

Inteligência artificial e viés no judiciário: uma análise constitucional sob a lente empírica

Artificial intelligence and bias in the judiciary: a constitutional analysis through an empirical lens

Rayssa de Souza Gargano¹

Klever Paulo Leal Filpo²

Carmen Caroline Ferreira do Carmo Nader³

RESUMO

O crescente uso da inteligência artificial (IA) no sistema de justiça suscita debates acerca da compatibilidade entre automação decisória e garantias fundamentais. Este artigo realiza uma análise jurídico-constitucional, com enfoque empírico-documental, sobre os riscos de enviesamento algorítmico e seus impactos sobre direitos fundamentais processuais. A pesquisa⁴, de natureza qualitativa, combina revisão bibliográfica crítica com levantamento empírico de dados institucionais acerca da utilização de sistemas de IA no Judiciário brasileiro, além de examinar o caso paradigmático do sistema COMPAS, empregado nos Estados Unidos. O estudo conta com recursos da CAPES e da FAPERJ. Argumenta-se que a ausência de transparência e auditabilidade algorítmica pode comprometer o devido processo legal, a imparcialidade judicial e a fundamentação das decisões. Conclui-se pela necessidade de aperfeiçoamento das políticas públicas de governança algorítmica no Brasil, de modo a assegurar que a inteligência artificial atue como instrumento de apoio à jurisdição, e não como fator de reprodução de desigualdades estruturais.

Palavras-chave: Inteligência Artificial. Governança Algorítmica. Enviesamento Algorítmico. Devido Processo Legal.

¹ Mestra em direito pela Universidade Católica de Petrópolis (Prosuc/Capes), professora e pesquisadora na área de precedentes e Inteligência Artificial. Advogada Sócia no escritório Gargano & Santos. Presidente Comissão Especial de Mentoria Jurídica da OAB/RJ - 1 Subseção de Nova Iguaçu. Pesquisadora do Grupo Interdisciplinar de Pesquisa sobre Administração de Conflitos (GIPAC/UCP-CNPq) e Observatório das reformas processuais destinadas à solução de demandas seriais e ações coletivas. Professora e Preceptora Cível da Universidade Unigranrio - Afya. Professora do CEPEP/UERJ na área de inteligência artificial e do Curso Adonai. E-mail: rayssagargano@gmail.com.

² Doutor em Direito. Professor do Programa de Pós-graduação em Direito da Universidade Católica de Petrópolis e do Programa de Pós-graduação em Direito, Ambiente e Desenvolvimento da Universidade Federal Rural do Rio de Janeiro. Coordenador do Grupo Interdisciplinar de Pesquisa sobre Administração de Conflitos (GIPAC/UCP-CNPq), Pesquisador do INCT/InEAC. Jovem Cientista do Nosso Estado – FAPERJ. E-mail: klever.filpo@yahoo.com.br.

³ Doutoranda em Direito pela UVA. Mestre em Direito pela UCP. Docente. Coordenadora do Núcleo de Práticas Jurídicas da Afya Unigranrio Nova Iguaçu. Advogada. Coordenadora Acadêmica da ESA da OAB de Nova Iguaçu. Superintendente Jurídica da Secretaria da Mulher de Nova Iguaçu. E-mail: carmen-caroline@gmail.com.

⁴ A presente pesquisa contou com recursos da FAPERJ e do CNPq, por meio do Programa de Apoio à Fixação de Jovens Doutores no Brasil – 2023 (bolsista Mirel Legrá Fleitas – PPGD/UCP), bem como com bolsa de Mestrado PROSUC/CAPES (bolsista Rayssa Gargano – PPGD/UCP). Os resultados da pesquisa têm gerado discussões e trabalhos apresentados e publicados em eventos nacionais e internacionais, incluindo o Congresso Nacional do CONPEDI, realizado em São Paulo, em 2025.

ABSTRACT

The growing use of artificial intelligence (AI) in the justice system has sparked debates regarding the compatibility between automated decision-making and fundamental guarantees. This article conducts a constitutional-legal analysis, with an empirical-documentary focus, of the risks of algorithmic bias and its impacts on procedural fundamental rights. The qualitative research combines a critical literature review with an empirical survey of institutional data concerning the use of AI systems within the Brazilian Judiciary, as well as an examination of the paradigmatic case of the COMPAS system employed in the United States. The study was developed with funding from CAPES and FAPERJ. It is argued that the lack of algorithmic transparency and auditability may compromise due process of law, judicial impartiality, and the duty to provide reasoned decisions. The article concludes that there is a need to improve public policies on algorithmic governance in Brazil in order to ensure that artificial intelligence functions as a tool supporting judicial activity rather than as a mechanism for reproducing structural inequalities.

Keywords: Artificial Intelligence. Algorithmic Governance. Algorithmic Bias. Due Process of Law.

INTRODUÇÃO

A incorporação de sistemas de inteligência artificial ao funcionamento do Poder Judiciário constitui fenômeno crescente e irreversível. Ferramentas algorítmicas vêm sendo utilizadas para triagem processual, agrupamento temático, elaboração de minutas e, em alguns contextos internacionais, até mesmo para auxiliar decisões judiciais de natureza penal.

O presente estudo propõe examinar a intersecção entre políticas públicas, inteligência artificial e proteção de direitos fundamentais processuais, tomando como referência empírica o sistema COMPAS, utilizado nos Estados Unidos para avaliação de risco de reincidência criminal, e o panorama atual de implementação de IA no Judiciário brasileiro.

Trata-se de pesquisa qualitativa, de natureza exploratória e analítica, desenvolvida por meio de revisão bibliográfica crítica e análise empírico-documental. A etapa empírico-documental concentrou-se no levantamento e exame de dados institucionais sobre a utilização de sistemas de inteligência artificial no Judiciário brasileiro, com base em relatórios públicos, atos normativos e pesquisas acadêmicas, visando à construção de um panorama descritivo-analítico orientado à identificação de riscos de enviesamento algorítmico e de déficits de transparência e auditabilidade. Em complemento, procede-se à análise do caso paradigmático do sistema COMPAS, empregado nos Estados Unidos, como referência comparativa para a discussão dos impactos sobre garantias fundamentais processuais, especialmente o devido processo legal, a imparcialidade judicial e a fundamentação das decisões.

O estudo articula literatura nacional e estrangeira sobre governança algorítmica com levantamento de dados institucionais relativos ao uso de inteligência artificial nos tribunais brasileiros. A abordagem adotada é jurídico-constitucional, examinando os impactos da automação decisória à luz das garantias do devido processo legal, da imparcialidade judicial e da vedação ao retrocesso em matéria de direitos fundamentais.

O problema central consiste em indagar em que medida a incorporação de sistemas algorítmicos ao sistema de justiça pode comprometer garantias constitucionais, especialmente quando tais sistemas operam com baixa transparência e limitada possibilidade de controle externo.

Nesse contexto, impõe-se reconhecer que a racionalidade algorítmica, embora revestida de aparência técnica e neutralidade matemática, não se encontra imune às influências estruturais dos dados que a alimentam e das escolhas normativas que orientam sua programação. Sistemas de inteligência artificial são treinados com base em conjuntos de dados historicamente constituídos, os quais podem refletir desigualdades sociais, raciais e econômicas já presentes na sociedade. Quando tais padrões são incorporados a mecanismos decisórios estatais, corre-se o risco de transformar vieses estruturais em decisões institucionalmente legitimadas, com impacto direto sobre direitos fundamentais.

Além disso, a opacidade dos algoritmos, especialmente quando desenvolvidos por empresas privadas sob a proteção do segredo comercial, dificulta o exercício pleno do contraditório e da ampla defesa, uma vez que as partes não dispõem de meios adequados para compreender, questionar ou auditar os critérios subjacentes às decisões automatizadas. A ausência de mecanismos claros de explicabilidade e *accountability*⁵ compromete não apenas a fundamentação das decisões judiciais, mas também o próprio controle democrático da atividade jurisdicional, exigindo, portanto, reflexão crítica acerca das políticas públicas que regulam a governança algorítmica no âmbito do Poder Judiciário.

INTELIGÊNCIA ARTIFICIAL NO JUDICIÁRIO BRASILEIRO: PANORAMA INSTITUCIONAL E ALGUNS DADOS EMPÍRICOS

⁵ *Accountability* é um conceito oriundo da teoria política e da administração pública que se refere aos deveres de prestação de contas, transparência e responsabilização dos agentes públicos e das instituições perante a sociedade. No contexto da governança algorítmica, o termo designa a obrigação de assegurar a auditabilidade dos critérios, dados e parâmetros empregados por sistemas de inteligência artificial, possibilitando o controle institucional e social das decisões automatizadas. A *accountability* não se limita à publicidade dos atos, abrangendo também mecanismos efetivos de revisão, correção e responsabilização em casos de erros, discriminações ou violações de direitos fundamentais. Trata-se, portanto, de elemento indispensável para a compatibilização entre a inovação tecnológica e os princípios do Estado Democrático de Direito.

A transformação digital do Judiciário brasileiro foi intensificada a partir da pandemia de Covid-19, consolidando ambiente propício à adoção de tecnologias de automação e inteligência artificial. A necessidade de garantir a continuidade da prestação jurisdicional em contexto de distanciamento social acelerou processos de virtualização já em curso, ampliando o uso de sistemas eletrônicos e fomentando a incorporação de ferramentas algorítmicas voltadas à racionalização do acervo processual.

Pesquisa empírica conduzida por Salomão e Tauk (2023) identificou mais de quarenta iniciativas de uso de IA em diferentes órgãos do Judiciário, incluindo Conselho Nacional de Justiça (CNJ), Supremo Tribunal Federal (STF), Superior Tribunal de Justiça (STJ), Tribunal Superior do Trabalho (TST), Tribunais Regionais Federais (TRFs), Tribunais de Justiça (TJs) e Tribunais Regionais do Trabalho (TRTs). O levantamento evidencia que a inteligência artificial já integra parcela relevante da infraestrutura institucional do Judiciário brasileiro, deixando de constituir experiência pontual para se tornar instrumento estrutural de gestão processual.

Sistemas como Victor, desenvolvido no âmbito do STF em parceria com a Universidade de Brasília, destinam-se à triagem de recursos extraordinários e à identificação de temas com repercussão geral. No STJ, os sistemas Sócrates e Athos auxiliam na organização e classificação de processos repetitivos, promovendo maior eficiência na tramitação. Tais ferramentas demonstram que o uso da IA no Brasil concentra-se, predominantemente, em funções de apoio administrativo e organização processual e não, ao menos formalmente, em decisões preditivas de natureza penal.

Segundo dados divulgados pelo Sistema de Automação da Justiça (SAJ), a taxa média de acerto na leitura automatizada de processos alcança aproximadamente 86% (Softplan, 2022). Esses sistemas operam por meio de técnicas de processamento de linguagem natural e mineração de dados, sendo capazes de identificar padrões textuais e classificar processos com base em parâmetros previamente definidos. Apesar da expressiva taxa de acurácia declarada, tais indicadores não esclarecem como se distribuem eventuais erros nem se há impactos diferenciados sobre determinados grupos sociais.

O levantamento documental evidencia aspectos estruturais que merecem reflexão crítica no âmbito da governança algorítmica do Poder Judiciário brasileiro. Observa-se, inicialmente, a inexistência de padronização nacional quanto à transparência algorítmica, uma vez que não há modelo uniforme de divulgação dos critérios técnicos utilizados pelos

sistemas implementados nos diferentes tribunais. Cada órgão adota práticas próprias, sem que se verifique diretriz vinculante que assegure publicidade clara e acessível acerca do funcionamento das ferramentas empregadas.

Além disso, constata-se escassez de informações públicas sobre os parâmetros técnicos, os conjuntos de dados e as metodologias de treinamento utilizadas pelos sistemas de inteligência artificial, o que dificulta a realização de auditorias independentes e o controle externo qualificado. Essa limitação compromete a possibilidade de verificação da existência de eventuais vieses estruturais ou impactos desproporcionais sobre determinados grupos sociais.

Por fim, verifica-se a inexistência de obrigação generalizada de comunicação às partes quanto ao uso de ferramentas automatizadas no suporte à decisão judicial. Em regra, não há menção expressa, nos atos processuais, à eventual utilização de sistemas algorítmicos na triagem, organização ou apoio à análise do caso concreto. Tal cenário pode fragilizar o contraditório substancial, na medida em que impede que os jurisdicionados tenham ciência plena dos instrumentos tecnológicos que influenciam, direta ou indiretamente, a formação do convencimento judicial.

Adicionalmente, observa-se que, embora a Resolução n.º 332/2020 do CNJ estabeleça diretrizes para o uso de inteligência artificial no Poder Judiciário, a regulamentação concentra-se em princípios gerais, como transparência, governança e respeito aos direitos fundamentais, sem instituir mecanismos concretos e vinculantes de auditabilidade técnica periódica. Tal cenário revela assimetria entre a velocidade de implementação tecnológica e o ritmo de consolidação normativa.

Embora o Brasil não utilize modelos preditivos penais nos moldes do COMPAS, que será analisado adiante neste artigo, a incorporação crescente de IA aos fluxos processuais já constitui realidade consolidada. A literatura sobre governança algorítmica (Coutinho, 2024) alerta que mesmo sistemas destinados a funções administrativas podem influenciar indiretamente a tomada de decisão judicial, especialmente quando estruturam previamente o universo informacional acessado pelo magistrado. Assim, a análise empírico-institucional demonstra que a discussão sobre enviesamento algorítmico não se restringe a modelos explicitamente decisórios, mas alcança também ferramentas de organização e filtragem processual.

Diante desse panorama, impõe-se a necessidade de vigilância institucional contínua, bem como o desenvolvimento de políticas públicas capazes de assegurar transparência, explicabilidade e *accountability* na utilização de inteligência artificial no sistema de justiça brasileiro.

ENVIESAMENTO ALGORÍTMICO E O RISCO DE MITIGAÇÃO DE DIREITOS

O incremento do uso de máquinas para substituir ações e decisões humanas é fruto de um longo processo de construção histórica. Em diferentes épocas, observou-se crescimento significativo em algumas áreas, enquanto em outras os avanços foram mais modestos. Contudo, o que se revela de maneira inequívoca é que cada um desses esforços, independentemente de sua magnitude, contribuiu para moldar a realidade contemporânea. Essa trajetória conduziu a um cenário em que a tecnologia desempenha papel central na sociedade, transformando a maneira como interagimos, trabalhamos e tomamos decisões.

A inteligência artificial pode ser concebida como um campo do conhecimento associado à linguagem e à inteligência, ao raciocínio, à aprendizagem e à resolução de problemas (Kaufman, 2019, p. 19). Gradualmente, inclusive no Brasil, essa tecnologia vem sendo incorporada pelos tribunais para desempenhar variadas tarefas, incluindo o apoio à movimentação processual e, mais recentemente, à prolação de decisões judiciais.

Na medida em que a inteligência artificial chega às barras dos tribunais, surgem questões relevantes a serem consideradas. No Estado Democrático de Direito, a atuação do Poder Judiciário é pautada por garantias como a imparcialidade do juiz e o devido processo legal (Santoro, 2016). Considerando especialmente a imparcialidade judicial, sua atuação não deve suscitar dúvidas quanto à sua neutralidade (Baptista, 2020, p. 215-216).

Ao se examinarem os vieses algorítmicos, é possível identificá-los sob duas perspectivas. A primeira, e mais imediata, relaciona-se à origem da programação do sistema. Indaga-se se um grupo de pessoas foi responsável por sua elaboração. Além disso, considerando o uso decisório da inteligência artificial, questiona-se se existe capacidade jurídico-legal para que programadores da área de tecnologia da informação filtrem demandas e exerçam julgamento. Mesmo que haja um grupo de magistrados envolvido na criação ou no treinamento da IA, como se dá a composição desse grupo? Quais são os critérios utilizados para o seu recrutamento? Ele é formado por representantes de grupos heterogêneos, com diferentes perspectivas e linhas de pensamento? Essas e outras questões

contribuem para alimentar o debate sobre as potencialidades e os riscos do uso da inteligência artificial pelas cortes de justiça.

Para exemplificar, no Brasil, a Justiça do Trabalho orienta-se por diversos princípios, entre os quais se destaca o princípio da proteção ao trabalhador, que se desdobra na garantia de direitos como a irredutibilidade salarial e a indisponibilidade dos direitos trabalhistas, aspectos especialmente relevantes em um país marcado pela desigualdade social. A partir desse exemplo, pode-se refletir acerca das possíveis perplexidades que poderão surgir à medida que um sistema de julgamento baseado em inteligência artificial, mas desenvolvido ou treinado sob outra perspectiva jurídica, cultural ou mesmo em tradição jurídica distinta, venha a ser empregado para decidir processos trabalhistas no Brasil.

O mesmo raciocínio pode ser aplicado às relações de consumo, uma vez que os consumidores brasileiros são especialmente protegidos pelo Código de Defesa do Consumidor, que admite, entre outros aspectos, significativa intervenção do Estado-juiz em contratos privados, característica que não necessariamente está presente em outros sistemas jurídicos nos quais a inteligência artificial tenha sido treinada.

Sarlet (2009) estabelece forte relação entre os direitos fundamentais sociais e a vedação ao retrocesso, no contexto brasileiro, compreendida em sentido amplo como toda e qualquer forma de proteção dos direitos fundamentais diante de medidas do poder público, com destaque para a atuação do legislador e do administrador, que tenham por objetivo a supressão ou mesmo a restrição desses direitos, sejam eles de natureza social ou não. Segundo o autor (2009, p. 120-121):

De outra parte, segue sendo necessária uma preocupação permanente com a consolidação e manutenção pelo menos dos níveis de proteção social mínimos, onde e quando alcançados, nas várias esferas da segurança social e da tutela dos direitos fundamentais compreendidos em toda a sua amplitude, inclusive como condição para a funcionalidade da própria democracia e sobrevivência do Estado Constitucional. Especialmente tendo em conta a instabilidade econômica que se verifica em escala global, assim como em virtude dos efeitos perversos da globalização, em particular no plano econômico, não se pode simplesmente negligenciar a relevância do reconhecimento de uma proibição de retrocesso como categoria jurídico-constitucional, ainda mais quando a expressiva maioria das reformas não dispensa mudanças no plano das políticas públicas e da legislação. Com efeito, dentre os diversos efeitos perversos da crise e da globalização econômica (embora não se possa imputar à globalização todas as mazelas vivenciadas na esfera social e econômica), situa-se a disseminação de políticas de flexibilização e até mesmo supressão de garantias dos trabalhadores (sem falar no crescimento dos níveis de desemprego e índices de subemprego), redução dos níveis de prestação social, aumento desproporcional de contribuições sociais por parte dos participantes do sistema de proteção social,

incremento da exclusão social e das desigualdades, entre outros aspectos que poderiam ser mencionados.

Na sociedade da informação, é necessário conciliar os benefícios provenientes da inovação e do avanço tecnológico com a proteção das liberdades individuais, dos direitos fundamentais e das garantias processuais. Essas medidas, de natureza precaucional, têm como objetivo atender ao duplo propósito de impulsionar a economia e proteger o cidadão. Contudo, no momento, percebem-se alguns riscos. Um primeiro exemplo é a ausência de transparência em relação aos algoritmos utilizados nos sistemas atualmente empregados no Brasil pelos tribunais. Por outro lado, tampouco são conhecidas as consequências decorrentes do uso contínuo dos dados das partes pelos sistemas automatizados. Esses, entre outros aspectos, recomendam maior debate sobre o tema em diferentes instâncias decisórias e, especialmente, no âmbito acadêmico.

A falta de transparência no compartilhamento dos algoritmos que fundamentam decisões automatizadas utilizadas no Judiciário levanta preocupações relevantes. Esse questionamento é crucial, pois a não divulgação dessas informações pode comprometer os direitos dos usuários da justiça, especialmente daqueles pertencentes a grupos vulneráveis. Tais práticas podem contribuir para a perpetuação de desigualdades sociais, sublinhando a necessidade de reformulação das políticas públicas que regem o uso da inteligência artificial no sistema judicial.

O CASO COMPASS

O software COMPAS, que significa *Correctional Offender Management Profiling for Alternative Sanctions* (Perfil de Gerenciamento de Criminoso Correcional para Sanções Alternativas), foi desenvolvido pela empresa Northpointe com o intuito de avaliar a probabilidade de um indivíduo reincidir em determinada conduta criminosa. O sistema cria níveis de avaliação por meio de escalas que indicam a probabilidade de risco e de reincidência geral. A escala criada, conforme informado pela empresa, leva em conta construções psicológicas e comportamentais que indicam alta relevância para reincidência em crimes e tendência a carreiras criminosas (Northpointe, 2015).

Esse sistema, utilizado na justiça criminal dos Estados Unidos como ferramenta para auxiliar juízes na avaliação da probabilidade de reincidência criminal por parte de um réu, é amplamente referenciado como exemplo de viés algorítmico (Salomão; Tauk, 2023). De acordo com critérios de IA, há uma avaliação de modelagem comportamental utilizada

no pré-julgamento do COMPAS, baseada na teoria da aprendizagem social. Essa teoria sustenta que a maneira tradicional de compreender a aprendizagem ocorre por meio da modelagem de comportamento. Ou seja, uma vez que o comportamento é modelado e imitado, ele é reforçado em sequências de atitudes semelhantes, indicando que determinada ação tem maior probabilidade de ocorrer novamente (Northpointe, 2015), isto é, reincidências comportamentais podem ocorrer e seriam passíveis de previsão.

O COMPAS é utilizado para informar decisões sobre colocação, supervisão e gerenciamento de casos de pessoas que infringem a lei, com foco em fatores de risco e necessidades que afetam a reincidência (Northpointe, 2015). Consiste em escalas de risco preditivas e escalas de necessidades para identificar programas em áreas como emprego, moradia e abuso de substâncias, classificando o indivíduo nos seguintes níveis: Nível 1: interpretação básica, identificando áreas de necessidade; Nível 2: identificação de fatores criminogênicos inter-relacionados; e Nível 3: interpretação integrada, utilizando teorias criminológicas para explicar padrões de comportamento criminal (Northpointe, 2015, p. 04, tradução nossa). O sistema exerce diversas funções, entre elas a de sugerir o tipo de regime de cumprimento de pena ao qual o condenado deve ser submetido.

Um dos problemas apontados pelos críticos dessa “inteligência” é que os critérios e dados subjacentes à análise do comportamento do acusado não são acessíveis à defesa (Filpo; Gargano, 2025). Essa falta de transparência impede que o acusado compreenda plenamente as bases de sua condenação, configurando afronta ao devido processo legal e aos princípios fundamentais de justiça e igualdade aceitos em diversos países. No processo judicial, tal como regulamentado no Brasil, o juiz deve fundamentar suas decisões e explicitá-las, sob pena de torná-las frágeis por ausência de adequada fundamentação. Informar que determinada decisão foi sugerida pelo sistema, sem a devida explicitação dos critérios efetivamente utilizados, não atende a esse dever de fundamentação. Ademais, o juiz julga fatos ocorridos no passado, não lhe cabendo a atividade de prejulgamento, o que poderia implicar, inclusive, violação do dever de imparcialidade.

As políticas públicas que governam a implementação de tais ferramentas tecnológicas demandam aperfeiçoamentos, a fim de evitar a perpetuação de injustiças sociais, especialmente pela exclusão de grupos historicamente marginalizados do acesso ao conhecimento e à contestação de decisões que impactam suas vidas. É imperativo ampliar o debate e envolver mais atores nessa discussão: as Defensorias Públicas e a Advocacia,

por exemplo, para garantir a proteção dos direitos e assegurar um sistema de justiça mais equitativo e transparente.

Diversos estudos independentes examinaram a validade preditiva das escalas de risco do COMPAS em relação a grupos raciais e de gênero (Brennan, 2009). Esses estudos constataram que os modelos de reincidência apresentaram desempenho equivalente para homens afro-americanos e brancos na previsão de resultados de prisão em uma amostra de liberdade condicional. Um estudo anterior investigou a precisão preditiva do COMPAS para diferentes grupos étnicos, revelando resultados significativamente mais fracos para homens afro-americanos (Fass; Heilbrun; DeMatteo; Fretz, 2008). Esses resultados foram posteriormente questionados por outras pesquisas (Fass *et al.*, 2008), que consideraram a amostra muito pequena e desproporcional, o que comprometeria sua validade.

Vale mencionar um estudo realizado por Dressel e Farid (2018), que avaliaram o sistema COMPAS como ferramenta preditiva da probabilidade de reincidência, analisando dados inseridos (raça, histórico familiar de criminalidade, entre outros) e aplicando, a partir disso, diferentes níveis de pena. De modo geral, mesmo entre réus com antecedentes criminais, pessoas brancas foram avaliadas com menor risco de reincidência em comparação com pessoas negras.

A precisão geral do COMPAS para réus brancos é de 67,0%, apenas ligeiramente superior à precisão de 63,8% para réus negros (Dressel; Farid, 2018). Os erros cometidos pelo COMPAS, contudo, afetaram réus negros e brancos de maneira distinta: réus negros que não reincidiram foram incorretamente classificados como reincidentes em 44,9% dos casos, quase o dobro da taxa observada entre réus brancos (23,5%); por outro lado, réus brancos que reincidiram foram incorretamente classificados como não reincidentes em 47,7%, também quase o dobro da taxa verificada entre réus negros (28,0%) (Dressel; Farid, 2018).

Em síntese, segundo o estudo mencionado, as decisões pareceram impactar desproporcionalmente grupos vulneráveis. Trata-se de dados que colocam em discussão a eficácia da IA quando se busca construir um sistema de justiça mais justo e inclusivo, e não reprodutor de desigualdades sociais. Dessa forma, segundo Salomão e Tauk (2023), a experiência do COMPAS alerta para a importância de uma regulamentação cuidadosa e ponderada da inteligência artificial no Brasil, a fim de garantir o respeito aos direitos fundamentais e a realização da justiça de forma equitativa. Contudo, é importante ressaltar

que o caso COMPAS não é apresentado como realidade transplantável ao Brasil, mas como paradigma empírico de reflexão e alerta, para que não apenas suas potencialidades, mas especialmente suas limitações e possíveis efeitos discriminatórios, sejam considerados em um futuro marco regulatório.

REFLEXÕES NORMATIVAS

A análise do panorama brasileiro recortado neste artigo e da experiência norte-americana com o sistema COMPAS evidencia que a discussão acerca da inteligência artificial no sistema de justiça não pode ser reduzida a uma dicotomia simplista entre eficiência e atraso institucional. O que está em jogo não é apenas a modernização tecnológica do Judiciário, mas a preservação de garantias estruturantes do Estado Democrático de Direito diante de uma nova forma de racionalidade decisória: a racionalidade algorítmica.

Os dados empíricos apresentados demonstram que, embora a precisão estatística de determinados sistemas possa parecer satisfatória sob uma perspectiva quantitativa, a distribuição dos erros não se dá de maneira neutra. No caso do COMPAS, as taxas diferenciadas de falsos positivos e falsos negativos revelam que a automação decisória pode produzir impactos desproporcionais sobre grupos historicamente vulnerabilizados, como a população negra.

Tal constatação impõe reflexão acerca da distinção entre eficiência operacional e justiça material. Um sistema pode apresentar desempenho técnico adequado e, ainda assim, reproduzir desigualdades estruturais se os dados que o alimentam estiverem impregnados por padrões sociais discriminatórios. Nesse passo, cabe retomar a antiga distinção entre quantidade e qualidade das decisões judiciais (Filpo, 2016).

No contexto brasileiro, ainda que não se utilize formalmente modelo preditivo penal semelhante ao COMPAS, a incorporação crescente de sistemas de inteligência artificial aos fluxos processuais levanta questionamentos análogos. Ferramentas de triagem, classificação e organização de processos influenciam a forma como o magistrado acessa a informação e estrutura sua análise. Mesmo quando a decisão final permanece formalmente humana, o ambiente informacional previamente filtrado por algoritmos pode impactar o processo decisório. Assim, a discussão sobre enviesamento algorítmico não se

limita à substituição do juiz por uma máquina, mas alcança também os mecanismos de apoio invisíveis que moldam a racionalidade judicial.

A ausência de transparência quanto aos critérios técnicos utilizados nos sistemas implementados no Judiciário brasileiro agrava essa preocupação. A opacidade algorítmica, entendida como a impossibilidade de conhecer plenamente o funcionamento dos algoritmos, compromete o exercício do contraditório e da ampla defesa, pois impede que as partes compreendam e contestem eventuais influências automatizadas sobre o processo. No Brasil, a Constituição da República assegura o devido processo legal (art. 5º, LIV), o contraditório e a ampla defesa (art. 5º, LV) e a fundamentação das decisões judiciais (art. 93, IX), estabelecendo parâmetros normativos que não podem ser relativizados sob o argumento da eficiência tecnológica.

Além disso, a vedação ao retrocesso em matéria de direitos fundamentais impõe limite relevante à incorporação de novas tecnologias no âmbito jurisdicional. Se a adoção de sistemas algorítmicos implica redução concreta das garantias processuais ou aumento de assimetrias estruturais, estar-se-á diante de potencial regressão incompatível com a lógica constitucional. A inovação tecnológica, portanto, deve ser orientada por uma perspectiva de fortalecimento institucional, e não de substituição acrítica de garantias por métricas de desempenho, menos ainda de retrocesso em relação às conquistas constitucionais.

A reflexão sobre esses dados conduz à necessidade de formulação de políticas públicas de governança algorítmica que transcendam diretrizes genéricas e estabeleçam mecanismos efetivos de *accountability*, auditabilidade e explicabilidade (Doneda; Almeida, 2018). Não se trata de impedir o avanço tecnológico, mas de assegurar que a inteligência artificial atue como instrumento de aprimoramento da jurisdição, preservando a centralidade do julgamento humano e a individualização das decisões.

Em última análise, a experiência comparada e o panorama empírico brasileiro indicam que o verdadeiro desafio contemporâneo reside em harmonizar inovação tecnológica com compromisso constitucional. A inteligência artificial pode representar ferramenta valiosa para a racionalização do acervo judicial; contudo, sua implementação exige prudência institucional, controle democrático e constante monitoramento de impactos. Somente assim será possível evitar que algoritmos, concebidos sob a promessa de neutralidade, convertam-se em novos vetores de reprodução de desigualdades no sistema de justiça.

CONSIDERAÇÕES FINAIS

A análise desenvolvida ao longo deste estudo demonstra que a expansão dos meios tecnológicos no âmbito do Poder Judiciário constitui fenômeno estrutural e irreversível. A intersecção entre inteligência artificial e políticas públicas revela, contudo, que a incorporação de soluções algorítmicas não pode ocorrer dissociada do arcabouço constitucional que fundamenta o Estado Democrático de Direito. A inovação tecnológica deve ser compreendida como instrumento de aperfeiçoamento institucional, e não como fator de relativização das garantias processuais historicamente consolidadas.

A experiência empírica internacional, especialmente a análise do sistema COMPAS, evidencia que o enviesamento algorítmico pode produzir impactos desproporcionais sobre grupos historicamente vulnerabilizados. Ainda que a precisão estatística de determinados modelos pareça satisfatória sob perspectiva quantitativa, a distribuição desigual dos erros revela riscos concretos à igualdade material e ao devido processo legal, comprometendo a qualidade das decisões sob uma perspectiva principiológica. Tais dados reforçam a necessidade de revisão contínua das políticas públicas voltadas à governança algorítmica, de modo a assegurar transparência, responsabilidade institucional e respeito à dignidade da pessoa humana, conforme defendido pelos autores citados ao longo deste trabalho.

A realidade brasileira, embora distinta da norte-americana no que se refere ao uso de modelos preditivos penais, já evidencia a consolidação de ferramentas algorítmicas nos fluxos processuais. A pesquisa empírico-documental demonstrou que sistemas de triagem e classificação influenciam a estrutura informacional do processo judicial, o que exige atenção redobrada quanto à preservação do contraditório substancial e da fundamentação das decisões. A proteção dos direitos fundamentais deve permanecer como eixo central da formulação e implementação dessas tecnologias.

Nesse contexto, a transparência algorítmica emerge como requisito indispensável. É imperativo que haja clareza quanto aos critérios técnicos, às bases de dados e aos parâmetros utilizados nos sistemas empregados pelo Judiciário, bem como comunicação adequada às partes acerca do eventual uso de ferramentas automatizadas no suporte à decisão. A ausência de publicidade e explicabilidade compromete o controle democrático da atividade jurisdicional e fragiliza a legitimidade das decisões proferidas.

Não se ignora que as inovações tecnológicas podem contribuir significativamente para a redução do acervo processual e para a melhoria da eficiência administrativa do sistema de justiça. Em cenário marcado por elevada litigiosidade e escassez de recursos humanos, a inteligência artificial pode desempenhar papel relevante como instrumento de apoio à atividade jurisdicional, racionalizando tarefas repetitivas e ampliando a capacidade institucional de resposta.

Todavia, a substituição integral do julgamento humano por decisões automatizadas não se apresenta compatível com a complexidade das relações jurídicas e com a exigência constitucional de individualização das decisões. A atividade jurisdicional envolve apreciação contextualizada dos fatos, ponderação de princípios e consideração das especificidades das partes, dimensões que não podem ser integralmente capturadas por modelos algorítmicos baseados em padrões estatísticos.

Por conseguinte, impõe-se o enfrentamento da lacuna regulatória ainda existente quanto à auditabilidade, à accountability e à publicidade dos sistemas de inteligência artificial utilizados no Judiciário. A compreensão de como determinados atos processuais e decisões são produzidos constitui condição essencial para o exercício do contraditório e para a manutenção da confiança pública na Justiça.

Tal reivindicação está em conformidade com a Resolução n.º 332/2020 do Conselho Nacional de Justiça (CNJ), que dispõe sobre ética, transparência e governança na produção e no uso de inteligência artificial no Poder Judiciário, um importante ponto de partida, mas que ainda carece de aperfeiçoamentos na medida em que os tribunais se lançam, de forma crescente, à utilização de sistemas de IA. Espera-se que esses avanços sejam acompanhados de forma próxima pela advocacia, pelas demais funções essenciais à Justiça e pela academia, de modo a garantir um espaço democrático de debate sobre a temática, evitando que se torne um campo restrito a um decisionismo judicial.

As reformas nas políticas públicas de governança algorítmica mostram-se, portanto, fundamentais para assegurar que o sistema de justiça seja simultaneamente eficiente e equitativo. A proteção do devido processo legal em um ambiente tecnologicamente mediado exige abordagem multifacetada, que envolva transparência, controle institucional, formação adequada de magistrados e participação democrática de diferentes atores. Somente mediante tais salvaguardas será possível garantir que a inteligência artificial atue

como instrumento de fortalecimento da justiça, e não como mecanismo de reprodução de desigualdades estruturais.

REFERÊNCIAS

BAPTISTA, Bárbara Gomes Lupetti. A crença no princípio (ou mito) da imparcialidade judicial. **Revista de Estudos Empíricos em Direito**, v. 7, n. 2, p. 203-223, jun. 2020.

BOVENS, Mark. Analysing and assessing accountability: a conceptual framework. **European Law Journal**, v. 13, n. 4, p. 447-468, 2007.

BRENNAN, T.; DIETERICH, W.; EHRET, B. Evaluating the predictive validity of the COMPAS risk and needs assessment system. **Criminal Justice and Behavior**, v. 36, p. 21-40, 2009.

CONSELHO NACIONAL DE JUSTIÇA (CNJ). **Resolução nº 332, de 21 de agosto de 2020**. Brasília: CNJ, 2020.

COUTINHO, Luiza Loureiro. **Riscos em sistemas de inteligência artificial**: definição, tipologia, correlações, principiologia, responsabilidade civil e regulação. Salvador: Juspodivm, 2023.

CRESPO, Marcelo; SANTOS, Coroliano Aurélio de Almeida Camargo. Inteligência artificial, algoritmos e decisões injustas: é hora de revermos criticamente nosso papel em face da tecnologia. **Migalhas**, 2017. Disponível em: https://www.academia.edu/35023035/Intelig%C3%Aancia_artificial_algoritmos_e_decis%C3%B5es_injustas_C3%A9_hora_de_revermos_criticamente_nosso_papel_em_face_da_tecnologia_pdf. Acesso em: 29 jan. 2025.

DONEDA, Danilo; ALMEIDA, Virgílio A. F. O que é a governança de algoritmos. *In*: BRUNO, Fernanda *et. al.* **Tecnopolíticas da vigilância**: perspectivas da margem. Rio de Janeiro: Boitempo, 2018. p. 141-148.

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

EIERMANN, Martin. There is no devil's bargain between privacy and public health. **Foreign Affairs**, 2020. Disponível em: <https://www.foreignaffairs.com/articles/united-states/2020-04-13/there-no-devils-bargain-between-privacy-and-public-health>. Acesso em: 03 nov. 2024.

FASS, T.; HEILBRUN, K.; DEMATTEO, D.; FRETZ, R. The LSI-R and the COMPAS. **Criminal Justice and Behavior**, v. 35, p. 1095-1108, 2008.

FILPO, Klever Paulo Leal. **Mediação judicial**: discursos e práticas. Rio de Janeiro: Mauad X; FAPERJ, 2016.

FILPO, Klever Paulo Leal; GARGANO, Rayssa de Souza. Políticas públicas sobre inteligência artificial: refletindo sobre enviesamento algorítmico e proteção a direitos. *In: ENCONTROS NACIONAIS DO CONPEDI, XXXII CONGRESSO NACIONAL DO CONPEDI. Anais [...].* São Paulo: CONPEDI, 2025.

KAUFMAN, Dora. **A inteligência artificial irá suplantará a inteligência humana?** Barueri, SP: Estação das Letras e Cores, 2019.

MOSES, Lyria Bennett. How to think about law, regulation and technology: problems with ‘technology’ as a regulatory target law. **Innovation and Technology**, v. 5, n. 1, p. 1-20, 2013. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2464750. Acesso em: 03 out. 2024.

NORTHPOINTE. **Practitioner’s guide to COMPAS core.** 2015. Disponível em: <https://s3.documentcloud.org/documents/2840784/Practitioner-s-Guide-to-COMPAS-Core.pdf>. Acesso em: 22 out. 2024.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). Recommendation of the Council on Artificial Intelligence. **OECD Legal Instruments**, 2024. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em: 17 maio 2024.

PINTO, Henrique Alves. **A tripla fundamentação das decisões jurisdicionais pautadas em inteligência artificial.** Salvador: Juspodivm, 2023.

RIZER, Arthur; WATNEY, Caleb. Artificial intelligence can make our jail system more efficient, equitable, and just. **Texas Review of Law & Politics**, v. 23, n. 1, 2018. Disponível em: <http://www.questia.com/read/1P4-2193091511/artificial-intelligence-can-make-our-jail-system-more>. Acesso em: 10 set. 2024.

SANTORO, Antonio Eduardo Ramires; PAIVA, Livia de Meira Lima; MELCHIOR, Antonio Pedro. O princípio da imparcialidade na interceptação das comunicações telefônicas. *In: SANTORO, Antonio Eduardo Ramires; MADURO, Flávio Mirza (org.). Interceptação telefônica: os 20 anos da Lei nº 9.296/96.* Belo Horizonte: D’Plácido, 2016.

SARLET, Ingo Wolfgang. Notas sobre a assim designada proibição de retrocesso social no constitucionalismo latino-americano. **Revista do Tribunal Superior do Trabalho**, Brasília, v. 75, n. 3, jul./set. 2009.

SOFTPLAN PLANEJAMENTO E SISTEMAS S.A. **SAJ Digital.** Disponível em: <http://www.sajdigital.com/pesquisa-desenvolvimento/exemplos-inteligencia-artificial/>. Acesso em: 14 nov. 2024.

TAUK, Caroline Somesom; SALOMÃO, Luis Felipe. Inteligência artificial no Judiciário brasileiro: estudo empírico sobre algoritmos e discriminação. **Revista Diké (UESC)**, v. 22, n. 23, p. 2-32, jan./jun. 2023.

Acesso geral: <https://ojs.uva.br/index.php/revista-aquila>

Licença Copyleft: Atribuição-Usos não Comerciais-Vedada a Criação de Obras Derivadas

