

REVISTA MUNDO EM MOVIMENTO

Uninassau - Caruaru

2(1): 97-110, 2025

ISSN: 2966-2176

A FERRAMENTA DE DUPLO GUME: SUPERVISÃO HUMANA E VALIDADE DE ATOS JURISDICIONAIS ALGORÍTMICOS

THE DOUBLE-EDGED TOOL: HUMAN OVERSIGHT IN ALGORITHMIC JUDICIAL ACTS

Jean César Leite de Souza – Acadêmico

jeancezarldsouza@gmail.com

Eva de Azevedo Gomes – Advogada e Docente

evagomes100@gmail.com

RESUMO: Introdução: A crescente algoritmização do Poder Judiciário, embora promova eficiência, estabelece um campo de tensão devido ao risco de viés e **opacidade algorítmica** (*black box*), confrontando garantias constitucionais essenciais. O estudo investiga o regime de **responsabilidade jurídica** para "erros algorítmicos" e como garantir a **supervisão humana** para proteger direitos fundamentais. A pesquisa, documental e comparativa, focou na legislação nacional (**LGPD** e **Resolução nº 332/2020 do CNJ**) e internacional (GDPR e EU AI Act). Sustenta-se que a IA é uma **ferramenta de duplo gume**, e sua opacidade compromete a validade constitucional do ato jurisdicional. Conclui-se que a eficácia da IA depende intrinsecamente da transparência e da capacidade de fiscalização, sendo imperativo que o

arcabouço normativo incorpore o direito à explicação e a **revisão humana obrigatória** como garantias processuais essenciais.

PALAVRAS-CHAVE: Inteligência Artificial, Viés Algorítmico, Supervisão Humana, Responsabilidade Estatal, Validade Constitucional.

ABSTRACT: Introduction: The growing algorithmization of the Judiciary, although promoting efficiency, establishes a field of tension due to the risk of bias and **algorithmic opacity** (*black box*), directly confronting essential constitutional guarantees. The study investigates the legal **liability framework** for "algorithmic errors" and how to ensure **human oversight** to protect fundamental rights. The documentary and comparative research focused on national (**LGPD** and **CNJ Resolution nº 332/2020**) and international legislation (GDPR and EU AI Act). It is argued that AI is a **double-edged tool**, and its opacity compromises the constitutional validity of the judicial act. It is concluded that the effectiveness of AI intrinsically depends on transparency and the capacity for oversight, making it imperative that the normative framework incorporates the right to explanation and **mandatory human review** as essential procedural guarantees. **Keywords:** Artificial Intelligence; Algorithmic Bias; Human Oversight; State Responsibility; Constitutional Validity.

KEYWORDS: Artificial Intelligence; Algorithmic Bias; Human Oversight; State Responsibility; Constitutional Validity.

1. INTRODUÇÃO

A incorporação da Inteligência Artificial (IA) no sistema de justiça tem se expandido rapidamente, motivada pelo potencial de otimizar fluxos processuais e reduzir a morosidade judicial. Pesquisas demonstram que

tecnologias de automação podem aprimorar a gestão de acervos (Moura & Oliveira, 2022).

Essa expansão tecnológica é vista como uma modernização necessária para atender à crescente demanda por Justiça. Contudo, essa otimização não está isenta de conflitos com a ordem jurídica vigente. O jurista Richard Susskind (2019) adverte sobre a necessidade de aliar a inovação com a preservação das garantias processuais, estabelecendo a premissa de que a tecnologia deve servir ao Direito, e não o contrário.

O sistema de justiça tende a incorporar tecnologias capazes de acelerar procedimentos, mas essa introdução não pode comprometer direitos constitucionalmente assegurados, sob pena de minar a própria legitimidade da Justiça. O desafio é usar a tecnologia para servir à justiça, e não permitir que a tecnologia determine o que a justiça é." (Susskind, 2019, citado em, adaptado).

Contudo, a utilização de sistemas algoritmizados em decisões judiciais levanta preocupações fundamentais sobre imparcialidade e direitos fundamentais. Como afirma Pasquale (2015), os algoritmos tendem a operar como "caixas-pretas", nas quais nem sempre é possível compreender como decisões são geradas.

A dificuldade em escrutinar a lógica interna dos sistemas complexos de *machine learning* é o que a literatura denomina opacidade algorítmica. Essa característica coloca em xeque o princípio basilar do Direito Público brasileiro, que exige a motivação clara de todos os atos decisórios (art. 93, IX, CF/88).

Além disso, a opacidade é o campo fértil para a proliferação do viés algorítmico. A opacidade se agrava diante do uso de dados historicamente enviesados (O’Neil, 2016).

Cathy O’Neil (2016) desmistifica a neutralidade da matemática ao rotular esses sistemas como *Weapons of Math Destruction* (Armas de Destruição Matemática), mostrando que os modelos, ao serem treinados com dados pretéritos, tendem a reproduzir padrões históricos de exclusão. Em contextos jurídicos, essa replicação pode levar à discriminação algorítmica, o que colide diretamente com os princípios da isonomia (art. 5º, *caput*, CF/88) e do devido processo legal (art. 5º, LIV, CF/88).

Diante deste quadro de risco, o arcabouço normativo se move no sentido de exigir maior *accountability*. A Lei Geral de Proteção de Dados (LGPD), em seu art. 20, garante o direito do titular à revisão de decisões automatizadas, enfatizando a relevância da intervenção humana. De forma complementar, o Regulamento Geral de Proteção de Dados (GDPR) europeu consagra o conceito de “direito à explicação”. No Brasil, o Conselho Nacional de Justiça (CNJ), por meio da Resolução nº 332/2020, já estabeleceu diretrizes éticas e de governança que condicionam o uso da IA à supervisão humana e à indelegabilidade da função judicial, integrando a regulamentação doméstica ao debate internacional do EU AI Act (alto risco).

Portanto, a presente investigação se propõe a debater a IA não apenas como um avanço tecnológico, mas como uma “ferramenta de duplo gume” no Direito. O foco se concentra em analisar os riscos, os desafios regulatórios e,

fundamentalmente, a imprescindibilidade da supervisão humana como requisito inafastável para garantir a validade dos atos jurisdicionais automatizados.

2. A INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA DE DUPLO GUME NO DIREITO

A IA aplicada ao Judiciário opera em uma zona limítrofe entre eficiência e risco. Segundo Susskind (2019), o sistema de justiça tende a incorporar tecnologias capazes de acelerar procedimentos, mas essa introdução não pode comprometer direitos constitucionalmente assegurados.

A metáfora do "duplo gume" reflete a realidade de que a mesma tecnologia que promete desafogar tribunais pode, inadvertidamente, perpetuar ou intensificar injustiças. A otimização processual é o lado da promessa; a discriminação algorítmica é o lado do perigo. A reprodução de vieses ocorre porque os algoritmos são treinados em dados históricos que carregam consigo as marcas de desigualdades sociais preexistentes, especialmente no contexto de um país com profundas disparidades, como o Brasil.

Barocas e Selbst (2016) demonstram que a discriminação pode emergir de forma natural em modelos estatísticos.

De fato, a discriminação algorítmica não depende de intenção maliciosa do programador, mas da própria lógica de minimização de erros estatísticos do sistema. Se um grupo social historicamente marginalizado está super-representado em dados negativos (por exemplo, prisões ou condenações), o algoritmo aprenderá que ser membro desse grupo é uma variável preditiva de risco, perpetuando o ciclo.

A discriminação pode surgir em modelos estatísticos de maneira natural, mesmo quando não há intenção de violar a lei. Em muitos casos, as decisões algorítmicas acabam sendo baseadas em *proxies* para características sensíveis – como raça ou gênero – o que perpetua padrões discriminatórios. A chave é entender que, mesmo se uma característica sensível for removida, outras variáveis correlacionadas irão substituí-la, mantendo o viés de forma sub-reptícia." (Barocas & Selbst, 2016, p. 675, adaptado).

O problema se aprofunda no âmbito judicial. Pesquisas nos EUA revelaram que sistemas como o COMPAS demonstravam vieses raciais na determinação de risco de reincidência (Angwin et al., 2016).

O caso do sistema COMPAS, que erroneamente classificava réus negros como sendo de maior risco de reincidência do que réus brancos com histórico criminal similar, tornou-se um marco global na discussão sobre a falibilidade e o perigo da IA na Justiça. Esse estudo de caso reforça o argumento de que a mera aplicação estatística do passado não pode ser confundida com Justiça e que, no Brasil, a adoção de tecnologias similares deve ser tratada como uma questão de alta periculosidade social.

Embora o Brasil não use exatamente o mesmo sistema, o risco potencial demonstra a urgência de auditoria. No contexto nacional, qualquer sistema que utilize a IA para prever risco, auxiliar na triagem de processos ou subsidiar decisões de mérito deve ser submetido a um rigoroso escrutínio para que não viole a garantia de tratamento igual perante a lei, conforme exigido pela Constituição Federal. A urgência da auditoria, portanto, é um reflexo direto do dever estatal de garantir a isonomia processual e material.

2.1. OPACIDADE ALGORÍTMICA E A NECESSIDADE DE TRANSPARÊNCIA

A literatura classifica algoritmos complexos de *machine learning* como opacos. Essa opacidade não é apenas uma barreira técnica, mas um impedimento ao exercício pleno da cidadania e da fiscalização jurisdicional. A dificuldade em escrutinar a lógica interna dos sistemas complexos de *machine learning* é o que a literatura denomina opacidade algorítmica.

Essa opacidade contraria a exigência de motivação e fundamentação das decisões judiciais (art. 93, IX, CF/88). O Artigo 93, IX, da Constituição Federal estabelece uma garantia pétrea do devido processo legal: todos os julgamentos dos órgãos do Poder Judiciário serão públicos, e fundamentadas todas as decisões, sob pena de nulidade. A aplicação de um algoritmo que não permite rastreabilidade, explicabilidade ou justificação do raciocínio subjacente, por mais eficiente que seja, anula essa garantia.

Todos os julgamentos dos órgãos do Poder Judiciário serão públicos, e fundamentadas todas as decisões, sob pena de nulidade, podendo a lei limitar a presença, em determinados atos, às próprias partes e a seus advogados, ou somente a estes, em casos nos quais a preservação do direito à intimidade do interessado no sigilo não prejudique o interesse público à informação." (Constituição Federal, Art. 93, IX, de 1988).

A violação não é apenas formal. A falta de transparência impede o controle de constitucionalidade e legalidade, tornando impossível verificar se houve a aplicação de viés racial, de gênero ou socioeconômico, contrariando o art. 5º da Carta Magna. A transparência, nesse contexto, deve ser entendida como o direito à explicabilidade (direito à explicação), permitindo que o indivíduo afetado compreenda os fatores que levaram o algoritmo a sugerir ou tomar determinada decisão.

A LGPD prevê, no art. 20, o direito à revisão, mas a lei não regula especificamente a utilização da IA pelo Estado no âmbito judicial. Para Wachter e Mittelstadt (2019), a transparência é condição para a legitimidade. Assim, o desafio regulatório brasileiro é preencher essa lacuna. A transparência exige que o código e os dados de treinamento utilizados sejam auditáveis por órgãos independentes ou pelo próprio sistema de Justiça, de modo a garantir que a legitimidade do ato judicial não seja perdida para a complexidade tecnológica.

2.2. O REGIME DE RESPONSABILIDADE JURÍDICA DOS ERROS ALGORÍTMICOS

A definição de responsabilidade é um desafio. Doneda e Almeida (2021) sugerem que a responsabilidade deve ser distribuída.

O uso da IA pelo Poder Judiciário, sendo uma atividade tipicamente estatal, submete-se ao regime da responsabilidade civil objetiva prevista na Constituição Federal. O erro, o viés ou a discriminação proveniente de um algoritmo que auxilia (ou determina) um ato jurisdicional pode ser enquadrado como uma falha na prestação do serviço público, independentemente da prova de culpa dos agentes envolvidos.

As pessoas jurídicas de direito público e as de direito privado prestadoras de serviços públicos responderão pelos danos que seus agentes, nessa qualidade, causarem a terceiros, assegurado o direito de regresso contra o responsável nos casos de dolo ou culpa." (Constituição Federal, Art. 37, §6º, de 1988).

A ausência de supervisão humana constitui falha grave, podendo configurar responsabilidade civil objetiva do Estado (CF/88, art. 37, §6º). A falha ocorre não apenas quando o algoritmo produz um resultado errôneo, mas quando o Estado não institui os mecanismos de controle adequados para prevenir e corrigir o viés e a opacidade. A Resolução nº 332/2020 do CNJ reforça esse dever de cuidado, ao exigir a manutenção da intervenção humana e a observância de princípios éticos, tornando a violação de tais normas um indicativo de falha na prestação do serviço público judicial. O erro algorítmico, nesse sentido, é transformado em um dano jurídico por omissão na fiscalização e na garantia do devido processo legal.

A doutrina brasileira aponta que a atribuição da responsabilidade se torna complexa devido à multiplicidade de atores. Doneda e Almeida (2021) defendem o modelo de responsabilidade distribuída, onde o foco não recai em identificar um único culpado, mas em garantir a reparação do dano e a

accountability (responsabilidade e prestação de contas por ações, decisões e resultados, tanto em nível individual quanto organizacional e governamental) do sistema como um todo. A classificação europeia de sistemas judiciais como de “alto risco” reforça a tese de que o risco inerente à IA na justiça exige um padrão de cuidado mais rigoroso por parte do Estado. O debate, portanto, migra da culpa individual para a falha sistêmica.

2.3 SUPERVISÃO HUMANA COMO CONDIÇÃO DE VALIDADE DOS ATOS JURISDICIONAIS AUTOMATIZADOS

A supervisão humana não é apenas recomendável: ela é essencial.

A tese central do trabalho sustenta que a intervenção e a revisão humana não são medidas de aprimoramento, mas sim requisitos inafastáveis para a validade constitucional dos atos judiciais assistidos por IA. O objetivo principal da supervisão é garantir que o juízo final permaneça sob a égide da discricionariedade, da imparcialidade e da fundamentação humana. O algoritmo não pode substituir o juiz, mas apenas auxiliá-lo, pois a função judicial é indelegável. A necessidade de supervisão é formalizada pela Resolução nº 332/2020 do CNJ, que estabelece o uso ético da IA e, crucialmente, proíbe a substituição da atividade judicial por sistemas automatizados.

O conceito de *due process* algorítmico exige que, mesmo com o uso de software para suporte decisório, os padrões de legalidade,

verificação e justificação sejam rigorosamente mantidos. A opacidade algorítmica (*black box*) é um dos maiores impeditivos à conformidade constitucional. Em consonância com o princípio da indelegabilidade, a Resolução nº 332/2020 do CNJ estabelece que o uso da inteligência artificial não pode resultar na substituição de magistrados e servidores no exercício de suas atribuições e deve sempre garantir a intervenção humana para revisão das decisões, reforçando o caráter instrumental da tecnologia." (Citron, 2014, adaptado, e CNJ, Res. 332/2020, adaptado).

Pasquale (2020) afirma que o algoritmo não pode substituir o juiz. A ausência de revisão humana, portanto, viola diretamente o devido processo legal e o contraditório, pois impede o exercício pleno do direito de defesa e de questionamento dos parâmetros aplicados.

Dessa forma, a supervisão humana deve ser constante, qualificada e instituída como requisito normativo. A mera ratificação *pro forma* de uma recomendação algorítmica pelo magistrado, sem escrutínio real da lógica subjacente ou dos riscos de viés, não satisfaz o dever de fundamentação e configura a falha do serviço público. A revisão humana, já prevista pelo CNJ e reforçada como um direito legal para sistemas de alto risco no Marco Legal da IA pós-2024, torna-se um mandato legal com o objetivo explícito de prevenir e minimizar riscos, exigindo que os supervisores possam intervir no sistema.

3. CONSIDERAÇÕES FINAIS

A utilização da IA no Judiciário é um avanço, no entanto, a ausência de supervisão traz riscos graves que devem ser considerados. A pesquisa

demonstra que a Inteligência Artificial, embora otimize a celeridade processual, traz consigo um risco inerente de viés e opacidade que pode comprometer a validade constitucional dos atos jurisdicionais. Este risco deriva da replicação de dados históricos enviesados e da consequente ameaça aos princípios da isonomia e da fundamentação das decisões.

O estudo evidencia que algoritmos podem reproduzir desigualdades. A responsabilidade por eventuais "erros algorítmicos" recai sobre o Estado, sob o regime da responsabilidade civil objetiva (Art. 37, §6º, CF/88), especialmente quando há omissão na implementação de mecanismos de fiscalização humana. A opacidade e o viés algorítmico, ao impedirem a rastreabilidade do processo decisório, configuram falha sistêmica na prestação do serviço público.

Conclui-se que a IA não deve ser proibida, mas sua adoção exige forte comprometimento com princípios constitucionais. O desafio imposto pela IA, como ferramenta de duplo gume, não reside em sua proibição, mas na redefinição do papel do Direito e do agente decisor frente à ferramenta. O arcabouço normativo, incluindo a Resolução 332/2020 do CNJ, deve urgentemente incorporar o direito à explicação e a revisão humana obrigatória como garantias processuais essenciais para assegurar a justiça e a proteção dos direitos fundamentais contra o viés algorítmico.

A supervisão humana é condição de validade jurídica. A intervenção humana qualificada deve ser instituída como um mandato legal, garantindo que a tecnologia seja sempre um instrumento de apoio, e não uma autoridade

decisória final. A validade do ato jurisdicional algorítmico está, portanto, condicionada ao atendimento de três garantias inegociáveis para sistemas de alto risco: Direito à Explicação, Direito à Contestação e Revisão Humana Obrigatória.

REFERÊNCIAS

ANGWIN, J.; LARSON, J.; MATTU, S.; KIRCHNER, L. (2016). Machine Bias. **ProPublica**. Disponível em: <https://www.propublica.org/article/machine-bias> Acesso em: 21 nov. 2025.

BAROCAS, S.; SELBST, A. (2016). Big Data's Disparate Impact. **California Law Review**, 104(3), 671-732.

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

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018** (Lei Geral de Proteção de Dados Pessoais - LGPD). Brasília, DF: Presidência da República.

CITRON, D. K. (2014). Technological due process. **Washington University Law Review**, 85(6).

CONSELHO NACIONAL DE JUSTIÇA (CNJ). **Resolução nº 332, de 21 de agosto de 2020**. Dispõe sobre a política judiciária nacional de tratamento de dados e uso de inteligência artificial no Poder Judiciário. Brasília, DF: CNJ. DONEDA, D.; ALMEIDA, V. (2021). **Governança de Algoritmos e Proteção de Dados**. Rio de Janeiro: FGV. Disponível em: <https://bibliotecadigital.fgv.br> Acesso em: 23 nov. 2025.

EUROPEAN COMMISSION. (2021). **Proposal for an Artificial Intelligence Act (EU AI Act)**. Bruxelas: European Commission. Disponível em: <https://artificialintelligenceact.eu>. Acesso em: 23 nov. 2025.

MOURA, P.; OLIVEIRA, D. (2022). *Justiça Digital e Inovação Tecnológica no Judiciário Brasileiro*. São Paulo: RT.

O'NEIL, C. (2016). **Weapons of Math Destruction**. New York: Crown. Disponível em: <https://weaponsofmathdestructionbook.com> Acesso em: 23 nov. 2025.

PASQUALE, F. (2015). **The Black Box Society**. Harvard University Press. Disponível em: <https://www.hup.harvard.edu/books/9780674970847> Acesso em: 16 nov. 2025.

PASQUALE, F. (2020). **New Laws of Robotics**. Harvard University Press. Disponível em: <https://www.hup.harvard.edu/books/9780674975224> Acesso em: 20 nov. 2025.

SUSSKIND, R. (2019). **Online Courts and the Future of Justice**. Oxford University Press. Disponível em: <https://global.oup.com/academic/product/online-courts-and-the-future-of-justice> Acesso em: 20 nov. 2025.

WACHTER, S.; MITTELSTADT, B.; FLORIDI, L. (2017). Why a right to explanation of automated decision-making does not exist. **International Data Privacy Law**. Disponível em: <https://academic.oup.com/idpl/article/7/2/76/3860948> Acesso em: 22 nov. 2025.