

IX ENCONTRO VIRTUAL DO CONPEDI

**DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS
III**

Todos os direitos reservados e protegidos. Nenhuma parte destes anais poderá ser reproduzida ou transmitida sejam quais forem os meios empregados sem prévia autorização dos editores.

Diretoria - CONPEDI

Presidente - Profa. Dra. Samyra Haydêe Dal Farra Naspolini - FMU - São Paulo

Diretor Executivo - Prof. Dr. Orides Mezzaroba - UFSC - Santa Catarina

Vice-presidente Norte - Prof. Dr. Jean Carlos Dias - Cesupa - Pará

Vice-presidente Centro-Oeste - Prof. Dr. José Querino Tavares Neto - UFG - Goiás

Vice-presidente Sul - Prof. Dr. Leonel Severo Rocha - Unisinos - Rio Grande do Sul

Vice-presidente Sudeste - Profa. Dra. Rosângela Lunardelli Cavallazzi - UFRJ/PUCRio - Rio de Janeiro

Vice-presidente Nordeste - Prof. Dr. Raymundo Juliano Feitosa - UNICAP - Pernambuco

Representante Discente: Prof. Dr. Abner da Silva Jaques - UPM/UNIGRAN - Mato Grosso do Sul

Conselho Fiscal:

Prof. Dr. José Filomeno de Moraes Filho - UFMA - Maranhão

Prof. Dr. Caio Augusto Souza Lara - SKEMA/ESDHC/UFMG - Minas Gerais

Prof. Dr. Valter Moura do Carmo - UFERSA - Rio Grande do Norte

Prof. Dr. Fernando Passos - UNIARA - São Paulo

Prof. Dr. Edinilson Donisete Machado - UNIVEM/UENP - São Paulo

Secretarias

Relações Institucionais:

Prof. Dra. Claudia Maria Barbosa - PUCPR - Paraná

Prof. Dr. Heron José de Santana Gordilho - UFBA - Bahia

Profa. Dra. Daniela Marques de Moraes - UNB - Distrito Federal

Comunicação:

Prof. Dr. Robison Tramontina - UNOESC - Santa Catarina

Prof. Dr. Liton Lanes Pilau Sobrinho - UPF/Univali - Rio Grande do Sul

Prof. Dr. Lucas Gonçalves da Silva - UFS - Sergipe

Relações Internacionais para o Continente Americano:

Prof. Dr. Jerônimo Siqueira Tybusch - UFSM - Rio Grande do Sul

Prof. Dr. Paulo Roberto Barbosa Ramos - UFMA - Maranhão

Prof. Dr. Felipe Chiarello de Souza Pinto - UPM - São Paulo

Relações Internacionais para os demais Continentes:

Profa. Dra. Gina Vidal Marcilio Pompeu - UNIFOR - Ceará

Profa. Dra. Sandra Regina Martini - UNIRITTER / UFRGS - Rio Grande do Sul

Profa. Dra. Maria Claudia da Silva Antunes de Souza - UNIVALI - Santa Catarina

Educação Jurídica

Profa. Dra. Viviane Coêlho de Séllos Knoerr - Unicuritiba - PR

Prof. Dr. Rubens Beçak - USP - SP

Profa. Dra. Livia Gaigher Bosio Campello - UFMS - MS

Eventos:

Prof. Dr. Yuri Nathan da Costa Lannes - FDF - São Paulo

Profa. Dra. Norma Sueli Padilha - UFSC - Santa Catarina

Prof. Dr. Juraci Mourão Lopes Filho - UNICHRISTUS - Ceará

Comissão Especial

Prof. Dr. João Marcelo de Lima Assafim - UFRJ - RJ

Profa. Dra. Maria Creusa De Araújo Borges - UFPB - PB

Prof. Dr. Antônio Carlos Diniz Murta - Fumec - MG

Prof. Dr. Rogério Borba - UNIFACVEST - SC

D597

Direito, governança e novas tecnologias III [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Luiz Alberto Pereira Ribeiro; Regina Vera Villas Boas; Paulo Campanha Santana – Florianópolis: CONPEDI, 2026.

Inclui bibliografia

ISBN: 978-65-5274-590-3

Modo de acesso: www.conpedi.org.br em publicações

Tema: Constituição, Processo e Justiça Social: o papel da jurisdição constitucional na efetivação de direitos

2. Direito. 3. Governança e novas tecnologias. IX Encontro Virtual do CONPEDI (2: 2026: Florianópolis, Brasil).

CDU: 34



Conselho Nacional de Pesquisa
e Pós-Graduação em Direito Florianópolis
Santa Catarina – Brasil
www.conpedi.org.br

IX ENCONTRO VIRTUAL DO CONPEDI

DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS III

Apresentação

Texto de apresentação do livro do GT 52 - DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS III

O “IX Encontro Virtual do CONPEDI – Sociedade Científica do Direito no Brasil” foi realizado de forma virtual entre os dias 22 e 26 de junho de 2026. Este evento exemplar foi um marco de excelência acadêmica e colaboração científica, reunindo pesquisadores e estudiosos de diversas áreas do Direito.

Destacamos especialmente o Grupo de Trabalho – GT 52, intitulado “Direito, Governança e Novas Tecnologias III”, que se destacou pela profundidade e relevância dos temas abordados. Sob a coordenação dos Professores- Drs. Luiz Alberto P. Ribeiro (UEL e PUC/PR), Paulo Campanha (Mackenzie/Brasília) e Regina Vera Villas Bôas (PUC/SP), o GT proporcionou um espaço privilegiado para a discussão de questões fundamentais no campo do direito, governança e das novas tecnologias.

Neste GT foram apresentados trabalhos de elevada qualidade e importância crítica, sob os seguintes títulos:

1 MODERAÇÃO DE CONTEÚDO, GOVERNANÇA DE PLATAFORMAS DIGITAIS E DIREITOS FUNDAMENTAIS: PARÂMETROS CONSTITUCIONAIS PARA O BRASIL.

2 MERCADOS DE CARBONO E O NOVO SISTEMA BRASILEIRO DE COMÉRCIO DE EMISSÕES (SBCE).

6 GOVERNANÇA E REGULAÇÃO DA INFRAESTRUTURA PARA VEÍCULOS ELÉTRICOS EM CONDOMÍNIOS NO BRASIL: RISCOS, RESPONSABILIDADE CIVIL E DESIGUALDADE DE ACESSO

7 ENTRE O VIÉS HUMANO E O ALGORÍTMICO: DESAFIOS DA RACIONALIDADE PROCESSUAL NA ERA DA INTELIGÊNCIA ARTIFICIAL.

8 ENTRE A CELERIDADE E A EFICÁCIA: OS DESAFIOS DA AUTOMAÇÃO NO INSTITUTO NACIONAL DO SEGURO SOCIAL.

9 EMPRESAS TRANSNACIONAIS, RELAÇÕES DE TRABALHO E GOVERNANÇA DIGITAL NAS CONTRATAÇÕES PÚBLICAS: ANÁLISE CRÍTICA À LUZ DA LEI Nº 14.133/2021.

10 ECONOMIA DOS DADOS E GOVERNANÇA ALGORÍTMICA: PERFILAMENTO E VULNERABILIDADE DIGITAL DO CONSUMIDOR NA SOCIEDADE INFORMACIONAL.

11 DEMÔNIO DE LAPLACE, PREVISIBILIDADE E CONTROLE: PROBLEMÁTICAS EPISTEMOLÓGICAS, ÉTICAS E DE JUSTIÇA NAS DECISÕES JURÍDICO ADMINISTRATIVAS BASEADAS EM BIG DATA, ALGORITMOS E INTELIGÊNCIA ARTIFICIAL.

12 DIREITO FUNDAMENTAL À PROTEÇÃO DE DADOS E A TUTELA ESPECIAL NO TRATAMENTO DOS DADOS SENSÍVEIS: DA ESFERA DA PRIVACIDADE À DIGNIDADE DA PESSOA HUMANA E AUTODETERMINAÇÃO INFORMATIVA.

13 DIREITO PROCESSUAL E HOMOTRANSFOBIA ALGORÍTMICA: POR UM DEVER DE CUIDADO JURISDICIONALIZADO.

16 AUTOMAÇÃO DAS DECISÕES JUDICIAIS E OS LIMITES CONSTITUCIONAIS DA INTELIGÊNCIA ARTIFICIAL NO ESTADO DEMOCRÁTICO DE DIREITO.

17 AUTORIA E ORIGINALIDADE NA ERA DA INTELIGÊNCIA ARTIFICIAL: DESAFIOS JURÍDICOS DA PROTEÇÃO AUTORAL DIANTE DAS CRIAÇÕES AUTOMATIZADAS E DA AUTOPUBLICAÇÃO DIGITAL.

18 CAPITALISMO DE VIGILÂNCIA E DIREITOS FUNDAMENTAIS SOCIAIS: OS LIMITES JURÍDICOS DA EXPLORAÇÃO ECONÔMICA DE DADOS PESSOAIS NO BRASIL.

19 CAPITALISMO DE VIGILÂNCIA E HIPERCONSUMO: REFLEXÕES SOBRE A CRISE DA AUTONOMIA NA SOCIEDADE DIGITAL.

20 ADVOCACIA PÚBLICA E GOVERNANÇA INFORMACIONAL NA ADMINISTRAÇÃO PÚBLICA: PROTEÇÃO DE DADOS, POLÍTICAS PÚBLICAS E CONTROLE DE JURIDICIDADE.

A qualidade dos trabalhos apresentados neste GT 52 foi notável, refletindo o compromisso dos participantes com a pesquisa de alto nível e a inovação acadêmica. As contribuições dos estudiosos trouxeram insights significativos e promoveram um debate enriquecedor sobre os desafios contemporâneos e as perspectivas futuras nessas áreas cruciais que envolvem o Direito.

O IX Encontro Virtual do CONPEDI não apenas consolidou seu papel como um canal de referência no cenário acadêmico nacional e internacional, como também reafirmou o compromisso com a qualidade científica e a excelência na produção do conhecimento jurídico.

Os coordenadores do GT 52 deixam aqui abraços fraternos e cumprimentos a todos os participantes do CONPEDI e aos leitores dos presentes textos.

Prof. Dr. Luiz Alberto P. Ribeiro (UEL e PUC/PR)

Prof. Dr. Paulo Campanha (Mackenzie/Brasília)

Profa. Dra. Regina Vera Villas Bôas (PUC/SP)

ENTRE O VIÉS HUMANO E O ALGORÍTMICO: DESAFIOS DA RACIONALIDADE PROCESSUAL NA ERA DA INTELIGÊNCIA ARTIFICIAL

BETWEEN HUMAN AND ALGORITHMIC BIAS: CHALLENGES OF PROCEDURAL RATIONALITY IN THE AGE OF ARTIFICIAL INTELLIGENCE

Marcella Carneiro Holanda ¹

Lucas Antunes Santos ²

Resumo

O artigo analisa a crescente incorporação de sistemas de inteligência artificial no Poder Judiciário, impulsionada pela busca por eficiência diante da alta litigiosidade e da necessidade de maior celeridade processual. Problematisa-se o fenômeno da dupla opacidade, caracterizado pela interação entre limitações cognitivas humanas e a falta de transparência algorítmica, circunstância que gera riscos à fundamentação das decisões judiciais e à legitimidade democrática da jurisdição. O objetivo consiste em investigar meios capazes de garantir a racionalidade e a legitimidade da atividade jurisdicional assistida por tecnologia, especialmente diante da tensão entre inovação tecnológica, devido processo legal e proteção dos direitos fundamentais. A metodologia emprega uma abordagem transdisciplinar, articulando conceitos da Teoria do Direito, Neurociência, Psicologia Cognitiva e Tecnologia. Os resultados indicam que a utilização acrítica de algoritmos pode reproduzir desigualdades estruturais sob uma aparência de neutralidade técnica e criar uma autoridade ilegítima. Conclui-se pela necessidade de um modelo de governança híbrida e de uma racionalidade processual tecnodemocrática, na qual a tecnologia atue como suporte sob estrita supervisão humana e auditoria contínua, assegurando a subordinação da lógica algorítmica aos direitos fundamentais e aos princípios democráticos que orientam a prestação jurisdicional.

Palavras-chave: Inteligência artificial, Legitimidade democrática, Poder judiciário, Racionalidade processual, Viés algorítmico

Abstract/Resumen/Résumé

decisions and for the democratic legitimacy of jurisdiction. The objective is to investigate mechanisms capable of ensuring the rationality and legitimacy of technology-assisted judicial activity, especially in light of the tension between technological innovation, due process of law, and the protection of fundamental rights. The methodology adopts a transdisciplinary approach, articulating concepts from Legal Theory, Neuroscience, Cognitive Psychology, and Technology. The findings indicate that the uncritical use of algorithms may reproduce structural inequalities under an appearance of technical neutrality and create an illegitimate authority. The article concludes that there is a need for a hybrid governance model and a techno-democratic procedural rationality, in which technology operates as a support tool under strict human supervision and continuous auditing, ensuring the subordination of algorithmic logic to fundamental rights and to the democratic principles that guide judicial activity.

Keywords/Palabras-claves/Mots-clés: Artificial intelligence, Democratic legitimacy, Judiciary, Procedural rationality, Algorithmic bias

1 Introdução

A promessa de uma justiça mais célere, eficiente e objetiva tem impulsionado a crescente incorporação de sistemas de Inteligência Artificial (IA) e aprendizado de máquina (*machine learning*) no Poder Judiciário global, inclusive, o brasileiro. Contudo, a transição do julgamento humano para o auxílio ou a delegação algorítmica não elimina a falibilidade da decisão, pois ela transmuta a natureza do erro e obscurece suas origens.

O avanço tecnológico ocorre em um cenário de litigiosidade crescente, segundo o relatório Justiça em Números 2025, o Judiciário brasileiro finalizou 2024 com 80,6 milhões de processos pendentes, mesmo após registrar produtividade recorde de 44,8 milhões de processos baixados. Os números evidenciam a pressão institucional por soluções que aumentem a eficiência e reduzam o acúmulo processual, reforçando a busca por automação e sistemas inteligentes.

Esse fenômeno jurídico contemporâneo encontra-se, portanto, diante de um impasse fundamental, qual seja, conciliar a racionalidade processual e a legitimidade democrática das decisões judiciais, considerando a tensão entre vieses pelos algoritmos e a necessidade de transparência e fundamentação das decisões nas democracias contemporâneas.

Nesse contexto, a própria política de transformação digital do Judiciário revelou essa tendência, conforme a Pesquisa CNJ/PNUD (2025) sobre inteligência artificial no Judiciário 2024, o programa Juízo 100% Digital já abrange 79,6% das unidades judiciárias, e há atualmente 397 Núcleos de Justiça 4.0, voltados à tramitação eletrônica e ao uso intensivo de dados.

No entanto, tais iniciativas não apenas modernizam a gestão processual, mas também criam um terreno institucional propício para a expansão de sistemas baseados em inteligência artificial, inserindo a tecnologia no centro da atividade jurisdicional.

Diante dessa crescente dependência de ferramentas tecnológicas, o presente artigo parte da premissa de que a decisão judicial na era da IA enfrenta o risco de uma “dupla opacidade”.

A primeira camada dessa opacidade é cognitiva e humana. Ao contrário do ideal formalista de um juiz puramente racional que aplica a lei aos fatos de maneira mecânica, a psicologia cognitiva e a economia comportamental, notadamente os estudos de Kahneman e Tversky (1974) demonstram que o julgamento humano opera sob uma racionalidade limitada. Magistrados, enquanto seres humanos, são suscetíveis a heurísticas (atalhos mentais) e vieses cognitivos que distorcem a avaliação das provas e a aplicação do direito, além de sofrerem a

influência de fatores extralegais, como o cansaço e a variação de humor, que geram "ruído" e inconsistência nas sentenças.

A segunda camada de opacidade é computacional e algorítmica. A introdução de algoritmos promete mitigar o ruído humano através da padronização, mas frequentemente introduz novas formas de discriminação e falta de transparência. Sistemas de *data mining* e aprendizado de máquina não são neutros; eles "aprendem" com dados históricos que podem refletir preconceitos passados ou desigualdades estruturais, cristalizando discriminações sob uma aparência de objetividade técnica.

Além disso, muitos desses sistemas operam como "caixas-pretas" (*black boxes*), onde a correlação entre os dados de entrada e o resultado final é complexa demais para ser explicada em termos de causalidade humana ou jurídica, desafiando o imperativo constitucional da motivação das decisões.

A relevância desse problema é ampliada pelo fato de que o uso de IA no Judiciário brasileiro já é significativo, a pesquisa de 2024 indica que 45,8% dos tribunais utilizam ferramentas de IA generativa e que somente no ano de 2024 foram desenvolvidos 98 novos projetos de inteligência artificial. Portanto, discutir vieses algorítmicos e legitimidade democrática é analisar uma realidade plenamente instalada no sistema judicial.

O contexto de adoção dessas tecnologias é de expansão acelerada. Nos Estados Unidos, ferramentas como o COMPAS são utilizadas para prever reincidência criminal e auxiliar na dosimetria da pena, embora estudos demonstrem que tais sistemas podem não ser tão precisos e apresentem disparidades raciais significativas, punindo desproporcionalmente réus negros.

Diante deste cenário, o objetivo deste artigo é cruzar a teoria democrática da decisão, que exige que o provimento jurisdicional seja fruto de um discurso racional, justificável e comunicativamente válido, na esteira de Habermas (1997), com os riscos ético-técnicos da IA.

Para tanto, verifica-se na literatura crítica sobre algoritmos de Barocas e Selbst (2016) um alerta para o impacto desproporcional em grupos vulneráveis, a falácia da neutralidade matemática, afirmada por O'Neil (2016), e a necessidade de mecanismos de *accountability* que vão além da simples transparência do código-fonte, buscando uma regularidade processual auditável, conforme defendido por Kroll et al. (2017).

Nessa perspectiva, a pergunta central que orienta esta pesquisa é como garantir a racionalidade e a legitimidade das decisões judiciais quando assistidas por inteligência artificial diante da tensão entre vieses humanos, vieses algorítmicos e a exigência democrática de transparência, motivação e controle?

O que se pretende demonstrar é que a legitimidade da jurisdição algorítmica não reside na substituição do juiz pela máquina, nem na rejeição ludista da tecnologia, mas na implementação de um sistema de governança híbrida.

Este sistema deve integrar a capacidade da IA de reduzir o "ruído" humano com a reserva humana indispensável para a valoração ética e a justificação hermenêutica, assegurando que a tecnologia atue sob estrita supervisão e que seus critérios sejam traduzíveis em uma linguagem acessível ao debate público e ao controle jurisdicional.

Desse modo, para responder à pergunta, adota-se uma abordagem transdisciplinar, articulando conceitos da Teoria do Direito, da Neurociência e da Tecnologia. A análise se divide em três etapas subsequentes. A primeira examina as limitações da racionalidade judicial humana (vieses e ruídos); a segunda analisa a natureza dos vieses algorítmicos e a opacidade técnica; e a terceira propõe caminhos para uma integração responsável que preserve o devido processo legal e a fundamentação das decisões.

2 Racionalidade processual democrática e limites da decisão judicial

A compreensão da decisão judicial na contemporaneidade exige o reconhecimento de que o direito moderno opera sob uma tensão constitutiva entre a facticidade da sua imposição e a validade de sua legitimação.

Por essa razão, a racionalidade processual, não se exaure na lógica formal ou na subsunção automatizada, pelo contrário, ela demanda um agir comunicativo capaz de transmutar o poder administrativo em autoridade legítima através do discurso. Ainda assim, a despeito das dificuldades, a inserção de tecnologias no judiciário deve ser escrutinada sob a ótica dessa complexa teia de validação normativa, integridade interpretativa e limitações cognitivas humanas.

Embora exista uma tensão entre facticidade e validade como paradigma para o discurso, é crucial observar que a legitimidade da decisão judicial em um Estado Democrático de Direito não repousa apenas na legalidade formal, mas também, depende da qualidade do processo comunicativo que a antecede. Nesse sentido, Jürgen Habermas (1997) elucida que, no direito moderno, a facticidade da imposição do direito pelo Estado interliga-se com a força de um processo de normatização do direito, que tem a pretensão de ser racional, por garantir a liberdade e fundar a legitimidade.

Essa tensão é vital, pois o direito não opera apenas como um sistema de coerção, mas depende, essencialmente, da suposição de legitimidade que fundamenta a obediência a suas normas. Nesse sentido, o processo judicial deve ser compreendido como um espaço

institucionalizado de discursos de aplicação. Pois, ao contrário de uma visão solipsista da decisão, a teoria do discurso pressupõe que a racionalidade da jurisdição depende da satisfação de condições comunicativas. Conforme observa Habermas (1997, p. 286):

No discurso racional, nós supomos condições comunicativas as quais: previnem uma quebra não motivada da argumentação; garantem a liberdade da escolha de temas e a inclusão das melhores informações e argumentos, através do acesso universal à argumentação e da participação simétrica nela, com igualdade de chances; neutralizam qualquer coerção que possa influir sobre o processo de entendimento a partir de fora ou que decorra desse mesmo entendimento. A única “coerção” permitida é a obrigação da busca cooperativa da verdade.

Ou seja, os discursos jurídicos não podem se mover auto-suficientemente num universo hermeticamente fechado do direito vigente, ele precisa manter-se aberto a argumentos de outras procedências, especialmente a argumentos pragmáticos, éticos e morais.

Portanto, a racionalidade processual democrática exige que a decisão seja fruto de um debate em que a força do melhor argumento prevaleça. Diante disso, a tecnologia, ao intervir nesse cenário, não pode suprimir a dimensão pragmática do processo de fundamentação, sob pena de erodir a aceitabilidade racional das decisões judiciais. Afinal, para Habermas (1997, p. 32 e 28) a validade do direito não é um dado estático, mas epistêmico ("a validade (*Gültigkeit*) tem de ser entendida epistemicamente como validade que se mostra para nós (*Geltung*)").

No que tange à estrutura da argumentação e a ponderação de princípios, é fundamental observar que para operacionalizar essa racionalidade discursiva no âmbito da decisão judicial, Robert Alexy (2008) fornece uma estrutura dogmática a qual fundamenta que a decisão judicial não opera apenas no nível das regras (mandamentos definitivos), mas essencialmente no nível dos princípios.

Seguindo essa lógica, Alexy (2008) estabelece que princípios são normas que ordenam que algo seja realizado na maior medida possível dentro das possibilidades jurídicas e fáticas existentes, configurando-se, portanto, mandamentos de otimização.

Diante de um cenário de colisão de direitos fundamentais, inerente aos casos complexos, a decisão não se resolve pela invalidade de uma norma, como ocorreria no conflito de regras, mas pelo sopesamento. Dessa forma, a racionalidade do processo é garantida pela lei de colisão e pela lei do sopesamento. Pois, conforme Alexy (2008, p. 593) "Quanto maior for o grau de não-satisfação ou de afetação de um princípio, tanto maior terá que ser a importância da satisfação do outro".

Sob essa ótica, a racionalidade processual, depende, inevitavelmente, da capacidade do julgador de justificar a prevalência de um princípio sobre outro em condições concretas.

Afinal, a discricionariedade judicial não é arbitrária, mas estrutural e epistêmica, e surge quando é incerta a cognição daquilo que é obrigatório, proibido ou facultado em virtude dos direitos fundamentais. Portanto, é justamente o controle racional dessa incerteza que legitima a jurisdição constitucional.

Contudo, é preciso destacar que a fundamentação judicial exige mais do que o sopesamento, ela requer também, integridade. Habermas (1997) ao analisar a obra de Ronald Dworkin, destaca a pretensão de que a decisão judicial satisfaça simultaneamente a segurança jurídica e a aceitabilidade racional. Conforme o autor, Dworkin introduz a ideia de uma interpretação construtiva, onde o juiz (idealizado na figura de Hércules) deve reconstruir racionalmente o direito vigente para encontrar a resposta que melhor se justifique à luz da moral política da comunidade.

Seguindo a interpretação de Habermas (1997) sobre a teoria dworkiniana, uma decisão jurídica de um caso particular só é correta, quando se encaixa num sistema jurídico coerente. A propósito, a integridade impede que o direito seja tratado como um agregado de decisões isoladas ou discricionárias, cabendo ao juiz, portanto, justificar uma ordem jurídica concreta em seus elementos essenciais, de tal modo que nela se encaixem todas as decisões tomadas em casos singulares, como se fossem componentes coerentes.

Dessa forma, a tecnologia aplicada ao judiciário deve, portanto, ser capaz de preservar essa coerência narrativa e principiológica, evitando a fragmentação do direito em meros dados estatísticos desconexos da história institucional.

Importa ressaltar que, não obstante o ideal normativo de racionalidade discursiva (Habermas) e da pretensão de correção (Alexy, Dworkin), a decisão judicial é um ato humano, sujeito a falhas cognitivas.

E essa construção teórica da racionalidade jurídica encontra limites quando confrontada com as condições reais do julgamento humano. De acordo com Marden e Wykrota (2018) o modelo tradicional de decisão, baseado na figura de um julgador capaz de depurar emoções, controlar vieses e operar por pura lógica, cria uma “zona de risco hermenêutica” ao pressupor uma racionalidade que o cérebro humano não é capaz de sustentar de modo contínuo. O juiz não decide a partir de um ponto arquimediano de neutralidade, ele é um ser biológico, historicamente situado, emocionalmente modulável e cognitivamente limitado.

E a psicologia comportamental, por meio de Daniel Kahneman e Amos Tversky (1974) demonstra que o julgamento sob incerteza é frequentemente mediado por heurísticas, ou seja, atalhos mentais que, embora econômicos, levam a erros sistemáticos e previsíveis.

Para os referidos autores, três heurísticas principais afetam o julgamento: (i) a representatividade, ou seja, quando a probabilidade de um evento é avaliada pelo grau em que ele se assemelha a uma classe ou processo, o que acarreta, muitas vezes, na ignorância da probabilidade prévia (taxa base); (ii) disponibilidade, pela qual a frequência ou probabilidade de um evento é avaliada pela facilidade com que instâncias ou ocorrências vêm à mente, podendo, inclusive, levar juízes a superestimar riscos baseados em casos recentes ou emocionalmente salientes; e (iii) ancoragem e ajuste, em que estimativas são feitas partindo de um valor inicial (âncora) que é ajustado para produzir a resposta final. Como alertam os estudiosos, ajustes são tipicamente insuficientes, significando que a decisão final pode ser indevidamente influenciada por um ponto de partida arbitrário.

Nessa busca por soluções, Cass Sunstein propõe mecanismos para mitigar esses vieses e promover a deliberação. Como bem destaca Habermas (1997), Sunstein sugere que os tribunais desenvolvam estratégias interpretativas para promover a deliberação no governo, a exemplo de exigir uma análise fundamentada (*reasoned analysis requirement*) quando a deliberação parece estar ausente. O controle racional da decisão judicial, portanto, implica reconhecer e mitigar esses vieses cognitivos, uma função que a publicidade e a motivação das decisões visam assegurar.

Para complementar a análise da racionalidade judicial, é imperativo integrar a dimensão neurobiológica da mente humana. Nesse campo, António Damásio (2011) demonstra que a racionalidade não é um processo desencarnado, ela depende intrinsecamente da regulação biológica e dos sentimentos.

Segundo o autor, a consciência e a mente emergem da regulação da vida (homeostase): "A mente consciente emerge na história da regulação da vida". Isso significa que a decisão racional humana é influenciada por marcadores somáticos, isto é, sinais baseados em emoções que ajudam a selecionar opções vantajosas antes mesmo que o raciocínio lógico opere plenamente. Nas palavras de Damásio (2011, p. 20) "Quando conteúdos pertencentes ao *self* ocorrem no fluxo da mente, provocam o aparecimento de um marcador [...] Esses sentimentos geram uma distinção entre *self* e *não self*".

No contexto judicial, isso implica que a valoração não é apenas um cálculo lógico, mas sim um processo enraizado no valor biológico de sobrevivência e bem-estar. Como bem descreve Damásio (2011, p. 59) "o valor biológico guia e colore de maneira natural, por assim dizer, quase tudo que ocorre em nosso cérebro é riquíssimo em mente e consciência". Dessa forma, a decisão judicial, portanto, envolve um sujeito (o juiz) dotado de um *self* que testemunha e sente, algo fundamentalmente distinto de um processamento algorítmico.

Para Damásio (2011, p. 6-7), notadamente a ausência desse substrato biológico em sistemas artificiais é um ponto crítico, pois a mente humana possui um "*self*" que atua como testemunha e proprietário dos processos mentais. "Consciência não é meramente estar acordado [...]. Acordar significou ter de volta minha mente..., agora comigo nela, cômico tanto da propriedade (a mente) como do proprietário (eu)", finaliza o autor.

Ora, a máquina opera sem essa subjetividade sentida. A racionalidade democrática pressupõe justamente essa subjetividade capaz de sentir a gravidade da decisão sobre a vida alheia.

Portanto, constata-se que a racionalidade processual democrática é uma ideia teórica complexa que entrelaça a validade discursiva das normas (Habermas), a estrutura lógica dos princípios (Alexy), a integridade interpretativa (Dworkin) e a realidade cognitiva e biológica do julgador (Kahneman, Damásio).

Em última análise, a decisão judicial não é um mero *output* lógico, trata-se de um ato comunicativo de um sujeito biológico (o juiz) inserido em uma comunidade política, sujeito a vieses, mas constrangido por procedimentos que visam a correção e a justiça. Sob essa ótica, a introdução de sistemas automatizados deve ser avaliada quanto à sua capacidade de preservar, ou ameaçar, essa arquitetura, especialmente a dimensão da justificação inteligível e a responsabilidade moral intrínseca ao sujeito humano que sente e decide.

3 Inteligência artificial no judiciário e a emergência do viés algorítmico

A introdução de sistemas de inteligência artificial (IA) e modelos estatísticos no Poder Judiciário é frequentemente celebrada sob a promessa de eficiência, neutralidade e objetividade aos julgamentos. Todavia, uma análise crítica da literatura especializada revela que a automação não elimina a subjetividade humana nem os preconceitos sociais preexistentes. Pelo contrário, existe o risco real de obscurecê-los sob um manto de tecnicidade, o que, paradoxalmente, amplia suas consequências nocivas.

Diante desse cenário, este capítulo se propõe a investigar de que maneira a IA, ao ser alimentada por dados históricos e operar mediante modelos opacos, pode reestruturar e amplificar desigualdades, gerando o que se pode denominar de uma "autoridade ilegítima" baseada em previsões enviesadas.

Essa crença na suposta objetividade dos algoritmos é desafiada pela constatação de que modelos matemáticos não são representações neutras da realidade. Sobre o tema, Cathy O'Neil (2016) define muitos desses algoritmos como "Armas de Destruição Matemática" (WMDs - *Weapons of Math Destruction*), caracterizados por três elementos principais:

opacidade, escala e dano. Segundo a autora, modelos são, em essência, opiniões embutidas na matemática.

Dentre essas características, a opacidade assume um papel central. Grande parte desses algoritmos operam como verdadeiras "caixas-pretas" (*black boxes*), cujos funcionamentos internos são invisíveis para todos, exceto para seus criadores.

No contexto judicial, essa falta de transparência impede que os indivíduos afetados, sejam réus, litigantes ou cidadãos sob vigilância, compreendam ou contestem as decisões tomadas a seu respeito. O'Neil (2016) argumenta que, embora esses sistemas sejam comercializados como justos e objetivos porque não envolvem humanos preconceituosos vasculhando pilhas de papel, eles frequentemente codificam preconceitos, mal-entendidos e vieses humanos em sistemas de *software* que gerenciam nossas vidas.

Além da opacidade, a escalabilidade (*scale*) esses modelos permitem que um erro ou um viés, que antes estaria restrito ao julgamento de um único indivíduo, seja aplicado industrialmente a milhares ou milhões de pessoas. O dano (*damage*) ocorre quando esses sistemas punem desproporcionalmente os pobres e oprimidos, criando ciclos perniciosos que reforçam a desigualdade.

Para tanto, é fundamental compreender que a discriminação gerada por algoritmos não exige, necessariamente, uma intenção maliciosa por parte dos programadores. Conforme explicam Barocas e Selbst (2016) o processo de mineração de dados (*data mining*) pode discriminar de formas não intencionais, mas devastadoras, reproduzindo padrões históricos de preconceito.

Afinal, a qualidade do algoritmo é diretamente proporcional à qualidade dos dados que o alimentam, seguindo a máxima da ciência da computação, "lixo entra, lixo sai" (*garbage in garbage out*). Se os dados de treinamento (*training data*) refletem decisões anteriores contaminadas por preconceito, o algoritmo aprenderá a reproduzir esse preconceito, conferindo-lhe um verniz de legitimidade imparcial.

Segundo Barocas e Selbst (2016), o impacto discriminatório na mineração de dados decorre de mecanismos técnicos específicos inerentes ao processo. Inicialmente, isso ocorre na definição da variável alvo. É nesse momento que a tradução subjetiva de um problema real para uma questão computável, como, por exemplo, a determinação do que constitui um bom funcionário pode sistematicamente desfavorecer classes protegidas.

Além disso, o modelo pode ser comprometido por dados de treinamento enviesados, se a amostragem refletir preconceitos históricos ou uma supervisão desproporcional (como o policiamento excessivo em certas áreas), o algoritmo aprenderá e reproduzirá essas distorções,

tratando a maior incidência de registros como prova de maior criminalidade.

Por fim, é crucial considerar que, mesmo quando atributos sensíveis são deliberadamente excluídos, o sistema pode identificar *proxies*, ou variáveis substitutas, que operam como codificações redundantes (*redundant encodings*) da classe protegida, a exemplo de códigos postais correlacionados a raça, perpetuando, assim, a discriminação sob uma aparência de neutralidade.

Essa dinâmica reflete o que Safiya Noble (2021) aponta como discriminação algorítmica estrutural, quando o sistema não apenas falha, mas funciona exatamente para manter as hierarquias sociais existentes, "descobrimo" que membros de grupos protegidos possuem taxas sistematicamente inferiores em critérios valorizados pelo modelo, simplesmente porque a sociedade já os colocou em desvantagem.

A aplicação prática desses conceitos no judiciário pode ser ilustrada por sistemas de avaliação de risco criminal (*criminal risk assessment*), como o COMPAS e o LSI-R. O'Neil (2016) critica severamente esses modelos, observando que eles utilizam dados como o histórico de emprego, nível educacional e antecedentes criminais de familiares para calcular o risco de reincidência. Ao penalizar indivíduos por circunstâncias de sua vida e ambiente, e não apenas por seus crimes, esses modelos criminalizam a pobreza e criam um ciclo de respostas nocivas ao processo.

Pesquisas empíricas elaborada por Dressel e Farid (2018) demonstraram que o *software* comercial COMPAS não é mais preciso na predição de reincidência do que pessoas sem especialização em justiça criminal. Pelo contrário, o sistema demonstrou vieses raciais significativos, em que réus negros que não reincidiram foram incorretamente classificados como de alto risco em uma taxa quase duas vezes maior que a de réus brancos (44,9% vs. 23,5%), enquanto réus brancos que reincidiram foram incorretamente classificados como de baixo risco muito mais frequentemente do que réus negros.

Além disso, a integridade dos *datasets* é fundamental. Buolamwini e Gebru (2018) demonstraram no estudo *Gender Shades* que algoritmos de análise facial comercializados por grandes empresas (IBM, Microsoft, Face++) apresentaram taxas de erro drasticamente maiores para mulheres de pele escura (até 34,7%) em comparação com homens de pele clara (0,8%). Isso ocorre porque os *benchmarks* e dados de treinamento são majoritariamente compostos por indivíduos de pele clara. Se tecnologias de visão computacional forem integradas a sistemas de vigilância e justiça sem auditoria, o risco de identificação errônea e injustiça para populações vulneráveis é imenso.

Nesse sentido, a implementação de sistemas de IA no judiciário, a exemplo do que já ocorre no judiciário brasileiro, projeto VICTOR no STF, o Logos, recentemente lançado pelo STJ, e sistemas de classificação no TJCE, embora possuam o objetivo de modernizar a análise e elaboração de conteúdos judiciais, visando trazer mais celeridade e eficiência na produção de decisões, traz à tona também o risco do viés da automação (*automation bias*).

Almeida (2021) define este viés como a tendência humana de favorecer, sem questionar, as evidências ou sugestões fornecidas por sistemas automatizados.

Em um cenário de sobrecarga de trabalho, juízes e servidores podem utilizar a automação como um atalho cognitivo (heurística), aceitando as recomendações algorítmicas para economizar recursos mentais.

Skitka et al. (1999) alertam que isso pode levar a "erros de comissão", onde os humanos seguem uma diretiva incorreta da máquina porque falham em verificar outras informações disponíveis.

No contexto judicial, se um algoritmo de triagem sugere que um recurso não tem repercussão geral ou que um réu apresenta alto risco, o magistrado pode acatar essa sugestão acriticamente, abdicando de seu dever de julgamento individualizado e fundamentado.

Kleinberg et al. (2016) alertam que é matematicamente impossível satisfazer simultaneamente todas as condições de justiça (calibração, equilíbrio para a classe positiva e equilíbrio para a classe negativa) quando as taxas-base de reincidência diferem entre grupos. Portanto, a confiança cega na objetividade matemática de um sistema pode ignorar as escolhas normativas e os *trade-offs* inerentes à sua construção.

Portanto, é importante ressaltar que a inteligência artificial no judiciário não elimina o viés humano, ela apenas o camufla com tecnologia, uma vez que, ao processar dados históricos contaminados por desigualdades estruturais por meio de modelos opacos, a IA corre o risco de perpetuar injustiças sob a aparência de ciência e neutralidade. O'Neil (2016) adverte que esses modelos definem sua própria realidade e usam-na para justificar seus resultados.

Nesse cenário, a emergência do viés algorítmico pode criar, portanto, um risco de "autoridade ilegítima", na medida que decisões que afetam liberdades e direitos fundamentais sejam delegadas a sistemas que não possuem capacidade moral, não prestam contas de sua lógica interna (*black box*) e falham em tratar os indivíduos com a equidade exigida pela justiça. Diante disso, a automação acrítica da jurisprudência, sem a devida supervisão humana consciente dos limites e vieses da máquina, ameaça transformar o devido processo legal em uma linha de montagem de desigualdades.

4 Caminhos para conciliar racionalidade democrática e inovação tecnológica

Diante dos riscos de reprodução de desigualdades e da emergência de uma "autoridade ilegítima" impulsionada por sistemas automatizados, esse capítulo explora caminhos para conciliar a eficiência tecnológica com a racionalidade democrática, sugerindo um modelo de governança que assegure a supremacia dos direitos fundamentais sobre a lógica de otimização algorítmica.

A busca pela eficiência, embora legítima, não pode suplantiar a racionalidade democrática. Utilizando como exemplo o cenário atual do judiciário brasileiro, que reflete uma virada tecnológica sem precedentes, o Relatório Justiça em Números 2025 mostra que 99,4% dos casos novos ingressaram em formato eletrônico e 94,8% do acervo tramita digitalmente. Essa digitalização massiva, aliada à produtividade recorde de 44,8 milhões de processos baixados (um aumento de 19,9% em relação ao ano anterior), criam um terreno fértil para a implementação da IA, mas exige, concomitantemente, um modelo de governança que assegure a supremacia dos direitos fundamentais sobre a lógica de otimização algorítmica.

No entanto, a resposta jurídica ao avanço da IA não pode se limitar a ajustes pontuais, ela exige uma regulação baseada em direitos. Sobre essa necessidade, Cathy O'Neil (2016) argumenta que devemos impor valores humanos a esses sistemas, mesmo que isso custe eficiência. Como exemplo, a autora cita o modelo europeu de proteção de dados, que exige consentimento (*opt-in*) e proíbe a reutilização de dados para fins não autorizados, como um exemplo de regulação que limita o poder de corretores de dados e protege a cidadania.

No contexto trabalhista, Barocas e Selbst (2016) reforçam que a legislação atual da União Europeia, é muitas vezes insuficiente para lidar com a discriminação não intencional da mineração de dados, pois os modelos podem discriminar corretamente do ponto de vista estatístico (*proxy discrimination*), reproduzindo desigualdades estruturais.

Isso sugere a necessidade de regulações que vão além da proibição do tratamento desigual intencional (*disparate treatment*) e o enfrentamento do impacto desproporcional (*disparate impact*) de forma substantiva, possivelmente exigindo que operadores adotem práticas alternativas menos discriminatórias ou corrijam proativamente seus dados.

Já que a opacidade dos sistemas (*black box*) é uma barreira central à justiça, o relatório *Litigating Algorithms* (2018) destaca que a batalha pela transparência é fundamental, pois em casos envolvendo benefícios de saúde e avaliação de professores, o acesso ao código-fonte e aos dados de treinamento foi crucial para identificar erros de codificação e falhas

lógicas que prejudicavam cidadãos vulneráveis.

Nesse contexto, a transparência não é apenas técnica, mas funcional. Buolamwini e Gebru (2018) defendem a necessidade de relatórios de precisão por subgrupos demográficos (*intersectional accuracy reports*) para evitar que altas taxas de acerto global mascarem falhas graves em grupos específicos.

Além disso, Dressel e Farid (2018) demonstraram que classificadores lineares simples e transparentes (usando apenas duas variáveis, idade e condenações anteriores) podem ser tão precisos quanto algoritmos proprietários complexos (como o COMPAS), sugerindo que a complexidade opaca nem sempre é necessária para a eficácia.

Diante desse quadro, sugere-se que, para mitigar o risco de desumanização, a IA deve ser posicionada estritamente como ferramenta de suporte, e jamais como decisora final. Entretanto, essa forma de uso também exige cautela.

Almeida (2021) alerta para o risco do viés da automação, um fenômeno no qual os seres humanos tendem a favorecer, sem questionamentos, as evidências geradas pela tecnologia. Essa confiança excessiva e acrítica pode ocorrer porque, muitas vezes, o julgamento humano profundo consome muita energia e tempo de processamento cerebral, levando os indivíduos a aceitarem a resposta automatizada como um atalho para economizar esforço cognitivo.

Desta forma, para que modelos híbridos (juiz + IA) funcionem democraticamente, a supervisão humana não pode ser uma formalidade (*rubber stamp*). É necessário que o sistema cognitivo humano (Sistema 2, analítico e lento) seja ativado para revisar as intuições rápidas fornecidas pelo sistema automatizado (Sistema 1), garantindo que a decisão leve em conta nuances contextuais que a máquina ignora.

Assim, a implementação de sistemas automatizados no setor público requer a estruturação de salvaguardas processuais rigorosas para mitigar danos e garantir a transparência.

Fundamentalmente, isso passa pelo direito à contestação e ao devido processo legal (*Due Process*). Conforme o relatório do *AI Now Institute* (2018), estratégias de litígio bem-sucedidas utilizaram argumentos constitucionais para desafiar a opacidade de algoritmos que suprimiam benefícios sem notificação ou explicação adequadas, estabelecendo que a proteção de segredos comerciais não pode se sobrepor ao direito do público de compreender a lógica das decisões estatais.

Para complementar essa transparência, Cathy O'Neil (2016) preconiza a realização de auditorias algorítmicas independentes, nas quais auditores externos e acadêmicos avaliam os

modelos, tratando-os como 'caixas-pretas' para mensurar impactos desproporcionais e revelar vieses ocultos, em vez de confiar na autoavaliação dos criadores da tecnologia.

Os mecanismos de governança ética e participação são vitais para o desenvolvimento decisório, a exemplo dos comitês de acompanhamento, focando em segurança, transparência ética e proteção de dados. Tais órgãos podem fornecer dados empíricos em tempo real, permitindo que magistrados superem heurísticas ou intuições iniciais e fundamentem suas decisões em evidências concretas e atualizadas sobre o funcionamento das instituições.

No entanto, a Pesquisa CNJ/PNUD (2025) sobre Uso de Inteligência Artificial no Poder Judiciário em 2024 revela um desafio estrutural: a inexperiência ou falta de mão de obra especializada é o principal obstáculo para o desenvolvimento de projetos de IA, citado por 29,9% dos órgãos. Se os tribunais enfrentam barreiras de expertise interna, aumenta-se o risco de dependência de caixas-pretas tecnológicas sem a devida auditabilidade, dificultando a transparência funcional necessária para identificar vieses.

Nesse contexto, a Resolução CNJ nº 615/2025 vem estabelecer diretrizes para governança, o desenvolvimento e a auditoria de soluções de IA, buscando mitigar riscos éticos evidenciados pela prática atual.

Dados recentes mostram que 50% dos tribunais e conselhos ainda não possuem diretrizes internas específicas sobre o uso da IA generativa e 57,6% indicam que seus profissionais utilizam essas ferramentas por meio de contas pessoais, fora de ambientes controlados. Assim, a ausência de governança compromete a segurança jurídica e a proteção de dados, reforçando a necessidade de comitês de acompanhamentos que forneçam dados empíricos em tempo real para fundamentar decisões.

Diante desse quadro de alta digitalização e volume processual massivo em todo o judiciário nacional, propõe-se um modelo de governança híbrida para mitigar o risco de desumanização e o viés da automação, segundo o qual o julgador, pressionado por metas de produtividade, pode aceitar acriticamente a sugestão da máquina, não posicionando a IA estritamente como ferramenta de suporte, como deve ser.

Neste contexto, a teoria de Habermas (1997) oferece um alicerce teórico para enfrentar a "mudez" dos algoritmos. O autor faz uma distinção entre a ação estratégica (orientada para o sucesso/fins) e a ação comunicativa (orientada para o entendimento). Algoritmos operam fundamentalmente na lógica estratégica, otimizando meios para atingir fins predefinidos (eficiência, lucro, reincidência), sem capacidade de questionar a validade normativa desses fins. Como exemplo, a redução do acervo processual no judiciário brasileiro, que caiu 5,3% no último ano, e a legitimidade do direito dependem da ação

comunicativa. Ou seja, o sucesso numérico do Judiciário ao reduzir o estoque processual deve ser acompanhado pela garantia de que o poder administrativo automatizado permaneça subordinado ao poder comunicativo.

Para Habermas (1997), a legitimidade do direito depende do processo discursivo em que cidadãos livres e iguais podem validar normas via argumentos racionais. Tecnologias que não argumentam, mas apenas calculam probabilidades, ameaçam substituir essa racionalidade comunicativa por imperativos sistêmicos, como o dinheiro e poder administrativo.

Todavia, manter a racionalidade comunicativa exige que o poder administrativo, agora automatizado, permaneça vinculado e subordinado ao poder comunicativo gerado na esfera pública. O direito deve funcionar como um transformador que traduz as demandas da vida para a linguagem sistêmica, sem permitir que a lógica sistêmica colonize a justiça.

Sob essa ótica, propõe-se um modelo de racionalidade processual tecnodemocrática que, longe de rejeitar a inovação, submete a tecnologia às condições do discurso democrático.

Esse modelo se fundamenta, primeiramente, na intersubjetividade de sua construção, reconhecendo que a definição de variáveis-alvo não é uma atividade neutra, mas uma tradução subjetiva de problemas sociais em código, devendo ser tratada como uma escolha normativa sujeita à deliberação pública, pois, conforme Cathy O'Neil, modelos são essencialmente opiniões embutidas em matemática.

Em segundo lugar, exige-se revisão crítica aprofundada, a fim de evitar que os sistemas se tornem armas de destruição matemática que operam em ciclos de respostas perniciosas, definindo sua própria realidade sem correção externa. É imperativo estabelecer mecanismos de apelação e auditoria contínua que permitam contestar a caixa-preta.

Por fim, busca-se uma “higienização” decisória democrática, inspirada na distinção entre ruído e viés de Kahneman et al., em que o sistema deve reduzir a variabilidade indesejada (ruído) mediante diretrizes consistentes, mas sem impor uma rigidez algorítmica que elimine a discricionariedade necessária para a equidade e para a evolução de valores morais, preservando o espaço do julgamento humano no qual a regra fixa se mostra insuficiente.

5 Conclusão

A presente pesquisa buscou responder à indagação central sobre como garantir a racionalidade e a legitimidade das decisões judiciais quando assistidas por inteligência artificial, diante da tensão constitutiva entre os vieses humanos, os vieses algorítmicos e as exigências democráticas de transparência e motivação. A investigação demonstrou que a

resposta a esse problema não reside na substituição da subjetividade do juiz pela suposta objetividade da máquina, nem na rejeição da tecnologia, mas na implementação de um sistema de governança híbrida que submeta a lógica estratégica da eficiência algorítmica à racionalidade comunicativa do processo democrático.

Os resultados alcançados evidenciam que a decisão judicial na era digital enfrenta uma "dupla opacidade", a cognitiva, inerente à racionalidade limitada humana, e a computacional, derivada das "caixas-pretas" algorítmicas. Em resposta ao problema proposto, conclui-se que a legitimidade da jurisdição contemporânea depende da capacidade institucional de transmutar o poder administrativo dos algoritmos em autoridade legítima, assegurando que a tecnologia atue como suporte para reduzir ruídos, sem, contudo, usurpar a reserva humana de valoração ética e justificação hermenêutica.

A revisitação dos objetivos específicos permite aferir o percurso investigativo. O primeiro objetivo, voltado ao exame das limitações da racionalidade judicial humana, foi atingido ao integrar as contribuições da psicologia cognitiva e da neurociência, confirmando que o julgamento humano é biologicamente situado e suscetível a heurísticas e marcadores somáticos, desmistificando o ideal de neutralidade absoluta.

O segundo objetivo, focado na dissecção da natureza dos vieses algorítmicos, foi cumprido ao demonstrar, através da literatura crítica, que modelos matemáticos não são neutros, mas opiniões codificadas que podem reproduzir desigualdades estruturais e operar como "armas de destruição matemática".

Por fim, o terceiro objetivo foi alcançado ao propor caminhos para uma integração responsável, culminando na formulação do modelo de racionalidade processual tecnodemocrática, que exige a subordinação da tecnologia às condições do discurso e à implementação de mecanismos de auditabilidade e revisão humana.

Não obstante os resultados obtidos, o tema permanece aberto e sujeito a constantes reconfigurações, dada a velocidade da evolução tecnológica. É imperativo reconhecer as limitações deste estudo.

Do ponto de vista metodológico, a pesquisa adotou uma abordagem teórico-bibliográfica e documental, o que, embora tenha permitido a construção de um arcabouço conceitual sólido, limitou a análise empírica profunda sobre o funcionamento interno dos algoritmos específicos atualmente em uso nos tribunais brasileiros, devido à opacidade e proteção de propriedade intelectual que muitas vezes cercam essas ferramentas.

Contextualmente, a análise baseou-se em dados e normativas recentes, como os relatórios do CNJ de 2024 e 2025 e a Resolução nº 615/2025, capturando um momento de

transição institucional que pode sofrer alterações rápidas na prática forense.

Teoricamente, a aplicação da teoria do discurso de Habermas a sistemas não-humanos apresenta desafios ontológicos que exigem contínua refinação para evitar antropomorfismos indevidos.

Diante dessas constatações e lacunas, sugere-se análise de longo prazo sobre a eficácia da Resolução CNJ nº 615/2025 na mitigação do "viés da automação" e na promoção de uma cultura de transparência funcional, assegurando que o Judiciário digital permaneça, em última instância, um espaço de realização de justiça humana.

6 Referências

AI NOW INSTITUTE. **Litigating Algorithms**: Challenging Government Use of Algorithmic Decision Systems. New York: AI Now Institute, 2018. Disponível em: <https://ainowinstitute.org/litigatingalgorithms.pdf>.

ALEXY, Robert. **Teoria dos Direitos Fundamentais**. Tradução de Virgílio Afonso da Silva. 2. ed. São Paulo: Malheiros, 2008.

ALMEIDA, Gabriela Perissinotto de. **Vieses e heurísticas na tomada da decisão judicial**. Orientador: Sérgio Nojiri. Trabalho de Conclusão de Curso (Graduação em Direito) – Faculdade de Direito de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2021.

BAROCAS, Solon; SELBST, Andrew D. Big Data's Disparate Impact. **California Law Review**, Berkeley, v. 104, n. 3, p. 671-732, 2016.

BRASIL. Conselho Nacional de Justiça. **Justiça em números 2025**. Brasília: CNJ, 2025. Disponível em: <https://justica-em-numeros.cnj.jus.br/>. Acesso em: 10 dez. 2025.

BRASIL. Conselho Nacional de Justiça. **Pesquisa inteligência artificial no Judiciário 2024**: resumo executivo. Brasília: CNJ, 2025.

BRASIL. Conselho Nacional de Justiça. **Resolução nº 615, de 2025**. Institui diretrizes para o desenvolvimento, a governança e o uso de Inteligência Artificial no Poder Judiciário. Brasília: CNJ, 2025.

BUOLAMWINI, Joy; GEBRU, Timnit. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *In*: CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2018, New York. **Proceedings...** New York: PMLR, 2018. v. 81, p. 1-15.

DAMÁSIO, António R. **E o cérebro criou o Homem**. Tradução de Laura Teixeira Motta.

São Paulo: Companhia das Letras, 2011.

DRESSEL, Julia; FARID, Hany. The accuracy, fairness, and limits of predicting recidivism. **Science Advances**, Washington, v. 4, n. 1, eaao5580, jan. 2018. DOI: 10.1126/sciadv.aao5580.

DWORKIN, Ronald. **Levando os Direitos a Sério**. Tradução de Nelson Boeira. São Paulo: Martins Fontes, 2002.

HABERMAS, Jürgen. **Direito e democracia**: entre facticidade e validade. Tradução de Flávio Beno Siebeneichler. Rio de Janeiro: Tempo Brasileiro, 1997. v. 1.

KLEINBERG, Jon; MULLAINATHAN, Sendhil; RAGHAVAN, Manish. Inherent Trade-Offs in the Fair Determination of Risk Scores. **arXiv preprint arXiv:1609.05807**, 2016.

KROLL, Joshua A. et al. Accountable Algorithms. **University of Pennsylvania Law Review**, Philadelphia, v. 165, p. 633-705, 2017.

MARDEN, Carlos; WYKROTA, Leonardo Martins. Neurodireito: o início, o fim e o meio. **Revista Brasileira de Políticas Públicas**, Brasília, v. 8, n. 2, p. 48-63, ago. 2018.

O'NEIL, Cathy. **Weapons of Math Destruction**: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.

SILVA, Camila Narici da; STEWART, Michael López; CUNHA FILHO, Marcio Camargo. Ruído: examinando erros no julgamento humano para além do viés. **Revista Brasileira de Sociologia do Direito**, [S.l.], v. 9, n. 2, p. 128-136, maio/ago. 2022.

TVERSKY, Amos; KAHNEMAN, Daniel. Judgment under Uncertainty: Heuristics and Biases. **Science**, New Series, v. 185, n. 4157, p. 1124-1131, 1974.