

I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL III

PAULO CEZAR DIAS

VALTER MOURA DO CARMO

FERNANDO GALINDO AYUDA

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

I61

Inteligência Artificial: Desafios da Era Digital III [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Fernando Galindo Ayuda, Paulo Cezar Dias, Valter Moura do Carmo. – Florianópolis: CONPEDI, 2025.

Inclui bibliografia

ISBN: 978-65-5274-097-7

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

Tema: Inteligência Artificial e Sustentabilidade na Era Transnacional

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Internacionais. 2. Inteligência Artificial. 3. Desafios da Era Digital. I

International Experience Perugia – Itália. (1: 2025 : Perugia, Itália).

CDU: 34



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

I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL III

Apresentação

APRESENTAÇÃO

O I International Experience – Perúgia – Itália foi realizado nos dias 28, 29 e 30 de maio de 2025, com o tema "Inteligência Artificial e Sustentabilidade na Era Digital". O Grupo de Trabalho (GT) "Inteligência Artificial: Desafios da Era Digital III" ocorreu nos dias 29 e 30 de maio, nos períodos vespertinos, na Universidade de Perúgia.

O GT destacou-se não apenas pela qualidade dos trabalhos apresentados, mas também pelo nível acadêmico dos autores — doutores, mestres, professores pesquisadores e seus alunos pós-graduandos. O evento também proporcionou um importante espaço de interlocução internacional, contando com a participação de renomados juristas e professores de instituições estrangeiras, como os Professores Doutores Roberto Cippitani (Universidade de Perúgia) e Fernando Galindo (Universidade de Zaragoza – Espanha), que enriqueceram os debates e contribuíram para o sucesso da atividade.

Foram apresentados 15 (quinze) artigos, os quais foram objeto de intenso debate presidido pelos coordenadores e enriquecido pela participação ativa do público presente na Faculdade de Direito de Perúgia – ITÁLIA.

A apresentação dos trabalhos permitiu discussões atualizadas e profícuas sobre temas como inteligência artificial, uso de dados pessoais, dever de informação, riscos e interações tecnológicas. As abordagens trataram dos desafios enfrentados pelas diversas linhas de pesquisa jurídica no estudo do futuro da regulação no Brasil, dos abusos relacionados à inteligência artificial e das possíveis soluções para a proteção de dados em um mundo globalizado.

As temáticas incluíram: tecnologias relacionadas a fake news, deepfakes e bots; compliance; a consideração do elemento humano na aplicação da I.A. nas decisões judiciais; a inteligência artificial como ferramenta de proteção no sistema de justiça criminal; o consentimento informado e o uso de dados pessoais; regulamentação e governança da I.A.; precarização do governo digital e aplicação da inteligência artificial em distintos setores jurídicos.

A seguir, apresenta-se a relação dos trabalhos que compõem este Grupo de Trabalho, acompanhados de seus respectivos autores:

1. CAPACIDADE ARTIFICIAL DAS MÁQUINAS E A EXIGÊNCIA DE TRANSFORMAÇÕES NA MANEIRA DO SABER DE PROFISSIONAIS, de Fernanda Conceição Pohlmann.
2. AI, VOCÊ ESTÁ AÍ? O PANORAMA JURÍDICO RELATIVO À (AUTO) IDENTIFICAÇÃO DA INTELIGÊNCIA ARTIFICIAL, de Gabriel Siqueira Eliazar de Carvalho, André Fortes Chaves e Marcello Silva Nunes Leite.
3. DEMOCRACIA EM REDE: O PAPEL DA INTELIGÊNCIA ARTIFICIAL E DOS ALGORITMOS NA LIBERDADE DE EXPRESSÃO E NO PLURALISMO POLÍTICO, de Kennedy da Nobrega Martins, Alexandre Manuel Lopes Rodrigues e Jadgleison Rocha Alves.
4. INTELIGÊNCIA ARTIFICIAL E DIREITOS FUNDAMENTAIS: DESAFIOS E TENSÕES NA ERA DIGITAL, de Jesualdo Eduardo de Almeida Junior e Gustavo Roberto Dias Tonia.
5. INTELIGÊNCIA ARTIFICIAL E DEMOCRACIA: O PERIGO DA MANIPULAÇÃO DE INFORMAÇÕES, de Claudia Maria da Silva Bezerra e Luiz Eduardo Simões de Souza.
6. INFLUÊNCIAS DO REALISMO JURÍDICO E O USO DA INTELIGÊNCIA ARTIFICIAL NA ELABORAÇÃO DE DECISÕES JUDICIAIS NO BRASIL: VIESES COGNITIVOS E HEURÍSTICAS NO PROCESSO DECISÓRIO, de Kerry Barreto, Fausto Santos de Moraes e Júlia Regina Bassani Caus.
7. CRITÉRIOS QUANTITATIVOS PARA A MENSURAÇÃO DE RESULTADOS NO JUÍZO 100% DIGITAL: RISCOS PARA A QUALIDADE DA PRESTAÇÃO JURISDICIONAL NO BRASIL, de Orides Mezzaroba, José Renato Gaziero Cella e Lia Loana Curial Oliva.
8. AS PROVAS DIGITAIS NO PROCESSO CIVIL E O (DES)CABIMENTO DA CADEIA DE CUSTÓDIA, de Jesualdo Eduardo de Almeida Junior e Gustavo Roberto Dias Tonia.
9. A REVOLUÇÃO DA INTELIGÊNCIA ARTIFICIAL NOS GABINETES JUDICIAIS: EFICIÊNCIA COM GARANTIAS CONSTITUCIONAIS, de Lisbino Geraldo Miranda do Carmo, Deise Neves Nazaré Rios Brito e Jimmy Souza do Carmo.

10. GENEALOGIA E INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL PARA ELABORAÇÃO DE UM ONOMÁSTICO DOS IMIGRANTES ITALIANOS QUE DESENVOLVERAM O SUL DO ESTADO DE SANTA CATARINA DE 1877 A 1897, de Júlio Cesar Cancellier de Olivo.

11. A REDE-LAB COMO INOVAÇÃO NA POLÍTICA ANTILAVAGEM DE CAPITAIS NO BRASIL, de Lorryne Souza Galli e Matheus Felipe de Castro.

12. ARMAS AUTÔNOMAS LETAIS: OS IMPACTOS DA INTELIGÊNCIA ARTIFICIAL PARA OS DIREITOS HUMANOS E SUA CONSEQUENTE REGULAMENTAÇÃO, de Alexandre Gonçalves Ribeiro e Renata Mantovani de Lima.

13. A INTELIGÊNCIA ARTIFICIAL COMO FERRAMENTA ESSENCIAL NA ELUCIDAÇÃO DE CRIMES SEXUAIS PRATICADOS COM VIOLÊNCIA CONTRA A MULHER, de Eneida Orbage de Britto Taquary e Catharina Orbage de Britto Taquary Berino.

14. A INTELIGÊNCIA ARTIFICIAL NO DIREITO PENAL: AVANÇOS, DESAFIOS E IMPACTOS NA INVESTIGAÇÃO E NO SISTEMA JUDICIAL, de Eneida Orbage de Britto Taquary, Bianca Cristina Barbosa de Oliveira e Tiago de Lima Mascarenhas Santos.

15. ENTRE CÓDIGOS E DIREITOS: UMA ANÁLISE CONSTITUCIONAL DA PROTEÇÃO DE DADOS PESSOAIS NO CONTEXTO DA INTELIGÊNCIA ARTIFICIAL, de Lisbino Geraldo Miranda do Carmo, Deise Neves Nazaré Rios Brito e Paulo Henrique da Silva Costa.

Por fim, os organizadores e coordenadores do Grupo de Trabalho "Inteligência Artificial: Desafios da Era Digital III" parabenizam e agradecem aos autores pelos valiosos trabalhos apresentados, cuja leitura certamente contribuirá para o aprofundamento do debate acadêmico e científico na área.

Prof. Dr. Fernando Galindo - Universidad de Zaragoza - Espanha

Prof. Dr. Valter Moura do Carmo – PPGPJDH - ESMAT e UFT

Prof. Dr. Paulo Cezar Dias – Centro Universitário Eurípides de Marília - SP

**INFLUÊNCIAS DO REALISMO JURÍDICO E O USO DA INTELIGÊNCIA
ARTIFICIAL NA ELABORAÇÃO DE DECISÕES JUDICIAIS NO BRASIL: VIESES
COGNITIVOS E HEURÍSTICAS NO PROCESSO DECISÓRIO**

**INFLUENCES OF LEGAL REALISM AND THE USE OF ARTIFICIAL
INTELLIGENCE IN MAKING JUDICIAL DECISIONS IN BRAZIL: COGNITIVE
BIASES AND HEURISTICS IN THE DECISION MAKING PROCESS**

Kerry Barreto ¹
Fausto Santos de Moraes ²
Júlia Regina Bassani Caus ³

Resumo

O uso da Inteligência Artificial (IA) na elaboração de decisões judiciais no Brasil pode ser benéfica, pois dá mais agilidade e precisão na análise processual. Contudo, suscita preocupações em relação à manutenção das “sugestões” de máquina equivocadas, injustas ou discriminatórias, em decorrência de vieses cognitivos e heurísticas humanas quanto de sua apreciação. O objetivo geral deste artigo é analisar o impacto dos vieses cognitivos e das heurísticas no uso da IA no Poder Judiciário brasileiro, questionando como a IA pode afetar a racionalidade jurídica das decisões. Para tanto, foram delineados três objetivos específicos: (a) investigar a perspectiva do realismo jurídico norte-americano sobre o processo decisório, (b) compreender o uso da IA no Poder Judiciário brasileiro, e (c) identificar problemas relacionados ao uso da IA, como vieses de máquina e o viés humano confirmatório. A pesquisa é bibliográfica, com uso do método dedutivo. Desse modo, é imprescindível garantir que a implementação da IA pelo Poder Judiciário seja concretizada com rigor ético e técnico, cientes dos julgadores da necessidade de tomarem decisões de modo racional.

Palavras-chave: Vieses cognitivos, Heurística, Inteligência artificial, Poder judiciário, Decisão judicial

Abstract/Resumen/Résumé

The use of Artificial Intelligence (AI) in the preparation of judicial decisions in Brazil can be

“suggestions”, due to cognitive biases and human heuristics regarding their assessment. The general objective of this article is to analyze the impact of cognitive biases and heuristics on the use of AI in the Brazilian Judiciary, questioning how AI can affect the legal rationality of decisions. To this end, three specific objectives were outlined: (a) investigate the perspective of North American legal realism on the decision-making process, (b) understand the use of AI in the Brazilian Judiciary, and (c) identify problems related to the use of AI, such as machine biases and human confirmatory bias. The research is bibliographic, using the deductive method. Therefore, it is essential to ensure that the implementation of AI by the Judiciary is carried out with ethical and technical rigor, with judges aware of the need to make rational decisions.

Keywords/Palabras-claves/Mots-clés: Cognitive biases, Heuristics, Artificial intelligence, Judiciary, Legal decision

INTRODUÇÃO

O uso da inteligência artificial (IA) é cada vez mais comum em todos os aspectos da vida humana. Especificamente na elaboração de decisões judiciais no Brasil, ela tem se apresentado como uma ferramenta promissora, inclusive para otimizar o trabalho, oferecendo maior agilidade e precisão na análise dos processos. Dessa forma, a IA vem ajudando magistrados e advogados na triagem dos processos, na pesquisa de jurisprudência e na própria identificação e parametrização de padrões nas decisões. Contudo, seu uso levanta alguns questionamentos e reflexões no sentido da preservação da racionalidade jurídica e preocupação em relação a vieses cognitivos e heurísticas nas decisões.

Dessa forma, o objetivo geral do presente artigo é analisar, com base na linha idiossincrática da teoria realista do direito e em estudos sobre o processo mental de tomada de decisões, as implicações do uso de inteligência artificial generativa (IAG) na elaboração de decisões judiciais.

A partir do objetivo geral, foram delineados três objetivos específicos: 1. analisar a perspectiva do realismo jurídico norte-americano, mais especificamente, a sua linha idiossincrática, sobre o processo de tomada de decisões pelo judiciário, integrando os avanços das pesquisas da neurociência e da psicologia cognitiva, especificamente no que se refere ao funcionamento do Sistema 1 e Sistema 2 da mente humana e aos vieses cognitivos e às heurísticas envolvidos; 2. investigar o uso da IA pelo Poder Judiciário brasileiro, compreendendo suas aplicações, benefícios e desafios no processo de tomada de decisões; e 3. identificar e analisar os principais problemas no uso da IA no Poder Judiciário, como o “engessamento” do pensamento humano, os vieses de máquina, as limitações das heurísticas de programação, as alucinações de IA e o viés humano confirmatório, além de avaliar as medidas propostas para mitigar os riscos do uso da IA no processo de tomada de decisão.

O presente estudo é elaborado com base em pesquisa bibliográfica e o método empregado é o dedutivo.

Acredita-se que a corrente do realismo jurídico norte-americana contribui em demonstrar que fatores extrajurídicos estão presentes na tomada de decisões judiciais, entre os quais, atualmente, se constata as “sugestões” fornecidas por ferramentas de IA. A linha idiossincrática do realismo, atrelada aos mais atuais estudos de neurociência e psicologia cognitiva, propicia embasamento teórico para testes empíricos sobre vieses cognitivos confirmatórios das “sugestões” da IA. Ainda, se verifica que o uso da IA pelo Poder Judiciário

mostra-se como uma realidade recente e demandará especial atenção dos operadores do direito. A implementação da IA no Poder Judiciário requer que seja acompanhada de um rigoroso controle ético e técnico, propiciando que a tecnologia realmente seja utilizada de maneira mais justa, transparente e capaz de promover uma justiça imparcial e mais equitativa.

1. O REALISMO JURÍDICO NORTE-AMERICANO E O PROCESSO DE TOMADA DE DECISÕES

Como contraponto à teoria formalista do direito, diversos juristas norte-americanos no início do século XX desenvolveram teorias que se alicerçavam em elementos factuais para explicar aspectos do fenômeno jurídico. Shecaira e Struchiner (2016, p. 05) apontam que todos compartilham um “espírito antimetafísico”. Já Billier e Maryioli (2005) classificam a grande profusão de proposições do chamado realismo norte-americano em quatro diferentes teses: 1. tese da concepção instrumentalista ou funcionalista do direito; 2. tese da indeterminação da indeterminação das regras jurídicas; 3. tese do behaviorismo ou comportamentalismo e, 4. tese do anticonceptualismo. Com relação aos estudos da decisão judicial, Shecaira e Struchiner (2016, p. 01) consideram que, ao contrário das teorias do direito, as teorias da decisão judicial procuram esclarecer “[...] como os juízes trabalham, isto é, como eles raciocinam, que tipo de material eles utilizam para a construção das suas decisões etc”. Seguindo este parâmetro, em que pese a falta de homogeneidade não permitir enquadrar as diversas teses do realismo norte-americano como uma teoria jurídica ou mesmo se extrair de seus pensamentos uma teoria da decisão judicial realista, Shecaira e Struchiner (2011) acreditam que alguns dos componentes daquela corrente buscaram formular uma teoria descritiva da decisão judicial.

De acordo com Brando (2013), Oliver Wendell Holmes Jr., componente da Suprema Corte Norte-Americana no século XIX, é apontado como aquele de quem surgiu o embrião do realismo jurídico, apresentando um posicionamento que rompia com o formalismo que imperava em sua época, algo a que os realistas apontam como formalismo jurídico, jurisprudência mecânica ou teoria formalista do direito. Segundo Holmes Jr. (1881, *apud* Brando, 2013), ao apreciar como um julgador chega a suas conclusões, seria necessário considerar sua experiência profissional, seus conhecimentos extralegais e suas opções políticas; e o resultado seria a criação de novas normas jurídicas, estabelecendo um novo precedente, não vinculado necessária ou logicamente com o direito então em vigor. Sua visão se opunha ao entendimento formalista de que as mudanças jurídicas na *Common Law* decorriam de um

procedimento lógico e dedutivo pelo qual um novo precedente apenas aclararia precedentes anteriores ou outras normas já contidas no arcabouço jurídico.

A obra de Holmes Jr., contudo, não é enquadrada dentro do realismo jurídico. Os intelectuais desenvolvedores daquela corrente de pensamento concordavam com ele que das normas e precedentes não se poderia extrair diretamente o conteúdo de decisões judiciais, mas dissentiam em relação ao que efetivamente leva à conclusão do julgador. De igual turno, excetuada a posição extremada de Jerome Frank, os realistas norte-americanos não defendiam uma indeterminação absoluta do direito e uma total imprevisibilidade das conclusões decorrentes da atividade jurisdicional (Brando, 2013).

No que diz respeito ao processo de tomada de decisão, para Leiter (2010), os realistas se dividem em duas linhas de pensamento: 1. a idiossincrática, que defende que o juiz chega à sua conclusão através de um palpite ou uma intuição (*hunch*) sobre o que seria o resultado mais justo para o caso concreto, formulando posteriormente uma justificativa (*post hoc*) para a decisão; e 2. a sociológica, que argumenta que as decisões judiciais embora decidam o caso concreto de forma relativamente semelhante e previsível, são influenciadas por forças sociais.

Nessa perspectiva, Schauer (2009) defende que de ambas as linhas (idiossincrática e sociológica) se poderia extrair a existência de duas proposições: 1. a maioria dos juízes tem uma inclinação para chegar a um determinado resultado antes de consultar os materiais jurídicos; e, 2. o juiz raramente terá dificuldade de encontrar a roupagem jurídica para a decisão tomada no vasto, complexo e amplamente indeterminado universo de materiais jurídicos existentes.

O entendimento da linha idiossincrática do realismo parece ser o melhor instrumental teórico a ser utilizado para responder à indagação de como os juízes decidem. Segundo tal corrente, o julgador, ao apreciar o caso concreto, chegaria inicialmente a um palpite ou intuição de forma automática, num passo seguinte elaborando o texto da decisão judicial, sentença ou voto com elementos de uma racionalização *post hoc* (Frank, 1930/2009).

Obtempera Hart (1983) que as normas jurídicas expressas em regras seriam internalizadas tanto por cidadãos, quanto por autoridades oficiais, que as observariam de forma automática ou para evitar um litígio judicial que a parte considera que perderia. Nesse sentido, os casos levados aos tribunais seriam representativos de apenas uma parcela dos assim chamados “casos difíceis”, surgindo como resultado de um “efeito seletivo”. Por outro lado, Hart (1983) reconhece que a linguagem, em alguns casos, pode apresentar uma vagueza potencial em decorrência de sua textura aberta. Assim, mesmo uma regra jurídica descreva de

forma clara uma conduta, o contexto pode tornar impreciso o direcionamento sobre como agir. Neste cenário, segundo Hart (1983), abrir-se-ia um campo de discricionariedade para o juiz.

Em síntese, essas reflexões fazem questionar a natureza do direito e contextualizam a sua constante evolução, refletindo a ideia de que o direito é moldado tanto por fatores internos das normas quanto por influências externas, o que desafia a ideia de um sistema jurídico fechado.

Schauer (2009) aduz que não se pode ignorar a importância da visão sugerida pelo realismo jurídico, que defende uma apreciação do processo de tomada de decisão judicial a partir de uma perspectiva externa (extrajurídica), sujeita a um rigoroso teste empírico. A hipótese das duas proposições comuns aos teóricos do realismo deve, então, ser revisitada, não apenas em situações de decisões judiciais tomadas por humanos, como também nos casos de juízes e tribunais se valendo de ferramentas de IA preditiva ou generativa como auxiliar na análise do caso concreto ou mesmo na elaboração do texto decisório, questão que será tratada neste artigo.

Sobre o tema da atividade judicial sob um enfoque dos teóricos do realismo jurídico, Horta e Costa (2017) constataram que quanto maior a margem de discricionariedade do julgador, maior é a aferição de influências extrajurídicas, o que pode ser apurado não apenas nos “casos difíceis”, mas também nos casos considerados “fáceis” ou corriqueiros, pelo que não será levada em consideração, neste trabalho, a crítica de Hart de que o realismo traria tese relevante apenas no que concerne ao julgamento dos “casos difíceis”.

Ao refletir sobre o realismo e suas implicações, fica claro que a análise das decisões não pode ficar restrita às normas formais, devendo levar em conta os fatores extrajudiciais que podem tender a decisão dos magistrados. Assim, a introdução da IA levanta novas questões no quesito de como essas influências externas podem impactar na (im)parcialidade e na consistência das decisões.

No ano de 2021, a Resolução nº 423, do CNJ, alterou a Resolução nº 75, que trata dos concursos para a magistratura em âmbito nacional através da Resolução nº 423, de 2021, para incluir como temas obrigatórios para as provas, entre outros, economia comportamental, heurística e vieses cognitivos, a percepção de justiça e processo cognitivo de tomada de decisão (CNJ, 2021).

Em apoio a esta mudança, Horta (2021, n.p.) sustentou que “o ‘justo’ é frequentemente o auto-interesse remodelado como um bom discurso de justificação”, e que “[p]sicólogos sabem há décadas que temos um “ponto cego” para nossos próprios vieses (*bias blind spot*)”.

Juízes, desembargadores, ministros de tribunais superiores, juízes auxiliares, assessores e oficiais de gabinetes, ou seja, todos os envolvidos na elaboração das decisões judiciais, sentenças ou acórdãos, compartilham de uma característica fundamental: são seres humanos. Por mais que busquem se manter racionais e concentrados ao analisar petições, documentos, perícias, provas testemunhais, depoimentos e quaisquer outros documentos, legislação, jurisprudência e obras doutrinárias, os julgadores, como qualquer ser humano, estão sujeitos a influências que escapam do seu controle, sendo levados a conclusões que surgem de forma autônoma e automática. O pensamento puramente racional é um engodo e todos tendem a acreditar ter pleno controle de suas deliberações mentais, quando, na realidade, são impactados por variados processos cognitivos que funcionam além da percepção consciente.

Nessa perspectiva, relevante que se analise não apenas como a mente humana funciona, mas os impactos aferíveis de vieses de percepção nos julgamentos jurídicos. Kahneman (2012) desenvolveu uma série robusta de estudos empíricos que elucidam como a mente humana desenvolve o pensamento e chega a conclusões. O autor percebeu que o ser humano não atua apenas racionalmente. Na verdade, ele acredita que a mente racional estaria a todo momento sob a ilusão de ter o controle dos próprios pensamentos, opiniões e ações.

Como aponta Kahneman (2012, n.p.), os vieses são erros sistemáticos produzidos pela mente humana e que podem ser constatados de forma empírica, pois “se repetem de forma previsível em circunstâncias particulares”. A todo instante, o cérebro humano gera impressões e intuições, criando pensamentos e tomando decisões silenciosas em nossas cabeças, mas não nos damos conta disto. Assim, Kahneman (2012) aduz que, quando nos deparamos com questões complexas e difíceis, respondemos automaticamente a uma questão mais fácil, sem nem ao menos percebermos está substituição, e para esse fenômeno damos o nome de heurística intuitiva.

A vida mental do ser humano, segundo Kahneman (2012) pode ser descrita como um conjunto de dois sistemas que trabalham em conjunto: um sistema rápido, chamado de Sistema 1; e um lento, chamado de Sistema 2. O Sistema 1 atua com rapidez e de modo automático, solicitando pouco ou nenhum esforço mental, sem haver qualquer percepção de um controle voluntário do indivíduo; já o Sistema 2 requer atenção e é chamado a atuar sempre que há necessidade de uma atividade mental mais trabalhosa, como no caso de cálculos complexos (Kahneman, 2012). Desse modo, o autor entende que quando utilizado este último sistema, tem-se uma percepção subjetiva de que o ser humano se concentra e exerce uma atividade ou uma

escolha. Apesar de o ser humano acreditar que está sempre atuando de forma assertiva, é o sistema rápido e automático que o leva a tomar a maior parte das suas decisões.

O normal é que o Sistema 1 gere intuições, bem como uma solução especializada ou uma resposta heurística, que são confirmadas pelo Sistema 2. Em alguns casos, o ser humano necessita se concentrar para chegar a respostas usando um meio mais lento, deliberado e trabalhoso de pensar, ou seja, usando o Sistema 2. Como esclarece Kahneman (2012, n.p.), “somos inclinados a superestimar quanto compreendemos sobre o mundo e subestimar o papel do acaso nos eventos”, decorrendo está confiança em nossas percepções do mundo e julgamentos pessoais de uma “certeza ilusória da percepção tardia”. O Sistema 1 sugere pensamentos e ações a todo momento, e uma função primordial do Sistema 2 é justamente monitorar e controlar tais sugestões, seja consentindo que o Sistema 1 expresse diretamente seu comportamento, seja suprimindo ou modificando as manifestações do Sistema 1 (Kahneman, 2012).

As operações da memória associativa – uma vasta rede de ideias, dispostas como nódulos e ligadas umas às outras, desenvolvidas pelo Sistema 1, geram uma hipótese e conduzem a um “viés de confirmação” (*confirmation bias*). O Sistema 2, ao ser confrontado com tal hipótese, desenvolve uma busca por evidências confirmadoras, algo a que se dá o nome de estratégia de teste positivo (*positive test strategy*). O Sistema 2 testa a hipótese não tentando a refutar, mas procurando dados e informações que mais provavelmente se apresentem como compatíveis com as crenças que se tenha no momento. É como se o Sistema 1 “enganasse” o Sistema 2, com seu viés de confirmação, criando uma condição favorável à aceitação das sugestões por ele produzidas. O Sistema 2 acaba por exagerar na probabilidade de ocorrência de eventos extremos e improváveis, em consequência de tal viés (Kahneman, 2012, n.p.).

O processo decisório humano, seja consciente ou inconsciente, ocorre de forma constante. Desse modo, as decisões podem variar desde escolhas simples e automatizadas, até decisões mais complexas e deliberadas. Embora influenciadas por fatores racionais e estruturados, essas decisões também são permeadas por elementos intuitivos e instintivos que operam no nível inconsciente. Além das evidências factuais disponíveis, o processo decisório humano é fortemente influenciado por aprendizados prévios e experiências acumuladas. Entretanto, a limitação da capacidade de raciocínio humano impede que todas as variáveis de um caso sejam analisadas de forma racional e exaustiva (Korteling; Toet, 2020).

Para contornar essas limitações, Korteling e Toet (2020) sintetizam que o ser humano desenvolveu heurísticas, ou seja, mecanismos que funcionam como atalhos mentais para

facilitar a tomada de decisões de maneira mais ágil. Embora úteis em situações específicas, as heurísticas também são suscetíveis a aplicações inadequadas, o que pode gerar vieses cognitivos e frequentemente comprometem a precisão de julgamentos que exigem elevado rigor analítico (Korteling; Toet, 2020).

Além disso, os vieses cognitivos também podem exercer uma influência significativa sobre o raciocínio jurídico, como de resto em todo processo de tomada de decisão pelo ser humano. Uma forma como o desvio cognitivo se apresenta é o viés de confirmação, caracterizado pela busca e interpretação seletiva de informações que reforçam crenças prévias e compromete a imparcialidade dos julgamentos, conduzindo à subestimação de dados contraditórios, ou seja, o sujeito dará maior valor às suas ideologias e crenças sobre como as coisas funcionam do que às evidências que lhe são apresentadas (Tabak; Amaral, 2018). Sobre o tema, Andrade (2019) aponta que novas informações a que o sujeito tenha acesso serão por ele valorizadas de maneira diferente conforme confirmem ou não o que ele pensa.

Kahneman (2012) esclarece que o viés de ancoragem se reflete na tendência de decisões subseqüentes serem influenciadas desproporcionalmente por informações iniciais, mesmo que irrelevantes. No Direito, de modo peculiar, esse viés da ancoragem é notável em avaliações judiciais que se baseiam em argumentos iniciais das partes ou em recomendações preliminares fornecidas por sistemas de IA que realizam predições de reincidência. Esses sistemas podem oferecer uma âncora cognitiva que limita a capacidade de revisão crítica pelos operadores jurídicos, especialmente quando estes não possuem formação técnica suficiente para avaliar as limitações dessas ferramentas (Kahneman, 2012).

Faz-se necessária uma abordagem reflexiva e crítica do Direito brasileiro para que as transformações tecnológicas não comprometam a legitimidade do sistema judicial. O estudo das heurísticas e dos vieses cognitivos no Direito é indispensável para compreender as potencialidades e os desafios impostos pela interação entre tecnologia e governança jurídica. O enfrentamento dessas questões exige um compromisso ético com os valores fundamentais do Estado Democrático de Direito, além de uma abordagem interdisciplinar que integre as contribuições das ciências humanas e sociais.

2. USO DE INTELIGÊNCIA ARTIFICIAL (IA) COMO FERRAMENTA AUXILIAR PARA A ELABORAÇÃO DE DECISÕES JUDICIAIS NO BRASIL

Não existe um consenso sobre o conceito de IA. Russel e Norvig (2013) abordam o tema sob quatro perspectivas: 1. pensar como humano, simulando as funções cognitivas; 2. agir como humano, imitando seus comportamentos e interações; 3. pensar racionalmente, utilizando-se de princípios como a lógica formal; e 4. agir racionalmente, ou seja, atuar de forma autônoma e racional, com base em objetivos e visando a maximização de resultados. No quesito assimilação de conteúdo, tanto Searle (1998), quanto Penrose (1994) esclarecem que, do ponto de vista da semiótica, a linguagem opera em três níveis: sintático, semântico e pragmático, sendo que os algoritmos são desenvolvidos de forma limitada apenas aos dois primeiros, pelo que os que existem atualmente chegam a uma forma de entendimento e não de compreensão.

Visualizando o contexto atual, salienta-se que o ser humano ainda não vive e nem experimenta uma realidade de IA forte, com sistemas capazes de realizar tarefas cognitivas de modo semelhante à mente humana, ou seja, com consciência, compreensão, capacidade de aprendizado geral, criatividade e total domínio da linguagem natural em seus três níveis. O que temos é somente a chamada IA fraca, básica (Boeing; Rosa, 2020).

O que capacita máquinas a gerar textos e interagir com a linguagem humana são algoritmos desenvolvidos com técnicas de Processamento de Linguagem Natural (*Natural Language Processing* – NLP). Basicamente, o que algoritmos fazem é calcular estatisticamente qual a aproximação de um termo ou expressão a partir de ocorrências num determinado contexto. Tendo por ponto inicial a agregação de vários contextos, o algoritmo apreende características do termo considerado, as quais permitiriam chegar a similaridades de significado entre palavras e entre conjuntos de palavras (Russel; Norvig, 2013). O que entendemos como parecido ou diferente, para a máquina seria a probabilidade maior ou menor de o termo aparecer num contexto. Ao conjunto de técnicas de NLP que transformam palavras, expressões e frases em vetores matemáticos com representação numérica dá-se o nome de *word embedding*. Em outras palavras, o *word embedding* seria uma forma de facilitar o aprendizado de máquina, através da representação dos dados de forma mais eficiente, se valendo de vetores (Boeing; Rosa, 2020).

Então, embora a inteligência artificial generativa (IAG) possa produzir textos com base em comandos que lhe são fornecidos, considerando dados colocados à sua disposição, ela não é capaz de um raciocínio como o cérebro humano, isto é, não tem compreensão e criatividade,

em que pese o resultado fornecido se assemelhar em muito ao que seria criado por um humano. Como pondera Cathy O’Neil (2021, n.p.), “computadores, a despeito de todos os seus avanços em linguagem e lógica, ainda têm muitos problemas com conceitos” e, ao trabalharem com termos abstratos, o máximo que conseguem é se utilizar de aproximações.

Boeing e Rosa (2020), em “Ensinando um Robô a Julgar”, trabalham tanto com a possibilidade de a máquina assumir o caráter de decisor, quanto ser meramente uma ferramenta à disposição do juiz humano, hipótese que tratam como “corrobótica”. Os referidos autores apontam três tipos de aprendizado de máquina que podem ocorrer no âmbito do Poder Judiciário: 1. robô-classificador, 2. robô-relator e 3. robô-julgador.

O robô-classificador, do qual o projeto Victor, do Supremo Tribunal Federal, seria exemplo, teria a função de localizar materiais úteis para que julgadores humanos fundamentem suas decisões (Boeing; Rosa, 2020, p. 95).

Quanto ao robô-relator, Boeing e Rosa (2020, p. 98) aduzem que ele seria utilizado para “extrair e condensar informações relevantes de um ou mais documentos”. Para os autores supramencionados, a sua capacidade seria não apenas de localizar documentos similares, mas também diferenciar, em cada peça processual, o que diz respeito à descrição dos fatos, legislação, jurisprudência e estruturas argumentativas, podendo inclusive elaborar decisões.

Tanto no caso do robô-relator, quanto do robô-classificador, para Boeing e Rosa (2020) seria apenas ferramentas auxiliares ao juiz humano no processo decisório, ao qual ainda caberia os elementos finais de redação (ou talvez refazer por completo o texto). Por conta do viés confirmatório e das heurísticas humanas, “o juiz, como tendência, concordará com a máquina, seja por conveniência, seja pelo fato de que uma decisão amparada por um algoritmo terá menores chances de ser revista” (Boeing; Rosa, 2020, p. 99). Dentro desse contexto, ainda, Boeing e Rosa (2020, p. 99) asseveram que, ao aceitar a sugestão (texto elaborado) do algoritmo, a intervenção humana é mínima e, [a]o passo que o modelo aprimora sua acurácia, a tendência é que o juiz se torne uma espécie de ‘canal de entrega’ de decisões geradas por robôs”.

Já quanto ao robô-julgador, Boeing e Rosa (2020) destacam que neste último caso, o algoritmo cria um resultado que será considerado como a própria decisão judicial e poderia ser empregada, por exemplo, nos casos de demandas repetitivas, para apreciação de requisitos de admissibilidade recursal ou para estabelecer condenação em custas e honorários nos recursos improvidos dos juizados especiais. O humano atuaria apenas como instância revisora, no caso de a parte discordar do resultado.

Partindo dessa premissa, desataca-se que o uso da IA tem se expandido de modo significativo no Poder Judiciário brasileiro, em especial com a aplicação dos modelos mencionados (robô-classificador e o robô-relator), que auxiliam os magistrados na organização de dados e elaboração das decisões, consequentemente agilizando os processos. Todavia, necessitam ainda da atuação humana. No entanto, o robô-julgador, que seria capaz de criar decisões ainda enfrenta desafios éticos e jurídicos. A IA vem com o intuito de melhorar a eficiência, mas ainda é imprescindível que o magistrado mantenha a sua função a fim de garantir a justiça e respeitar os direitos pleiteados em cada demanda.

Em relatório divulgado pela Fundação Getúlio Vargas (FGV), denominado “Tecnologia Aplicada à Gestão dos Conflitos no Âmbito do Poder Judiciário com Ênfase em Inteligência Artificial”, consta relação de um total 64 projetos de inteligência artificial, seja preditiva ou generativa, já implementados ou ainda em desenvolvimento, em 47 diferentes tribunais brasileiros. Relevante apontar que o estudo considerou exclusivamente ferramentas de inteligência artificial, desenvolvidas tanto por equipes internas dos tribunais, quanto em parcerias com universidades, empresas privadas ou outros órgãos públicos, excluindo iniciativas que devem ser consideradas apenas como sistemas de tecnologia da informação (Salomão, 2020). A esses projetos se somam situações em que magistrados, desembargadores e ministros de tribunais superiores e suas assessorias se valem de ferramentas como o conhecido ChatGPT para a elaboração de textos e cujo emprego é de difícil constatação. Aliás, as pesquisas realizadas são anteriores à disponibilização do ChatGPT ao público, o que se deu em novembro de 2022.

Dessa forma, Salomão (2020) sintetiza que, com o emprego da IA, espera-se aumentar a produtividade do Judiciário brasileiro, que, por exemplo, no ano de 2020 contava com cerca de 78,7 milhões de ações tramitando. Neste contexto, Salomão (2020, p. 69) ainda salienta que os objetivos que estariam sendo atendidos pela IA são vários, como o aumento na “otimização de atendimentos aos advogados e ao público, maior segurança; automação de atividades; melhor gestão dos recursos humanos para a atividade-fim do Judiciário; e aumento da celeridade na tramitação processual”.

Em pesquisa mais recente, Tauk *et al* (2023) dividiram 64 modelos de IA encontrados em quatro grupos. No primeiro grupo, estariam os sistemas que auxiliam na atividade-meio do judiciário, como a administração, gestão de recursos e de pessoal. Já no segundo grupo, estariam os modelos que “se destinam à automação dos fluxos de movimentação do processo e das atividades executivas de apoio aos juízes por meio da execução de tarefas pré-determinadas”,

estando relacionada à atividade-fim da prestação jurisdicional (Tauf *et al*, 2023, p. 11). Sequencialmente, no terceiro, estariam os sistemas que “colaboram na elaboração de minutas com conteúdo decisório de sentença, votos ou decisões interlocutórias” (Tauf *et al*, 2023, p. 12). E por fim, no quarto grupo estariam as “iniciativas relacionadas a formas adequadas de resolução de conflitos, em que se usam informações de processos similares para amparar as partes na busca da melhor solução” (Tauf *et al*, 2023, p. 13).

Tauf *et al* (2023, p. 13) apontam que, dos quatro grupos de modelos de IA, não se verificou “a interpretação de textos legais, a elaboração de argumentação jurídica e, muito menos, a tomada de decisão pela máquina”, em que pese no terceiro grupo se constatar o emprego de IA para a elaboração de textos de decisão, o que se enquadra no modelo de robô-relator, já exemplificado anteriormente.

Em um estudo do Conselho Nacional de Justiça (2024, p. 52), apurou-se que 49,4% dos magistrados e 49,5% dos servidores do poder judiciário já haviam utilizado ou tido experiência com Inteligência Artificial Generativa (IAG) na elaboração de textos seja no curso da execução de seus trabalhos, sejam em outras atividades profissionais ou acadêmicas. Dentre as ferramentas mais utilizadas, foram citados o ChatGPT, o Copilot, o Gemini e o Bing AI.

Dessa forma, a utilização da IA no processo decisório do Poder Judiciário tem várias nuances. Ela auxilia com a agilidade na análise de grandes volumes de dados e na redução na carga de trabalho dos julgadores e de seus assessores, podendo contribuir para uma maior consistência nas decisões. Contudo, apresenta desafios significativos, especialmente no quesito de garantir que os algoritmos sigam uma parametrização justa e argumentação jurídica apropriada, minimizando ao máximo e, se possível, evitando os vieses e discriminação. Ainda, cabe destacar que outro desafio relevante está condicionado na preservação da imparcialidade e na ética, haja vista a IA não poder substituir a análise e o julgamento humano.

3. OS PROBLEMAS DA IAG E AS MEDIDAS VISANDO MINIMIZAR EFEITOS DELETÉRIOS DO USO DA IA COMO FERRAMENTA DE AUXÍLIO NA TOMADA DE DECISÕES PELOS JUÍZES NO BRASIL

Em que pese o aparente ganho de objetividade alcançada pelo emprego da IAG no processo de tomada de decisão no âmbito do Poder Judiciário, tal ferramenta não está imune a falhas. Não cabe uma análise acurada de como elas ocorrem, mas em linhas gerais, seguindo Boeing e Rosa (2020), pode-se dizer o principal problema verificado são os vieses de máquina,

os quais decorreriam ou seriam potencializados: (a) por um equívoco na interpretação da realidade pelos criadores ou usuários da IA; (b) pela escolha de dados de treinamento; (c) na formulação das perguntas feitas à IA; e (d) na interpretação das respostas fornecidas pela IA.

Em relação aos elementos considerados para o treinamento da máquina, Boeing e Rosa (2020) consideram que o relevante é a forma como as perguntas são feitas à IAG e não necessariamente a qualidade ou quantidade dos dados utilizados. A falsa sensação de “completude de informação” cria uma cegueira quanto à informação indisponível e não permite prever situações sociais mais complexas.

Um outro problema da inteligência artificial são as heurísticas de programação, estratégia de dividir as tarefas realizadas pelo algoritmo para, por exemplo, empregar menos capacidade de processamento ou reduzir o conjunto de dados considerados, quando seu universo é muito grande. Com essa tática, para se chegar a um resultado considerado consistente, de forma relativamente eficiente, desiste-se de obter um resultado ótimo ou ao menos o melhor resultado que se poderia esperar (Russell, 2013). As heurísticas quase nunca são de conhecimento dos usuários e sua configuração demanda alta especialização, do contrário podendo levar a resultados imprecisos.

Modelos de Linguagem de Grande Escala (*Large Language Models* - LLMs) como o ChatGPT ou outro *chatbot* de IAG, ao operarem vetores, podem concluir pela existência de padrões ou considerarem objetos que são inexistentes ou imperceptíveis para a mente humana, chegando a resultados sem sentido ou totalmente imprecisos, as chamadas “alucinações” (IBM, 2023).

Dahl, Magesh, Suzgun e Ho (2024) apontam uma pesquisa realizada nos Estados Unidos, explorando como LLMs respondem a questionamentos jurídicos, vindo a confrontar posteriormente os resultados com critérios legais verossímeis, chegando à conclusão de que eles podem levar a repostas equivocadas. Do resultado dos estudos, os autores supramencionados sintetizam e extraem que: 1. LLMs apresentaram altos graus de alucinação ao serem questionados sobre casos judiciais federais: ChatGPT 4, 58%; ChatGPT 3.5, 69%, PaLM 2, 72%; e Llama 2, 88%; e 2. LLMs não conseguiriam identificar que estão gerando alucinações legais (Dahl; Magesh; Suzgun; Ho, 2024).

Já Araújo, Gabriel e Porto (2024, p. 98) destacam outro ponto importante: a dificuldade de explicar como os algoritmos processam os dados para chegar a uma conclusão, a chamada “opacidade”. Em situações em que são empregadas técnicas de aprendizado de máquina (*machine learning*), os autores mencionados consideram que a opacidade tende a se

potencializar, pois nem mesmo os desenvolvedores da IA conseguirão compreender como o algoritmo operou (Araújo; Gabriel; Porto, 2024).

Conforme Boeing e Rosa (2020), o modelo de IA nos moldes de um robô-classificador se demonstra com maior potencialidade de criar no juiz a falsa sensação de que todo o material relevante para o julgamento do caso concreto está à disposição. A máquina também poderia corroborar preconcepções decorrentes dos dados com que foi alimentada, talvez decisões prévias do próprio magistrado. Isto levaria a um viés de confirmação, quando da apresentação do resultado. Outro ponto a ser considerado, conforme apontam Boeing e Rosa (2020) é que, à medida que os resultados forem sendo validados, a ferramenta tende a confirmar seu grau de acurácia, levando a uma realidade de retroalimentação, em que as decisões validadas também serão consideradas para sugestões futuras. O algoritmo terá mais dados similares e as sugestões divergentes vão se tornar cada vez menos comuns, simplificando e tendendo a padronizar os resultados.

No caso do robô-julgador, Boeing e Rosa (2020) alertam que, à medida que o modelo se torna mais assertivo e tem o poder de decidir sozinho, há um risco de o magistrado se tornar um canal de entrega para decisões geradas pela IA. Ao atuar como instância revisora, contudo, o juiz tem que estar atento para não ser afetado por informações seletivas fornecidas por outros algoritmos, ou seja, deve realizar a revisão sem se valer de outras ferramentas de IA. A retroalimentação do modelo em que a IA é o próprio julgador, tende a ser ainda mais perigosa, pela rapidez com que os processos serão resolvidos e tornados fundamentos de novas decisões automáticas (Boeing; Rosa, 2020).

Em síntese, é nítido destacar que a IA auxilia no processo de decisão no Poder Judiciário, todavia, ela levanta questões complexas sobre a confiabilidade das decisões automatizadas, sendo imprescindível a revisão dos magistrados. Embora a IA tenha o intuito de aumentar a eficiência e a padronização das decisões, é necessário que os magistrados tenham responsabilidade de revisar com cautela os resultados dessas decisões. Para tanto, se torna fundamental a adoção de medidas que visem minimizar os efeitos deletérios do uso da IA, pois seu uso indiscriminado tende a acarretar riscos para a amplificação dos vieses e a padronização excessiva nas decisões, o que compromete a individualização dos julgamentos e a garantia de justiça em cada caso concreto.

Outro efeito que se pode prever é que o emprego da IA levará a uma maior uniformidade nas decisões judiciais, com potencial tanto de um certo “engessamento” do pensamento humano, quanto à perpetuação de preconceitos sociais e escolhas culturais, em

prejuízo de grupos marginalizados ou minoritários (Ferrari, 2023). Por isso, ao jurisdicionado deve ser assegurado o conhecimento de assistência de IA na tomada de decisões, como ela foi elaborada e os dados utilizados em seu treinamento, isto é, para a predição ou geração de resultados. Também deve ser permitido o acompanhamento das soluções empregadas a fim de conferir se os resultados obtidos estão de acordo com o que era esperado e respeitam os Direitos Fundamentais.

Como dito anteriormente, a sociedade ainda não vive numa realidade de IA forte, com sistemas capazes de realizar tarefas cognitivas de modo semelhante à mente humana, ou seja, com consciência, compreensão, capacidade de aprendizado geral, criatividade e total domínio da linguagem natural e da semântica. Russel e Norvig (2023) defendem que o que se tem à disposição, atualmente, são ferramentas de IA fraca, focada em tarefas preditivas ou de geração de textos, sons, imagens e vídeos. Assim, juízes, desembargadores e ministros de tribunais superiores não podem se valer de IA de forma desarrazoada, sem apreciar as nuances do caso concreto, sem terem conhecimento dos argumentos jurídicos considerados pelo algoritmo na elaboração do texto decisório ou sem tomarem as cautelas necessárias para não permitir, por exemplo, que uma alucinação de máquina venha a ser mantida (Russel; Norvig, 2023).

Nesse aspecto, Araújo, Gabriel e Porto (2020, p. 105) sustentam ser um imperativo ético que a IA generativa seja utilizada como ferramenta apenas por pessoas que dominam o conhecimento e podem tanto produzir decisões sua utilização, como validar de modo adequado o resultado por ela apresentado.

Como as tecnologias não são neutras, tanto a programação da IA quanto os dados utilizados podem ser tendenciosos. Como sustentam Araújo, Gabriel e Porto (2020, p. 97-98), algoritmos de IA podem refletir preconceitos de seus desenvolvedores ou discriminação encontrada nos dados usados no seu treinamento, sendo, “[...] essencial evitar a reprodução de vieses e discriminação presentes nos dados, por meio de análises cuidadosas e pré-processamento dos dados para mitigar esses vieses”.

Visando justamente compatibilizar a implantação da mecanismos de IA no Poder Judiciário com o respeito aos direitos fundamentais, bem como manter coerência no processo de tomada de decisão, o CNJ editou no ano de 2020 a Resolução nº 332, estabelecendo no *caput* do Art. 7º diretrizes que as decisões judiciais devem seguir com o uso da IA, dentre eles podendo-se destacar “eliminar ou minimizar [...] os erros de julgamento decorrentes de preconceitos”.

Ainda, com a finalidade de se afastar vieses discriminatórios, foi estabelecido pela Resolução nº 332, por exemplo: 1. a necessidade de prévia homologação do modelo de IA, visando a identificação de preconceitos ou generalizações em seu desenvolvimento, gerando vieses discriminatórios em seu funcionamento (artigo 7º, § 1º); 2. a adoção de medidas corretivas no caso de se verificar tendências discriminatórias no uso da ferramenta de IA (artigo 7º, § 2º); 3. a descontinuidade de utilização da ferramenta de IA caso não haja possibilidade de eliminação do viés discriminatório (artigo 7º, § 3º); 4. obrigatoriedade de composição de equipes de pesquisa, desenvolvimento e implantação de soluções de IA orientadas pela diversidade, considerando aspectos como gênero, raça, etnia, cor, orientação sexual, pessoas com deficiência e geração (artigo 20, “caput”); 5. a vedação a que estudos, pesquisas, ensino e treinamentos de IA fiquem subordinados a sectarismo capaz de direcionar seu desenvolvimento ou resultados (artigo 21, III); e 6. o desestímulo ao uso de modelos de IA, em especial a sugestão de modelos de decisões preditivas (artigo 23, “caput”) (CNJ, 2020).

Em complementação, Requião (2024, p. 14) sustenta que “o sucesso do uso da IA no judiciário depende de constante análise para enfrentamento de vieses”, o que poderia ser feito pela criação dos chamados “*red teams*”, equipes multidisciplinares que poderiam “verificar problemas na programação, como analisar fatores éticos e jurídicos”.

Em suma, o uso da IA pelo Poder Judiciário traz tanto benefícios como desafios. Desse modo, é fundamental que o uso da IA seja acompanhada de rigoroso controle, transparência, monitoramento constante e revisão humana, para então se tornar verdadeiramente aliada ao ser humano e não um obstáculo. A IA deve ser vista como uma ferramenta de apoio e agilidade para o Poder Judiciário, contudo, só será possível perfectibilizar essa mão dupla de apoio com constante atualização dos mecanismos de controle capazes de garantir que a IA contribua de forma positiva tanto para o fortalecimento do próprio Poder Judiciário bem como produza confiança, transparência e ética para a sociedade, estando os julgadores cientes da necessidade de garantirem um processo decisório que decorra de uma construção dentro da racionalidade jurídica.

CONSIDERAÇÕES FINAIS

As teses realistas apresentadas possuem um elemento comum que é a consideração de fatores não apenas racionais, em termos de lógica formal e consciência, para o ato de tomada de decisão. Essa visão aumenta o campo de entendimento para uma resposta prática sobre como

os juízes decidem, especialmente, considerando a possibilidade de vieses no processo de tomada de decisão.

Com o crescente emprego da IA pelo Poder Judiciário, o que vem sendo potencializado pela universalização e acesso aos modelos de linguagem grande (LLM, na sigla em inglês), um tipo específico de inteligência artificial generativa (IAG), como ferramenta de auxílio na elaboração de textos decisórios, ainda mais se faz relevante considerar as teses realistas como suporte de estudo dos assim chamados vieses de máquina e do próprio viés humano confirmatório daquilo que é produzido pelo algoritmo da IAG.

É verdade que a decisão judicial só pode ser considerada pelo seu texto, sendo que intuições, crenças, inclinações e vieses dos julgadores dificilmente estarão expressamente apresentados. Apesar disso, a tecnologia permite que na programação da IA utilizada no processo decisório seja possível programar travas que impeçam, por exemplo, vieses discriminatórios. Contudo, com vistas na disseminação do uso de LLMs, cabe ao magistrado, na condição de supervisão humana, revisar todas as minutas decisórias sugeridas pela IA como forma de controle contra vieses e injustiças.

Nesse sentido, como foi apresentado, uma série de novos problemas, em especial os vieses de máquina, que repercutem estereótipos, preconceitos e discriminação humanos são preocupações recorrentes com o empregado da tecnologia. De modo especial, a edição da Resolução nº 332 de 2020 pelo CNJ, regulamentando o uso da IA pelo Poder Judiciário, é uma medida salutar e, se respeitada, permitirá uma minimização das consequências deletérias do uso desarrazoado da IA. De qualquer modo, é necessária atenção dos juristas no sentido de exercer fiscalização dos modelos de IA empregados pelo Poder Judiciário e dos resultados por eles gerados, evitando equívocos e injustiças na apreciação do caso concreto.

REFERÊNCIAS

ANDRADE, Flávio da Silva. A tomada da decisão judicial criminal à luz da psicologia: heurísticas e vieses cognitivos. **Revista Brasileira de Direito Processual Penal**, Porto Alegre (RS), v. 5, n.1, p. 507-540, 2019. Disponível em: <https://bdjur.stj.jus.br/server/api/core/bitstreams/85c01f77-975d-4a70-b060-6d4a0c6221a3/content>. Acesso em 10 dez. 2024.

ARAÚJO, Valter Shuenquener de; GABRIEL, Anderson de Paiva; PORTO, Fábio Ribeiro. **‘Juízo 100% digital’ e transformação tecnológica da Justiça no século XXI: Novo modelo de trabalho utiliza todo o potencial que a tecnologia pode fornecer ao Poder Judiciário**. [S.l.: s.n., s.d.]. Disponível em: <https://www.jota.info/opiniao-e-analise/colunas/juiz-hermes/juizo-100-digital-e-transformacao-tecnologica-da-justica-no-seculo-xxi>. Acesso em: 10 jan. 2025.

BILLIER, Jean-Cassien; MARYIOLI, Aglaé. **As teorias antiformalistas**. Tradução de Maurício de Andrade. Barueri (SP): Manole, 2005.
Disponível em: <https://pt.scribd.com/document/280065755/Historia-Da-Filosofia-Do-Direito-Jean-Cassien-Billier-e-Aglae-Maryioli>. Acesso em: 22 dez. 2024.

BOEING, Daniel Henrique; ROSA, Alexandre Moraes da. **Ensinando um robô a julgar: pragmática, discricionariedade, heurísticas e vieses no uso de aprendizado de máquina no judiciário**. Florianópolis (SC): Emais Editora e Livraria Jurídica, 2020.

BOLONHA, Carlos; FORTES, Pedro Rubim Borges; LUCAS, Daniel de Souza. Realismos jurídicos, direito em ação e o estudo empírico das instituições. **REI - Revista Estudos Institucionais**: Journal of Institutional Studies, Rio de Janeiro, v. 6, n. 1, p. i-vii, jan./abr. 2020. Disponível em: https://www.researchgate.net/publication/340889131_REALISMOS_JURIDICOS_DIREITO_EM_ACAO_E_O_ESTUDO_EMPIRICO_DAS_INSTITUICOES. Acesso em: 12 dez. 2024.

BRANDO, Marcelo Santini. **Como decidem os juízes?** Uma investigação da teoria realista da decisão judicial a partir das contribuições das ciências cognitivas e da psicologia moral. Orientador: Noel Struchiner. 2013. 112 f. Dissertação (Mestrado em Direito) – Faculdade de Direito, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2013. Disponível em: https://www.dbd.puc-rio.br/pergamum/tesesabertas/1112604_2013_pretextual.pdf. Acesso em: 10 jan. 2025.

CONSELHO NACIONAL DE JUSTIÇA. **O uso da inteligência artificial generativa no Poder Judiciário brasileiro: relatório de pesquisa**. Brasília, DF: CNJ, 2024. Disponível em: <https://www.cnj.jus.br/wp-content/uploads/2024/09/cnj-relatorio-de-pesquisa-iag-pj.pdf>. Acesso em: 02 jan. 2025.

CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 332, de 29 de abril de 2020**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Judiciário e dá outras providências. Brasília, DF: CNJ, 2020. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/atos-normativos?documento=4715#art23>. Acesso em: 28 dez. 2024.

CONSELHO NACIONAL DE JUSTIÇA. **Resolução nº 423, de 05 de outubro de 2021**. Dispõe sobre a ética, a transparência e a governança na produção e no uso de Inteligência Artificial no Judiciário e dá outras providências. Brasília, DF: CNJ, 2021. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/4147>. Acesso em: 03 jan. 2025.

DAHL, Matthew; MAGESH, Varun; SUZGUN, Mirrac; HO, Daniel E. Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models. **Journal of Legal Analysis (forthcoming)**, [S.l.], v. 16, n. 1, p. 64-93, 25 abr. 2024. Disponível em: <https://arxiv.org/abs/2401.01301>. Acesso em: 02 jan. 2025.

FERRARI, Isabela. **Discriminação algorítmica e poder judiciário: limites à adoção de sistemas de decisões algorítmicas no Judiciário brasileiro**. Florianópolis: Emais editora, 2023.

HART, Herbert Leonel Adolphus. **Essays in Jurisprudence and Philosophy**. Oxford (Reino Unido): Oxford University Press, 1983. Disponível em:

<https://doctoradohumanidades.files.wordpress.com/2014/02/text-filosofc3ada-h-l-a-hart-essays-in-jurisprudence-and-philosophy-1984.pdf>. Acesso em: 18 dez. 2024.

HORTA, Ricardo Lins. **Quais os desafios e as oportunidades da decisão sobre concurso para juiz?** [S.l.]: Conjur, 05 out. 2021. Disponível em: <https://www.conjur.com.br/2021-out-05/horta-heuristicas-vieses-concursos-magistratura/>. Acesso em: 03 jan. 2025.

HORTA, Ricardo de Lins e; COSTA, Alexandre Araújo. Das teorias da interpretação à teoria da decisão: por uma perspectiva realista acerca das influências e constrangimentos sobre a atividade judicial. **Revista Opinião Jurídica**, Fortaleza (CE), v. 15, n. 20, p. 271-297, 12 jul. 2017. Disponível em: <https://periodicos.unichristus.edu.br/opiniaojuridica/article/view/1387/470>. Acesso em: 30 dez. 2024.

IBM. **What are AI hallucinations?** [S.l.]: IBM, 01 set. 2023. Disponível em: <https://www.ibm.com/think/topics/ai-hallucinations>. Acesso em 02 jan. 2025.

KAHNEMAN, Daniel. **Rápido e devagar**: duas formas de pensar. Tradução de Cássio de Arantes Leite. Rio de Janeiro: Objetiva, 2012. *E-book* Kindle.

KORTELING, Johan Egbert Hans; TOET, Alexander. Cognitive Biases. In: DELLA SALA, Sergio (ed.). **Encyclopedia of behavioral neuroscience**. Reference Module in Neuroscience and Biobehavioral Psychology. Amsterdã (Países Baixos)-Edimburgo (Reino Unido): **Elsevier Science Direct**, 2020. Disponível em: https://www.researchgate.net/publication/338403182_Cognitive_biases_Section_published_in_the_Encyclopedia_of_Behavioral_Neuroscience_Elsevier_Science_2021_httpsdoiorg101016B978-0-12-809324-524105-9. Acesso em: 10 jan. 2025.

LEITER, Brian. American Legal Realism. In: PATTERSON, Dennis (Org). **A companion to philosophy of law and legal theory**. 2ª ed. Malden (EUA): Blackwell Publishing, 2010. Disponível em: <http://ndl.ethernet.edu.et/bitstream/123456789/27748/1/22.pdf>. Acesso em: 18 dez. 2024.

O'NEIL, Cathy. **Algoritmos de Destruição em Massa**: Como o *Big Data* aumenta a desigualdade e ameaça à democracia. Tradução de Rafael Abraham. 1ª ed. Santo André, SP: Editora Rua do Sabão, 2020. *E-book* Kindle.

PENROSE, Roger. **Shadows of the mind**: a search for missing Science of consciousness. Oxford: University Press, 1994. Disponível em: https://altexploit.wordpress.com/wp-content/uploads/2017/07/roger-penrose-shadows-of-the-mind_-a-search-for-the-missing-science-of-consciousness-oxford-university-press-1994.pdf. Acesso em: 18 dez. 2024.

REQUIÃO, Maurício de Sant'Ana. Inteligência artificial, vieses cognitivos e decisões judiciais. **Revista do Programa de Pós-Graduação em Direito da UFBA**, Salvador (BA), v. 34, n. 2, jul/dez 2024. Disponível em: <https://periodicos.ufba.br/index.php/rppgd/article/view/63797/36206>. Acesso em: 03 jan. 2025.

RUSSELL, Stuart Jonathan; NORVIG, Peter. **Inteligência artificial**. Tradução de Regina Célia Simille. Rio de Janeiro, RJ: Elsevier, 2013.

SALOMÃO, Luis Felipe. **Inteligência Artificial: tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário brasileiro**. Rio de Janeiro, RJ: FGV, 2020. Disponível em: <https://bdjur.stj.jus.br/handle/2011/156490>. Acesso em: 10 dez. 2024.

SCHAUER, Frederick. **Thinking like a lawyer: a new introduction to legal reasoning**. Cambridge, EUA: Harvard University Press, 2009. Disponível em: https://codolc.com/books/Thinking_Like_a_Lawyer.pdf. Acesso em: 23 dez. 2024.

SEARLE, John R. **O mistério da consciência**. Tradução de André Yuji Pinheiro Uema e Vladmiir Safatle. São Paulo, SP: Paz e Terra, 1998.

SHECAIRA, Fabio; STRUCHINER, Noel. Realismo Jurídico. In: TRAVESSONI, Alexandre (Org.) **Dicionário de Teoria e Filosofia do Direito**. São Paulo, SP: Editora LTr, 2011.

SHECAIRA, Fabio; STRUCHINER, Noel. **Teoria da argumentação jurídica**. Rio de Janeiro, RJ: Contraponto, 2016.

STANOVICH, Keith. E.; WEST, Richard F. Individual differences in reasoning: Implications for the rationality debate? **Behavioral and Brain Sciences**, [S.l.], v. 23, n. 5, 09 abr. 2001, p. 645-665. Disponível em: <https://pt.scribd.com/document/87286377/Individual-Difference-in-Reasoning>. Acesso em: 10 jan. 2025.

TABAK, Benjamin Miranda; AMARAL, Pedro Henrique Rincon. Vieses cognitivos e desenhos de políticas públicas. **Revista Brasileira De Políticas Públicas**. [S.l.], v. 8, n. 2, 2018, p. 472-491. Disponível em: <https://eppg.fgv.br/sites/default/files/5278-24114-2-pb-7-25.pdf>. Acesso em: 10 dez. 2024.

TAUK, Caroline Somesom; NUNES, Dierle; BRAGANÇA, Fernanda; LOSS, Juliana; COLEHO, José Leovigildo; BRAGA, Renata. Inteligência Artificial: tecnologia aplicada à gestão de conflitos no âmbito do Poder Judiciário brasileiro. In: SALOMÃO, Luis Felipe; TAUKE, Caroline Somesom; LEME, Elton; LOSS, Juliana. (orgs). **Inteligência Artificial**. 3. ed. Rio de Janeiro: FGV, 2023. Disponível em: <https://conhecimento.fgv.br/publicacao/inteligencia-artificial-tecnologia-aplicada-gestao-dos-conflitos-no-ambito-do-poder>. Acesso em 22 dez. 2024.