

A INCORPORAÇÃO DE TECNOLOGIAS DE INTELIGÊNCIA ARTIFICIAL PELO SUPREMO TRIBUNAL FEDERAL

THE ADOPTION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES BY THE
BRAZILIAN SUPREME FEDERAL COURT

LA ADOPCIÓN DE TECNOLOGÍAS DE INTELIGENCIA ARTIFICIAL POR PARTE
DEL TRIBUNAL SUPREMO FEDERAL DE BRASIL

André Luiz Ache Mansur¹, Tércio Rodes Magalhães Silva², Lara Wasilewski³

DOI: 10.54899/dcs.v23i91.5831

Recibido: 08/05/2026 | Aceptado: 03/06/2026 | Publicación en línea: 10/06/2026.

RESUMO

Este trabalho analisa os impactos da adoção de tecnologias de Inteligência Artificial (IA) pelo Supremo Tribunal Federal (STF), considerando seus reflexos na efetividade da prestação jurisdicional e na preservação de direitos e garantias fundamentais. O objetivo central é compreender como essas ferramentas contribuem para a celeridade processual e influenciam a qualidade das decisões e a segurança jurídica. Para tanto, buscou-se: identificar as principais soluções de IA utilizadas pelo STF, avaliar benefícios e desafios decorrentes de sua implementação e examinar implicações na fundamentação e previsibilidade das decisões. A pesquisa adota abordagem qualitativa e exploratória, baseada em análise documental de normativas, jurisprudência e relatórios institucionais, complementada por revisão bibliográfica sobre IA aplicada ao direito e inovação no Judiciário. Os resultados indicam que, desde 2018, o STF tem investido em sistemas como Victor, RAFA 2030, VitóriaIA e MARIA, voltados à classificação de processos e à sistemática de precedentes, com ganhos relevantes em eficiência. Contudo, verificou-se que o uso da IA na Suprema Corte impõe riscos à manutenção de direitos e garantias fundamentais como contraditório, ampla defesa, juiz natural e transparência, exigindo cautelas e regulamentação rigorosa. Nesse sentido, destaca-se a Resolução CNJ nº 615/2025 como instrumento normativo moderno para orientar sua aplicação. Conclui-se que, no atual estágio tecnológico, a IA deve ser utilizada como ferramenta auxiliar, sem substituir a atuação de magistrados e demais servidores em atividades sensíveis, como a prolação de decisões interlocutórias ou definitivas, assegurando os benefícios dessa modernização sem o comprometimento de direitos e garantias fundamentais.

Palavras-chave: Direitos e Garantias Fundamentais. Modernização Institucional. Precedentes.

¹ Mestre em Direitos Humanos e Democracia pelo Programa de Pós-Graduação em Direito da Universidade Federal do Paraná (PPGD - UFPR), Curitiba, Paraná, Brasil. E-mail: andremansuradv@gmail.com

Orcid: <https://orcid.org/0009-0003-0840-6785>

² Especialista em Teologia Sistemática pela Faculdade Batista de Minas Gerais (FBMG), Centro Universitário do Paraná (UNIPA - UNICESUMAR), Curitiba, Paraná, Brasil. E-mail: terciordes@gmail.com

³ Mestre em Educação pelo Programa de Pós-Graduação da Universidade Tuiuti do Paraná (PPGED/UTP), Centro Universitário do Paraná (UNIPA/UNICESUMAR), Curitiba, Paraná, Brasil. E-mail: laraleski@yahoo.com.br

ABSTRACT

This study analyzes the impacts of adopting Artificial Intelligence (AI) technologies by the Brazilian Supreme Federal Court (STF), considering their effects on the effectiveness of judicial services and the preservation of fundamental rights and guarantees. The main objective is to understand how these tools contribute to procedural speed and influence decision quality and legal certainty. To achieve this, the research sought to identify the main AI solutions implemented by the STF, assess benefits and challenges arising from their use, and examine implications for reasoning and predictability of decisions. The study adopts a qualitative and exploratory approach, based on documentary analysis of regulations, case law, and institutional reports, complemented by a literature review on AI applied to law and judicial innovation. The results indicate that, since 2018, the STF has invested in systems such as Victor, RAFA 2030, Vitória, and MARIA, aimed at classifying cases and managing precedents, achieving significant efficiency gains. However, it was found that the use of AI in the Supreme Court poses risks to the maintenance of fundamental rights and guarantees, such as adversarial proceedings, due process, natural judge, and transparency, requiring strict regulation. In this regard, CNJ Resolution No. 615/2025 stands out as a modern normative instrument to guide its application. The study concludes that, at the current technological stage, AI should be used as an auxiliary tool, without replacing human participation in sensitive activities, ensuring modernization benefits without compromising fundamental rights and guarantees.

Keywords: Fundamental Rights and Guarantees. Institutional Modernization. Precedents.

RESUMEN

Este estudio analiza las repercusiones de la adopción de tecnologías de inteligencia artificial (IA) por parte del Tribunal Supremo Federal de Brasil (STF), teniendo en cuenta sus efectos sobre la eficacia de los servicios judiciales y la protección de los derechos y garantías fundamentales. El objetivo principal es comprender cómo estas herramientas contribuyen a la agilidad procesal e influyen en la calidad de las decisiones y la seguridad jurídica. Para ello, la investigación se propuso identificar las principales soluciones de IA implementadas por el STF, evaluar los beneficios y retos derivados de su uso, y examinar las implicaciones para el razonamiento y la previsibilidad de las decisiones. El estudio adopta un enfoque cualitativo y exploratorio, basado en el análisis documental de normativas, jurisprudencia e informes institucionales, complementado con una revisión bibliográfica sobre la IA aplicada al derecho y la innovación judicial. Los resultados indican que, desde 2018, el STF ha invertido en sistemas como Victor, RAFA 2030, Vitória y MARIA, destinados a clasificar los casos y gestionar los precedentes, logrando importantes ganancias de eficiencia. Sin embargo, se constató que el uso de la IA en el Tribunal Supremo plantea riesgos para el mantenimiento de los derechos y garantías fundamentales, como el principio de contradicción, el debido proceso, el juez natural y la transparencia, lo que requiere una regulación estricta. En este sentido, la Resolución n.º 615/2025 del CNJ destaca como un instrumento normativo moderno para orientar su aplicación. El estudio concluye que, en la etapa tecnológica actual, la IA debe utilizarse como herramienta auxiliar, sin sustituir la participación humana en actividades sensibles, garantizando los beneficios de la modernización sin comprometer los derechos y garantías fundamentales.

Palabras clave: Derechos y Garantías Fundamentales. Modernización Institucional. Precedentes.



Esta obra está bajo una [Licencia Creative Commons Atribución- NoComercial 4.0 Internacional](https://creativecommons.org/licenses/by-nc/4.0/)

INTRODUÇÃO

A incorporação de tecnologias de Inteligência Artificial (IA) pelo Supremo Tribunal Federal (STF) representou um marco na modernização do Poder Judiciário brasileiro. O uso dessas ferramentas busca aprimorar a eficiência e a celeridade na análise e no julgamento dos processos, especialmente diante do elevado volume de demandas que chegam à Corte. No entanto, o uso da IA no STF também levanta questões sobre transparência, controle decisório e impactos na função jurisdicional, tornando-se um tema relevante para o estudo do direito e da tecnologia.

O problema de pesquisa que norteia este trabalho consiste em investigar quais são os impactos do uso da Inteligência Artificial pelo STF na efetividade da prestação jurisdicional e na segurança jurídica das decisões proferidas. Em outras palavras, busca-se compreender de que maneira a adoção dessas tecnologias influencia a celeridade processual, a qualidade das decisões e a preservação dos princípios fundamentais do direito.

Para responder a essa questão, formulam-se as seguintes hipóteses, que também constituem os objetivos da pesquisa. O objetivo geral é analisar o impacto da implementação da Inteligência Artificial no STF, considerando suas contribuições para a eficiência do tribunal e suas implicações no processo decisório. Para atingir esse objetivo, estabelecem-se os seguintes objetivos específicos: (a) identificar as principais ferramentas de IA utilizadas pelo STF e suas finalidades; (b) avaliar os benefícios e desafios decorrentes da adoção dessas tecnologias na prática judicial; e (c) examinar os impactos da IA na fundamentação e na previsibilidade das decisões do tribunal.

Metodologicamente, a pesquisa adota uma abordagem qualitativa e exploratória, com base na análise documental de normativas, jurisprudência e relatórios institucionais sobre o uso da IA no STF. Além disso, será realizada uma revisão bibliográfica sobre os conceitos de Inteligência Artificial aplicados ao direito, bem como estudos sobre inovação e transformação digital no Judiciário. A fundamentação teórica terá como base autores que discutem os limites da automação no processo judicial e os impactos da IA na tomada de decisões jurídicas.

Por meio dessa investigação, pretende-se contribuir para a compreensão do papel da

Inteligência Artificial no STF e seus desdobramentos para a efetividade da Justiça, observando em que medida a tecnologia pode ser um instrumento de otimização processual sem comprometer garantias fundamentais do ordenamento jurídico.

O QUE É INTELIGÊNCIA ARTIFICIAL (IA)?

A IA revela-se como uma ferramenta revolucionária de transformação para o mundo contemporâneo, atuando em praticamente todos os contextos da vida humana, o que gradativamente vem ocorrendo nos últimos cinquenta anos (Vigliar, 2023). Russel e Norvig (2022) afirmam não haver uma única abordagem para definir a Inteligência Artificial, mas há ao menos quatro escolas nas Ciências da Computação que podem atribuir à IA o desenvolvimento de computadores: 1) que ajam de forma humana; ou 2) que pensem de forma humana; ou 3) que pensem racionalmente; ou 4) que ajam racionalmente. Apesar da diversidade de abordagens, Vigliar (2023) apresenta um conceito simplificado de como a Inteligência Artificial pode comumente ser apresentada:

De uma forma concisa e abrangente definimos inteligência artificial como a capacidade de sistemas e máquinas, através de informações preestabelecidas, diagnosticar melhores decisões no tocante as definições de acordo com cada possibilidade, independente de definição prévia (Vigliar, 2023, p.92).

A Inteligência Artificial, como o nome sugere, é uma “inteligência” não humana, criada por humanos para atender as demandas que lhe serão preestabelecidas também pelos humanos e constituída, em última análise, por um conjunto de códigos – programas e algoritmos – de computador e dados constituídos por informações, indexadores e modelos estatísticos de aprendizado de máquina.

Conforme Russel e Norvig (2022), uma das formas mais bem-sucedidas para se entender o que é a IA foi o chamado Teste de Turing, no qual se propôs uma abordagem pragmática suportada no desempenho obtido por um sistema computacional em um certo teste de inteligência comportamental.

Alan Turing, em seu famoso ensaio “Computing Machinery and Intelligence” (1950), sugeriu que, em vez de perguntar se as máquinas podem pensar, devemos perguntar se as máquinas podem passar por um teste de inteligência comportamental, que veio a ser chamado teste de Turing. O teste consiste em fazer um programa desenvolver uma conversa (via mensagens digitadas) com um interrogador por 5 minutos. O

interrogador deve então descobrir se teve a conversação com um programa ou uma pessoa; o programa passa pelo teste se enganar o interrogador durante 30% do tempo. Para Turing, o ponto principal não foram os detalhes exatos do teste, mas sim a ideia de medir a inteligência pelo desempenho sobre algum tipo de tarefa comportamental aberta, em vez de especulação filosófica (Russel; Norvig, 2022, p. 892).

De fato, para que o computador fosse considerado apto no Teste de Turing, diversas capacidades atualmente identificadas como Inteligência Artificial deveriam nele estar presentes, de acordo com Russel e Norvig (2022): o processamento de linguagem natural, para viabilizar a comunicação em linguagem humana; representação de conhecimento, para armazenar o que se sabe e ou se apreende; raciocínio automatizado, que permita responder perguntas e formular conclusões; e o aprendizado de máquina, no sentido de prover adaptabilidade a novas circunstâncias propostas, a detectar padrões e a projetar conhecimentos e possibilidades a partir deles.

Russel e Norvig (2022) afirmam que é da autoria conjunta de Warren McCulloch e Walter Pitts, em 1943, o primeiro trabalho a ser reconhecido como Inteligência Artificial. Segundo esses autores, McCulloch e Pitts,

[...] inspirados pelo trabalho de modelagem matemática de Nicolas Rashevsky (1936, 1938), orientador de Pitts, basearam-se em três fontes: o conhecimento da fisiologia básica e da função dos neurônios no cérebro; uma análise formal da lógica proposicional criada por Russell e Whitehead; e a teoria da computação de Turing. Esses dois pesquisadores propuseram um modelo de neurônios artificiais, no qual cada neurônio se caracteriza por estar “ligado” ou “desligado”, com a troca para “ligado” ocorrendo em resposta à estimulação por um número suficiente de neurônios vizinhos. O estado de um neurônio era considerado “equivalente, em termos concretos, a uma proposição que definia seu estímulo adequado”. Por exemplo, eles mostraram que qualquer função computável podia ser calculada por alguma rede de neurônios conectados e que todos os conectivos lógicos (E, OU, NÃO etc.) podiam ser implementados por estruturas de redes simples. McCulloch e Pitts também sugeriram que redes definidas adequadamente seriam capazes de aprender (Russel; Norvig, 2022, p. 15).

Esses autores afirmam ainda que os principais marcos da história da Inteligência Artificial poderiam ser resumidos nas contribuições dos vencedores do Prêmio Turing. Cada um deles produziu contribuições científicas fundamentais para as aplicações em IA que, hoje, estão acessíveis aos consumidores. Segundo Russel e Norvig (2022), esses personagens são:

[...] Marvin Minsky (1969) e John McCarthy (1971) pela definição dos fundamentos do campo com base na representação e no raciocínio; Ed Feigenbaum e Raj Reddy (1994) pelo desenvolvimento de sistemas especialistas, que codificam o conhecimento humano para resolver problemas do mundo real; Judea Pearl (2011) pelo desenvolvimento de técnicas de raciocínio probabilístico que lidam com a incerteza de um modo baseado em princípios; e, finalmente, Yoshua Bengio, Geoffrey Hinton e Yann LeCun (2019) por

tornar o “aprendizado profundo” (redes neurais multicamadas) uma parte crítica da computação moderna (Russel; Norvig, 2022, p. 15).

Ainda que o ramo da Inteligência Artificial não seja tão novo, mais recentemente, uma categoria de aplicações em IA – a Inteligência Artificial Generativa (IAG) - eclodiu e popularizou-se de tal modo a revolucionar muito daquilo que se denomina trabalho de natureza intelectual. Conforme transcrição do Evento “Impactos da Inteligência Artificial no Constitucionalismo Contemporâneo”, realizado em agosto de 2024, nas palavras do Excelentíssimo Senhor Ministro Luís Roberto Barroso, presidente do STF naquele momento,

Como se esse conjunto de novidades já não fosse suficiente, vem agora a Inteligência Artificial. Na verdade, a Inteligência Artificial já vem de longe, embora o impacto mais perceptível que ela tenha trazido, pelo menos no mundo intelectual, tenha sido a Inteligência Artificial Generativa, essa do ChatGPT, que é capaz de gerar conteúdo: texto, sons e imagens, e isso é uma transformação muito profunda. A Inteligência Artificial, dito singelamente, é um conjunto de softwares, de programas de computador, que transferem para os computadores, para máquinas, capacidades cognitivas e decisórias humanas. Isso é basicamente o que a Inteligência Artificial faz. Só que, hoje em dia, no atual estado da arte, a Inteligência Artificial não tem consciência de si própria, não sabe o que é certo e o que é errado, o que é bom e o que é ruim, não tem sequer senso comum (Brasil, 2024, p.12).

De fato, as aplicações de IA são hoje onipresentes e oferecem diversos benefícios à vida humana pela possibilidade de armazenamento e processamento de um conjunto de informações, semelhantemente à forma humana, muito além da capacidade do cérebro humano e em alta velocidade (Brasil, 2024, p.12). Por outro lado, existem os riscos, como a utilização para fins bélicos, massificação na circulação da desinformação da maldade, a exemplo do Deepfake, que é a “possibilidade de alguém me colocar aqui falando coisas que eu nunca disse, sem que seja possível detectar a olho nu que isso está acontecendo” (Brasil, 2024, p.13).

MOTIVAÇÕES E BENEFÍCIOS DO USO DE IA NO ÂMBITO DO JUDICIÁRIO

A crescente complexidade das demandas judiciais e o acúmulo de processos nos tribunais brasileiros têm impulsionado a busca por soluções tecnológicas que promovam maior eficiência e racionalidade na prestação jurisdicional. Essencialmente, o que se busca é dar maior celeridade à resolução dos conflitos e ampliar o acesso à justiça.

Nesse cenário, as ferramentas de Inteligência Artificial (IA) se mostram promissoras para auxiliar as mais diversas tarefas operacionais do judiciário, especialmente na triagem de

processos, na elaboração de minutas e na análise de grandes volumes de dados, ou seja, a IA possibilita a automatização de tarefas repetitivas e burocráticas, liberando servidores técnicos, analistas e magistrados para atividades de maior complexidade cognitiva. Conforme destaca o relatório do Conselho Nacional de Justiça (CNJ),

No campo do Direito, a IA tem encontrado uma série de aplicações, tanto na prática dos escritórios de advocacia quanto nos tribunais, para classificação e acompanhamento de processos, síntese e extração de informações relevantes de documentos, buscas inteligentes de conteúdo, em particular jurisprudência, predição de decisões e análises de desempenho dos tribunais e geradores automáticos de minutas e documentos (petições, despachos etc.) (Conselho Nacional de Justiça, 2024, p. 14).

Como exemplo, temos o caso do Supremo Tribunal Federal que, nesse diapasão, tem se utilizado da Inteligência Artificial para a produção de minutas auxiliares à prestação jurisdicional, tais como: resumos de votos, resumos de relatórios de ministros em Recursos Extraordinários (Res), em Recursos Extraordinários com Agravo (AREs) e o artefato de análise inicial de processos de reclamação, produtos que colaboram com o aumento da eficiência dos profissionais envolvidos e da qualidade dos elementos processuais (Brasil, 2025).

Outro importante motivador para a adesão de tecnologias de IA se ancora na constante busca pela modernização institucional no âmbito dos governos. A digitalização dos serviços públicos e as políticas de governo eletrônico têm alcançado os sistemas de justiça, os quais têm sido constantemente atualizados com vistas ao atendimento das exigências de uma sociedade cada vez mais conectada (Rustamzade; Bagirov; Huseynov, 2024).

Mais recentemente, o surgimento no mercado da espécie de tecnologia de Inteligência Artificial denominada IA Generativa (IAG) – a exemplo de soluções como o ChatGPT da OpenAI – elevou sobremaneira as expectativas e as demandas pelo uso da IA no âmbito do judiciário, inclusive pela possibilidade de seu uso para produção de atos processuais decisórios, uma vez que

os chatbots de IAG utilizam Large Language Models (LLMs) para gerar respostas aos prompts submetidos pelos usuários em seus sistemas. Usuários inserem textos ou fazem requisições em linguagem natural — os chamados prompts — para que os chatbots de IAG respondam perguntas ou realizem tarefas, como sintetizar documentos ou sistematizar em tabelas, extrair excertos relevantes, produzir redações, poemas, letras de músicas, receitas culinárias, realizar recomendações etc. Caso necessário, os usuários podem refinar os prompts inseridos para receber respostas mais adequadas aos seus interesses. (Conselho Nacional de Justiça, 2024, p. 27).

Tais fatos evidenciam que o uso progressivo das tecnologias de Inteligência Artificial no

Poder Judiciário não se configura mera tendência, mas uma realidade. Nesse sentido, é imperiosa a avaliação criteriosa das formas de aplicação dessas tecnologias, em especial no Supremo Tribunal Federal, e suas implicações e riscos na administração da justiça e na manutenção dos direitos e garantias fundamentais.

O USO DA INTELIGÊNCIA ARTIFICIAL NO PODER JUDICIÁRIO BRASILEIRO

Conforme relatou o CNJ (2024), no âmbito do judiciário, a tecnologia tem sido implantada de maneira progressiva a fim de otimizar as atividades processuais direcionadas ao acréscimo de eficiência ao sistema judiciário, no intuito de aprimoramento da gestão processual e, ao mesmo tempo, a redução da morosidade de suas intervenções.

Desde o fim do século passado, os tribunais têm sido gravemente afetados pelas intervenções da informatização, tendo como marco a promulgação da Lei n. 11.419/2006, que dispõe sobre a informatização do processo judicial. Essa etapa de digitalização, inclusive, configurou-se como importante fundamento para a futura adoção da IA no Poder Judiciário.

Em nada difere o histórico da aplicação da Inteligência Artificial no âmbito do Poder Judiciário – tanto para atividades-meio quanto para atividades finalísticas (Conselho Nacional de Justiça, 2024), de modo que uma avaliação dos normativos do Conselho Nacional de Justiça que tratam do tema pode dar uma razoável noção desse progresso.

Primeiramente, temos a Resolução n. 332/2020 e a Portaria n. 271/2020 como primeiro marco regulatório adotado pelo judiciário brasileiro. Conforme o relatório do CNJ, sob título “O uso da inteligência artificial generativa no Poder Judiciário brasileiro” (Conselho Nacional de Justiça, 2024), tais normativos, no intuito de aproveitar sinergias e identificar riscos à prática judicial, trouxeram

[...] regras para a documentação no CNJ, na Plataforma Sinapses, que armazena, controla o versionamento, a distribuição e a auditoria dos modelos de IA empregados nos tribunais brasileiros, e trazem também regras sobre governança ética de IA, para assegurar seu desenvolvimento responsável, voltado para a gestão de riscos a direitos fundamentais apontados anteriormente (Conselho Nacional de Justiça, 2024, p. 16).

Mais adiante, o fenômeno da Inteligência Artificial Generativa (IAG) e os riscos decorrentes de seu uso levaram o CNJ a estabelecer grupo de trabalho para revisão e atualização da Resolução n. 332/2020, no intuito de “[...] adequá-la aos avanços da tecnologia, principalmente às IAGs, além de propor a realização de pesquisa empírica para apurar o uso dessas ferramentas

nos tribunais [...]” (Conselho Nacional de Justiça, 2024, p. 17).

Recentemente, o CNJ, em substituição à Resolução n. 332/2020, publicou a Resolução Nº 615 de 11/03/2025, com o propósito de estabelecer “[...] diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário” (Conselho Nacional de Justiça, 2025). Tal resolução conta com 47 artigos, sendo o principal e mais recente regulamento para o uso da IA no Poder Judiciário.

Segundo o CNJ (2024), no âmbito do judiciário “[...] há 147 sistemas de IA comunicados ao Sistema Sinapses” (Conselho Nacional de Justiça, 2024, p. 22). Tais aplicações direcionam-se tanto às atividades-meio quanto às finalísticas que envolvem a prática jurídica.

O CNJ (2024) lista a categoria de aplicações de IA no âmbito do Poder Judiciário, conforme o seguinte: a) Automação de tarefas repetitivas: eficiência operacional e economia de tempo dos servidores e das servidoras; b) Suporte à decisão e eficiência operacional: auxílio a decisões e redução do tempo de tramitação dos processos; c) Melhoria em serviços de atendimento: uso de IA em balcões virtuais, chatbots e tradução de “juridiquês” para linguagem comum; d) Otimização de processos administrativos: IA aplicada em licitações, estratégias administrativas e resolução de problemas administrativos; e) Análise aprofundada de documentos e precedentes: melhoria da análise jurídica e da identificação de litispendência e demandas predatórias; f) Apoio à tomada de decisão judicial: auxílio a magistrados em minutas de decisão e julgamentos; g) Concentração em análises processuais: foco em análises qualificadas e contribuições diretas para atividades judicantes; h) Minimização de deficiências processuais: redução de erros e ineficiências no processo judicial; i) Eficiência na prestação jurisdicional: agilização do processamento judicial, especialmente em anos eleitorais (Conselho Nacional de Justiça, 2024, p. 22 e 23).

Como se vê, o Poder Judiciário tem avançado no uso da IA implementando sistemas inteligentes voltados à otimização dos serviços e maior viabilidade da eficiência administrativa.

PRINCIPAIS FERRAMENTAS DE IA UTILIZADAS PELO STF E SUAS FINALIDADES

Diante de um cenário de grande volume de processos, o Supremo Tribunal Federal tem tido um papel de vanguarda no que tange à adoção de aplicações de Inteligência Artificial no Poder Judiciário brasileiro. Abordam-se, nos parágrafos seguintes, as principais ferramentas de

IA utilizadas pelo Supremo identificadas neste trabalho.

Segundo Peixoto (2020), em abril de 2018, o STF lançou o “Projeto de Pesquisa & Desenvolvimento de aprendizado de máquina (machine learning) sobre dados judiciais das repercussões gerais do Supremo Tribunal Federal – STF”, em parceria com a Universidade de Brasília (UnB), responsável pela execução, cujo produto final – uma solução de IA – recebeu o nome de Victor, que se tratou de “[...] homenagem à Victor Nunes Leal, Ministro do STF entre 1960 e 1969, responsável pela sistematização da jurisprudência do STF em Súmula, prática que facilitou a aplicação de precedentes judiciais aos recursos [...]” (Peixoto, 2020, p. 10).

Sendo uma solução baseada na capacidade de aprendizado de máquina ou *machine learning*, Peixoto (2020) relata ainda que “a IA do Victor busca analisar texto do processo para classificá-lo em algum tema reconhecido de repercussão geral” (Peixoto, 2020, p. 3). Já o Ministro Luís Roberto Barroso tornou públicos maiores detalhes a respeito do funcionamento do Victor, como segue:

Nós temos o programa Victor, que é um programa, um sistema que identifica entre os recursos extraordinários que chegam, aqueles que já foram objeto de repercussão geral, para os quais, portanto, já tem uma solução definida pelo Supremo. De modo que, aquele recurso não deve sequer subir, apenas mandar que o tribunal de onde veio o recurso aplique a tese que nós já afirmamos (Brasil, 2024, p.14).

Em 2022, na gestão do Ministro Luiz Fux, o STF desenvolveu a RAFA 2030 - Redes Artificiais Focadas na Agenda 2030, a qual adota as capacidades de redes neurais com comparação semântica e cujo funcionamento visa “[...]apoiar a classificação de processos na Corte de acordo com os Objetivos de Desenvolvimento Sustentável (ODS) da Agenda 2030 da Organização das Nações Unidas (ONU)” (Brasil, 2022).

Ainda na busca do aperfeiçoamento operacional da repercussão geral e da sistemática de precedentes judiciais, em 2023, o STF, na gestão Rosa Weber, por meio de sua “Assessoria de Inteligência Artificial (AIA), em parceria com a Secretaria de Tecnologia da Informação (STI) e a Secretaria de Gestão de Precedentes do STF” (Brasil, 2023), elaborou a solução denominada VitórIA, cujo propósito é agrupar “[...] processos por similaridade de temas para identificação de novas controvérsias” (Brasil, 2023). Segundo o Ministro Luís Roberto Barroso, VitórIA é o programa que agrupa os recursos por temas, para que a gente saiba qual tema está tendo maior incidência para pinçá-lo e dar repercussão geral, para tentar diminuir a litigiosidade naquela matéria (Brasil, 2024, p.14).

Por fim, o Supremo Tribunal Federal (STF) lançou em 16/12/2024 a ferramenta MARIA

(Módulo de Apoio para Redação com Inteligência Artificial), sua primeira tecnologia baseada em Inteligência Artificial Generativa (Brasil, 2024), cuja ênfase é a produção automatizada de diversos tipos de texto e será aplicada para resumir votos, resumir relatórios em processos recursais e realizar a análise inicial de processos de reclamação.

Conforme divulgado, quanto ao resumo de votos, “a MARIA pode gerar automaticamente minutas de ementas, com o resumo do entendimento do ministro sobre a matéria em questão”. A ferramenta ainda é capaz de “[...] resumir relatórios de ministros em Recursos Extraordinários (REs) e em Recursos Extraordinários com Agravo (AREs)” (Brasil, 2025). Conforme afirmou o Ministro Barroso, “Nos processos das classes de recurso extraordinário e recurso extraordinário com agravo, a MARIA automatiza a elaboração de relatórios processuais essenciais para as decisões judiciais” (Brasil, 2024).

No que concerne à análise inicial de processos de reclamação, a solução MARIA foi concebida para poder analisar o teor da petição inicial e responder a questionamentos que direcionam o estudo inicial da peça.

PRINCIPAIS RISCOS INERENTES AO USO DE IA NO SUPREMO TRIBUNAL FEDERAL

A aplicação da Inteligência Artificial no Poder Judiciário, e em especial no Supremo Tribunal Federal, embora promissora, traz consigo riscos significativos que precisam ser cuidadosamente compreendidos e mitigados. Rustamzade, Bagirov e Huseynov advertem que a “[...]sensibilidade dos processos, restrições legais, garantia da independência do judiciário, design do sistema, boa experiência do usuário e alta interoperabilidade[...]”⁴ (Rustamzade; Bagirov; Huseynov, 2024, p. 166, tradução nossa) são aspectos relevantes a demonstrar que a adoção de IA não pode ser feita de forma acrítica ou desregulada.

Essa é uma discussão que merece ser realizada com profunda análise, a fim de entender, no contexto da IA, sobre a ética no uso dessas tecnologias no âmbito dos tribunais. Como a IA é algo novo e, de certo modo, complexo para muitas pessoas, torna-se prudente verificar sobre como deve ser utilizada dentro dos parâmetros constitucionais aceitos.

Além dos desafios típicos ao uso da IA, como privacidade, explicabilidade e viés, existem

4 At the same time, the introduction of AI into the justice system is associated with risks and challenges dictated by the following system features: sensitivity of processes, legal restrictions, ensuring the independence of the judiciary, system design, good user experience, and high interoperability.

riscos no campo da ética relevantes ao nosso estudo. Conforme explica Blackman:

Quase todas as discussões sobre ética de IA começam com a condenação da existência da IA enviesada, inexplicável e violadora da privacidade, ao mesmo tempo em que se expressa indignação com as empresas que as desenvolvem. Essas pessoas estão abordando a questão do conteúdo. A conversa, então, imediatamente gira em torno de técnicas ou de ferramentas que os desenvolvedores podem usar para mitigar riscos éticos. A suposição implícita nesse eixo é que entendemos os problemas de Conteúdo bem o suficiente para começar a trabalhar na mitigação dos riscos. Entretanto, essa suposição também é falsa. Mesmo aqueles que passam muito tempo falando e até trabalhando no campo da ética de IA têm muita dificuldade em entender o que deve ser feito, porque estão tentando construir Estrutura em torno de algo que ainda acham gosmento, difuso e subjetivo. Essas pessoas estão fazendo mitigação de risco ético de IA e, embora possam saber muito sobre IA e sobre mitigação de risco, não sabem muito sobre ética (Blackman 2024, p. 16).

Nessa mesma linha, Vigliar (2023) reflete sobre o uso da IA no Poder Judiciário brasileiro:

Qual o limite para a utilização e substituição da inteligência e supervisão humana por máquinas previamente dotadas de capacidade de superar as vicissitudes que envolvem decisões e atos na esfera judicial, cabendo ao Estado regulamentar as novas aplicações de modo que os princípios da dignidade da pessoa humana sejam respeitados e aplicados no Poder Judiciário brasileiro (Vigliar, 2023, p.89).

Certamente um dos limites a ser resguardado são as garantias e direitos fundamentais dos cidadãos brasileiros na prestação jurisdicional e, juntamente nessa esteira, requer-se a otimização do tempo e de recursos para a resolução de conflitos por meio de acertadas decisões (Vigliar, 2023).

Emerge saber qual é o limite para o uso de uma inteligência automatizada no lugar da humana ou em auxílio a ela. Assim, discorre-se adiante os principais pontos de atenção identificados na aplicação de inteligência artificial no Poder Judiciário, em especial, na Suprema Corte, como a possibilidade de erros, as limitações técnicas da IA e o “viés de automação”; o requisito da explicabilidade; o problema do “viés algorítmico”; o risco de delegação indevida de certas tarefas exclusivas do magistrado à IA; autoria e responsabilização; violação de privacidade e confidencialidade das informações; e a necessidade de observância do requisito da preservação de direitos e garantias fundamentais.

A possibilidade de erro é uma das importantes características presentes no uso das tecnologias de Inteligência Artificial. Esse fato, em determinados casos de uso, pode tornar a aplicação da IA extremamente sensível ou eventualmente contraindicada. Ainda que robustos, tais sistemas são projetados para produzir os melhores resultados possíveis a partir do

processamento de grandes volumes de dados (base de aprendizado) e mecanismos de correlação e padronização estatísticos e probabilísticos.

Para melhor proveito possível, deve-se compreender os resultados não como substancialmente verdadeiros, mas que, a partir do algoritmo e de suas entradas e bases de dados, mostraram-se mais probabilisticamente corretos.

Essa natureza estocástica da IA implica que os resultados não são necessariamente corretos ou verificáveis. É o que muito bem elucida o CNJ ao afirmar que, quanto ao comportamento dos algoritmos de IA Generativa, “o resultado gerado é apenas aquilo que a ferramenta prevê ser a combinação mais provável de palavras subsequentes a partir do texto formulado no prompt” (Conselho Nacional de Justiça, 2024, p. 30). Assim, *contrariu sensu*, algoritmos de IA não estão imunes a falhas, imprecisões ou interpretações equivocadas, já que possibilidade de erros é algo inerente a essa tecnologia.

Um típico erro encontrado nas IAs chamadas Generativas é o fenômeno das “alucinações”, respostas geradas pela IA que têm plausibilidade, sendo, porém, factualmente incorretas ou inexistentes (Conselho Nacional de Justiça, 2024). No âmbito do Poder Judiciário, alucinações de IA já se fizeram gravemente presentes. Tal risco inclusive ocasionou incidente noticiado na imprensa envolvendo o Tribunal Regional da 1.^a Região, em que sentença elaborada com o uso do ChatGPT fez referência a precedentes inexistentes, gerando constrangimento às partes e aos advogados, com repercussão negativa para o Poder Judiciário (Conselho Nacional de Justiça, 2024, p. 31).

Há ainda um outro risco comum que é o problema dos “vazamentos semânticos”, ou seja, quando a IA interpreta equivocadamente termos técnicos ou jurídicos. Exemplifica o CNJ que: “sequências que tomam o termo ‘expirar’, ligado no campo jurídico a prazos de validade, gerando sequências de texto ligadas a expirar como parte do processo respiratório” (Conselho nacional de Justiça, 2024, p. 30). Ainda que sutis, os “vazamentos semânticos” têm o potencial de comprometer a coerência e a validade de uma decisão judicial, especialmente em áreas do Direito ricas em termos técnicos próprios como: o direito tributário, penal ou administrativo.

Outra característica das IAs que contribui para a ocorrência de erros é a incapacidade dos algoritmos de IA em realizar inferências lógicas ou de senso comum. As IAs são cognitivamente limitadas, ou seja, os “sistemas baseados em LLMs não realizam raciocínio correspondente ao humano, apenas correlacionam estatisticamente sequências de símbolos” (Conselho nacional de Justiça, 2024, p. 31).

Assim, mesmo que textos gerados pela IA pareçam linguisticamente corretos, juridicamente consistentes e convincentes, é possível que os tais careçam de fundamentação lógica ou de aderência ao caso concreto. A IA não detém a capacidade humana de ponderar, interpretar e decidir com base em valores jurídicos e princípios constitucionais. Como afirma o CNJ, em seu relatório de pesquisa sobre o uso da IA Generativa no judiciário brasileiro,

Também em função de serem probabilísticos os modelos que embasam sistemas de IAG, tais ferramentas podem apresentar falhas em aplicações mais refinadas e sofisticadas de domínios técnicos, além de serem incapazes de realizar inferências lógicas ou de senso comum. Ainda que o texto produzido possa passar tal impressão, sistemas baseados em LLMs não realizam raciocínio correspondente ao humano, apenas correlacionam estatisticamente sequências de símbolos (palavras, trechos de textos). Por consequência, IAGs não fazem propriamente distinção entre fatos, contraposição de opiniões, deliberações ou exercício de juízo, compreensão ou análise, muito embora o resultado gerado pareça mimetizar esses processos do pensamento humano (Conselho Nacional de Justiça, 2024, p. 30 e 31).

Assim, por todo o exposto, é evidente que a possibilidade de erros e as limitações técnicas aliadas ao viés de automação – o qual consiste na aceitação e emprego de suas saídas sem questionamento e com confiança cega no sistema de IA (Conselho nacional de Justiça, 2024, 2024) - podem implicar em graves erros que comprometam a credibilidade institucional e provoquem danos concretos às partes envolvidas nos processos judiciais.

O uso de IA em casos de uso relacionados à prestação jurisdicional exige a implementação de mecanismos de revisão e validação humanas, de modo que se possa identificar e corrigir previamente eventuais falhas que tenham o condão de gerar efeitos jurídicos negativos.

A explicabilidade é uma condição necessariamente relevante para o uso da IA no âmbito do judiciário. A todo momento, a prestação jurisdicional demanda, seja na dimensão material ou processual do direito, os fundamentos jurídicos daquilo que se pede ou que se resolve, ou seja, explicações. Conforme Reid Blackman, as explicações são profundamente importantes quando queremos avaliar se um determinado tratamento é fundamentado, seja do ponto de vista ético, regulatório ou jurídico (Blackman, 2024, p. 78).

A literatura especializada, porém, tem apontado que os modelos de IA, especialmente os baseados em aprendizado de máquina, operam com tão elevado número de variáveis e de complexas relações de caráter probabilístico e estatístico que, conforme relata Blackman (2024), tornam-se, em muitos casos, ininteligíveis ao homem médio, na medida em que ele procure descrever em linguagem puramente humana os passos tomados pelo algoritmo para a entrega de determinado resultado, ainda que, em teoria, possam ser compreendidos os fundamentos

matemáticos que embasaram toda essa complexidade.

Em um contexto democrático, a legitimidade das decisões judiciais depende da sua inteligibilidade. São princípios como da motivação das decisões, do devido processo legal, do contraditório e da ampla defesa que tornam essencial a demanda por explicações humanamente inteligíveis, seja determinado procedimento realizado por um ser humano, seja por uma máquina.

Se os resultados produzidos pela IA não puderem ser explicados em termos compreensíveis, ela não poderá ser legitimamente utilizada como instrumento de produção de decisões, uma vez que os fundamentos que subsidiaram seus resultados não poderão ser combatidos ou eventuais erros e alucinações evidenciados. Assim, podemos afirmar que uma solução de IA aplicada ao judiciário deve permitir que possam ser dadas legítimas explicações aos resultados produzidos, pois uma boa explicação demonstrará respeito, permitirá que o usuário da IA ou as pessoas afetadas por ela tomem decisões embasadas, possibilitando que avaliem se as regras da decisão são justas, boas, razoáveis ou corretas (Blackman, 2024, p. 79).

Assim, resultados não explicáveis também afrontam a necessidade premente de transparência na atuação do Poder Judiciário e a confiança pública no sistema de justiça. Isso significa que, sem que haja clareza de como, quando e por que determinada IA produziu certos resultados, o exercício pleno de direitos e garantias fundamentais pelas partes – tais como o direito ao contraditório e à ampla defesa - resta prejudicado. Como alertam Nogueira e Moraes,

[...] há, contudo, um fator a considerar, pois as decisões tomadas por humanos são impugnáveis, podendo delimitar os fatores que ensejaram determinada resposta, o que não ocorreria com os algoritmos e ferramentas de inteligência artificial, não possuindo a clareza e a transparência necessárias para além de seus programadores (Nogueira; Moraes, 2023, p. 99).

O problema do “viés algorítmico” é um dos aspectos que mais despertam preocupações quando se cogita o uso da IA no âmbito Poder Judiciário. Por viés algorítmico considera-se a tendência dos algoritmos de IA de reproduzir acriticamente tendências, preconceitos e padrões - ainda que representem desigualdades sociais - presentes nos dados utilizados para o seu treinamento. Esse fenômeno pode ocorrer no caso em que o modelo é treinado com ampla base de documentos disponíveis na Internet e em redes sociais, que podem deter vieses e estereótipos quanto a grupos ou populações minorizadas, sendo mais difícil, em relação a sistemas de aprendizado desenvolvidos especificamente para determinada aplicação, fazer a curadoria para eliminar tais vieses (Conselho Nacional de Justiça, 2024, p. 36 e 37).

Quando tal tendência se configura, determinados procedimentos do sistema de justiça

automatizados pela IA podem ser contaminados a ponto de acentuar desigualdades estruturais e reforçar parcialidades, comprometendo a igualdade perante a lei. O viés também pode se manifestar na repetição de decisões, ou seja, quando “[...] um determinado entendimento pode ser inscrito no algoritmo e prejudicar um réu que poderia ser absolvido por meio de outro entendimento” (Martins; Sá, 2023, p. 142), eventualmente mais adequado ao caso concreto.

Em outras palavras, o viés algorítmico pode se manifestar na consolidação de determinados entendimentos de tal forma a impedir o surgimento de raciocínios jurídicos diversos e melhormente aplicáveis a determinados casos concretos e, portanto, mitigando a pluralidade interpretativa. Uma vez treinadas para seguirem decisões outrora firmadas, as soluções de IA reforçariam a jurisprudência dominante. Tal fato pode necessariamente dificultar a avaliação de teses inovadoras e a proteção de direitos de grupos minoritários.

Reid Blackman (2024), em sua obra “Máquinas Éticas”, propõe a supervisão humana para mitigação de viés nos seguintes termos:

1. Deve haver um indivíduo, ou melhor, um conjunto de indivíduos com relevantes conhecimentos éticos, legais e de negócios que determine quais (se existentes) das métricas matemáticas de justiça são apropriadas para o caso de uso específico em questão.
2. Precisamos de um indivíduo ou, idealmente, um conjunto de indivíduos que tenha a experiência relevante para selecionar estratégias apropriadas de mitigação de viés. As tentativas de identificar e de mitigar possíveis vieses em modelos devem começar antes de treinar o modelo e, idealmente, antes de determinar a aparência e a origem dos conjuntos de dados de treinamento.
3. Deve-se incluir um advogado ao determinar as técnicas apropriadas de mitigação de viés (Blackman, 2024, p. 60).

Assim, a prevenção do viés deve ser seriamente pensada desde a concepção dos sistemas de IA para aplicação na prestação jurisdicional. O risco de delegação indevida de certas tarefas exclusivas do magistrado à IA configura-se outra questão relevante, em especial, no que tange à manutenção de garantias e direitos fundamentais. É a possibilidade de que a IA seja utilizada para substituir, e não meramente auxiliar, o magistrado humano. Tal inovação toca o âmago finalístico da prestação jurisdicional, a saber, a decisão judicial. Tamanha é sua importância e sensibilidade que, conforme lecionam Martins e Sá, nela reside a garantia de que o usuário do sistema de justiça, amparado pelos princípios constitucionais como o devido processo legal e do contraditório e ampla defesa, possa efetivamente garantir os seus direitos conquistados na lide, diante do arcabouço probatório que prove ser ele o detentor do direito posto em julgamento (Martins; Sá, 2023, p. 127).

O juiz humano, ainda que sujeito a falhas, durante o processo decisório, desenvolve

modelo de pensamento e discurso racional dialeticamente, ou seja, por meio do confronto interativo de argumentos objetivos ou subjetivos levantados: os dados obtidos nas provas documentais ou testemunhais; a compreensão particular na aplicação, ponderação e valoração de princípios do direito; a forma como interpretar e aplicar a lei aplicável ao caso concreto; sua experiência de vida; usos e costumes sociais; e, principalmente, a sensibilidade e a empatia. Tudo isso na busca da resolução da demanda, pacificação dos ânimos e consecução dos direitos e garantias fundamentais. É função do juiz “[...]entregar a justiça com equidade, sem dar margens a injustiças que podem comprometer a sua missão constitucional” (Martins; Sá, 2023, p. 133).

Seria possível à IA simular plenamente essa capacidade decisória exigida do juiz humano? Não! É fato que os algoritmos atuais não permitem decisões autônomas que dependam de tamanha complexidade. A IA, como sistema estocástico, ou seja, cujo funcionamento é meramente baseado em probabilidades, aleatoriedades e adaptabilidade a dados complexos e variáveis, é desprovida de capacidade reflexiva e, portanto, insensível às peculiaridades e minúcias dos casos concretos.

Uma máquina apenas repete o que ela foi ensinada, não é um sistema que reflete sobre decisões. Se uma determinada comarca muda de juiz e ele possui entendimento diverso ao anterior em relação a temas similares, a IA se fosse programada apenas por um entendimento, estaria prejudicada. Esse fato demonstra que as decisões tomadas pelas inteligências artificiais dificilmente seriam neutras em questões que envolvam mais de um entendimento, pois, conforme explicitado anteriormente, elas se baseariam, para tomar as suas decisões, nos dados que já estariam dispostos em seus códigos por meio do abastecimento realizado pelos desenvolvedores. Assim, essa seria uma desvantagem das decisões tomadas exclusivamente pela IA (Martins; Sá, 2023, p. 142).

Essa limitação é especialmente preocupante em casos que envolvam direitos fundamentais, minorias ou situações de vulnerabilidade, nas quais a sensibilidade humana é indispensável. É evidente que a justiça não pode ser reduzida a uma operação estatística. Porém, há uma tendência natural na delegação excessiva de tarefas decisórias às máquinas, o que representa fator de risco ao processo decisório. Faz-se necessário assentar a compreensão de que

os modelos de IA não são dotados de representação do conhecimento jurídico ou capacidade de realização de inferências típicas do raciocínio jurídico. Ou seja, a tecnologia empregada não simula a capacidade humana de interpretação, construção de conceitos jurídicos, argumentação e realização de inferências práticas a partir de normas jurídicas ou éticas (Maranhão, J.; Abrusio, J.; Almada, 2021, Apud Conselho Nacional de Justiça, 2024, p. 15).

Acrescente-se ainda o fato da ausência ou reduzida transparência e explicabilidade nos

sistemas de IA. Conforme tratamos anteriormente, em regra, os algoritmos operam como “caixas-pretas”, dificultando a compreensão dos critérios utilizados para se chegar à determinada conclusão. Essa condição vai de encontro ao dever constitucional de fundamentação das decisões judiciais, previsto no art. 93, IX, da Constituição Federal. Como destaca Avelar, a fundamentação das decisões judiciais deve ser entendida como um imperativo do sistema constitucional, assegurando a legitimidade do exercício do poder jurisdicional e fortalecendo a confiança no Poder Judiciário e nos valores democráticos” (Avelar, 2025).

Nesse mesmo contexto, temos a emblemática decisão lavrada em sede do Habeas Corpus de No. nº 74.073/RJ, de relatoria do ministro Celso de Mello, a qual leciona que a fundamentação dos atos decisórios qualifica-se como pressuposto constitucional de validade e eficácia das decisões emanadas do Poder Judiciário, de tal modo que a inobservância do dever imposto pelo art. 93, IX, da Carta Política, precisamente por afetar a legitimidade jurídica dessas deliberações estatais, gera, de maneira irremissível, a sua própria nulidade (BRASIL, STF, HC 74.073/MG, Rel. Min. Celso de Mello, 1997, p. 4).

Assim, a delegação total da produção de decisões à IA, sem ao menos supervisão do juiz humano, tendo em vista ausência de explicabilidade típica dessas soluções, é prejudicial à necessidade de adequada fundamentação jurídica, fato que, por si mesmo, implicaria na sua nulidade, sob pena de flagrante ofensa aos princípios do contraditório e da ampla defesa. Sobre a necessidade de adequada fundamentação das decisões judiciais, afirmou o Ministro Celso de Mello:

Tenho para mim que se reveste de nulidade o ato decisório, que, descumprindo o mandamento constitucional que impõe a qualquer Juiz ou Tribunal o dever motivar a sentença ou o acórdão, deixa de examinar, com sensível prejuízo pra o réu, fundamento relevante em que se apoia a defesa técnica do acusado (Brasil, STF, HC 74.073/MG, Rel. Min. Celso de Mello, 1997, p. 5).

Em outras palavras, uma decisão tomada exclusivamente pela máquina teria o condão de comprometer os princípios e garantias fundamentais assentadas na Constituição Federal, tais como: o juiz natural, o princípio do livre convencimento motivado do juiz, o princípio da devida motivação das decisões, os princípios do contraditório e da ampla defesa, o princípio da transparência, e, por conseguinte, a própria legitimidade do processo. Faz-se necessário ressaltar que

[...] o propósito do uso da IA em processos judiciais não é substituir os juristas, mas

fornece apoio na identificação de problemas e na indicação das prováveis consequências dos fatos do caso, oferecendo avaliações computadorizadas automatizadas que determinem integralmente o resultado de uma decisão ou, ao menos, influenciem-na. O uso da IA não representa uma decisão de retirar o poder das mãos humanas, mas sim uma nova forma de exercê-lo, colocando a justiça no centro de sua atividade consciente.⁵ (Rustamzade; Bagirov; Huseynov, p. 170, tradução nossa).

A autoria e a responsabilização dos tomadores de decisão judiciais são importantes pilares do Estado Democrático de Direito. O juiz é o agente responsável por fundamentar e assinar a decisão, sendo possível identificar sua autoria e, se necessário, questioná-la por meio dos instrumentos processuais adequados.

A introdução da Inteligência Artificial no processo decisório tem o condão de tornar esse vínculo de responsabilidade precário e difuso, levantando sérias preocupações quanto à *accountability* das decisões automatizadas. Nesse sentido, diversos dilemas poderiam ser verificados, tais como: 1) em que medida o magistrado poderia ser responsabilizado sobre eventuais equívocos em uma decisão assistida pela IA?, 2) como questionar uma decisão cuja lógica interna é ininteligível e a autoria atribuída a uma máquina?

As reflexões a respeito dos aspectos autoria e responsabilização podem se tornar problemática ao se considerar a cadeia de atores envolvidos no desenvolvimento de sistemas de IA, tais como: programadores, analistas de sistemas, analistas de requisitos, engenheiros de dados, engenheiros de IA, consultores, fornecedores de tecnologia etc. Qual seria o grau de responsabilização de cada um deles?

Assim, em observâncias aos direitos e garantias fundamentais dos jurisdicionados, as decisões eventualmente produzidas pela IA ou com o seu auxílio devem necessariamente passar pelo crivo do magistrado, o qual, após revisão, correção e validação de seu conteúdo, passa a assumir plenamente sua autoria. A IA assim passa a ter o papel de mera ferramenta auxiliar. Conforme lecionam Rustamzade, Bagirov e Huseynov, o trabalho da IA na redação de decisões judiciais é impossível de forma separada de um juiz humano; os algoritmos existentes não permitem que a IA produza ‘decisões de máquina’ de forma independente, sem a devida avaliação e qualificação jurídica por um juiz humano⁶ (Rustamzade; Bagirov; Huseynov, 2024, p. 174,

5 This leads to the conclusion that the purpose of using AI in court cases is not to replace lawyers, but to provide support by identifying problems and indicating the likely consequences of the facts of the case, providing automated computerised assessments that either fully determine the outcome of a decision or at least influence it. The use of AI does not represent a decision to take power out of human hands, but rather a new way of using human power by making justice central to its conscientious activity.

6 On the other hand, the work of AI in drafting court decisions is impossible separately from a human judge; existing algorithms do not allow AI to make “machine decisions” (court acts) independently, without providing legal assessment and qualification by a human judge.

tradução nossa).

A possibilidade de violação de privacidade e confidencialidade das mais diversas informações se configura como um risco comum quando se pensa na utilização de IA no Poder Judiciário, uma vez que a operacionalização envolve necessariamente o tratamento de informações confidenciais que incidentalmente poderiam ser expostas ou utilizadas indevidamente.

Uma primeira categoria de informações confidenciais são os dados pessoais sensíveis, conforme especificado na Lei Geral de Proteção de Dados, como segue:

Art. 5º Para os fins desta Lei, considera-se: I - dado pessoal: informação relacionada a pessoa natural identificada ou identificável; II - **dado pessoal sensível: dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural;** (Brasil, 2018, art. 5º, grifo nosso).

Quando refletimos na prestação jurisdicional, a confidencialidade cristaliza-se no instituto do segredo, que consiste na restrição de acessos a informações constantes dos autos a determinados atores. A respeito do segredo de justiça, preceitua o Código de Processo Civil, em seu artigo 189:

Art. 189. Os atos processuais são públicos, todavia tramitam em segredo de justiça os processos: I - em que o exija o interesse público ou social; II - que versem sobre casamento, separação de corpos, divórcio, separação, união estável, filiação, alimentos e guarda de crianças e adolescentes; III - em que constem dados protegidos pelo direito constitucional à intimidade; IV - que versem sobre arbitragem, inclusive sobre cumprimento de carta arbitral, desde que a confidencialidade estipulada na arbitragem seja comprovada perante o juízo. § 1º **O direito de consultar os autos de processo** que tramite em segredo de justiça e de pedir certidões de seus atos **é restrito às partes e aos seus procuradores.** § 2º **O terceiro que demonstrar interesse jurídico pode requerer ao juiz certidão do dispositivo da sentença, bem como de inventário e de partilha resultantes de divórcio ou separação** (Brasil, 2015, art. 189, grifo nosso).

O uso indevido dessas informações sujeitas à confidencialidade poderia ocorrer, por exemplo, nos casos em que as fabricantes das soluções de IA utilizassem tais informações para o treinamento e aperfeiçoamento de seus modelos de IA, o que claramente violaria o direito à privacidade e a confidencialidade. Na prática, essa hipótese, exemplifica o CNJ, poderia acontecer no uso inadvertido de soluções de IA Generativa por servidores, magistrados ou advogados:

As requisições feitas por meio de uso de prompts às ferramentas podem incluir dados pessoais ou mesmo conteúdo confidencial da organização. Tais informações podem ser coletadas pela empresa que disponibiliza o sistema e, em alguns casos, pode ser usada para retreinamento, fazendo com que o conteúdo confidencial possa ser incorporado ao modelo e “reproduzido” em outros usos por outros usuários (Conselho Nacional de Justiça, 2024, p. 35).

Ainda que certas soluções de IA permitam a seus desenvolvedores ou administradores ativarem bloqueios à utilização dos dados informados nos prompts para retreinamento das soluções (Conselho Nacional de Justiça, 2024), resta ainda a possibilidade de que servidores, magistrados e advogados recorram às soluções gratuitas na internet que não guardem as mesmas cautelas.

Assim, a preservação da confidencialidade no judiciário quanto ao uso de IA passa necessariamente por processos, políticas e controles que assegurem sua utilização segura, sobretudo, a conscientização dos usuários quanto a esse dever.

Especificamente, no caso de uma corte constitucional como o Supremo Tribunal Federal, agravam-se as preocupações de que tal uso possa implicar em eventuais violações a direitos e garantias fundamentais, sobretudo em uma instância da qual não cabe recurso. Assim, a aplicação da IA deve ser fundada nos princípios que fundam o estado democrático de direito.

A introdução da IA nos processos judiciais deve estar fundamentada, sobretudo, nos princípios do Estado de Direito, do respeito aos direitos humanos fundamentais, do respeito à honra e à dignidade, da igualdade perante a lei e o tribunal, da proporcionalidade, da competitividade entre as partes, da transparência, da imparcialidade e da equidade.⁷ (Rustamzade; Bagirov; Huseynov, 2024, p. 175, tradução nossa).

Nessa mesma direção, o Ministro Luís Roberto Barroso, durante o evento “Impactos da Inteligência Artificial no Constitucionalismo Contemporâneo”, expôs a necessidade de, no desenvolvimento e uso de solução de IA no Poder Judiciário brasileiro, assegurar a proteção de direitos fundamentais como a privacidade, a liberdade de expressão, a proteção à democracia.

Mas há valores que nós não podemos perder de vista: os direitos fundamentais, como a privacidade, como a liberdade de expressão, temperado com a livre iniciativa que também é um valor importante; tem a proteção à democracia com os riscos da desinformação e dos ataques consertados contra a democracia; e tem uma regulação sobre a governança [...] (Brasil, STF, 2024, p.13).

⁷ The introduction of AI into judicial proceedings should be based primarily on the principles of the rule of law, respect for fundamental human rights, respect for honour and dignity, equality before the law and the court, proportionality, competitiveness of the parties, transparency, impartiality and fairness.

Conforme exposto nos tópicos anteriores, existem diversos aspectos típicos das tecnologias de Inteligência Artificial que oferecem tanto oportunidades para melhoria da prestação jurisdicional quanto risco de violações de direitos e garantias fundamentais.

O uso inadvertido dessas tecnologias oferece riscos no aspecto da segurança da informação, ou seja, de violações da privacidade e confidencialidade de dados sensíveis e daqueles sujeitos a sigilo de qualquer natureza, em especial, o próprio segredo de justiça.

Apesar dos avanços e possibilidades que as tecnologias de IA oferecem, é necessário compreender que elas estão sujeitas a erros, não sendo isso demérito, mas característica intrínseca a um sistema estocástico. Devido a essa característica, a certas aplicações, a exemplo da tomada de decisões, não é indicada a aplicação da IA sem a supervisão do magistrado. De outro modo, é real a possibilidade do comprometimento dos requisitos de coerência e validade da decisão judicial.

Quando focamos na automação de decisões judiciais pela IA, o requisito da explicabilidade se torna indispensável, pois decisões incompreensíveis, ininteligíveis, imotivadas, cujas teses não são coordenadas ou não podem ser claramente identificadas, são, por si só, um obstáculo à apresentação de antíteses correspondentes ou, até mesmo, ao combate de eventuais erros. Sem explicabilidade, decisões produzidas por IA são uma grave afronta aos princípios do devido processo legal, do contraditório e ampla defesa e da transparência, bem como do dever constitucional de fundamentar as decisões.

Semelhantemente, há que se ater à questão do “viés algorítmico” com a devida atenção. O desenvolvimento das soluções de IA para a Suprema Corte devem ser desenvolvidas de modo a não reproduzir preconceitos e padrões de desigualdades sociais, bem como permitirem o surgimento de novas teses jurídicas. De outro modo, princípios como o da dignidade da pessoa humana, o da pluralidade de ideias, e do contraditório e ampla defesa restariam eventualmente esvaziados.

De fato, o uso exclusivo de IA para automatização de mecanismos de natureza decisória, sejam elas interlocutórias ou definitivas de mérito, tem o poder de atacar princípios como do juiz natural, do livre convencimento motivado do juiz, da devida motivação de decisões, entre tantos outros e, especialmente, da própria legitimidade do processo judicial.

Ademais, faz-se necessário que o juiz revise, corrija e assuma a autoria em definitivo da decisão, de modo que possa ser devidamente combatida ou eventualmente mantida, no que couber. Nessa perspectiva, o Ministro Barroso, durante solenidade de lançamento do sistema

MARIA, afirmou: “Nada do que a gente tem feito para agilizar a jurisdição dispensa o trabalho e a responsabilidade do juiz” (Brasil, 2025).

Ora, direitos e garantias fundamentais são elementos basilares da própria ideia de estado democrático de direito, razão pela qual as devidas cautelas devem ser adotadas no uso da IA para que eles sejam, de fato, salvaguardados.

O USO SUSTENTÁVEL DA IA NO JUDICIÁRIO: DIRETRIZES EUROPÉIAS E NACIONAIS

O uso sustentável da IA nos órgãos do judiciário tem sido objeto de atenção das autoridades ao redor do mundo. Conforme o CNJ (2024), iniciativas de legislação em diferentes países têm adotado abordagem baseada em riscos³², com imposições de obrigações de governança conforme a categoria de riscos das aplicações, de modo a comprometer os agentes desenvolvedores e aplicadores de inteligência artificial com as melhores práticas de gestão e mitigação dos riscos associados à tecnologia (Conselho Nacional de Justiça, 2024, p. 20).

A Europa, por exemplo, tem se destacado por adotar uma abordagem ética e criteriosa com base em riscos, merecendo destaque a “Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu Ambiente”, adotada pela CEPEJ na sua 31.^a reunião plenária, ocorrida em Estrasburgo, nos dias 3 e 4 de dezembro de 2018 (Concil of Europe, 2018). Essa Carta estabelece cinco princípios a serem seguidos para o desenvolvimento de solução baseadas em IA nos sistemas judiciais, a saber:

1. Princípio dos direitos fundamentais: Garantir que a concepção e a aplicação de instrumentos e serviços de inteligência artificial sejam compatíveis com os direitos fundamentais, incluindo o direito à proteção dos dados pessoais
2. Princípio da não discriminação: Prevenir especificamente o desenvolvimento ou a intensificação de qualquer discriminação entre indivíduos ou grupos de indivíduos
3. Princípio da qualidade e da segurança: No que diz respeito ao tratamento de decisões e dados judiciais, utilizar fontes certificadas e dados intangíveis com modelos elaborados em ambiente tecnológico seguro
4. Princípio da transparência, imparcialidade e equidade: Tornar os métodos de tratamento de dados acessíveis e compreensíveis, autorizar auditorias externas
5. Princípio "sob controle do usuário": Impedir uma abordagem prescritiva e garantir que os usuários sejam atores informados e controlem suas escolhas (Concil of Europe, 2018).

A Carta ainda estabelece um conjunto de boas práticas para cada um dos princípios que devem ser seguidas no desenvolvimento das soluções em IA. A regulamentação europeia tem servido de referencial para outras jurisdições, especialmente a brasileira. Tal influência

evidencia-se na Resolução n. 332/2020 do CNJ, a qual, como primeiro normativo adotado no âmbito do judiciário brasileiro, referência explicitamente o conteúdo da “Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu Ambiente” um dos motivadores considerados.

O mesmo espírito se observa na Resolução n. 615/2025 do CNJ, publicada como uma substituta da Resolução n. 332/2020, que estabelece em seu art. 3º os seguintes princípios:

Art. 3º O desenvolvimento, a governança, a auditoria, o monitoramento e o uso responsável de soluções de IA pelos tribunais têm como princípios: I – a justiça, a equidade, a inclusão e a não-discriminação abusiva ou ilícita; II – a transparência, a eficiência, a explicabilidade, a contestabilidade, a auditabilidade e a confiabilidade das soluções que adotam técnicas de inteligência artificial; III – a segurança jurídica e a segurança da informação; IV – a busca da eficiência e qualidade na entrega da prestação jurisdicional pelo Poder Judiciário, garantindo sempre a observância dos direitos fundamentais; V – o devido processo legal, a ampla defesa e o contraditório, a identidade física do juiz e a razoável duração do processo, com observância das prerrogativas e dos direitos dos atores do sistema de Justiça; VI – a prevenção, a precaução e o controle quanto a medidas eficazes para a mitigação de riscos derivados do uso intencional ou não intencional de soluções que adotam técnicas de inteligência artificial; VII – a supervisão humana efetiva, periódica e adequada no ciclo de vida da inteligência artificial, considerando o grau de risco envolvido, com possibilidade de ajuste dessa supervisão conforme o nível de automação e impacto da solução utilizada; e VIII – a oferta, pelos tribunais e suas escolas, de capacitação contínua para magistrados e servidores sobre riscos da automação, vieses algorítmicos e análise crítica dos resultados gerados por IA (Conselho Nacional de Justiça, 2025).

A nosso sentir, dispõe o judiciário brasileiro de excelentes referências técnico-normativas no sentido da preservação de direitos e garantias fundamentais em face aos riscos inerentes ao uso da IA, as quais, se bem observadas, têm o condão de nortear um uso ético, transparente e responsável. O desafio, por ora, consiste em trazê-los à prática nos projetos e soluções a serem desenvolvidos.

CONCLUSÃO

A Inteligência Artificial (IA) consiste em um gênero de tecnologias que têm sido amplamente utilizadas nos mais diversos ramos da atividade humana e, não diferentemente, no Poder Judiciário brasileiro, uma vez que têm se mostrado promissoras no suporte à automatização de tarefas operacionais, repetitivas e intensivas no uso de recursos humanos, com o potencial de aumentar a eficiência da prestação jurisdicional. Um outro motivador à adoção dessas tecnologias tem sido a busca constante pela modernização institucional.

Desde 2018, Supremo Tribunal Federal tem investido em soluções de IA. O sistema Victor, desenvolvido pela UNB, inaugurou essa era, classificando e agrupando processos similares em sede de repercussão geral na sistemática de precedentes. Em seguida, soluções como a RAFA 2030, Vitória e MARIA foram desenvolvidas.

Por outro lado, o uso da IA no Poder Judiciário e sobretudo no caso da Suprema Corte, última instância recursal, oferece riscos à manutenção de direitos e garantias fundamentais. A depender de como as tecnologias de IA sejam aplicadas, princípios como do contraditório e ampla defesa, do juiz natural, da transparência, do livre convencimento motivado do juiz, entre outros, podem restar prejudicados ou esvaziados.

Ora, direitos e garantias fundamentais são elementos basilares da própria ideia de estado democrático de direitos, razão pela qual as devidas cautelas devem ser adotadas no uso da IA a fim de que eles sejam, de fato, salvaguardados. Nesse sentido, o CNJ tem regulamentado o seu uso, cujo instrumento mais recente é a Resolução N. 615/2025, que, a nosso ver, consiste de moderno instrumento para nortear o uso da IA no Supremo Tribunal Federal.

Sendo assim, no estado da arte em que a IA se encontra, é fundamental que ela seja vista como ferramenta auxiliar à prestação jurisdicional, mas nunca como um substituto do ser humano em atividades sensíveis do órgão julgador, das quais destacamos o proferir decisões interlocutórias ou definitivas de mérito.

REFERÊNCIAS

AVELAR, Daniel Ribeiro Surdi de. **IA e a fundamentação das decisões: desafios e perspectivas à luz da atualização da Resolução CNJ 332/20**. Consultor Jurídico, São Paulo, 1 mar. 2025. Disponível em: <https://www.conjur.com.br/2025-mar-01/ia-e-a-fundamentacao-das-decisoes-judiciais-desafios-e-perspectivas-a-luz-da-atualizacao-da-resolucao-cnj-332-20/>. Acesso em: 4 out. 2025.

BLACKMAN, Reid. **Máquinas Éticas: seu guia conciso para uma IA totalmente imparcial, transparente e respeitosa**. Rio de Janeiro: Editora Alta Books, 2024. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9788550822426/>. Acesso em: 30 mar. 2025.

BRASIL. **Lei n.º 13.105, de 16 de março de 2015**. Institui o Código de Processo Civil. **Diário Oficial da União**, seção 1, Brasília, DF, 17 mar. 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13105.htm. Acesso em: 20 out. 2025.

BRASIL. **Lei n.º 13.709, de 14 de agosto de 2018**. Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet). **Diário Oficial da União**, seção 1, Brasília, DF, 15 ago. 2018. Disponível em:

https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em: 20 out. 2025.

BRASIL. STF. **Habeas Corpus n.º 74.073/RJ**. Relator: Min. Celso de Mello. Brasília, DF, 1997. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=74968>. Acesso em: 4 out. 2025.

BRASIL. Supremo Tribunal Federal. **Impactos da inteligência artificial no constitucionalismo contemporâneo** [recurso eletrônico]. Brasília: STF, Secretaria de Altos Estudos, Pesquisas e Gestão da Informação, 2024. Disponível em: https://www.stf.jus.br/arquivo/cms/codi/anexo/Impactos_da_IA_no_Constitucional_Contemporaneo_2024_2703251.pdf. Acesso em: 21 set. 2025.

BRASIL. Supremo Tribunal Federal. **Ministra Rosa Weber lança robô Vitória para agrupamento e classificação de processos**. Brasília: STF, 2023. Disponível em: <https://noticias.stf.jus.br/postsnoticias/ministra-rosa-weber-lanca-robo-vitoria-para-agrupamento-e-classificacao-de-processos>. Acesso em: 20 out. 2025.

BRASIL. Supremo Tribunal Federal. **STF lança Maria, ferramenta de inteligência artificial que dará mais agilidade aos serviços do Tribunal**. Brasília: STF, 2025. Disponível em: <https://noticias.stf.jus.br/postsnoticias/stf-lanca-maria-ferramenta-de-inteligencia-artificial-que-dara-mais-agilidade-aos-servicos-do-tribunal>. Acesso em: 20 out. 2025.

BRASIL. Supremo Tribunal Federal. **STF lança RAFA, ferramenta de inteligência artificial para classificar ações na Agenda 2030 da ONU**. Brasília: STF, 2022. Disponível em: <https://noticias.stf.jus.br/postsnoticias/stf-lanca-rafa-ferramenta-de-inteligencia-artificial-para-classificar-acoes-na-agenda-2030-da-onu>. Acesso em: 9 nov. 2025.

CONCIL OF EUROPE. **Carta Europeia de Ética sobre o Uso da Inteligência Artificial em Sistemas Judiciais e seu Ambiente**. Estrasburgo: Comissão Europeia para a Eficiência da Justiça (CEPEJ), 2018. Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0>. Acesso em: 20 out. 2025.

CONSELHO NACIONAL DE JUSTIÇA. **O uso da inteligência artificial generativa no Poder Judiciário brasileiro: relatório de pesquisa**. Brasília: CNJ, 2024. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2024/09/cnj-relatorio-de-pesquisa-iag-pj.pdf>. Acesso em: 6 out. 2025.

CONSELHO NACIONAL DE JUSTIÇA. **Resolução n.º 615, de 11 mar. 2025**. Estabelece diretrizes para o desenvolvimento, utilização e governança de soluções desenvolvidas com recursos de inteligência artificial no Poder Judiciário. **Diário da Justiça Eletrônico do CNJ**, n. 54, p. 2–17, 14 mar. 2025. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/6001>. Acesso em: 18 out. 2025.

MARTINS, Alan Rocha; SÁ, Valdir Rodrigues de. **Inteligência artificial e a decisão judicial: benefícios e riscos à democracia**. In: VIGLIAR, José Marcelo M. (org.). **Inteligência artificial: aspectos jurídicos**. São Paulo: Almedina, 2023. p. 125–148. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9786556279091/>. Acesso em: 3 out. 2025.

NOGUEIRA, Luiz Henrique da Silva; MORAES, Ricardo André Barros de. **As vantagens, riscos e desafios na utilização da inteligência artificial no Poder Judiciário brasileiro**. In: VIGLIAR, José Marcelo M. (org.). **Inteligência artificial: aspectos jurídicos**. São Paulo: Almedina, 2023. p. 89–124. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9786556279091/>. Acesso em: 30 mar. 2025.

PEIXOTO, Fabiano Hartmann. **Projeto Victor: relato do desenvolvimento da Inteligência Artificial na Repercussão Geral do Supremo Tribunal Federal**. *Revista Brasileira de Inteligência Artificial e Direito*, Brasília, v. 1, n. 1, jan./abr. 2020, p. 1–20. ISSN 2675-3156.

RUSSELL, Stuart J.; NORVIG, Peter. **Inteligência Artificial: uma abordagem moderna**. 4. ed. São Paulo: GEN LTC, 2022. E-book. ISBN 9788595159495. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9788595159495/>. Acesso em: 18 out. 2025.

RUSTAMZADE, Aykhan; BAGIROV, Anar Ramiz; HUSEYNOV, Sahil Zahir. **Justice and Artificial Intelligence**. *Revista Jurídica Portucalense*, [s. l.], p. 163–184, 2024. DOI: 10.34625/issn.2183-2705(36)2024.ic-7. Disponível em: <https://revistas.rcaap.pt/juridica/article/view/36997>. Acesso em: 8 set. 2025.

VIGLIAR, José Marcelo M. **Inteligência artificial: aspectos jurídicos**. São Paulo: Almedina, 2023. E-book. ISBN 9786556279091. Disponível em: <https://app.minhabiblioteca.com.br/reader/books/9786556279091/>. Acesso em: 21 ago. 2025.