I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL II

EUDES VITOR BEZERRA

CINTHIA OBLADEN DE ALMENDRA FREITAS

Copyright © 2025 Conselho Nacional de Pesquisa e Pós-Graduação em Direito

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 II [Recurso eletrônico on-line] organização CONPEDI
Coordenadores: Cinthia Obladen de Almendra Freitas, Eudes Vitor Bezerra. – Florianópolis: CONPEDI, 2025.

Inclui bibliografia
ISBN: 978-65-5274-096-0
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 II

Apresentação

O I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA, com temática “Inteligência Artificial e Sustentabilidade na Era Transnacional”, realizado no período de 28 a 30 de maio de 2025 na Universidade degli Studi di Perugia – Itália, reuniu centenas de pesquisadores, professores e estudantes de Programas de Pós-Graduação em Direito (PPGD) do Brasil, da Itália e de outras nações.

Com submissões de trabalhos, o GT - INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL surpreendeu pela quantidade de trabalhos submetidos, tendo sido subdividido em quatro subgrupos. Assim, esta apresentação refere-se aos trabalhos submetidos, selecionados e, efetivamente, apresentados e discutidos no GT - INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL II.

Os trabalhos apresentados foram organizados em dois blocos distintos pelas temáticas centrais dos artigos, permitindo uma unidade de discussões e reflexões. No primeiro bloco, os trabalhos trataram de temas referentes à aplicação de sistemas de IA em: Educação, Meio Ambiente, Planejamento Sustentável e Cidades Inteligentes, Trabalho, Poder Judiciário e Medicina e Saúde. As discussões envolveram de modo primordial os riscos advindos da aplicação de sistemas de IA nestas áreas, permitindo reflexões sobre: a) Educação: personalização do ensino, padronização excessiva do aprendizado, a mercantilização da educação e o uso inadequado de dados sensíveis de estudantes; b) Meio Ambiente: aplicações de sistemas de IA na governança ambiental, riscos e responsabilidade jurídica, regulação; c) Planejamento Sustentável e Cidades Inteligentes: sistemas de IA no planejamento urbano e mudanças climáticas e, também, viés adultocêntrico nas cidades inteligentes; d) Trabalho: plataformas digitais, subordinação algorítmica, precarização do trabalho humano, jornadas extensas, remuneração variável, ausência de direitos trabalhistas e ambiente de trabalho estressante devido à vigilância constante dos algoritmos; e) Poder Judiciário: democratização da justiça e exclusão digital, celeridade processual, transparência e explicabilidade, minutas automatizadas e dignidade humana, juízo humano versus decisão automatizada; f) Medicina e Saúde: formação médica, diagnósticos, simulações clínicas, desinformação em saúde. Percebeu-se que a temática de Inteligência Artificial desenvolverá cada vez mais um papel preponderante no desenvolvimento e na sustentabilidade de um ecossistema tecnológico, o qual precisa estar fundamentado em princípios jurídicos para que os desafios da Era Digital sejam enfrentados e os riscos mitigados. Deste modo,

considerando-se como premissa que a regulação de sistemas de IA deve ser guiada por quatro elementos fundamentais: transparência, não discriminação, responsabilidade e segurança jurídica; as discussões foram produtivas e permitiram compreender que tais elementos são essenciais para garantir que o uso de sistemas de IA respeite os direitos fundamentais e promova justiça social. E, ainda, há que se pontuar que os sistemas de IA não poderão apenas contemplar aspectos técnicos, mas também precisarão estar atentos aos aspectos jurídicos, éticos, sociais, culturais e ambientais.

No segundo bloco, os trabalhos trataram de aspectos relacionados à interação entre Inteligência Artificial e os direitos fundamentais, abordando questões como personalidade jurídica, proteção de dados, ética algorítmica, direitos da personalidade, inclusão social, reconhecimento facial e riscos processuais no uso de IA na advocacia e na pesquisa jurídica. Foram analisados os desafios da ausência de atribuição de personalidade jurídica à inteligência artificial na reforma do Código Civil brasileiro, bem como a proteção de dados em holdings familiares a partir de uma análise comparativa entre a LGPD e o GDPR. Discutiu-se a ética em IA, com foco em transparência e justiça algorítmica, além da proteção jurídica dos ciborguêses e as complexas inter-relações entre direitos da personalidade e desenvolvimento tecnológico.

As discussões também abordaram a regulamentação da inteligência artificial na União Europeia, com destaque para a garantia de acesso pleno e igualdade para pessoas com deficiência, segundo o AI Act. Questões relacionadas ao reconhecimento facial nos estádios de futebol brasileiros também foram objeto de estudo, com ênfase nos riscos de criminalização seletiva e nos impactos sobre os direitos humanos nas arenas esportivas.

Por fim, os trabalhos exploraram os riscos jurídicos associados ao uso da inteligência artificial na advocacia e os posicionamentos dos tribunais brasileiros sobre a matéria, além de proporem uma análise teórica e recomendações práticas para a utilização metodologicamente adequada da IA comercial na pesquisa jurídica.

Felizes pela variedade de temas de pesquisa, os coordenadores do GT - INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL II convidam a todas e todos para a leitura na íntegra dos artigos.

Cinthia Obladen de Almendra Freitas – Pontifícia Universidade Católica do Paraná (PUCPR) – cinthia.freitas@pucpr.br

Eudes Vitor Bezerra – Universidade Federal do Maranhão (UFMA) – eudesvitor@uol.com.br

DECISÕES AUTOMATIZADAS E COGNIÇÃO HUMANA: O PAPEL DA IA NO PROCESSO DECISÓRIO JUDICIAL

AUTOMATED DECISIONS AND HUMAN COGNITION: THE ROLE OF AI IN THE JUDICIAL DECISION-MAKING PROCESS

Sergio Nojiri [1]
Luiz Guilherme da Silva Rangel [2]

Resumo

O artigo, além de apresentar conceitos básicos para compreensão do que se trata a inteligência artificial (IA), discute o emprego atual dessa tecnologia no âmbito dos sistemas de justiça brasileiro e internacional. A partir de uma revisão bibliográfica, foram avaliados os impactos que esses sistemas algorítmicos representam para o processo decisório com a constatação dos aspectos positivos de seu uso, especialmente sob o prisma da produtividade, mas também ressalvas quanto ao seu uso indiscriminado. Nesse aspecto, ressalta-se a possibilidade de replicação de preconceitos por parte desses sistemas e a falta de transparência na métrica utilizada pelos algoritmos. Ambos, viés da máquina e opacidade algorítmica, representam obstáculos importantes ao uso de inteligência artificial no Direito, pois podem conflitar com garantias processuais e constitucionais .

Palavras-chave: Inteligência artificial, Processo decisório, Cognição humana, Vieses de máquina, Opacidade algorítmica

Abstract/Resumen/Résumé

In addition to presenting basic concepts for understanding what artificial intelligence (AI) is, this article discusses the current use of this technology within the Brazilian and international justice system. Based on a literature review, the impacts that these algorithmic systems represent for the decision-making process were assessed, noting the positive aspects of their use, especially from the perspective of productivity, but also reservations regarding their indiscriminate use. In this regard, the possibility of replication of prejudices by these systems and the lack of transparency in the metrics used by the algorithms are highlighted. Both machine bias and algorithmic opacity represent significant obstacles to the use of artificial intelligence in Law, as they can conflict with procedural and constitutional guarantees.

Keywords/Palabras-claves/Mots-clés: Artificial intelligence, Decision-making, Human cognition, Machine bias, Algorithmic opacity

[1] Graduação, Mestrado e Doutorado em Direito pela Pontifícia Universidade Católica de São Paulo. LivreDocente e Professor pela Faculdade de Direito de Ribeirão Preto (USP). Juiz Federal.

[2] Graduado pela Faculdade de Direito de Ribeirão Preto da USP. Advogado.

Introdução

O surgimento da expressão “Inteligência Artificial” remonta a meados da década de 1950, e foi proferida pela primeira vez em uma palestra no *Dartmouth College* por John McCarthy, um professor universitário americano cuja pesquisa acadêmica estava relacionada à lógica simbólica e a avanços computacionais (NUNES; MARQUES, 2021, p. 711). Naquela época o termo Inteligência Artificial (IA) estava atrelado à mimetização do comportamento humano, o que se confirma, por exemplo, no comportamento dos *chatbots* criados logo em sequência, como, por exemplo, o ELIZA, criado por Joseph Weizenbaum em 1966 (GORDON, 2023). No entanto, as IAs modernas buscam fundamentalmente a exploração e resolução de problemas de forma mais célere, e não necessariamente em copiar a linguagem natural humana em uma determinada situação prosaica, como eram as primeiras IAs.

Por mais que as IAs pioneiras tenham galgado êxito em compreender a linguagem humana e resolver tarefas mais simples, elas não são necessariamente inteligentes, compreendendo aqui a inteligência como um conceito amplo. As capacidades humanas envolvem uma série de fatores que mesmo hoje ainda são misteriosos para as áreas da psicologia e da neurociência. Já em 1976, Joseph Weizenbaum, apresentava a aparente correlação entre inteligência humana e artificial e, diante do ser humano enfrentar problemas que nenhuma máquina enfrenta, ele processa a informação de maneira distinta dos computadores (WEIZENBAUM, 1976, p. 203).

Assim, mesmo considerando a capacidade de realização de operações cada vez mais complexas por IA, Rômulo Soares Valentini afirma que “máquinas não precisam emular a totalidade dos elementos que compõem o cérebro humano [...] para resolverem tarefas complexas”(VALENTINI, 2013), ou seja, por maior que seja a tendência de aproximar as duas entidades, não é possível fazer uma equiparação, pois existem elementos que compõem a inteligência humana, como os sentimentos, que ainda não são passíveis de reprodução por esses sistemas informáticos.

Por essa razão é que o foco atual dos estudos em IA não é o de simplesmente simular o comportamento humano, mas de encontrar sistemas de resolução de problemas de formas cada vez mais eficientes, o que, a depender da perspectiva, pode significar um impacto tanto positivo (como no uso para realização de tarefas administrativas de um tribunal) quanto negativo para a sociedade (como a possível perda de postos de trabalhos desses cargos administrativos).

1. Mecanismo de funcionamento e modos de aprendizagem das IAs

Os elementos cruciais para o funcionamento de um sistema de Inteligência Artificial são: I - os sensores (*input*); II - a lógica operacional (algoritmo); e III - os atuadores (*output*), ou seja, as informações de entrada, a computação desses dados e a resposta obtida pelo cômputo dessas informações com base no que a IA foi programada para fazer.

Suponha que é passada a atividade de selecionar bolas em uma imagem cheia de objetos com várias formas a um *software* (comando). Há em sua memória uma coletânea de dados dos formatos de vários objetos (informação de entrada - *input*), além de um algoritmo treinado para reconhecer quais deles são bolas excluindo os outros formatos e, por fim, a seleção de todas as bolas que estavam na imagem (informação de saída - *output*).

O exemplo supracitado trata da primeira abordagem de IA, o *Good Old-Fashioned AI* (GOFAI), ou ainda, IA simbólica, um modelo de programação mais simples que vigorou (majoritariamente) entre meados da década de 1950 e meados da década de 1980. Ele foi preparado para rodar nos sistemas dos primeiros computadores digitais de natureza binária e serial (BODEN, 2014, p. 89). A característica principal deste tipo de linguagem de programação é a de possuir uma menor autonomia em relação ao programador e trabalhar com pesquisa heurística, oferecendo sempre uma resolução exata ou aproximada do que se quis passar e tendo como base um conjunto finito de regras. Margaret A. Boden classifica a IA simbólica como " um espaço de pesquisa: um conjunto de possibilidades (definido por um conjunto finito de regras generativas), no qual a solução se encontra - e no qual deve ser encontrada" e exemplifica citando IAs capazes de realizar movimentos possíveis no xadrez, ou capazes de escrever textos (não necessariamente coerentes), considerando uma gramática e um vocabulário especificados pelo programador (BODEN, 2014, p. 90).

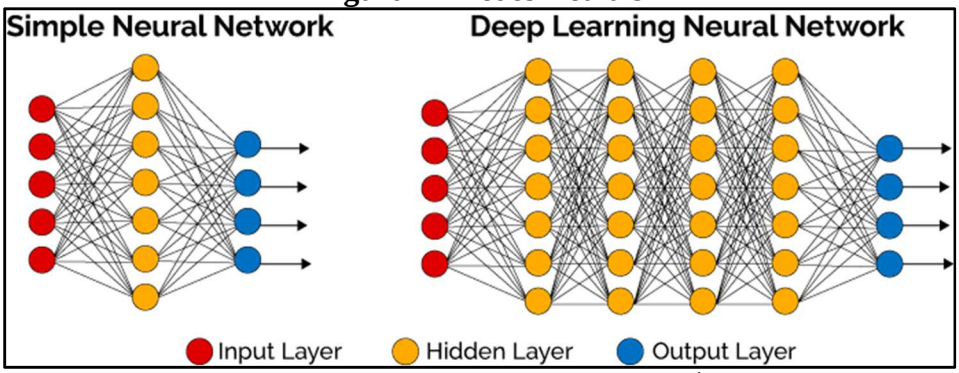
Posteriormente à predominância do modelo GOFAI, as pesquisas em Inteligência Artificial passaram por momentos de arrefecimento, primeiro entre 1974 e a década de 1980 e, depois, entre 1987 a meados da década de 1990, o que ficou conhecido como *AI winters*. Esses momentos são explicados devido ao corte no financiamento de pesquisas em IA por parte do governo americano, que liderava os estudos em computação. Assim, as pesquisas com *machine learning* e redes neurais (*deep learning*), embora já existissem desde a década de 1960, passaram a ganhar destaque apenas após a década de 1990, principalmente devido ao surgimento de modelos matemáticos mais complexos, redução dos custos de produção de equipamentos tecnológicos e maior quantidade de dados disponíveis para a análise (KONTSEVOI, 2023).

Assim, compreende-se o *machine learning* como os sistemas cujos algoritmos aprendem de forma automatizada com base nos dados inseridos pelo programador. Em outras palavras, o software aprende a identificar padrões e realizar generalizações com base nas informações pregressas sem que o programador precise inserir novos dados. A problemática existente nesses sistemas é que, embora não precisem ser alimentados toda vez que lhe for dada uma nova tarefa diferente de sua programação, e seja mais autônomo do que as IAs iniciais, o sistema fica dependente da qualidade e quantidade do *input data* (NUNES; MARQUES, 2021, p. 715).

Em consequência, o *deep learning* possui a mesma lógica de processamento por redes neurais, constituindo assim, também, uma forma de aprendizado automático (ou aprendizado de máquina), mas nesse caso as redes neurais artificiais do programa apresentam mais camadas neurais (ao menos duas) e lidam com volumes maiores de dados.

A imagem abaixo exemplifica o modelo de rede neural simples à esquerda e um modelo mais complexo com várias camadas representativo do *deep learning* a direita:

Figura 1 - Redes Neurais



Fonte: Deep Learning Book, 2022[1]

Os sistemas de *machine learning* e *deep learning* possuem diversos modos de aprendizagem. Para os fins deste trabalho, merecem destaque os aprendizados supervisionado, não supervisionado e por reforço.

Algoritmos de aprendizado supervisionado são aqueles que "requerem um sinal de retorno (de fontes externas) para cada um dos nós de saída de uma rede para que a aprendizagem se concretize". Ou seja, os dados fornecidos externamente confirmam os dados de entrada, pois nesses sistemas a obtenção da resposta esperada está contida nas informações de entrada, positivando o aprendizado (SUN, 2014, p. 111). Um bom exemplo é o método de aprendizado

[1] Disponível em: <https://www.deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>. Acesso em: 18 jan 2025.

do *Chat-GPT* ou do *Google* Tradutor. Esses sistemas são alimentados com informações pelo programador, sejam elas verbetes do Wikipedia ou traduções de termos, e posteriormente são confirmadas ou negadas por meio de uma sinalização externa que indica se a resposta dada corresponde a um padrão de resposta esperado. Desta forma, tanto o *Chat-GPT* como o *Google* Tradutor aprendem que uma determinada informação é imprecisa ou uma palavra pode ter outra tradução em uma determinada passagem, respectivamente, através de fatores externos que dão um *feedback* para que as camadas ocultas de algoritmos sejam calibradas e a resposta seja aperfeiçoada.

Já em sistemas de aprendizado não supervisionado, os algoritmos independem de sinais de erro ou da distinção entre os nós de entrada, ocultos (internos) e de saída, já que o próprio sistema fica encarregado de agrupar, classificar, planejar, configurar e otimizar, como os programas dos bancos para reconhecer a normalidade das transações, por exemplo.

Ao contrário de como ocorre em máquinas de aprendizado supervisionado, as IAs não supervisionadas, como aquelas utilizadas para a elaboração de perfil de consumidores de uma loja para criar uma disposição de produtos que maximize as vendas, não existe uma dependência de rotulagem predefinida dos perfis de compra para atingir o objetivo pretendido. Assim, o aprendizado desse sistema será através da percepção de padrões e subsequente otimização utilizando as informações de entrada como referência e ajuste de perfil conforme a necessidade (HONDA; FACURE; YAOHAO, 2017).

Os algoritmos de aprendizado por reforço, por sua vez, são ensinados pelo clássico sistema de tentativa e erro com recompensas e punições, usualmente aplicado para adestramento de animais, ou para ensinar uma nova habilidade a um humano, como tocar violino, por exemplo. Isso se dá pela exploração do ambiente pelo robô ou software, e a cada uma de suas ações há uma valoração da atitude tomada que permite que ele dê respostas mais assertivas para o que é proposto. Assim, não é necessário um sinal de erro, apenas uma indicação sobre se a resposta dada é boa ou ruim, geralmente sob a forma de um sinal numérico (SUN, 2014, p. 112).

Esse método é usado geralmente em sistemas que aprendem a jogar jogos como o *AlphaGo* da *Alphabet Inc*., que usa o aprendizado por reforço para avaliar suas jogadas no tradicional jogo chinês, GO, e consequentemente, as aperfeiçoa. Isso ocorre porque a máquina experimenta diferentes movimentos e observa as recompensas associadas, tendo como objetivo aprender um conjunto de estratégias que aumente a recompensa cumulativa ao longo dos jogos (KHOS, 2017).

Muito embora a divisão dos métodos de aprendizado tenha se dado em categorias bem definidas, é importante esclarecer que muitas das IAs modernas geralmente se baseiam em mais de um modelo de aprendizagem, podendo incluir outras não mencionadas. A exemplo disso há os sistemas de reconhecimento de transações bancárias em que há uma mescla entre aprendizado supervisionado e não supervisionado.

2. Como as máquinas tomam decisões

Mecanismos de decisão vão desde algoritmos de recomendação de filmes e músicas em aplicativos de *streaming* a algoritmos que recomendam o cálculo de pena e permissão ou proibição de liberdade condicional a réus (como veremos adiante).

No tocante às IAs que tomam decisões, saber como elas funcionam depende da natureza das questões apresentadas e o resultado que se espera atingir. O problema apresentado pode ser de caráter dinâmico (que necessita de uma sequência de decisões) ou estático (que necessita de apenas uma decisão) e se ele é lógico/determinístico (se A, então B) ou estocástico (depende de dados estatísticos) (AMIR, 2014, p. 202).

Para tarefas dinâmicas como um jogo de xadrez, por exemplo, o algoritmo precisa tomar decisões sequenciais e usa o seu aprendizado por reforço para fazer a avaliação de quais decisões devem ser tomadas, movimento a movimento. Por outro lado, uma sentença judicial de natureza estática (uma decisão única e terminativa do processo na instância em que se encontra) ensejaria um mecanismo artificial diverso. Para Ronald Dworkin, por exemplo, "os processos judiciais suscitam, pelo menos em princípio, três diferentes questões: questões de fato, questões de direito e as questões interligadas de moralidade política e fidelidade" (DWORKIN, 2007, p. 5-6). Isso pode significar que questões legais não podem ser reduzidas ao lógico ou ao estocástico. Ainda que haja elementos de lógica ou de estatística nas questões judiciais, possivelmente os casos não são integralmente compostos por apenas um desses elementos.

Supor que a interpretação do direito é apenas lógica, "se ocorreu o fato A, então aplica-se a legislação B", é pressupor que as questões jurídicas possuem uma relação de causa e consequência matemática. Embora possa existir essa relação entre fatos e Direito, as interpretações nesse campo, para além de questões linguísticas, sintáticas ou semânticas, decorrem essencialmente de fatores humanos (DWORKIN, 2007, p. 39). Portanto, ainda que máquinas compreendam a linguagem natural, elas estão confinadas apenas ao texto, uma fração superficial do fenômeno linguístico, já que a aquisição da linguagem humana se dá por uma

interação afetiva, sociolinguística, dificilmente capturada por um algoritmo ou uma rede neural (FINGER, 2021).

Também é precipitado supor que a interpretação jurídica baseia-se essencialmente em dados estatísticos, ainda que a interpretação de um determinado dispositivo legal possa decorrer de casos semelhantes. Isso porque pode haver uma limitação de novas interpretações (ou a superação de precedentes - *overruling*), principalmente para casos excepcionais, os chamados *hard cases*.

Assim, ainda que uma máquina possa jogar xadrez muito bem, evitar colisões em carros autônomos e elaborar perfis de consumidores, parece existir um limite de capacidade para as IAs. No entanto, cabe acrescer que o *machine learning* e *deep learning* atuais são apenas os estágios iniciais das IAs, e que em estágios mais avançados elas poderão ter *mind theory*, que é a capacidade de discernir as diferentes entidades com as quais interagem, entendendo suas emoções, raciocínio, etc. Nesse sentido, o seu estágio avançado seria o *self-aware* IA, ou seja, a capacidade de ser autoconsciente (ARAS, 2020, p. 91).

3. Informatização e formas de uso da IA no Direito

O uso de tecnologia no Direito brasileiro remonta à virada do século, em 2001, com a criação do "Sistema Justiça", uma solução informatizada de integração entre os gabinetes e órgãos da jurisdição do Superior Tribunal de Justiça. Posteriormente, a Lei nº 11.419, de 19 de dezembro de 2006, passou a regular a informatização do processo, que resultou na extinção de demandas físicas (em papel) a partir de 2009. Anos depois, em 2017 foi introduzido o uso de IAs pelo Judiciário através do Projeto VICTOR do Supremo Tribunal Federal, desenvolvido em parceria com a Universidade de Brasília (UnB), que teve como finalidade a criação de um sistema para agrupar os recursos recebidos pelo tribunal de acordo com o tema de repercussão geral em que se encaixam.

Esse projeto foi o ponto de partida para que vários outros surgissem, como a parceria de cooperação científica entre a Universidade de São Paulo e o Tribunal de Justiça do Estado de São Paulo (JORNAL DA USP, 2019). Segundo levantamento feito pelo Conselho Nacional de Justiça, só em 2022 o Brasil já contava com 111 projetos de IA nessa seara - o que representa um aumento de 171% em relação ao levantamento realizado em 2021 -, a maioria desenvolvidos em parcerias semelhantes aos casos de São Paulo e do STF (MAEGI, 2022).

O sistema VICTOR possui uma capacidade de avaliar os recursos extraordinários que se enquadram em um dos temas que lhe foi ensinado a identificar de forma muito mais ágil do que os servidores do tribunal (TEIXEIRA, 2018). No entanto, o uso de IAs no campo do Direito

não se restringe ao Poder Judiciário, podendo ser observadas iniciativas do tipo em áreas da advocacia privada. Um estudo de 2017 realizado pelo *McKinsey Global Institute* aponta que é possível automatizar 23% do trabalho de um advogado, com um vigoroso aumento na eficiência e velocidade do serviço desempenhado (LOHR, 2017). Nesse sentido, muitos escritórios de advocacia brasileiros têm otimizado seus trabalhos com uso de IAs, como o Urbano e Vitalino de Recife que conta com a assistência da "Carol" que utiliza a plataforma WATSON da IBM (MIGALHAS, 2017).

Já no Judiciário, para além das tarefas administrativas e de triagem, é possível encontrar casos em outros países, como o do Supremo Tribunal Popular da China, que integrou IAs em seu sistema jurídico também em processos de tomada de decisão alimentado por tecnologia de *machine learning* que analisa automaticamente os processos judiciais em busca de referências, recomenda leis e regulamentos, além de redigir documentos e corrigir erros humanos (CHEN, 2022).

O curioso do caso chinês é que os juízes devem usar a IA de forma ampla e seguir as recomendações da máquina, caso contrário, devem apresentar uma fundamentação das razões pelas quais não usaram a sentença redigida pelo computador. Porém, a implementação em larga escala das IAs no Judiciário chines é criticada por parte dos juízes, que defendem que as expectativas em relação às IAs são muito altas e que o papel desempenhado pelos juízes deve ser mantido (CHEN, 2022). Assim, casos do *Beijing Internet Court*, em que os processos são julgados por uma IA, sem uma supervisão humana, levantam o questionamento de quais espaços as IAs devem tomar dentro de um contexto iminente de implementação dessa tecnologia (LYNN, 2019).

No hemisfério ocidental também não faltam exemplos de usos jurídicos com finalidades de tomada de decisão, como o sistema *Harm Assessment Risk Tool* (HART) no Reino Unido e o COMPAS nos Estados Unidos. Ambos são sistemas semelhantes que possuem algoritmos que traçam perfis criminais e auxiliam de policiais a juízes na predição de futuras ocorrências sugerindo-lhes, por exemplo, uma resposta positiva ou negativa a um pedido de custódia (RESEARCH, 2023). E mesmo no Brasil já é possível encontrar projetos de IAs que sugerem prolação de sentenças com base na leitura de documentos, como é o caso do sistema CLARA do Tribunal de Justiça do Rio Grande do Norte (CABRAL, 2021).

Desta forma, parece haver hoje dois modos muito claros de uso de inteligência artificial no campo do Direito, uma como *database* ou auxiliar e outra como IA generativa, capaz de gerar textos, imagens e outras mídias, que, neste caso específico, aconselha ou profere decisões judiciais.

4. Possíveis benefícios na implementação de IA no Judiciário

Com 81 milhões de processos pendentes de julgamento, que representam um gasto efetivo, descontadas as despesas com servidores inativos, de R$ 94,4 bilhões de reais (CNJ, 2023), seria desarrazoado abrir mão do uso da tecnologia como forma de obtenção de eficiência na prestação do serviço nos tribunais.

Diante dessa situação, é importante compreender o papel que a estatística desempenha nas atividades do Poder Judiciário, especialmente no processo decisório, ainda que os problemas enfrentados pelo Direito não sejam integralmente de natureza estocástica, como ponderado anteriormente.

A grande vantagem de se aplicar a habilidade estatística dos computadores digitais no processo decisório, especificamente, está contida no conceito de estatística inferencial, a saber, área da estatística que busca fazer "considerações a respeito do que não foi observado diretamente e fazer inferências sobre, por exemplo, o comportamento futuro de variáveis, a associação entre duas ou mais variáveis ou as características de uma parcela não observada da população" (NUNES, 2019, p. 54). Essa possibilidade de indução estatística dialoga com a capacidade das IAs de identificar padrões e realizar generalizações com base nas informações pregressas sem necessidade de nova programação, e isso de modo muito mais acelerado do que qualquer mente humana poderia fazer. Assim, não é desarrazoado pensar em uma IA no processo decisório capaz de proferir sentenças cuja base de dados são os precedentes de um determinado tribunal.

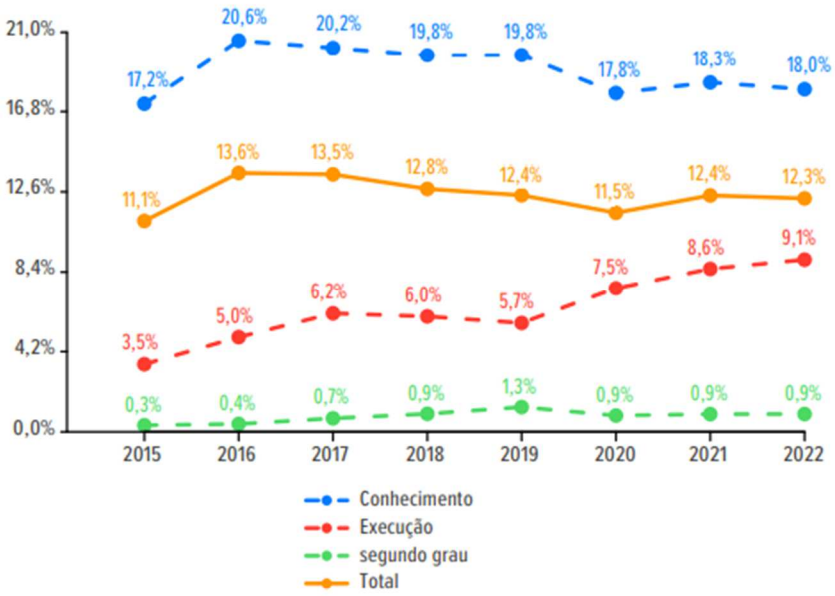
O Código de Processo Civil, em seu art. 925, parece trazer a previsibilidade na uniformização da jurisprudência. Um exemplo é o seu art. 927, que em seus incisos traz as ferramentas de Incidente de Resolução de Demandas Repetitivas (IRDR), os Incidentes de Assunção de Competência (IAC), bem como os Recursos Especiais (REsp) e Recursos Extraordinários (RE) repetitivos. Incrementar essa previsão legislativa com a habilidade estatística da inteligência artificial poderia tornar mais efetiva a replicação da decisão de um processo paradigma sobre os processos de mesmo tema.

Outra vantagem no uso de IAs em processo decisório reside no fato de que elas permitiriam as partes "prever" os resultados de um eventual julgamento (desde que os fatos relevantes e objeto da lide estejam claros), e assim, favorecer os meios de autocomposição e, por conseguinte, reduzir a demanda por acesso à justiça e, consequentemente, o contingente de processos (LUCON, 2022). Se uma das partes sabe que perderá uma determinada demanda, ela pode propor acordo, ou nem mesmo entrar com um processo, como faz a Procuradoria-Geral

da Fazenda Nacional com o PGFN *analytics* que “fornece aos procuradores as probabilidades de êxito de uma execução fiscal (baseado, por exemplo, em dados referentes a bens em nome do contribuinte a ser executado), indicando se deve ser ajuizada ou não” (MACHADO SEGUNDO, 2020).

Nesse aspecto, dar previsibilidade às lides menos complexas por meio de decisões automatizadas poderia estimular a resolução consensual, que embora tenha aumentado desde o advento do Novo Código de Processo Civil, ainda é baixa, como mostra o gráfico elaborado pelo Relatório Justiça em Números 2023 do CNJ :

Gráfico 1 - Série histórica do Índice de Conciliação



Fonte: Relatório Justiça em Números (CNJ, 2023)

5. Principais ressalvas quanto a aplicação de IA no Judiciário

As desvantagens do ingresso das IAs no Poder Judiciário parecem recair, especialmente, sobre a aplicação no processo decisório que realizam perfilamento de acusados (ou fins congêneres), pois os vieses de programação podem potencializar preconceitos estruturais, além da questão da opacidade dos algoritmos.

5.1. Vieses da máquina e preconceito estrutural

Existem vários relatos de casos de tratamento discriminatório realizados por máquinas, confirmando que as IAs não são completamente imparciais como alguns podem pensar, ou seja, não há uma neutralidade nos sistemas, uma vez que eles são programados por seres humanos e

os *inputs* (dados fornecidos) são adicionados por programadores também humanos e com base em informações que podem estar contaminadas por preconceito de gênero, raça, idade, etc.

Podemos falar, assim, na existência de um *algoritmo deturpado*, que são aqueles que apresentam padrões deturpados em sua formação. Sob a aparência de neutralidade, eles podem perpetuar vieses difíceis de serem percebidos, eventualmente até mesmo por seus programadores, resultando na legitimação, pela via tecnológica, de tratamentos desiguais e discriminatórios (ROQUE, 2021).

O enviesamento pode ocorrer em três momentos diversos: 1) no enquadramento do problema, 2) na coleta de dados e 3) em sua posterior preparação.

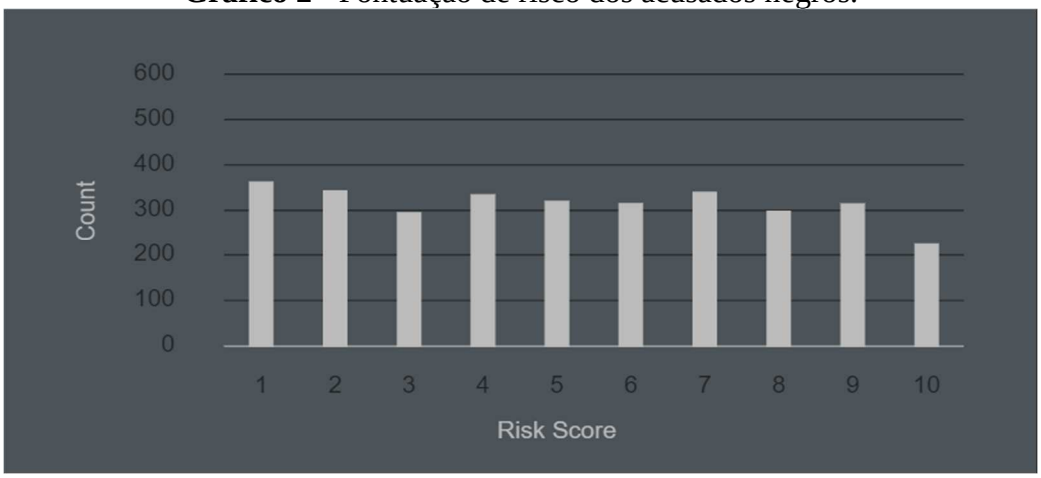
1) O enquadramento é definido como a fase em que os programadores estabelecem qual objetivo o algoritmo deve alcançar. Ocorre que, uma vez que se defina a atividade equivocadamente (ou de forma a não perseguir preceitos isonômicos), o sistema de aprendizado automático não vai espontaneamente usar parametros de valoração ético e perseguirá seu objetivo até o fim, sem medir o impacto da discriminação (HAO, 2019).

2) A coleta e preparação dos dados para o treinamento consiste na seleção de informações que a máquina usará como base para aprender e, em seguida, prepará-la para considerar ou não determinada coletânea de atributos para a realização da tarefa proposta. Assim, o enviesamento na fase de coleta pode ocorrer quando os dados inseridos não são representativos da realidade (por exemplo, a máquina tem como objetivo reconhecimento facial e é treinada apenas com rostos ocidentais brancos), ou quando os dados refletem preconceitos já existentes (como um software de recrutamento de um RH que seja treinado com contratações pregressas, que possivelmente refletem um padrão de contratação masculina).

Uma pesquisa realizada pela *ProPublica* entre os anos de 2013 e 2014 demonstrou como o sistema COMPAS reflete preconceitos estruturais de impactos danosos para a população afro-americana. O software desenvolvido pela empresa privada *Northpointe*, trabalha aplicando aos acusados (ou fichados) um questionário com perguntas como "Um dos seus pais foi alguma vez enviado para a prisão?", ou ainda, "Quantos dos seus amigos/conhecidos consomem drogas ilegalmente?", e a partir disso elabora um perfil de risco de reincidência criminal de 1 (baixo) a 10 (alto). A raça não é uma questão feita diretamente no perfilamento, mas as questões tangenciam temas relacionados à raça como o desemprego, a pobreza e a marginalização social. Nesse contexto, o COMPAS concluiu que jovens negros com histórico de exclusão obtiveram notas maiores de potencial ofensivo ou reicidencia, do que uma pessoa branca.

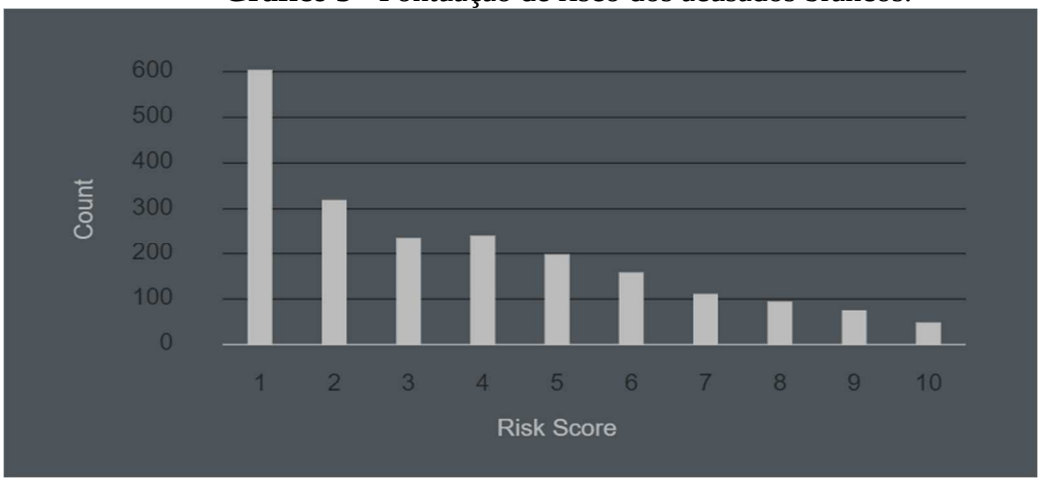
Sobre essa diferenciação na graduação de periculosidade de reincidência do *COMPAS*, a ProPublica elaborou dois gráficos interessantes que corroboram com a ideia de que a criação de perfil baseado nas perguntas mencionadas direcionam o sistema para se comportar de maneira discriminatória:

Grafico 2 - Pontuação de risco dos acusados negros.



Fonte: Análise da ProPublica dos dados do condado de Broward, FL (ANGWIN; LARSON; MATTU; KIRCHNER, 2016).

Gráfico 3 - Pontuação de risco dos acusados brancos.



Fonte: Análise da ProPublica dos dados do condado de Broward, FL (ANGWIN; LARSON; MATTU; KIRCHNER, 2016).

Um caso emblemático relatado no estudo da ProPublica compara situações envolvendo Brisha Borden, de 18 anos, e Vernon Prater, de 41 anos, ambos detidos no mesmo dia por crime semelhante e com dano em valor muito similar de US$80. Ambos tinham passagens anteriores, Vernon já havia sido preso por assalto a mão armada, enquanto Brisha por delitos de menor potencial ofensivo durante sua adolescência. Ao serem fichados no departamento de polícia de Fort Lauderdale, o sistema COMPAS deu a Brisha (uma jovem negra) grau de periculosidade 8, enquanto para Vernon (homem branco) o grau foi 3, no entanto,

após dois anos Vernon foi preso por roubar um galpão com milhares de dólares em produtos eletrônicos, enquanto que Brisha não retornou ao sistema de justiça.

A pesquisa da ProPublica mostra que o questionário que baseia o aprendizado da IA para auxiliar os juízes apresenta falhas graves que podem induzir maior penalização ou restrição de liberdade a pessoas negras. Esse é apenas um dos muitos que apresenta esse mesmo tipo de falha, restando claro que o *software* para avaliar reincidência em crimes e dar suporte na elaboração sobre condicional, reforça preconceitos humanos que já estão presentes na estrutura do Poder Judiciário.

A ProPublica, na mesma pesquisa, elaborou uma tabela sobre a falha nas previsões de reincidência do COMPAS, comparando pessoas negras e pessoas brancas:

Tabela 1 - Previsões falhas de reincidência criminal

	Brancos	Pretos
Ranqueados como alto risco, mas sem reincidência	23,5%	44,9%
Ranqueados como baixo risco, mas reincidentes	47,7%	28,0%

Fonte: Análise da ProPublica dos dados do condado de Broward, FL (ANGWIN; LARSON; MATTU; KIRCHNER, 2016).

Embora haja claramente um problema de enviesamento de *software*, há, por trás disso, um problema humano, que decorre de como os seres humanos convertem seus preconceitos (inconscientes ou não) em linguagem de programação. Por isso é preciso oferecer bases humanistas para os programadores de modo que esse problema seja evitado logo nas primeiras etapas do treinamento da máquina.

5.2. Opacidade dos algoritmos

Outra grande questão sobre a implementação de inteligências artificiais em âmbitos mais complexos da sociedade como o Direito é que, na eventualidade de deturpação algorítmica, não há muita clareza de como a máquina chegou a um determinado resultado (algumas vezes nem para os programadores). Sabe-se, é claro, que na ocorrência de um evento discriminatório causado por sistemas de computador há um evidente problema na alimentação e/ou preparação de dados do programa, mas não é sabido o cômputo que o levou a tal. Portanto, na implementação dessas tecnologias no Poder Judiciário, deve haver especial atenção para o potencial perigo que a opacidade algorítmica representa, pois isso pode implicar em sérias consequências ao devido processo legal.

De modo geral, o tratamento e armazenamento de dados e sua posterior computação são pouco claros o que acarreta uma situação de vulnerabilidade dos titulares de dados e uma relação de desequilíbrio entre desenvolvedores (ou programadores) e usuários (COSTA, 20121). Essa relação pode colocar os jurisdicionados em desvantagem, pois nem ele, nem seu patrono e, talvez, nem o magistrado, possuam conhecimento adequado para contestar a fórmula de como um determinado *input* chegou àquele *output*, ou mesmo para contestar os critérios de seleção e tratamento desses *inputs*.

Um bom exemplo é o caso *State v. Eric Loomis*, de 2016, em que Eric Loomis recorreu de uma sentença, na qual foi usado o sistema COMPAS, alegando a violação do direito ao devido processo legal. Isso porque os relatórios do COMPAS utilizam uma metodologia com base em dados probabilísticos dos quais não se sabe a métrica, por ser segredo comercial da *Northpointe*, empresa que o produz (STATE v. LOOMIS). Embora a corte de Wisconsin não tenha considerado o uso de algoritmos como violação de direitos, foi indicada a implementação de um aviso escrito para sinalizar aos magistrados a origem da avaliação. No entanto, apenas essa sinalização é insuficiente para alertar os magistrados do real risco do uso dessas ferramentas como parametrizador de decisões judiciais, já que o aviso é omisso quanto às críticas que se tem em relação às avaliações algorítmicas. Ele desconsidera a incapacidade técnica dos juízes em realizar o escrutínio dessas avaliações e não leva em conta a pressão interna e externa que existe no Judiciário para que os julgadores usem essas avaliações (STATE v. LOOMIS).

Ressalte-se que o cerne da questão não está necessariamente em informar ou não os magistrados que uma decisão sofreu incidência algorítmica parcial ou integral, mas na opacidade e suas consequências. Isso porque, no caso do aconselhamento na prolação da sentença, o viés apresentado pela máquina pode influenciar o juiz a validar o erro por meio do reforço de um preconceito que o julgador eventualmente já possua, ainda que inconscientemente. Já em casos de aplicação integral, coloca-se sob responsabilidade do magistrado a função de ratificar a decisão judicial e verificar vieses algorítmicos, algo que não lhe compete, posto que ele mesmo pode estar sujeito a elementos externos que o influenciam a decidir de uma maneira ou outra.

Na construção de uma decisão judicial tradicional, não “algoritmizada”, deve-se obedecer aos princípios da publicidade e da fundamentação, haja vista que para a realização da boa defesa os litigantes precisam conhecer o “itinerário lógico que o juiz percorreu para chegar a sua conclusão” (CALAMANDREI, 1995, p. 175). No caso do “juiz-robô”, deve haver

transparência e explicação de como ele chegou ao resultado final, com a demonstração de quais critérios utilizou?

Mesmo considerando questões de propriedade intelectual, relacionados a segredos comerciais, há opiniões favoráveis ao acesso dos mecanismos de tomada de decisão automatizados. O argumento é de que o esclarecimento da lógica da decisão algorítmica não pressupõe necessariamente a revelação do respectivo código-fonte. Ademais, argumenta-se que a própria Lei de Propriedade Industrial abre exceções ao segredo de indústrias no caso de ações judiciais (CANTARINI; GUERRA FILHO, 20220, p. 13) .

De toda sorte, no que diz respeito à regulação de IA, não há na legislação brasileira, especialmente na Lei Geral de Proteção de Dados Pessoais (LGPD), nada relacionado a algoritmos.

Uma alternativa à regulação estatal seria a autorregulação. O problema com os regimes de autorregulação é que geralmente eles impõem poucos limites às práticas corporativas (ZUBOFF, 2019, p. 321-322). Não deixa de ser ingênua a crença de que grandes empresas, especialmente as *big techs,* irão, diante de evidentes interesses econômicos, restringir seus lucros por questões éticas e morais.

5.3. Racionalidade das máquinas

Há no senso comum a ideia de que máquinas podem ser plenamente racionais e também imparciais, especialmente por conta do fator matemático e lógico de suas programações.

Por outro lado, na história da Filosofia do Direito não é incomum a noção de um Direito justo, baseado na razão. Exemplos podem ser facilmente encontrados em teorias do Direito natural, em trabalhos doutrinários de caráter dogmático e em dispositivos legais. No entanto, nos últimos tempos, inúmeras pesquisas têm demonstrado que juízes, cotidianamente, ainda que de maneira inconsciente, se valem de fatores emocionais e, portanto, não puramente racionais, para resolver questões de Direito. Trata-se, na verdade, de em uma mescla de razão e emoção no processo de tomada de decisão judicial (NOJIRI, 2021).

Assim, a perspectiva de uma razão jurídica como sinônimo de raciocínio lógico, além de equivocada, pode criar uma falsa expectativa de que uma IA pode substituir a figura humana do juiz já que ela pode preencher, até com mais efetividade, o requisito da racionalidade jurídica. Entretanto, o trabalho do juiz é muito mais complexo que a simples interpretação literal ou lógica das leis, uma vez que exige a compreensão de uma série de fatores externos que constituem a experiência e o processo cognitivo humanos.

Atualmente quase não há quem discorde de que fatores externos ao Direito como a ideologia política, o pensamento econômico e os sentimentos podem exercer influência sobre o modo de decisão do juiz. A questão que se coloca é como um sistema treinado para proferir decisões judiciais poderia tentar relacionar esses fatores extralegais a um caso concreto. Algoritmos seriam capazes de reproduzir sentimentos humanos que colaboram para uma decisão socialmente mais "justa"?

O caráter multifacetado do Direito pode exigir de seus operadores que tenham compreensão de que alguns aspectos da vida humana para que, no horizonte futuro, possa haver uma ressignificação de conceitos jurídicos. Sabemos que para essa tarefa é necessária a construção de um sólido arcabouço de experiências e conhecimentos humanos que possibilitem essa ressignificação que tanto importa ao Direito. Nossa dúvida recai na capacidade das máquinas em atingir esse objetivo, particularmente em razão de sua atuação baseada em dados, portanto, de caráter retrospectivo e não prospectivo.

6. Regulações Internacionais sobre o uso de IAs

Desde meados da segunda década do século XXI países do Norte Global com tradição em desenvolvimento tecnológico vêm buscando diretrizes para estabelecer padrões éticos para o desenvolvimento de robôs, para regulação de dados, tráfego de informações pela internet e, mais recentemente, para a operabilidade dos modelos de inteligência artificial. Muito embora esses códigos e diretrizes não possuam força normativa ou vinculante nos países em que foram estabelecidas, demonstram a crescente preocupação da sociedade.

Dentre as iniciativas, a mais antiga, de 2016, é a *Responsible Data Science.* De acordo com ela, o armazenamento e tratamento de dados das IAs devem seguir os seguintes princípios: a) *Fairness* (Justiça) a ideia de produção de ciência de dados de modo a evitar preconceitos; b) *Accuracy* (Acurácia) que produza resultados assertivos; c) *Confidentiality* (Confidencialidade) que garanta a privacidade dos titulares de dados; e d) *Transparency* (Transparência) e que produza resultados indiscutíveis no sentido de que as pessoas consigam inferir e interpretar corretamente os resultados (RESPONSIBLE DATA SCIENCE, 2016).

Em linha semelhante, em 2017, o *Association for Computer Machinery US Public Policy Council* (USACM), de forma mais específica, também definiu princípios orientando sistemas algorítmicos, no qual estão incluídas as inteligências artificiais. A declaração elaborada pelo USACM consiste em sete princípios (USACM, 2017): 1. Consciencialização; 2. Acesso e recurso; 3. Responsabilização; 4. Explicação; 5. Proveniência dos dados; 6. Auditabilidade; 7. Validação e testes.

E, mais recentemente, em 2019, a *World Comission on the Ethics of Scientific Knowledge and Technology* (COMEST) da UNESCO elaborou um estudo preliminar acerca da ética em inteligência artificial, em que também estabeleceu recomendações de uso e implementação dessa tecnologia. Nesse estudo, a comissão faz uma distinção entre declaração e recomendação, optando por fazer uma recomendação, pois reconhece que heterogeneidade de recomendações de parâmetros éticos de utilização de IA vêm sendo elaborados pelos mais diversos atores com os mais diversos interesses e pelo regramento de uma declaração a publicação e adesão seriam dificultadas (UNESCO, 2019). Assim, os princípios defendidos na recomendação buscam ser os mais genéricos possíveis e de acordo com o que as Nações Unidas defendem em linhas gerais, como: Direitos Humanos, Inclusividade, Bem-estar social, Autonomia, Explicabilidade, Transparência, Conhecimento e aprendizado (sobre IA para os cidadãos), Responsabilidade, Auditabilidade, Democracia, Boa governança e Sustentabilidade.

Esses exemplos de iniciativas internacionais apontam, dentre outras coisas, para questões já levantadas neste trabalho: os vieses e a opacidade dos sistemas algorítmicos. A implementação desses sistemas no processo decisório brasileiro precisa considerar essas recomendações éticas de modo a regulamentar modelos tecnológicos para que possa proporcionar aos jurisdicionados as melhores decisões judiciais possíveis.

Conclusões

Ao final das contas, a reflexão central que o presente estudo busca é sobre o papel da cognição humana nas decisões judiciais. Além disso, busca apresentar um breve panorama do papel do juiz em uma sociedade cada vez mais informatizada e que de certa forma equipara a cognição humana a uma questão lógico-matemática. Apesar disso, não se busca aqui vilanizar o processo decisório automatizado, pois há evidentes vantagens de cunho prático na implementação dessa tecnologia.

Em um cenário de legislação numerosa e complexa como a brasileira, com um vasto rol de precedentes, e um congestionamento de processos nos tribunais é inegável que o uso de inteligência artificial representa uma forma de aprimorar a prestação jurisdicional.

De outro lado, o desenvolvimento de tecnologias de IA envolve questões pouco claras para leigos e profissionais da área jurídica e criam problemas que podem afetar bens constitucionalmente assegurados como a igualdade e a transparência na fundamentação. Nesse sentido, o presente trabalho destaca dois pontos que afetam o processo decisório: a reprodução de preconceitos humanos decorrente de erro no treinamento do aprendizado (vieses da máquina) e a opacidade de cômputo dos algoritmos que dificultam o combate desses vieses.

As diversas discussões éticas apresentadas por entidades e associações da sociedade civil representam um norte para a implementação desses sistemas. Grande parte delas aponta para a necessidade de as decisões serem explicáveis, transparentes e com possibilidade de auditoria humana. Esses são apenas alguns dos desafios que os desenvolvedores de tecnologia devem enfrentar para um devido processo de aperfeiçoamento das máquinas.

No atual estado em que a tecnologia se encontra, o papel da IA no processo decisório deve se ater a temas menos complexos e que pouco exigem de uma perspectiva humana, que em geral são atividades massificadas e repetitivas, de identificação de casos semelhantes, organização de documentos, dentre outras atividades congêneres.

A IA representa para a história humana uma singularidade e autonomia sem precedentes (WEIZENBAUM, 1976) e propõe dilemas éticos que por vezes parecem irresolúveis. Em certo aspecto o aprendizado de máquina parece estar crescendo em progressão geométrica enquanto as discussões sobre sua singularidade e seus problemas parecem caminhar em progressão aritmética. Assim, a utilização de IAs em processo decisório demanda uma reflexão séria e crítica como maneira de se buscar parâmetros para a sua implementação.

Referências Bibliográficas

AMIR, Eyal. Reasoning and decision making. *In*: FRANKISH, Keith; RAMSEY, William M.(org.). **The Cambridge Handbook of Artificial Intelligence**. Cambridge: Cambridge University Press, 2014. Disponível em: <https://www.cambridge.org/core/books/cambridge-handbook-of-artificial-intelligence/reasoning-and-decision-making/30DFFAF6A0FF92E1D56664302DB41FDE>. Acesso em: 20 ago. 2024.

ANGWIN, Julia; LARSON, Jeff; MATTU, Suria; KIRCHNER, Lauren. Machine Bias. **ProPublica**, 23 mai 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. Acesso em: 02 out 2024.

ARAS, Vladmir. A inteligência artificial e o direito de ser julgado por seres humanos. *In*: PINTO, Henrique Alves. GUEDES, Jefferson Carús. CÉSAR, Joaquim Portes de Cerqueira. (coord.). **Inteligência artificial aplicada ao processo de tomada de decisões**. 1 ed. Belo Horizonte: Editora D'Plácido, 2020.

BODEN, Margaret A. GOFAI. *In*: FRANKISH, Keith; RAMSEY, William M.(org.). **The Cambridge Handbook of Artificial Intelligence**. Cambridge: Cambridge University Press, 2014. p. 89. Disponível em: <https://www.cambridge.org/core/books/cambridge-handbook-of-artificial-intelligence/gofai/FCF7D6DD921658FE8AE9F2A2B0FECBDD>. Acesso em: 18 ago. 2023.

CABRAL, Thiago Dias Delfino. Juízes-robôs? Notas sobre a utilização da inteligência artificial pelo Poder Judiciário. *In*: FEIGELSON, Bruno. BECKER, Daniel. RODRIGUES, Marco Antonio. (coord.) **Litigation 4.0**: O futuro do processo civil vis-à-vis as novas tecnologias. 1ª ed. São Paulo: Thomson Reuters Brasil, 2021.

CALAMANDREI, Piero. **Eles, os juízes, vistos por um advogado**. Tradução Eduardo Brandão. São Paulo: Martins Fontes, 1995.

CANTARINI, Paola. GUERRA FILHO, Willis Santiago. Levando os Direitos Fundamentais a sério: da necessária ponderação entre os Direitos Fundamentais envolvidos em perícia algorítmica face à proteção de segredo industrial. *In*: LAPORTA. Celeida M. Celentano. CANTARINI, Paola. GUERRA FILHO, Willis Santiago. (org.) **Direito e inteligência artificial**: fundamentos: vol. 3 - desafios e impactos sociais. Rio de Janeiro: Lumen Juris, 2022.
CHEN, Stephen. China's court AI reaches every corner of justice system, advising judges and streamlining punishment. **South China Morning Post**, 2022. Disponível em: <https://www.scmp.com/news/china/science/article/3185140/chinas-court-ai-reaches-every-corner-justice-system-advising>. Acesso em: 20 ago 2024.
CNJ. A era digital. **Superior Tribunal de Justiça**, 2023. Disponível em: <https://www.stj.jus.br/sites/portalp/Institucional/Historia/A-era-digital#:~:text=O%20processo%20eletr%C3%B4nico%20toma%20forma,e%20intelig%C3%AAncia%20%C3%A0s%20rotinas%20processuais.>. Acesso em: 23 ago 2024.
COSTA, Inês da Silva. A proteção da pessoa na era dos big data: a opacidade do algoritmo e as decisões automatizadas. **Revista Electrónica de Direito**. Porto, v. 24, n. 1, pp. 34-82. fev. 2021. Disponível em: <https://cij.up.pt/client/files/0000000001/4-ines-costa_1677.pdf>. Acesso em: 10 out 2024.
DWORKIN, Ronald. **O império do direito**. Tradução Jeferson Luiz Camargo. 3ª ed. São Paulo: Martins Fontes, 2007.
FINGER, Marcelo. Inteligência Artificial e os rumos do processamento do portugês brasileiro. **Estudos Avançados**. *[S. l.]*, v. 35, n. 101, 2021. Disponível em: <https://www.revistas.usp.br/eav/article/view/185030/171214>. Acesso em: 28 Ago. 2024.
GORDON, Rachel. ELIZA wins Peabody Award. **Computer Science & Artificial Intelligence Laboratory of the Massachusetts Institute of Technology**, 2022. Disponível em: <https://www.csail.mit.edu/news/eliza-wins-peabody-award>. Acesso em: 2 de ago. 2023.
HAO, Karen. This is how AI bias really happens - and why it's so hard to fix. **MIT Technology Review**, 2019. Disponível em: https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-bias-really-happensand-why-its-so-hard-to-fix/>. Acesso em: 02 out 2024.
HONDA, Hugo. FACURE, Matheus. YAOHAO, Peng. Os três tipos de aprendizado de máquina. **LAMFO-UnB**, 2017. Disponível em: <https://lamfo-unb.github.io/2017/07/27/tres-tipos-am/#:~:text=Outros%20exemplos%20de%20aplica%C3%A7%C3%B5es%20de,anomalias%20e%20visualiza%C3%A7%C3%A3o%20de%20dados.> Acesso em: 18 de ago. 2024.
KHOS, Greg. AlphaGo. Moxie Pictures e Reel As Dirt. **Netflix**, 2017.
KONTSEVOI, Boris. Artificial Intelligence: How to turn conversational AI into a success business. **Forbes**, 2023. Disponível em: <https://www.forbes.com/sites/forbestechcouncil/2023/04/05/artificial-intelligence-how-to-turn-conversational-ai-into-a-success-business/?sh=4e9018c265e5>. Acesso em: 18 de ago. 2024.
LOHR, Steve. A.I. is doing legal work. But won't replace lawyers, yet. **New York Times**, 2017. Disponível em: <https://www.nytimes.com/2017/03/19/technology/lawyers-artificial-intelligence.html>. Acesso em: 20 ago 2024.
LYNN, Bryan. Robot Justice: The rise of China's 'Internet Courts'. **Voa News**, 2019. Disponível em: <https://learningenglish.voanews.com/a/robot-justice-the-rise-of-china-s-internet-courts-/5201677.html>. Acesso em: 20 ago 2024.
MAEGI, Vanessa. Justiça 4.0: Inteligência Artificial está presente na maioria dos tribunais brasileiros. **Conselho Nacional de Justiça**, 2022. Disponível em: <https://www.cnj.jus.br/justica-4-0-inteligencia-artificial-esta-presente-na-maioria-dos-tribunais-brasileiros/>. Acesso em: 23 Ago 2024.

MIGALHAS. Escritório investe em inteligência artificial para auxiliar em serviços jurídicos. **Migalhas**, 2017. Disponível em: <https://www.migalhas.com.br/quentes/271124/escritorio-investe-em-inteligencia-artificial-para-auxiliar-em-servicos-juridicos>. Acesso em: 20 ago 2024.
NOJIRI, Sergio. **Emoção e intuição**: como (de fato) se dá o processo de tomada de decisão judicial. Belo Horizonte: Arraes Editores, 2021.
NUNES, Dierle. MARQUES, Ana Luiza Pinto Coelho. Decisão judicial e inteligência artificial: é possível a automação da fundamentação? In: NUNES, Dierle. LUCON, Paulo Henrique dos Santos, WOLKART, Erik Navarro. (org.). **Inteligência Artificial e direito processual**: os impactos da virada tecnológica no direito processual. 2 ed. Salvador: Editora JusPodivm, 2021.
RESPONSIBLE DATA SCIENCE INITIATIVE, 2016. **Responsible Data Science**. Disponível em: <https://redasci.org/home/program/>. Acesso em: 29 out. 2024.
RESEARCH. Helping police make custody decisions using artificial intelligence. **University of Cambridge**, 2023. Disponível em: <https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence>. Acesso em: 20 ago 2024.
ROQUE, Andre Vasconcelos. SANTOS, Lucas Braz Rodrigues dos. Inteligência artificial na tomada de decisões judiciais: três premissas básicas. **Revista Eletrônica de Direito Processual**. Rio de Janeiro, ano 15, v. 22, n. 1, p. 58-78. jan./abr. 2021. Disponível em: <https://www.e-publicacoes.uerj.br/index.php/redp/article/view/53537/36309>. Acesso em: 02 out 2024.
STATE v. LOOMIS. Wisconsin Supreme Court Requires Warning Before Use of Algorithmic Risk Assessment in sentencing, **Harvard Law Review**. Cambridge, v. 130, p. 1530-1537, Março, 2017. Disponível em: <https://harvardlawreview.org/2017/03/state-v-loomis/ Acesso em: 31.mai.2024.
SUN, Ron. Connectionism and neural networks. *In*: FRANKISH, Keith; RAMSEY, William M.(org.). **The Cambridge Handbook of Artificial Intelligence**. Cambridge: Cambridge University Press, 2014. Disponível em: <https://www.cambridge.org/core/books/cambridge-handbook-of-artificial-intelligence/connectionism-and-neural-networks/EB7A3A03C8A440689A6B4D94EA3727D7>. Acesso em: 18 ago. 2024.
TEIXEIRA, Matheus. STF investe em inteligência artificial para dar celeridade a processos. **Jota**, 2018. Disponível em: <https://www.jota.info/coberturas-especiais/inova-e-acao/stf-aposta-inteligencia-artificial-celeridade-processos-11122018>. Acesso em: 20 ago 2024.
UNESCO. Preliminary study on the ethics of artificial intelligence. **World Commission on the Ethics of Scientific Knowledge and Technology**, 2019. disponível em: <https://unesdoc.unesco.org/ark:/48223/pf0000367823>. Acesso em: 30 out. 2024.
JORNAL DA USP. USP e Tribunal de Justiça estudam parceria na área de inteligência artificial. **Jornal da USP**, 2019. Disponível em: <https://jornal.usp.br/institucional/usp-e-tribunal-de-justica-estudam-parceria-na-area-de-inteligencia-artificial/>. Acesso em: 20 ago 2024.
VALENTINI, Rômulo Soares. **Julgamento por computadores?** As novas possibilidades da juscibernética no século XXI e suas implicações para o futuro do Direito e do trabalho dos Juristas. Tese de Doutorado. Faculdade de Direito, Universidade Federal de Minas Gerais. Belo Horizonte. p. 52. Disponível em: <https://repositorio.ufmg.br/handle/1843/BUOS-B5DPSA>. Acesso em: 10 de ago 2024.
WEIZENBAUM, Joseph. **Computer power and human reason: from judgment to calculation**. Nova Iorque: W. H. Freeman and Company, 1976.
ZUBOFF, Shoshana. **Age of Surveillance Capitalism**: The Fight for a Human Future at the New Frontier of Power. New York: Public Affairs, 2019.