XIII ENCONTRO INTERNACIONAL DO CONPEDI MONTEVIDÉU - URUGUAI

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

EDSON RICARDO SALEME

EUDES VITOR BEZERRA

CINTHIA OBLADEN DE ALMENDRA FREITAS

Copyright © 2024 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
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

Membro Nato – Presidência anterior Prof. Dr. Raymundo Juliano Feitosa – UNICAP – Pernambuco

D597

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

[Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Edson Ricardo Saleme, Eudes Vitor Bezerra, Cinthia Obladen de Almendra Freitas – Florianópolis: CONPEDI, 2024.

Inclui bibliografia
ISBN: 978-85-5505-990-2
Modo de acesso: www.conpedi.org.br em publicações
Tema: ESTADO DE DERECHO, INVESTIGACIÓN JURÍDICA E INNOVACIÓN

1. Direito – Estudo e ensino (Pós-graduação) – 2. Governo digital. 3. Novas tecnologias. XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU (2: 2024 : Florianópolis, Brasil).

CDU: 34



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

XIII ENCONTRO INTERNACIONAL DO CONPEDI URUGUAI – MONTEVIDÉU

GOVERNO DIGITAL, DIREITO E NOVAS TECNOLOGIAS II

Apresentação

O conjunto de pesquisas que são apresentadas neste livro faz parte do Grupo de Trabalho de "DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS II", ocorrido no âmbito do XIII Encontro Internacional do CONPEDI, realizado entre os dias 18, 19 e 20 de setembro de 2024, na cidade de Montevidéu, Uruguai, promovido pelo Conselho Nacional de Pesquisa e Pós-Graduação em Direito – CONPEDI e que teve como temática central "Estado de Direito, Investigação Jurídica e Inovação".

Os trabalhos expostos e debatidos abordaram de forma geral distintas temáticas atinentes DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS, especialmente relacionadas aos principais desafios que permeiam a tecnologias jurídica, passando pela inteligência artificial, demais meios digitais, também apontando para problemas emergentes e propostas de soluções advindas de pesquisas em nível de pós-graduação, especialmente, Mestrado e Doutorado.

Os artigos apresentados no Uruguai trouxeram discussões sobre: Tecnologias aplicáveis aos tribunais, Governança digital e governo digital, Função notarial e novas tecnologias, Exclusão digital derivando tanto para exclusão social quanto para acesso à justiça, Eleições, desinformação e deepfake, cidades e TICs. Não poderiam faltar artigos sobre privacidade e proteção de dados pessoais, com atenção aos dados sensíveis, consentimento e LGPD, liberdade de expressão, censura em redes sociais, discriminação, herança digital, microtrabalho e o trabalho feminino, uso de sistemas de IA no Poder Judiciário e IA Generativa.

Destaca-se a relevância e artigos relacionados ao tema de Inteligência Artificial, tratando de vieses algorítmicos e do AI Act. E, ainda, aplicação de sistemas de IA ao suporte de pessoas com visão subnormal. Para além das apresentações dos artigos, as discussões durante o GT foram profícuas com troca de experiências e estudos futuros. Metodologicamente, os artigos buscaram observar fenômenos envolvendo Direito e Tecnologia, sem esquecer dos fundamentos teóricos e, ainda, trazendo aspectos atualíssimos relativos aos riscos que ladeiam as novas tecnologias, destacando os princípios e fundamentos dos direitos fundamentais

Considerando todas essas temáticas relevantes, não pode ser outro senão de satisfação o sentimento que nós coordenadores temos ao apresentar a presente obra. É necessário, igualmente, agradecer imensamente aos pesquisadores que estiveram envolvidos tanto na confecção dos trabalhos quanto nos excelentes debates proporcionados neste Grupo de Trabalho. Por fim, fica o reconhecimento ao CONPEDI pela organização e realização de mais um relevante evento internacional.

A expectativa é de que esta obra possa contribuir com a compreensão dos problemas do cenário contemporâneo, com o a esperança de que as leituras dessas pesquisas ajudem na reflexão do atual caminhar do DIREITO, GOVERNANÇA E NOVAS TECNOLOGIAS.

Prof. Dr. Edson Ricardo Saleme (UNISANTOS)

Prof. Dr. Eudes Vitor Bezerra (PPGDIR – UFMA)

Profª. Dra. Cinthia Obladen de Almendra Freitas (PPGD - PUCPR)

INTELIGÊNCIA ARTIFICIAL E ALGORITMOS DESIGUAIS: UMA ABORDAGEM JURISDICIONAL

ARTIFICIAL INTELLIGENCE AND UNEQUAL ALGORITHMS: A JURISDICTIONAL APPROACH

Jamile Sabbad Carecho Cavalcante [1]
Paulo José Pereira Carneiro Torres da Silva [2]
Ana Lucia Pazos Moraes [3]

Resumo

O artigo aborda o tema da discriminação ou racismo algoritmo no Poder Judiciário e a preocupação ética a respeito da utilização desses sistemas no Brasil sem a devida regulamentação. A questão colocada tem sua importância social e jurídica diante do avanço de inteligências artificiais no Poder Judiciário evidenciando a necessidade de se enfrentar as questões sensíveis, como a discriminação algorítmica, para que esse avanço seja feito respeitando os direitos fundamentais e garantias processuais. Além de refletir quais os riscos do desenvolvimento desses sistemas sem balizas éticas mínimas. O método de pesquisa será o indutivo, pois partiremos da análise da evidência concreta de um caso de discriminação algorítmica, o Compas utilizado no Poder Judiciário Norte Americano, para uma compreensão do instituto da discriminação algorítmica e o risco da utilização de sistemas de inteligência artificial no Poder Judiciário sem diretrizes éticas aptas a garantirem o resguardo dos direitos fundamentais e garantias processuais. A preocupação ética no âmbito do desenvolvimento de sistemas de inteligência artificial se concentra, principalmente, em três momentos na elaboração/arquitetura do sistema, no tratamento e inserção dos dados e na interpretação dos resultados pelo ser humano. A questão está intrinsecamente ligada à participação humana em todas as fases, enfatizando a absorção de discriminações e preconceitos enraizados socialmente pelos algoritmos.

Palavras-chave: Inteligência artificial, Discriminação algorítmica, Poder judiciário, Direitos fundamentais, Ética

Abstract/Resumen/Résumé

The article addresses the issue of algorithmic discrimination or racism within the Judiciary

---

[1] Mestre em Direito Público na linha de Pesquisa Direitos Fundamentais e Novos Direitos na UNESA/RJ, com bolsa Prosup/CAPES (2023). Pesquisadora especialista em Inteligência Artificial.

[2] Universidade Estácio de Sá - Professor do Programa de Mestrado e Doutorado em Direito da Universidade Estácio de Sá na disciplina Direito, Inovação e Tecnologia.

[3] Mestre em Direito Público na linha de Pesquisa Acesso à Justiça e Fundamentos do Processo na UNESA/RJ, com bolsa Prosup/CAPES (2023). Professora Titular de Direito da Universidade Unigranrio Afya.

and the ethical concerns regarding the use of these systems in Brazil without proper regulation. The issue is socially and legally significant given the advancement of artificial intelligence in the Judiciary, highlighting the need to tackle sensitive issues, such as algorithmic discrimination, to ensure this progress respects fundamental rights and procedural guarantees. It also reflects on the risks of developing these systems without minimum ethical guidelines. The research method will be inductive, starting with the analysis of concrete evidence from a case of algorithmic discrimination, Compas, used in the North American Judiciary, to understand the concept of algorithmic discrimination and the risk of using artificial intelligence systems in the Judiciary without ethical guidelines capable of safeguarding fundamental rights and procedural guarantees. Ethical concerns in the development of artificial intelligence systems focus mainly on three stages: the system's design/architecture, data processing and insertion, and human interpretation of the results. The issue is intrinsically linked to human participation in all phases, emphasizing the absorption of socially ingrained discriminations and biases by algorithms.

Keywords/Palabras-claves/Mots-clés: Artificial intelligence, Algorithmic discrimination, Judiciary, Fundamental rights, Ethics

INTRODUÇÃO

O movimento de digitalização[1] da vida está presente em toda a sociedade da informação[2] na qual estamos inseridos. Atualmente, cada vez mais decisões significativas são delegadas a sistemas de inteligência artificial no setor público e privado, desde a concessão, ou não, de um crédito pessoal à possibilidade de responder um processo em liberdade com pagamento de fiança. Decisões estas que impactam diretamente no exercício da cidadania das pessoas com claros reflexos em seus direitos fundamentais. Conforme traz o autor Hoffmann-Riem (2022, p. 13) "exemplos de tais usos são no governo eletrônico ou na administração da justiça (e-justiça, *Legal Technology*)"

O mundo digital on-line é visto, cada vez mais, como uma extensão da vida física no off-line, o maior exemplo disso atualmente é a chamada Internet das Coisas. Segundo o professor Hoffmann-Riem (2022, p. 29), para explicar melhor essa nova vivência do século XXI, podemos utilizar o termo "*on-life*", um mundo intermediário entre o *off-line* e o *on-line*. Para o autor, "neste mundo, os sistemas de computador podem, em grande parte, libertar as pessoas da necessidade de tomar decisões, ou seja, substituir decisões humanas".

Nesse sentido, Hoffmann-Riem (2022, p. 3) acrescenta que o avanço da inteligência artificial amplia de maneira considerável o potencial da digitalização e da

[1] No livro Teoria do Direito Digital, Wolfgang Hoffmann-Riem trabalha a questão da digitalização de diversos setores da vida na sociedade da informação, trazendo que: "O termo "digitalização" refere-se inicialmente apenas às tecnologias da informação específicas que processam dados digitais e às infraestruturas (software e hardware) criadas para as tecnologias digitais. No entanto, o termo também representa a mudança fundamental nas condições de vida desencadeadas pela sua utilização em todo o mundo. Permite a utilização de sistemas ciberfísicos para novos processos de produção em rede e automatizados (por exemplo, na indústria 4.0), alterações na forma como as pessoas vivem as suas vidas (por exemplo, na "casa inteligente"), a criação e utilização de redes sociais (como o Google ou o Facebook) e outros novos serviços de comunicação (por exemplo, mensagens instantâneas), bem como novos sistemas de vigilância por empresas privadas e agências governamentais". (Hoffmann-Riem, 2022, p. 1, tradução por Italo Fuhrmann)

[2] A sociedade da informação é um termo trabalhado pelo filósofo Byung-Chul Han em seus livros. Embora o autor não seja o responsável por formar este termo, é uma opção sistêmica trabalhar o termo a partir da sua perspectiva. No livro "Infocracia: digitalização e a crise da democracia", Han traz que: "Transparência significa a política do se tornar visível do regime da informação. Quem só faz alusão à política pública da informação de uma instituição ou pessoa, ignora seu alcance. A transparência é a coação sistêmica do regime de informação. O imperativo da transparência é: tudo deve estar disponível na condição de informação. Transparência e informação têm o mesmo significado. A sociedade da informação é a sociedade da transparência. O imperativo da transparência faz com que as informações circulem livremente. Não são as pessoas que são realmente livres, mas as informações. O paradoxo da sociedade de informação é: as pessoas estão aprisionadas nas informações. Afivelam elas mesmas os grilhões ao se comunicarem e ao produzirem informações. O presídio digital é transparente." (Han, 2022, p. 14; tradução brasileira: Gabriel S. Philipson)

utilização de *Big Data*[3], tendo a Pandemia de Covid-19 um importante papel condutor das transformações com consequências permanentes na aplicação dessas tecnologias. Essa transformação digital está inerente a todos os domínios da vida e traz consigo conquistas antes não experimentadas ou imaginadas, mas, também, desafios e perigos a serem enfrentados.

As pesquisas e as descobertas acerca da relação entre a utilização da tecnologia para tomada de decisões e o aumento das desigualdades sociais existentes ainda estão no início. Por isso, a pesquisadora Sofiya Noble (2021, p. 18) destaca em seu livro que "a inteligência artificial se tornará uma questão de direitos humanos de grande importância no século XXI."

De acordo com a *Organization for Economic Co-operation and Development* (*OECD),* a Inteligência Artificial é uma tecnologia com um propósito geral e com o potencial de melhorar o bem-estar, contribuir para uma atividade econômica global sustentável positiva, aumentar a inovação e a produtividade e auxiliar nos atuais desafios globais. Sendo assim, para a Organização um sistema de Inteligência Artificial (IA) é:

> um sistema baseado em máquina que pode, para fins explícitos ou implícitos, inferir, a partir do input que recebe, como gerar saídas como previsões, conteúdo, recomendações ou decisões que podem influenciar ambientes físicos ou virtuais. Os sistemas de IA são diferentes e variam em seus níveis de autonomia e adaptabilidade após a implantação (OECD, 2019).

Neste trabalho iremos utilizar o método de pesquisa dedutivo apoiado em revisão bibliográfica examinando artigos nacionais e internacionais que explorem a temática da ética algorítmica, definições, causas, diagnósticos apresentados para o fenômeno, a partir da perspectiva brasileira da temática observando o desenvolvimento da Inteligência Artificial no Poder Judiciário brasileiro com a ampla utilização de *Large Language Models* e sistemas generativos no âmbito legal.

A pesquisa tem como objetivo demonstrar os riscos envolvidos no desenvolvimento e uso de sistemas inteligentes no Poder Judiciário brasileiro, destacando o fenômeno da discriminação algorítmica. A questão colocada tem sua importância social e jurídica diante do avanço de inteligências artificiais no Poder Judiciário evidenciando a necessidade de se enfrentar as questões sensíveis, como a discriminação e a inteligibilidade das decisões algorítmicas para que esse avanço seja feito respeitando os direitos fundamentais e garantias processuais, como veremos a seguir.

[3] "O termo refere-se à dimensão e à diversidade dos dados que podem ser utilizados para aplicação das tecnologias digitais, bem como às várias possibilidades de as combinar e avaliar e de as tratar pelas autoridades públicas e privadas em diferentes contextos." (Hoffmann-Riem, 2022, p. 2)

1. INTELIGÊNCIA ARTIFICIAL E PODER JUDICIÁRIO

O desenvolvimento e a utilização de Inteligência Artificial no Poder Judiciário têm sido significativos ao redor do mundo nos últimos anos. Nas principais manchetes jornalísticas, encontramos questões como “juiz robô”, “automação dos julgamentos”, “robôs na advocacia” causando uma primeira impressão um tanto quanto equivocada acerca do futuro das profissões jurídicas, levando a crer que tão logo inexistirão advogados, juízes, enfim, juristas humanos e que as máquinas realizarão todas as atividades jurídicas desde petições simples à elaboração de acórdãos.

Muito embora, uma leitura mais atenta e uma maior compreensão do tema levem ao entendimento correto do atual estado da arte de sistemas inteligentes como sistemas assistivos no Poder Judiciário. Utilizados com objetivo de auxiliar as diversas atividades repetitivas que realizamos dentro das repartições públicas e ambientes privados, esses sistemas colaboram na movimentação da grande máquina do Judiciário, como pesquisa jurisprudencial, análise de casos idênticos, elaboração de métricas matemáticas de quantificação de danos ou de gratuidade de justiça, dentre outras tantas hipóteses.

Portanto, sistemas inteligentes estão sendo desenvolvidos para auxiliar na celeridade processual e na melhoria da qualidade da prestação jurisdicional, com foco em atividades repetitivas, tanto pelos tribunais como pelas instituições essenciais ao funcionamento da justiça: o Ministério Público, a Defensoria Pública e a Advocacia Pública e Privada.

No Brasil, o Tribunal de Contas da União (TCU, 2022) avaliou o estágio atual e perspectivas de uso dos sistemas de inteligência artificial na Administração Pública Federal, a fim de identificar os riscos associados, o impacto de um controle e a proposta para uma Estratégia Brasileira de Inteligência Artificial. Segundo o levantamento, mais de um terço das organizações federais estão no nível zero de maturidade em IA. Isso quer dizer que não usam ou planejam utilizar essa tecnologia. Outro terço se encontra no nível 1 e já conversam sobre sistemas inteligentes de maneira especulativa.

Diante desse quadro recente, o TCU (2022) tomará atitudes de supervisões futuras com a ajuda de um sistema de auditoria próprio a ser desenvolvido para fiscalizar IA e a elaboração de uma guia com parâmetros e riscos na contratação e desenvolvimento de soluções de IA.

Enquanto isso, o uso de plataformas de Inteligência Artificial no Poder

Judiciário, desde 2017, só cresce. O aumento visa maior celeridade processual, contribuindo não apenas com a redução na duração dos processos, mas também com maior qualidade da prestação jurisdicional ofertada, e, em última análise, proporciona o combate à morosidade da Justiça e ao grande acúmulo de processos na justiça brasileira.

O Brasil tem o maior sistema judiciário do mundo, composto por 92 tribunais (BREHM, *et.al*, 2020). Em 2023, o setor recebeu 35 milhões de casos novos (CNJ, 2023) e, em apenas quatro meses de 2024, já contabiliza a entrada de mais 7 milhões de processos (CNJ, 2024a). Segundo a atualização de dados do DataJud feita em 31 de maio de 2024, existem pouco mais de 82 milhões de processos pendentes (CNJ, 2024a), restando evidente o quanto a tecnologia pode contribuir com o enfrentamento à morosidade da Justiça e o grande acúmulo de processos.

De acordo com o levantamento mais recente feito em 2023 pelo Conselho Nacional de Justiça (CNJ, 2024b) e divulgado em 2024, 62 dos 94 órgãos do Poder Judiciário já possuem algum tipo de iniciativa de IA chegando a um total de 140 projetos de inteligência artificial desenvolvidos ou em desenvolvimento nos tribunais. O que representa um aumento de 26% em relação a pesquisa anterior de 2022. A maioria dos projetos, 86,4%, utiliza machine learning.

Esses projetos de IA têm principalmente os seguintes usos: busca de casos similares; classificação de documentos; automação; indexação de documentos digitalizados; sugestão e movimentos para despacho; padronização de legislação e jurisprudência e; consulta à legislação. O mesmo projeto pode ter mais de uma utilidade.

A rapidez do desenvolvimento de sistemas inteligentes no Poder Judiciário brasileiro é tanta que, em janeiro de 2023, apenas dois meses após o lançamento mundial do ChatGPT, o Tribunal de Justiça de Minas Gerais (TJMG, 2023) anunciou uma nova ferramenta de IA baseada no ChatGPT, o Sistema Assistente Virtual de Inteligência Artificial (SAVIA), que está em fase de desenvolvimento pelo tribunal, já disponível para testes. A ferramenta que utiliza o modelo de IA do GPT-3.5, desenvolvido pela OpenAI, tem como objetivo elaborar textos e documentos de forma similar aos produzidos pelo ser humano para auxiliar magistrados, servidores e colaboradores da área administrativa do tribunal.

A nova ferramenta do Tribunal de Justiça de Minas Gerais será testada para o desenvolvimento de atividades administrativas para futuramente ser adotada pelo tribunal para redação de e-mails, portarias, resoluções, relatórios e textos em geral. Na divulgação feita pelo próprio tribunal (TJMG, 2023) não há menção direta quanto à utilização da ferramenta para elaboração de decisões interlocutórias ou sentenças. É preciso

acompanhar os testes e o começo da implementação do sistema para entender quais serão as limitações praticadas no seu uso.

Com base nos dados do Observatório de Políticas de Inteligência Artificial da Organização para Cooperação e Desenvolvimento Econômico (OECD.AI, 2021), o Judiciário brasileiro conta com mais ferramentas de Inteligência Artificial que o judiciário de cada um desses países: União Europeia, Estados Unidos, Reino Unido, China e Índia. Essa quantidade significativa de sistemas de IA se dá pela multiplicidade de plataformas sendo desenvolvidas por cada tribunal da federação interessado. Não há uma comunicação entre os entes nem um estudo prévio sobre projetos existentes ou em elaboração nos demais tribunais para desenvolvimento conjunto ou importação da experiência.

Esse obstáculo já foi constatado em 2020 no estudo "O Futuro da IA no Sistema Judiciário Brasileiro (Mapeamento, Integração e Governança da IA)", desenvolvido por um grupo de mestrandos da Universidade da Columbia em parceria com o ITS-Rio (Instituto de Tecnologia e Sociedade do Rio de Janeiro), a pedido do Conselho Nacional de Justiça (CNJ). Entre as prioridades elencadas para o CNJ no estudo, está a elaboração de uma política diretiva clara para o uso de IA no sistema judiciário, com um conjunto de princípios e comandos eficazes para assegurar o uso seguro e ético (BREHM, *et.al*, 2020, p. 8).

Entretanto, existem outros desafios em relação aos sistemas de IA atuais. Dentre eles, a necessidade de implementação de um mecanismo de monitoramento e avaliação que assegure que a IA seja utilizada eticamente dentro do Poder Judiciário (BREHM, *et.al*, 2020, p. 9) pela existência de questões como a discriminação algorítmica e a opacidade[4] do algorítmico aliada à falta de explicabilidade[5] das decisões.

Para fins deste trabalho iremos considerar a definição de discriminação algorítmica extraída do *The Blueprint for an AI Bill of Rights* publicado em 2022, pela Casa Branca dos Estados Unidos da América, segundo o documento (*THE WHITE HOUSE*, 2022):

> A "Discriminação algorítmica" ocorre quando sistemas automatizados contribuem para tratamento ou impactos injustificados desfavoráveis a pessoas com base em sua raça, cor, etnia, sexo (incluindo gravidez, parto e condições médicas relacionadas, identidade de gênero, identificação intersex e orientação sexual), religião, idade, origem nacional, deficiência, status de veterano, informações genéticas ou qualquer outra classificação protegida por lei.

[4] "Coloquialmente, diz-se que um material é "opaco" quando não ele permite a passagem da luz de forma adequada, tornando difícil a visualização do que há em seu interior. Transpondo esta ideia para a presente análise, um modelo de IA opaco seria aquele em que não se consegue visualizar, com clareza, como e porque ele toma determinada decisão." (NUNES; ANDRADE, 2023, p. 7)

[5] "No contexto da IA, explicabilidade significa compreender melhor os motivos e detalhes por trás de uma decisão algorítmica." (NUNES; ANDRADE, 2023, p. 10).

> Dependendo das circunstâncias específicas, essa discriminação algorítmica pode violar as proteções legais. Neste contexto, o termo "discriminação algorítmica" assume esse significado (e não uma compreensão técnica de discriminação como distinguir entre itens). (tradução nossa)

Assim sendo, de maneira bem ampla e recente, a conceituação trabalhada nesse plano abarca a discriminação algorítmica exercida em desfavor de quaisquer minorias, identificadas até o momento, destacando que, a depender da aplicação dos sistemas, a discriminação pode violar os direitos desses sujeitos.

Um dos desafios mais expressivos na atualidade é a existência de vieses[6] algorítmicos em sistemas inteligentes. O viés algorítmico ocorre com o tratamento desigual constatado em sistemas de IA e uma das formas de combate é a transparência das decisões para que se possa compreender o que levou a máquina a chegar àquela decisão discriminatória e sanar esse vício.

As principais causas apontadas para o viés algorítmico são a estrutura social pré-existente, os dados enviesados, a construção da ferramenta e/ou a interpretação dos resultados.

Essa situação se caracteriza, principalmente, segundo os pesquisadores Bianca Kremer, Pablo Nunes e Thallita Lima (2023, p. 14-6) pelo potencial de absorção de todos as discriminações e preconceitos enraizados socialmente por parte dos algoritmos que são desenvolvidos por seres humanos, e, portanto, não são eticamente neutros, pelo contrário. Assim, os autores demostram os perigos éticos envolvidos nesta aplicação tecnológica:

> O potencial que os algoritmos têm de melhorar o bem-estar individual e social não elimina os riscos éticos envolvidos na sua utilização. Modelos de inteligência artificial e algoritmos são construídos por humanos e, portanto, não são eticamente neutros. Os dados que treinam esses algoritmos são montados, limpos, rotulados e anotados também por humanos. Em outras palavras, são seres humanos, imersos em uma sociedade complexa, que escrevem a receita (Kremer; Nunes; Lima, 2023, p. 15).

De todo modo, seja qual for o motivo do resultado discriminatório, fato é que, por maior que seja a nossa fé no *Big Data*, sistemas inteligentes poderão ser discriminatórios. O viés de cognição por si só pode gerar um prejuízo, mas aliado à falta de transparência ou motivação daquela decisão, esse prejuízo é majorado.

---

[6] "Podemos conceituar vieses (bias) como pesos desproporcionais a favor ou contra algo ou alguém. Uma decisão enviesada, ou tendenciosa, ganha contornos de unilateralidade, isto é, é composta pela visão de mundo, pela experiência, pelos valores e até mesmo pela intuição de um sujeito ou de um grupo em relação ao contexto no qual está inserido. As pessoas podem desenvolver vieses a favor ou contra um indivíduo, um grupo étnico e racial, uma orientação sexual ou identidade de gênero, uma nação ou povo, uma religião, uma classe social, um partido ou posicionamento político, uma ideologia, entre muitos outros elementos. Nessa direção, os vieses algorítmicos (*algorithmic bias*) são o fenômeno a partir do qual as pessoas incorporam sua visão de mundo e, não raras vezes, preconceitos às tecnologias." (KREMER; NUNES; LIMA, 2023, p. 27)

Segundo a professora Dora Kaufman (2016, p. 9) ética significa um tipo de comportamento adquirido ou conquistado, mas não instintivo. Dessa forma, se a Inteligência Artificial representa uma nova inteligência que coexiste com a inteligência humana, devemos ter uma nova ética que envolva conceitos éticos tradicionais com a elaboração de novos.

Portanto, a inteligência artificial apresenta uma série de desafios éticos diante do seu uso que devem ser trabalhados através de normativas ligadas à concepção, desenvolvimento, implantação e utilização de modelos de IA, de modo a garantir um combate efetivo à discriminação algorítmica, a opacidade e estabelecer *FTA* (*fairness, transparency e accountability*), assunto que iremos abordar neste trabalho.

Adiante partiremos para a análise da evidência concreta de um caso de discriminação algorítmica, o *Compas* utilizado no Poder Judiciário Norte Americano, para uma compreensão do instituto da discriminação algorítmica e o risco da utilização de sistemas de inteligência artificial no Poder Judiciário sem diretrizes éticas aptas a garantirem o resguardo dos direitos fundamentais e garantias processuais.

2. O CASE *COMPAS* E A DISCRIMINAÇÃO ALGORÍTMICA

O tratamento desigual já é amplamente debatido em diversos aspectos sociais e situações fáticas. Nessa parte do presente trabalho debateremos o tratamento desigual constatado em algoritmos de inteligência artificial, também chamado de viés algorítmico[7]. O problema central da utilização de algoritmos que contenham vieses discriminatórios é a sua utilização em larga escala.

No âmbito do Poder Judiciário (Larson, *et. al*, 2016), foram desenvolvidos *softwares* assistivos aos magistrados e oficiais de liberdade condicional nos Estados Unidos de avaliação de risco dos indivíduos para utilização no sistema prisional. O *COMPAS* (*Correctional Offender Management Profiling for Alternative Sanctions*) desenvolvido pela *Northpointe Inc*[8]*,* é um desses sistemas e está entre os mais utilizados no país. Esses *softwares* foram desenvolvidos com o objetivo de realizar uma avaliação

[7] Acerca do instituto da discriminação algorítmica recomendo a leitura dos livros "Algoritmos da Opressão: como o Google fomenta e lucra com o racismo" da autora Safiya Umoja Noble (2021) e "Algoritmos de destruição em massa: Como o *big data* aumenta a desigualdade e ameaça à democracia" da autora Cathy O'neil (2020).

[8] Segundo a ProPublica (2016b) a NorthPointe foi fundada em 1989 por Tim Brennan, professor de estatística na Universidade de Colorado e, Dave Wells, diretor de um programa de correções em Traverse City, Michigan. Brennan e Wells compartilham a adoração pelo o que Brennan chamou de "taxonomia quantitativa", a medição de traços de personalidade como inteligência, extroversão e introversão. Eles queriam desenvolver uma ferramenta que abordasse as principais teorias sobre as causas do crime e a batizaram de *Correctional Offender Management Profiling for Alternative Sanctions (COMPAS).*

da probabilidade dos membros do sistema prisional norte americano de reincidirem, ou seja, de cometerem novos crimes no período de dois anos após a referida prisão.

Nos Estados Unidos (Angwin, *et. al*, 2016) pontuações de avaliação de risco são cada vez mais comuns em tribunais de todo o país, sendo utilizadas em todas as fases do sistema de justiça criminal desde a quantificação da fiança até as decisões mais fundamentais sobre a liberdade dos réus. No Arizona, Colorado, Delaware, Kentucky, Louisiana, Oklahoma, Virgínia, Washington e Wisconsin, as avaliações dos réus são consideradas durante a elaboração das sentenças criminais. Essa aplicação é uma das mais preocupantes na utilização de *softwares* avaliativos de risco conforme será demonstrado com casos de utilização do *COMPAS*.

Uma pesquisa realizada pela ONG ProPublica (Larson, *et. al*, 2016), entre os anos de 2013 e 2014, no condado de *Broward* no Estado da Flórida nos Estados Unidos, analisou mais de 11.575 (onze mil, quinhentos e setenta e cinco) pessoas, que foram avaliadas na fase pré-julgamento, e comparou as taxas de reincidência previstas com as taxas reais destes réus. Cada réu recebe pelo menos três avaliações do sistema que variam de 1 a 10, sendo dez o maior risco: o risco de reincidência, o risco de violência e o risco de não comparecer. Os *scores* de 1 a 4 são considerados baixos, de 5 a 7 médios e de 8 a 10 altos.

A pesquisa da ProPublica (Larson, *et. al*, 2016) analisa somente as taxas de reincidência dos crimes comuns e dos crimes violentos, como homicídio, homicídio culposo, estupro forçado, roubo e agressão agravada (FBI, 2010), sendo essa chamada de taxa de reincidência violenta. Quando um conjunto completo dos crimes foi levado em consideração pelo sistema (Angwin, *et. al*, 2016), até mesmo contravenções, a precisão era equivalente a de um lançamento de moeda (o famoso cara ou coroa). Dos considerados propensos a reincidir, 61% foram presos por qualquer outro crime dentro de dois anos.

Em relação a reincidência violenta a assertividade do sistema cai para 20%. A previsão correta para réus brancos e negros foi aproximadamente a mesma taxa (59% para réus brancos e 63% para réus negros), mas os erros em relação a eles foram bem destoantes.

O algoritmo (Larson, *et. al*, 2016) ainda tem maior probabilidade de classificar erroneamente um réu negro como de maior risco do que um réu branco. Para investigar a distribuição de falsos positivos e falsos negativos entre as etnias, a ONG utilizou uma tabela de contingência presente no artigo do Exército da Salvação (Whiteacre, Kevin, 2006, *apud* Larson, *et. al*, 2016) conjuntamente aos dados da pesquisa feita pela ONG. Considerando *scores* diferentes de baixo como de maior risco, os réus negros são

rotulados pelo *software* quase o dobro de vezes dos réus brancos com risco de reincidir erroneamente (45% contra 23%), enquanto o erro em relação aos acusados brancos é *in bonam partem*, ou seja, também são acusados quase o dobro de vezes erroneamente, mas como de baixo risco de reincidência (48% contra 28%).

Também foi testada pela ProPublica (Larson, *et. al*, 2016) com alteração da definição de alto risco de maneira a incluir apenas o *score* alto do *COMPAS*, não englobando os médios nessa análise, obtendo um resultado ainda pior. Nesse cenário os réus negros tinham três vezes mais chances de serem falsamente classificados como de alto risco do que réus brancos, 16% contra 5%. O cenário se repete em relação à reincidência violenta.

Ainda que se controlasse crimes anteriores, reincidência futura, idade e gênero, os réus negros ainda teriam 45% a mais de probabilidade de receber pontuações de risco mais altas do que os réus brancos. Já no caso de reincidência violenta, o número é ainda maior, mesmo controlando esses fatores, os réus negros eram 77% mais propensos a receberem pontuações de risco mais altas. Desa forma, a pesquisa concluiu que o sistema classifica acusados negros como pessoas com maior probabilidade de reincidência somente por serem negros.

O *COMPAS*, segundo a ProPublica (Angwin, *et. al*, 2016) não avalia apenas o risco, mas também outras vinte e quatro "necessidades criminógenas" que se relacionam com as principais teorias da criminalidade, incluindo "personalidade criminosa", "isolamento social", abuso de substâncias" e "residência ou estabilidade".

A análise do *software COMPAS* (Larson, *et. al*, 2016) é feita a partir de 137 perguntas que são respondidas pelos acusados ao chegarem à delegacia de polícia ou extraídas dos seus registros criminais, incluindo questões como nível de escolaridade e empregabilidade. Não há perguntas especificas quanto à etnia, mas é possível chegar a isso através das respostas a essas e outras perguntas como com quantas pessoas mora e em qual bairro, quantos amigos utilizam drogas ilegais, se já se envolveu em brigas na escola, entre outras tantas. Além disso, o acusado precisa concordar ou discordar de afirmações (Angwin, *et. al*, 2016) como, por exemplo, se uma pessoa com fome poderia roubar e se ao perder a paciência se considera perigoso.

As perguntas selecionadas pelo sistema a serem feitas aos indivíduos são determinantes para gerar as várias pontuações de cada acusado, como a avaliação de risco de reincidência e de reincidência violenta, resultando no valor da fiança na maioria dos casos. Portanto, a escolha não foi feita aleatoriamente, muito embora a empresa não

divulgue publicamente os cálculos utilizados ou os cruzamentos feitos pelo sistema para chegar à pontuação de risco dos indivíduos a partir do questionário.

Segundo a pesquisa realizada pelos sociólogos americanos Aldon Morris e Vilna Bashi Treitler (2019, p. 28) para a compreensão da raça e desigualdade racial nos Estados Unidos da América, as disparidades raciais em termos de renda, riqueza, bairros de residência, educação, pobreza, encarceramento, saúde e expectativa de vida são gritantes e a partir desses dados dispares, é possível concluir que a realidade de raças desiguais, enraizada na escravidão racial, ainda permeia a sociedade norte-americana. E em relação aos efeitos da segregação racial nos EUA, ressaltam os sociólogos:

> A pobreza concentrada em bairros negros e em áreas degradadas, as quais lembram colônias internas, continua a ser um problema perturbador e persistente. Em tais localidades, residentes negros e pessoas de cor vivenciam altos índices de desemprego, escolas inferiores, alta criminalidade, intensa violência e altos níveis de encarceramento. (*Ibidem*, p. 26)

A pesquisadora Sofiya Noble trabalha no seu livro "Algoritmos de Opressão" (Noble, 2021, p.17) o ponto de vista do poder dos algoritmos nesta era neoliberalista, e novas formas de reforçar relações sociais opressivas por meio da tecnologia, praticando perfilamento social, que a autora denomina como demarcação tecnológica, ou, em inglês "*technological redlining*". Importa destacarmos a explicação da autora para o termo "*redlining*":

> "*Redlining*", literalmente, "aplicação de linha vermelha", é um termo utilizado nos Estados Unidos para se referir à prática de demarcar áreas urbanas, de acordo com critérios raciais e de classe, para limitar o acesso de moradores das regiões demarcadas a empréstimos financeiros e imobiliários. (Noble, 2021, p. 17)

Portanto, não existir uma pergunta específica quanto à etnia no *COMPAS* não tem grande significado. Afinal, como um sistema inteligente faz uma análise de grandes volumes de dados e identifica os padrões, nesse caso discriminatório em relação à população negra, o que pode acentuar essas disparidades raciais já existentes na sociedade e no sistema de justiça criminal.

Durante a pesquisa, a ProPublica (Larson, *et. al*, 2016) criou um modelo de regressão logística que considerou raça, idade, antecedentes criminais, reincidência futura, grau de acusação, sexo e idade para testar as disparidades raciais controlando esses fatores. O fator mais determinante para um *score* de risco alto foi a idade, réus com menos de 25 anos tinham 2,5 vezes mais chances de obter uma pontuação mais alta do que os infratores de meia-idade, mesmo controlando crimes anteriores, criminalidade futura, raça e gênero. Controlando os mesmos fatores, as acusadas do gênero feminino também tinham 19,4% mais chances de obter uma pontuação mais alta do que os de gênero masculino, muito embora os níveis de criminalidade feminino sejam mais baixos no geral.

Essa questão pode ser observada em um dos cases trabalhados na pesquisa (Angwin, *et. al*, 2016) para comparação de *scores* de dois casos semelhantes. O primeiro caso é da jovem Brisha Borden, de 18 anos, que estava atrasada para buscar sua afilhada na escola quando observou uma bicicleta e uma scooter soltas. Ela e a amiga pegaram os veículos para chegarem mais rápido à escola, mas logo notaram que eram grandes demais para os pequenos transportes que pertenciam a um menino de seis anos. Ao perceber a situação, uma mulher corre atrás dela dizendo que os veículos pertenciam ao seu filho. Elas largam e vão embora. Entretanto, já era tarde demais. Um vizinho que presenciou o assalto já tinha chamado a polícia, Borden e sua amiga foram presas e acusadas de roubo e furto qualificado dos itens, que foram avaliados em um total de US$80 (oitenta dólares).

O segundo caso é de um homem de 41 anos de idade, Vernon Prater, que foi pego por furtar US$86,35 (oitenta e seis dólares e trinta e cinco centavos) em ferramentas em uma loja vizinha da Home Depot. Prater já havia cometido outros delitos, já tinha sido condenado por assalto à mão armada e tentativa de assalto à mão armada, cumprindo cinco anos de prisão, além de outra acusação também por assalto à mão armada. Borden também tinha registro de delitos anteriores, mas cometidos quando era menor de idade.

Nesses casos semelhantes do condado de Broward na Flórida, a avalição de risco destes indivíduos foi oposta, Borden, que é mulher negra, foi classificada como de alto risco (pontuação 8) e Prater, que é um homem branco, de baixo risco (pontuação 3). Após se passar dois anos desta avaliação, Borden não foi acusada por nenhum novo crime e Prater cumpria pena de oito anos de prisão por invadir um depósito e roubar milhares de dólares em eletrônicos, ou seja, contrariando a previsão feit\ pelo *COMPAS*.

O condado de Broward luta contra a superlotação do sistema prisional e tem utilizado o *software* em audiências de pré-julgamento para ajudar a identificar quais réus eram de baixo risco e poderiam ser libertados para responderem em liberdade sob pagamento de fiança. No caso de Borden, ao ser considerada de alto risco, a sua fiança foi valorada pelo juiz para a quantia de mil dólares.

Segundo a ProPublica (Angwin, *et. al*, 2016), em 2014, Eric Holder, Procurador-Geral dos EUA, já demonstrava preocupação quanto à utilização de sistemas de avaliação de riscos nos tribunais, pela possibilidade de estarem reforçando preconceitos existentes na sociedade e impedindo uma justiça individualizada e igualitária. Na época, o Procurador solicitou que a Comissão de Sentenças dos EUA fizesse um estudo a respeito dos *scores* de risco, como o feito pela ProPublica no ano seguinte, o que não aconteceu. Na realidade, em contrariedade ao Procurador-Geral, o Instituto Nacional de Correções

do Departamento de Justiça dos Estados Unidos incentiva o uso de sistemas de avaliação de risco em todas as fases do processo criminal.

Em 2015, tramitou no Congresso norte-americano (*Congress*, 2015) um projeto de lei de Reforma e Correções de Sentenças que tinha como previsão o uso obrigatório desses sistemas em presídios federais. Nos debates estabelecidos na Comissão do Poder Judiciário foram acrescentadas questões importantes quanto à necessária certificação desses sistemas a ser feita pelo Instituto Nacional de Correções que comprove que as avaliações são baseadas em evidencias eficazes na redução ou mitigação do risco de reincidência dos infratores, sendo considerada para manutenção da classificação do indivíduo e a progressão de regime.

Cumpre observar ainda que os Estados Unidos é o país de maior população carcerária[9] do mundo contabilizando um total de 2,2 milhões de pessoas presas (*Bureau of Justice Statistics*, 2019) em 2019. No mesmo ano, os afro-americanos eram cerca de 13% da população dos EUA (BBC, 2021), mas representavam quase um terço da população carcerária do país, enquanto os brancos, por mais que representassem 60% da população total do país, correspondiam apenas a 30% da população carcerária. O que significa mil presidiários negros para cada 100 mil residentes negros no EUA, em comparação a 200 presidiários brancos para cada 100 mil americanos brancos. Uma diferença gritante.

Segundo a Pesquisa "*The Color of Justice: Racial and Ethnic Disparity in State Prisons*" do *The Sentencing Project* (2021) os americanos negros são encarcerados em prisões estaduais quase cinco vezes a mais que os brancos. Em nível nacional, um em cada 81 adultos negros está cumprindo pena em prisão estadual nos EUA, em nível estadual *Wisconsin* lidera ranking em encarceramento de pessoas negras, um em cada 36 negros de *Wisconsin* está preso.

De acordo com a pesquisa da ProPublica (Angwin, *et. al*, 2016), Wisconsin utiliza majoritariamente o *COMPAS* em cada etapa do sistema prisional, da sentença à liberdade condicional. O estado tem sido um dos principais na utilização do *score* de risco desenvolvido pela *Northpointe* em decisões de condenação, quando um acusado é condenado em qualquer lugar do estado, o Departamento de Correções anexa a avaliação do *software* ao relatório de pré sentença que é fornecido aos juízes.

Todos os dados corroboram a conclusão que a ProPublica (Angwin, *et. al*, 2016) traz na sua pesquisa de que o apelo dos *softwares* de *scores* de risco é nítido diante do

[9] Segundo a BBC, a definição da população carcerária dos EUA considera apenas presidiários condenados a mais de um ano em uma prisão federal ou estatual. (BBC, 2021)

fetichismo tecnológico e a fé cega no *Big Data.* Acredita-se que os computadores seriam mais assertivos do que os humanos. Por mais de dois séculos todas as decisões processuais nos EUA foram tomadas apenas por seres humanos guiados com base em seus instintos e vieses pessoais, idealmente os computadores poderiam trazer maior justiça e acertabilidade a esse processo do que já visto até então.

O objetivo inicial (Angwin, *et. al*, 2016) da maioria das ferramentas de avaliação de risco foi fornecer aos juízes informações sobre o tipo de tratamento que um indivíduo poderia precisar, desde tratamento de drogas até aconselhamento para saúde mental, ou seja, uma consideração do custo estatal para manutenção do indivíduo em liberdade sem reincidir. Muito embora, especialmente durante as audiências de sentença, ser considerado inelegível para tratamentos alternativos corresponda ao encarceramento.

Conforme levantado pelo estudo da ProPublica (Angwin, *et. al*, 2016), o Estado da Virgínia foi um dos pioneiros no uso do algoritmo na condenação de infratores não violentos em 2002 e utilizou o sistema alinhado a sua projeção original. Na época, tinha uma taxa de crescimento da população carcerária de 31% na década anterior e conseguiu diminuir para 5%, provando que é possível utilizar a ferramenta de avaliação de risco de modo favorável à população.

Em 2014 (Angwin, *et. al*, 2016), os juízes da Virgínia que usaram a ferramenta enviaram quase metade dos réus para penas alternativas à prisão. Em outras jurisdições, como o condado de Napa, na Califórnia, essas ferramentas de avaliação de risco são utilizadas para sugerir ao juiz um plano de liberdade condicional ou o tratamento mais apropriado àquele réu. Esses usos que acarretam na redução da taxa de encarceramento é o argumento central de quem defende a utilização dos *scores* de risco.

Entretanto, a ferramenta foi sendo utilizada de maneira discriminatória e vinculante por grande parte dos juízes quando baseiam suas decisões de tempo de reclusão de um acusado apenas no resultado do *software*. Ocorre que os réus raramente têm o contraditório oportunizado contra os resultados obtidos pelos sistemas de risco e ainda quando conseguem a possibilidade de contestar, não há uma ampla defesa, já que não há conhecimento acerca dos cálculos e correlações feitas pelo sistema a partir dos dados do acusado. O viés de cognição por si só pode gerar um prejuízo, mas aliado à falta de transparência ou motivação daquela decisão, este prejuízo é majorado.

Esse foi o caso do Eric Loomis, analisado pela ProPublica (Angwin, *et. al*, 2016), que foi condenado a seis anos de prisão em 2013 após furtar um veículo, evadir-se de um agente de trânsito e se envolver em um tiroteio. O réu já havia cometido o crime de

agressão sexual anteriormente, por isso, o *quantum* da pena de oito anos e seis meses de prisão não foi uma surpresa, mas sim a fundamentação da decisão.

A decisão foi fundamentada tão somente no *score* de Loomis como de alto risco dado pelo *COMPAS*, resultando no caso State v. Loomis pela violação dos seus direitos ao devido processo legal[10]. Embora tenha recorrido à Suprema Corte de Winsconsin requerendo que fosse relevado os critérios utilizados pelo *software* para chegar ao resultado do *score,* seu pleito foi negado afirmando que a pena seria a mesma se feita apenas por análise humana diante do seu crime e antecedentes. Não satisfeito com a decisão e de maneira a demonstrar que o seu direito à ampla defesa e contraditório havia sido violado recorreu à Suprema Corte Americana que negou o seu *writ de certiorari,* ou seja, teve seu caso inadmitido pela Corte.

O resultado do *software* também foi utilizado para fundamentar a quantificação da pena no caso do Zilly, em 2012. Conforme narra o caso, Zilly é um trabalhador da construção civil de 48 anos que realizou furtos em uma recaída para vender e comprar drogas pelo vício em metanfetamina. Zilly furtou um cortador de grama e algumas ferramentas e foi considerado como de alto risco de reincidência violenta.

O defensor público que atuou no caso de Zilly intimou Tim Brennan, um dos criadores do *COMPAS* como testemunha. Em seu testemunho, Brennan diz que o *software* não foi projetado para ser usado em sentenças e que ele mesmo não aprova a ideia de que o *COMPAS* seja a única evidência base de uma decisão. Nesse caso, após o depoimento de Brennan, o juiz reduziu a pena de Zilly de dois anos de prisão para 18 meses, pois acreditou que esse seria o quantitativo que ele teria dado se não tivesse se apoiado no *score* de risco.

O resultado dessa pesquisa realizada pela ProPublica (Larson, *et. al*, 2016) é a demonstração de violação do devido processo legal previsto na Quinta Emenda, no âmbito dos governos federais, e na Décima Quarta Emenda, no âmbito dos governos estaduais, ambas da Constituição dos Estados Unidos.

Com base na nossa Constituição Federal (1988), o devido processo legal está previsto no artigo 5º, LIV, "ninguém será privado da liberdade ou de seus bens sem o devido processo legal". Sendo assim, qualquer restrição à liberdade só poderá ocorrer mediante processo e julgamento previamente definidos em lei.

[10] Isso ocorre porque as técnicas de aprendizado automático de máquina não realizam um trabalho de interpretação. Na realidade, elas fazem um processamento automático onde a correlação de grandes quantidades de informação deve substituir a compreensão das reais causalidades de uma decisão. Não há uma tentativa de formalizar o raciocínio legal, apenas de antecipar as prováveis decisões de um juiz em situações semelhantes através de modelos matemáticos (CEPEJ, 2018).

Há um grande risco ao jurisdicionado quando esse resultado vinculante é tendencioso e causa um dano irreparável. Com isso, o inciso seguinte do artigo 5º, LV, traz o direito à ampla defesa e ao contraditório "aos litigantes, em processo judicial ou administrativo, e aos acusados em geral são assegurados o contraditório e a ampla defesa, com os meios e recursos a ela inerentes".

Os casos de utilização do *software* como única forma de fundamentação do *quantum* da sentença condenatória trata de afronta ao devido processo legal quando há uma falta de fundamentação da decisão contrariando essa necessidade estabelecida no artigo 93, IX da Constituição Federal. Isto porque o réu não poderá se defender de métricas e correlações feitas pelo sistema que são desconhecidas tanto pelo juiz como pelo acusado que tem seu direito ao contraditório e a ampla defesa negado neste contexto.

A fundamentação nos termos de um sistema operacional caracteriza violação à garantia constitucional de igualdade estabelecida no *caput* do artigo 5º da Constituição Federal e se mostra em desacordo ao querer do constituinte originário como um todo. No preâmbulo da Constituição Federal de 1988, o constituinte originário traz como centralidade da democracia brasileira o exercício dos direitos sociais e individuais, a igualdade e a justiça como valores supremos de uma sociedade fraterna, pluralista e sem preconceitos. Para além das garantias previstas no artigo 5º, viola a interpretação do próprio Estado Democrático de Direito brasileiro tendo em vista a redação do preâmbulo da Constituição Federal.

Conforme traz a professora Cathy O'Neil (2020, p. 13) os sistemas matemáticos demandam um *feedback*, ou seja, saber quando o resultado apresentado não foi o esperado, quando "saíram dos trilhos". A partir dos erros os estatísticos treinam seus modelos e os fazem cada vez mais inteligentes com maior nível de acurácia. Ao contrário, quando um sistema inteligente não é treinado para que entenda os erros como decisões ruins e escolha novos caminhos, pode ocorrer um ciclo tóxico de retroalimentação que o faz acreditar que estava certo e deve continuar a tomar essas decisões precipitadas como corretas.

O *case* do COMPAS é o mais notório e debatido em relação à discriminação algorítmica racial e de gênero no âmbito do Poder Judiciário, mas não foi o único caso de sistemas de inteligência artificial discriminatórios. Nos últimos anos foram veiculados diversos casos de sistemas de IA discriminatórios. Seja no âmbito da justiça, na saúde, na seleção, contratação e demissão de pessoas ou na concessão de crédito, o risco no uso de algoritmos para tomada de decisões públicas e privadas deve ser levado em conta.

CONCLUSÃO

Diante de todos os desafios existentes no movimento de digitalização dos processos e tribunais frente à realidade brasileira, a Inteligência Artificial surge no Poder Judiciário como uma forte aliada contribuindo, não apenas, com o menor tempo de duração dos processos, mas também com uma maior qualidade da prestação jurisdicional ofertada. Entretanto, é preciso refletir quais os parâmetros utilizados no seu desenvolvimento e aplicação para que seja assegurada a ética e a defesa dos direitos humanos.

Importa-se destacar que não há neutralidade ética na tecnologia, tendo em vista que em todos esses momentos há participação humana e, portanto, há potencial de absorção das discriminações e preconceitos enraizados socialmente por parte dos algoritmos. Questões discriminatórias, como o caso *Compas*, trazem riscos aos direitos fundamentais e a preocupação de como se dará a implementação destes sistemas no cotidiano do Poder Judiciário brasileiro sem uma regulação nacional.

Para enfrentar os desafios éticos, observamos ao longo do trabalho, precisa ser elaborada uma verdadeira política pública enquanto plano de governo para o desenvolvimento responsável da inteligência artificial no Brasil com métricas claras e delimitadas

Como traz o autor Hoffmann-Riem (2022, p. 6-8), todos os usos possíveis das tecnologias digitais e futuros desenvolvimentos estão sujeitos à afetação diante da necessidade de modulação por meio de lei. Para alcançar isso, o modo jurídico e sociopolítico de enxergar essas tecnologias deve ser ampliado para considerar todas as oportunidades e riscos da digitalização do Estado e da sociedade. Pensando, criteriosamente, em uma configuração legislativa que não impeça a inovação.

Pesquisar e regular a tecnologia é um dos grandes desafios do século XXI, tendo em vista a rapidez das mudanças tecnológicas. No espaço de tempo entre debater a regulamentação de maneira ampla e democrática e aprovar algum texto ou política pública a ser implementado, novas questões surgem que demandam análises detalhadas. Mas esse desafio deverá ser tão logo superado pelo Brasil com a aprovação de um marco legal para inteligência artificial, como aconteceu na União Europeia com a aprovação do texto final do *AI ACT* no dia 13 de março de 2024.

REFERÊNCIAS BIBLIOGRÁFICAS

ANGWIN, Julia; LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren. *Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica, 2016. Disponível em: <https://www.propublica.org/article/machine-bias-risk-* assessments-in-criminal-sentencing>. Acesso em 26 de ago. de 2023.

BREHM, Katie; HIRABAYASHI, Momori; LANGEVIN, Clara; MUÑOZCANO, Bernardo Riveira; SEKIZAWA, Katsumi; ZHU, Jiayi. (2020) O futuro da IA no Sistema Judiciário brasileiro: mapeamento, integração e governança da IA. Traduzido por: Matheus Drummond e Matheus de Souza Dipieri. ITS, 2020. Disponível em: <https://itsrio.org/wp- content/uploads/2020/07/TRADUÇÃO-The-Future-of-AI-in-the-Brazilian-Judicial- System.pdf>. Acesso em 10 de jul. de 2023.

COE. The CEPEJ European Ethical Charter on the use of artificial intelligence (AI) in judicial systems and their environment. ***CEPEJ****,* 2018. Disponível em: <https://rm.coe.int/carta-etica-traduzida-para-portugues-revista/168093b7e0#_ftn10>. Acesso em: 25 de out. de 2023.

CONSELHO NACIONAL DE JUSTIÇA. Estatísticas do Poder Judiciário. DataJUD: base nacional de dados do Poder Judiciário. CNJ, 2023. Disponível em: <https://painel-estatistica.stg.cloud.cnj.jus.br/estatisticas.html>. Acesso em 13 de jul. de 2023.

CONSELHO NACIONAL DE JUSTIÇA. Estatísticas do Poder Judiciário. DataJUD: base nacional de dados do Poder Judiciário. CNJ, 2024a. Disponível em: <https://painel-estatistica.stg.cloud.cnj.jus.br/estatisticas.html>. Acesso em 13 de maio de 2023.

CONSELHO NACIONAL DE JUSTIÇA. Resultados pesquisa IA no Poder Judiciário – 2024. **CNJ**, 2024b. Disponível em: <https://www.cnj.jus.br/programa-justica-4-0-divulga-resultados-de-pesquisa-sobre-ia-no-judiciario-brasileiro/> Acesso em 28 de maio de 2024.

HAN, Byung-Chul. **Infocracia: Digitalização e a crise da democracia.** Tradução de Gabriel S. Philipson. Petrópolis, RJ: Vozes, 2022.

HOFFMANN-RIEM, Wolfgang. **Teoria Geral do Direito Digital: transformação digital: desafios para o direito.** 2ª ed. Rio de Janeiro: Forense, 2022.

LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren; ANGWIN, Julia. *How we analyzed the COMPAS Recidivism Algorithm.* ProPublica, 2016. Disponível em: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>. Acesso em 10 de jun. de 2023.

KAUFMAN, Dora. **Inteligência artificial: questões éticas a serem enfrentadas.** IX Simpósio Nacional ABCiber, PUC São Paulo, 2016. Disponível em: <https://abciber.org.br/anaiseletronicos/wpcontent/uploads/2016/trabalhos/inteligencia_ artificial_questoes_eticas_a_serem_enfrentadas_ dora_kaufman.pdf.>. Acesso em 25 de nov de 2022.

KREMER, Bianca; NUNES, Pablo; LIMA, Thallita G. **Racismo Algorítmico.** Rio de Janeiro: CESeC, 2023. Disponível em: <https://opanoptico.com.br/wp-content/uploads/2023/11/PANOPT_panorama_algoritmoracista_211123.pdf>. Acesso em 23 de nov. de 2023.

MORRIS, Aldon; TREITLER, Bashi Vilna. O Estado racial da União: compreendendo raça e desigualdade racial nos Estados Unidos da América. Caderno CRH, Salvador, Centro de Recursos Humanos da UFBA, v. 32, n. 85, p. 15-31, jan./abr. Scielo, 2019. Disponível em: <https://www.scielo.br/journal/ccrh/about/#about>. Acesso em 27 de ago. de 2023.

NOBLE, Safiya Umoja. (2021). Algoritmos da Opressão: como o Google fomenta e lucra com o racismo. Trad. Felipe Damorim. Santo André: Editora Rua do Sabão.

OCDE. *Recommendation of the Council on Artificial Intelligence.* OCDE, 2019. Disponível em: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. Acesso em 16 de jun. de 2023.

OCDE.IA. *Database of national AI policies.* OCDE, 2021. Disponível em:<https://oecd.ai/en/>. Acesso em 17 de jun. de 2023.

Pause Giant AI Experiments: Na Open Letter. We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4. Future of Life, 2023. Disponível em: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>. Acesso em 20 de nov. de 2023.

PARLAMENTO EUROPEU. Lei da UE sobre IA: primeira regulamentação de inteligência artificial. **Parlamento Europeu**, 2023b. Disponível em: <https://www.europarl.europa.eu/news/pt/headlines/society/20230601STO93804/lei-da-ue-sobre-ia-primeira-regulamentacao-de-inteligencia-artificial>. Acesso em 27 de out. de 2023.

SIMONITE, Tom. *A Health Care Algorithm Offered Less Care to Black Patients: A study shows the risks of making decisions using data that reflects inequities in American society. Wired,* 2019. Disponível em: <https://www.wired.com/story/how-algorithm-favored-whites-over- blacks-health-care/>. Acesso em 10 de nov. de 2023.

TRIBUNAL DE CONTAS DA UNIÃO, Secom. TCU avalia uso de inteligência artificial pelo governo federal: levantamento do Tribunal avaliou o estágio atual e perspectivas de utilização de inteligência artificial (IA) nas organizações federais. Mais de um terço delas estão no nível zero de maturidade em IA. **TCU**, 2022. Disponível em: < https://portal.tcu.gov.br/imprensa/noticias/tcu-avalia-uso-de-inteligencia-artificial-pelo-governo-federal.htm>. Acesso em 12 de jul. de 2023.

TRIBUNAL DE JUSTIÇA DE MINAS GERAIS. TJMG apresenta SAVIA, nova ferramenta de inteligência artificial baseada no chatGPT. TJMG, 2023. Disponível em: <https://www.tjmg.jus.br/portal-tjmg/noticias/tjmg-apresenta-savia-nova-ferramenta-de-inteligencia-artificial.htm>. Acesso em 20 de out. de 2023.

UNITED STATES. Blueprint for an AI Bill of Rights: making automated systems work for the American people. ***The White House****,* 2022. Disponível em: <https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf>. Acesso em 21de set. de 2023.

O'NEIL, Cathy. **Algoritmos de destruição em massa: Como o *big data* aumenta a desigualdade e ameaça a democracia;** tradução Rafael Abraham. Santo André, SP: Editora Rua do Sabão, 2020. Título original: weapons of math destruction: how big data increases inequality and threatens democracy.