de Luís Antero Reto e Augusto Pinheiro. São Paulo: Edições 70, 2016.

BRASIL. Constituição da República Federativa do Brasil de 1988. Brasília, DF: Presidência da República, 1988.

BRASIL. Lei nº 13.105, de 16 de março de 2015. Código de Processo Civil. Brasília, DF: Presidência da República, 2015.

CARDOSO OV. O dever de fundamentação e o uso de inteligência artificial na redação de decisões judiciais. Revista de Estudos Jurídicos do STJ, 2026; 4(4): 367-391.

CONSELHO NACIONAL DE JUSTIÇA (Brasil). Comitê aprova medidas para ampliar a segurança da inteligência artificial no Judiciário. Brasília, DF, 28 maio 2026a.

CONSELHO NACIONAL DE JUSTIÇA (Brasil). Inteligência artificial: CNJ aprova orientações para a manutenção da segurança jurídica. Brasília, DF, 9 jun. 2026b.

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: Conselho Nacional de Justiça, 2025.

CONSELHO NACIONAL DE JUSTIÇA (Brasil). Resolução nº 674, de 25 de março de 2026. Altera o art. 15 da Resolução CNJ nº 615/2025. Brasília, DF: Conselho Nacional de Justiça, 2026c.

DWORKIN R. O império do direito. Tradução de Jefferson Luiz Camargo. 2. ed. São Paulo: Martins Fontes, 2007.

GRESHAKE K, et al. Not what you've signed up for: compromising real-world LLM-integrated applications with indirect prompt injection. arXiv, 2023.

NUNES PECEGO D, LOBATO COLLET JANNY TEIXEIRA R. Inteligência artificial no Judiciário: da opacidade à explicabilidade das decisões judiciais. Revista da Faculdade de Direito da UERJ, 2024; (43): 1-22. DOI: 10.12957/rfd.2024.87850.

SELBST AD, BAROCAS S. The intuitive appeal of explainable machines. Fordham Law Review, 2018; 87(3): 1085-1139.

SIQUEIRA DP, MORAIS FS, SANTOS MF. Inteligência artificial e jurisdição: dever analítico de fundamentação e os limites da substituição dos humanos por algoritmos no campo da tomada de decisão judicial. Seqüência: Estudos Jurídicos e Políticos, 2022; 43(91): 1-34. DOI: 10.5007/2177-7055.2022.e90662.

STRECK LL. Hermenêutica jurídica e(m) crise: uma exploração hermenêutica da construção do Direito. Porto Alegre: Livraria do Advogado, 1999.

STRECK LL, JUNG L. Hermenêutica e inteligência artificial: por uma alternativa paradigmática ao imaginário técnico-jurídico. Direito Público, 2024; 21(110). DOI: 10.11117/rdp.v21i110.7689.

SUPERIOR TRIBUNAL DE JUSTIÇA (Brasil). Sexta Turma identifica comando oculto em recurso para enganar sistema de IA e determina comunicação à OAB e ao MPF. Brasília, DF, 2 set. 2026a.

SUPERIOR TRIBUNAL DE JUSTIÇA (Brasil). STJ impõe multa por litigância de má-fé em recurso com prompt injection assinado por procurador municipal. Brasília, DF, 3 set. 2026b.

SUPERIOR TRIBUNAL DE JUSTIÇA (Brasil). Tentativas de uso de prompt injection no STJ serão investigadas. Brasília, DF, 20 maio 2026c.

TOLEDO C, PESSOA DA. O uso de inteligência artificial na tomada de decisão judicial. Revista de Investigações Constitucionais, 2023; 10(1): e237. DOI: 10.5380/rinc.v10i1.86319.