DOI: 10.5212/Emancipacao.v.25.2524134.025

Inteligência Artificial e Direitos Fundamentais no Brasil: Uma Análise sobre Vieses Algorítmicos, Transparência Decisória e o Projeto de Lei 2338/2023

Artificial Intelligence and Fundamental Rights in Brazil: An Analysis of Algorithmic Biases, Decision-Making Transparency and Bill 2338/2023

Barbara Guasque[*]

Resumo: O presente artigo tem por objetivo geral analisar criticamente os impactos da implementação de sistemas de inteligência artificial na esfera dos direitos fundamentais no Brasil, com foco nos vieses algorítmicos, na transparência dos processos decisórios e na adequação do Projeto de Lei brasileiro nº 2338/2023. Foram elencados como objetivos específicos os seguintes: a)mapear alguns modelos de IA utilizados no Brasil e no mundo, e que têm potencial de interferir em direitos fundamentais, como a privacidade, igualdade, e o devido processo legal; b) examinar os casos de vieses algorítmicos identificados em sistemas de IA, ilustrando suas consequências sobre grupos vulneráveis; c) avaliar o projeto de regulamentação brasileiro – o Projeto de Lei 2338/2023, e sua possibilidade de garantir um ambiente seguro e ético para o uso dessas tecnologias no Brasil. O presente estudo se justifica diante do crescimento acelerado do uso de IA em diversas esferas da vida social, econômica e jurídica no Brasil. A ausência de regulamentação robusta e o risco de perpetuação de discriminações, através de sistemas algorítmicos, demandam uma análise aprofundada e crítica para garantir que os direitos fundamentais sejam protegidos. A metodologia utilizada tem natureza exploratória e descritiva, abordando casos práticos e, também, como procedimento técnico, utilizou-se de pesquisa bibliográfica.
Palavras-chave: inteligência artificial; vieses algorítmicos; direitos fundamentais; regulamentação; Projeto de Lei 2338/2023

Abstract: The general objective of this article is to critically analyze the impacts of the implementation of artificial intelligence systems in the sphere of fundamental rights in Brazil, focusing on algorithmic biases, the transparency

[*] Pós-Doutora pela Universidade do Vale do Itajaí –UNIVALI. Doutora em Ciência Jurídica pela UNIVALI e Doutora em Direito pela Universidade de Alicante (Espanha). Mestre em Direito Econômico pela Pontifícia Universidade Católica do Paraná - PUC/PR. Professora na Pós-Graduação em Processo Civil da Univali. E-mail: barbaragiasque@hotmail.com

of decision-making processes and the adequacy of Brazilian Bill nº. 2338/2023. The following specific objectives were listed: a) mapping some AI models used in Brazil and around the world, and which have the potential to interfere with fundamental rights, such as privacy, equality, and due legal process; b) examine cases of algorithmic biases identified in AI systems, illustrating their consequences on vulnerable groups; c) evaluate the Brazilian regulatory - Bill 2338/2023, and its possibility of guaranteeing a safe and ethical environment for the use of these technologies in Brazil. The present study is justified given the accelerated growth in the use of AI in various spheres of social, economic and legal life in Brazil. The lack of robust regulation and the risk of perpetuating discrimination through algorithmic systems require an in-depth and critical analysis to ensure that fundamental rights are protected. The methodology used is exploratory and descriptive in nature, addressing practical cases and, also, as a technical procedure, bibliographic research was used.
Keywords: artificial intelligence; algorithmic biases; fundamental rights; regulation; Bill 2338/2023

Recebido em 27/10/2024. Aceito em 22/05/2025.

INTRODUÇÃO

O desenvolvimento da inteligência artificial (IA) tem promovido transformações profundas em diversas áreas da sociedade, influenciando desde aspectos cotidianos, como a recomendação de conteúdos em plataformas digitais, até esferas mais complexas, como a saúde, a segurança pública e a tutela jurisdicional. No entanto, junto às inegáveis vantagens trazidas por essas tecnologias, surgem desafios significativos, especialmente no que diz respeito a proteção de direitos fundamentais.

No Brasil, a incorporação acelerada de sistemas de IA em diferentes setores expõe a sociedade a riscos que ainda não estão totalmente compreendidos ou regulados. A falta de transparência nos processos decisórios algorítmicos e a possibilidade de perpetuação de preconceitos preexistentes, através dos chamados vieses algorítmicos, têm levantado preocupações éticas e jurídicas em escala global. Esses sistemas, ao se basearem em dados históricos que muitas vezes carregam desigualdades sociais, podem replicar e ampliar discriminações, afetando desproporcionalmente grupos vulneráveis.

Neste contexto, o presente artigo propõe uma análise sobre os impactos da implementação de sistemas de IA na esfera dos direitos fundamentais no Brasil. A discussão será norteada pela identificação de alguns modelos de IA já em uso no país e no mundo, especialmente aqueles que possuem o potencial de interferir diretamente em direitos como a privacidade, a igualdade e o devido processo legal. Além disso, o artigo se debruça sobre o Projeto de Lei nº 2338/2023, que busca estabelecer normas gerais para o desenvolvimento, implementação e uso responsável de sistemas de IA no Brasil.

A necessidade de um estudo detalhado sobre o tema se justifica tendo em vista a crescente utilização dessas tecnologias, em um cenário de regulação ainda insuficiente, que pode resultar em violações graves a direitos fundamentais. A compreensão dos riscos inerentes à IA e a análise da regulamentação são essenciais para garantir que os avanços tecnológicos ocorram de forma a respeitar e proteger os direitos fundamentais, preservando os valores democráticos que regem a sociedade.

Para tanto, o artigo se inicia trazendo breves explanações sobre IA e algumas aplicabilidades em diferentes setores, como saúde, segurança pública, mercado financeiro e de seguros e no Poder Judiciário. Em seguida o artigo irá tratar sobre os direitos fundamentais e os pontos de intersecção entre IA e ditos direitos. Na sequência, a análise se aprofundará nos vieses algorítmicos e seus impactos específicos sobre os direitos fundamentais, preparando o terreno para uma análise sobre o projeto de regulamentação brasileiro - Projeto de Lei 2338/2023 e os desafios futuros.

Do ponto de vista metodológico, este estudo adota uma abordagem qualitativa, de natureza exploratória e descritiva. Utiliza-se da técnica de pesquisa bibliográfica, baseada em doutrina especializada, artigos científicos, relatórios institucionais, legislação e projetos de lei. Complementarmente, realiza-se análise crítica de casos emblemáticos de aplicação de sistemas de inteligência artificial que evidenciam vieses algorítmicos e seus impactos sobre direitos fundamentais, com o objetivo de ilustrar os riscos concretos e contextualizar a necessidade de uma regulação adequada e eficaz. A escolha metodológica visa fornecer uma base teórica sólida para refletir sobre os desafios ético-jurídicos da IA e avaliar o Projeto de Lei nº 2338/2023 à luz da proteção constitucional dos direitos fundamentais no Brasil.

INTELIGÊNCIA ARTIFICIAL: BREVES EXPLANAÇÕES

Consoante Navas, a IA é um campo da ciência e engenharia, encarregado de compreender o comportamento inteligente do cérebro humano e, além disso, criar artefatos que simulem dito comportamento de maneira automatizada. Para a autora, a IA se ocupa de "emular as diversas capacidades do cérebro humano para apresentar comportamentos inteligentes, sintetizando e automatizando tarefas intelectuais" (Navas, 2017, p. 24).

Isto quer dizer que as máquinas, em alguma medida, imitam o processo cognitivo humano, após um processo de aprendizado baseado em dados que fornecem generalizações sobre dado assunto. Porém a IA alcançará sempre resultados superiores aos que poderia conseguir qualquer ser humano. Isto porque, o sistema não somente aceita, mas acessa uma quantidade de dados infinitamente maior que o cérebro humano. E, quanto mais dados, maior a possibilidade de relacioná-los, visualizar padrões ocultos e *insights* e, portanto, obter melhores resultados (Nieva Fenoll, 2018, p. 15).

A palavra-chave na IA é algoritmo (Nieva Fenoll, 2018, p. 15). Uma vez que um sistema de IA impende uma sequência de instruções que especifique as diferentes ações a serem executadas pelo computador a fim de resolver um determinado problema. Este esquema executivo contemplando as instruções, o caminho a ser percorrido, é desempenhado pela estrutura algorítmica (Navas, 2017, p. 24).

Um algoritmo é um esquema executivo, uma sequência de ações para resolver um problema ou responder uma questão. Assim, uma receita culinária é um algoritmo para humanos. Uma sequência de instruções para resolver o problema de como fazer determinada receita (Rodríguez, 2018, p. 109).

Todavia, computadores e máquinas precisam de algoritmos mais complexos, que os auxiliem a efetuar as tarefas definidas e que sempre produzam o mesmo resultado, com base nos mesmos parâmetros. Neste sentido se encontra um subconjunto específico de algoritmos usados para aprendizado de máquina (Rodríguez, 2018, pp. 109-111).

A particularidade destes algoritmos é que eles aprendem por conta própria, fazendo inferências a partir dos dados. Portanto, o aprendizado de máquina é a capacidade destas de aprender com os dados, identificando tendências e padrões em eventos aparentemente aleatórios. Acaso estes padrões se baseiem em séries de dados longas o suficiente e dotados de suficiente qualidade, podem ser usados para tentar prever o futuro "com uma base científica confiável" (Rodríguez, 2018, p. 133). Desta maneira, "o aprendizado de máquina olha para o passado para encontrar padrões e tentar prever o futuro". "Em essência, o aprendizado de máquina é pura predição", tendo por base, os dados e as experiências do passado (Rodríguez, 2018, p.105). Cabe ressaltar que a expressão "prever o futuro" diz respeito, na verdade, a estimativas probabilísticas: a inteligência artificial realiza apenas cálculos matemáticos e estatísticos sobre o que poderá ocorrer, com base em padrões extraídos de dados passados.

Esses algoritmos vêm sendo utilizados em diversas tarefas dentro do Poder Judiciário. Por exemplo, os sistemas Victoria do TJRJ e Elis do TJPE, que automatizaram o rito das execuções fiscais. Também o sistema Mandamus, que utiliza de técnicas de inteligência artificial para auxiliar na automação do processo de elaboração, distribuição e gerenciamento do cumprimento de mandados judiciais. Ainda, o mais conhecido deles, o Victor, utilizado no Supremo Tribunal Federal com o objetivo de otimizar a análise da Repercussão Geral (Da Rosa & Guasque, 2021, pp. 93-121).

Ainda que constituam ferramentas extremamente promissoras e profícuas, a IA possui limitações e, em muitos casos, a sua utilização impacta de maneira drástica direitos fundamentais. Torna-se, portanto, imprescindível a identificação dos impactos negativos que modelos enviesados podem provocar na sociedade e nos direitos fundamentais.

A preocupação se justifica porque os algoritmos tendem a refletir preconceitos já inerentes às sociedades, mas podem fazê-lo em uma escala potencialmente massiva e sem a devida supervisão, afrontando direitos fundamentais como a igualdade e a não discriminação, a liberdade, a autonomia e a privacidade, mormente em um cenário de ausência de regulamentação.

APLICAÇÕES DA INTELIGÊNCIA ARTIFICIAL NA SOCIEDADE

O desenvolvimento tecnológico sem precedentes experimentado nos últimos anos trouxe e incorporou ferramentas de inteligência artificial a todas as esferas da vida em sociedade. No cotidiano, os algoritmos indicam a melhor rota a ser seguida, recomendam músicas e filmes, sugerem compras direcionadas e conectam pessoas em redes sociais.

Na área da saúde, os algoritmos têm revolucionado a detecção precoce de doenças e o tratamento de pacientes (CB INSIGHTS, 2016). Ferramentas de IA são utilizadas para analisar grandes volumes de dados médicos, como imagens radiológicas, identificando padrões que possam indicar a presença de condições como o câncer — muitas vezes com precisão superior à dos métodos tradicionais. Além disso, esses algoritmos são empregados para prever os resultados de tratamentos e auxiliar na personalização das terapias, de acordo com as particularidades de cada paciente.

No mercado financeiro, a IA é amplamente utilizada para análise de risco, detecção de fraudes e otimização de investimentos. Algoritmos de aprendizado de máquina permitem que as instituições financeiras processem grandes quantidades de dados em tempo real, identificando padrões que revelam atividades suspeitas ou que apontam oportunidades lucrativas de investimento.

Empresas de seguros e instituições financeiras também têm empregado a IA para aprimorar a avaliação de riscos e a precificação de seguros, além de realizar análises de crédito mais precisas. Ferramentas de Big Data e aprendizado de máquina analisam vastos volumes de informações para prever o comportamento do consumidor, determinar o risco de inadimplência e, consequentemente, ajustar os prêmios de seguros (Gesser, 2023) e as taxas de juros de empréstimos (Malar, 2023).

A segurança pública é outra área em que a IA tem sido amplamente empregada. Ferramentas de reconhecimento facial e algoritmos preditivos — como os utilizados para prever a ocorrência e a localização de crimes — são exemplos de como a inteligência artificial vem sendo aplicada com o objetivo de aumentar a eficiência das operações policiais. No entanto, essas aplicações suscitam preocupações significativas quanto à privacidade e ao risco de discriminação.

Transcendendo aspectos individuais e coletivos da vida em sociedade, a utilização da inteligência artificial está presente também no Poder Judiciário. Até junho de 2022, 47 tribunais empregavam a IA em alguma(s) atividade(s) (FGV, 2022, p.10). No sistema judicial, a IA é aplicada para automatizar tarefas administrativas, como a triagem de processos, bem como para apoiar a tomada de decisões judiciais, com a promessa de aumentar a eficiência, conferir celeridade e reduzir a sobrecarga de trabalho no Judiciário.

A inteligência artificial já desempenha, portanto, um papel crucial em múltiplos setores da sociedade — da saúde à segurança pública, do sistema financeiro ao Judiciário. Com seu crescimento contínuo e sua implementação cada vez mais disseminada, torna-se essencial compreender não apenas suas capacidades e aplicações, mas também os desafios éticos e jurídicos decorrentes de seu uso. No próximo capítulo, serão analisados os impactos dessas aplicações sobre os direitos fundamentais, especialmente em contextos nos quais a transparência e a imparcialidade são exigências centrais.

DIREITOS FUNDAMENTAIS E INTELIGÊNCIA ARTIFICIAL

Os direitos fundamentais são os direitos básicos e inalienáveis garantidos a todos os indivíduos, reconhecidos e protegidos pelo ordenamento jurídico nacional.

Conforme Ferrajoli, os direitos fundamentais delimitam "a esfera do indecidível" no novo constitucionalismo e fazem ruir a concepção clássica de democracia como poder da maioria, determinando o ponto a partir do qual nenhum poder — nem mesmo o da maioria — está autorizado a decidir ou não decidir. Consubstanciam-se em limites democraticamente impostos para proteger o cidadão mediante regras prévias e universais, sejam elas de liberdade, que impõem proibições, ou sociais, que impõem obrigações ao legislador (Ferrajoli, 2011, p.29).

Tais direitos resguardam garantias resultantes de conquistas civilizatórias históricas, fruto de lutas e revoluções. Protegem não apenas os cidadãos contra os abusos do Estado, mas também asseguram direitos vitais a todos os indivíduos. Por essa razão, constituem a esfera do indecidível, expressando as cláusulas pétreas do ordenamento constitucional (Ferrajoli, 2011, p.29).

Os direitos fundamentais são essenciais à dignidade humana e abrangem, entre outros, o direito à vida, à liberdade, à privacidade, à igualdade e ao devido processo legal. No contexto brasileiro, esses direitos são assegurados pela Constituição Federal de 1988, a qual os estabelece como pilares do Estado Democrático de Direito.

A evolução tecnológica, especialmente com o advento da inteligência artificial (IA), impôs novos desafios à proteção desses direitos. A aplicação da IA em diversos setores da sociedade tem o potencial de impactar diretamente os direitos fundamentais, exigindo, portanto, uma análise cuidadosa e a criação de mecanismos legais capazes de assegurar que o avanço tecnológico ocorra de forma compatível com os princípios da justiça, da igualdade e da dignidade humana.

A seguir, analisam-se algumas das principais intersecções entre a IA e os direitos fundamentais.

Direito à Privacidade

O direito à privacidade, garantido no artigo 5º, inciso X, da Constituição Federal, mostra-se particularmente vulnerável no contexto da IA. As tecnologias de inteligência artificial frequentemente dependem da coleta, armazenamento e análise de grandes volumes de dados pessoais. Exemplos incluem sistemas de reconhecimento facial, que capturam e processam imagens de indivíduos sem seu consentimento explícito, e algoritmos de análise de crédito, que utilizam dados de comportamento online para definir a elegibilidade de empréstimos.

A *General Data Protection Regulation* (GDPR), na União Europeia, e a Lei Geral de Proteção de Dados (LGPD), no Brasil, representam tentativas de regulamentar o uso desses dados, impondo limites à sua coleta e exigindo maior transparência por parte das empresas. No entanto, a aplicação prática dessas normas no contexto da IA permanece desafiadora — tanto pelo caráter opaco e complexo dos algoritmos utilizados quanto por limitações na governança da internet e na implementação de mecanismos eficazes de fiscalização. Afinal, como é possível regular efetivamente grandes corporações tecnológicas que operam globalmente, por meio de regulamentações essencialmente locais? (Vega Iracelay, 2018, p.22).

Direito à Igualdade e à Não Discriminação

O princípio da igualdade, previsto no caput do artigo 5º da Constituição Federal, assegura que todos são iguais perante a lei, sem distinção de qualquer natureza. No entanto, sistemas de IA podem, inadvertidamente, perpetuar ou até intensificar desigualdades já existentes na sociedade. Isso ocorre, principalmente, por meio dos vieses algorítmicos: quando os sistemas são treinados com dados históricos tendenciosos, acabam por reproduzir preconceitos de raça, gênero, classe social, entre outros.

Exemplos notórios incluem sistemas de reconhecimento facial que apresentam maiores taxas de erro na identificação de pessoas negras (Buolamwini, 2018, p.1-15) e algoritmos de recrutamento que, ao processar currículos, tendem a privilegiar candidatos do sexo masculino em detrimento de candidatas mulheres (Dastin, 2018).

Essas práticas violam diretamente o direito à igualdade e impõem sérios obstáculos à promoção da justiça social, sobretudo em um contexto em que a inteligência artificial se torna cada vez mais presente nas dinâmicas sociais.

Direito ao Devido Processo Legal

O devido processo legal, previsto nos incisos LIV e LV do artigo 5º da Constituição Federal, assegura que ninguém será privado de sua liberdade ou de seus bens sem o respeito às garantias do contraditório, da ampla defesa e do julgamento imparcial. A introdução de sistemas de inteligência artificial no Judiciário, especialmente aqueles voltados à análise preditiva ou à automação de decisões processuais, levanta preocupações relevantes quanto à compatibilidade com esses princípios.

Nos Estados Unidos, o sistema COMPAS (*Correctional Offender Management Profiling for Alternative Sanctions*), utilizado para avaliar o risco de reincidência de réus, tornou-se emblemático nesse debate. Segundo estudo de Larson et al. (Larson et al, 2016), o sistema demonstrou vieses raciais, classificando pessoas negras como de maior risco de reincidência em comparação com pessoas brancas em situações equivalentes, comprometendo a isonomia e a justiça material.

Um dos principais problemas dos sistemas de IA aplicados à justiça reside em sua opacidade. Muitos deles operam como verdadeiras *caixas-pretas*, cujos processos internos de decisão não são acessíveis nem mesmo aos seus próprios desenvolvedores. Esse fenômeno, conhecido como *black box algorithms*, foi criticamente abordado por O'Neil (O'Neil, 2020, p.144), ao afirmar que, apesar de seu verniz técnico e matemático, os modelos algorítmicos são essencialmente construções humanas impregnadas de valores e julgamentos subjetivos.

Essa opacidade compromete a possibilidade de contestação efetiva das decisões automatizadas. Se a parte interessada não compreende os critérios que fundamentaram determinada decisão, sua capacidade de exercer o contraditório e apresentar defesa técnica adequada é substancialmente limitada. Essa assimetria informacional mina o princípio do *due process*, ao transferir poder decisório a sistemas que não se submetem aos mesmos critérios de publicidade e transparência exigidos do julgador humano (Floridi et al., 2018, p. 700).

Além disso, a ausência de mecanismos de explicação e revisão das decisões algorítmicas compromete a responsabilização (*accountability*), elemento essencial ao devido processo legal. A responsabilização garante que erros, discriminações ou abusos possam ser corrigidos por instâncias superiores ou independentes (Floridi et al., 2018, p. 700).

Para mitigar esses riscos, os sistemas de IA utilizados em contextos jurídicos e administrativos devem ser auditáveis, explicáveis e estar sujeitos à supervisão humana (Floridi et al., 2018, p. 700). A Recomendação da OCDE sobre Inteligência Artificial (OCDE, 2019) enfatiza a necessidade de que tais sistemas sejam transparentes e contenham mecanismos de contestação, especialmente quando envolvem decisões que afetam direitos fundamentais.

Segundo essa Recomendação, os agentes responsáveis pelo uso de IA devem comprometer-se com a transparência e a divulgação responsável dos sistemas, fornecendo, sempre que possível e útil, informações claras e acessíveis sobre as fontes de dados, os fatores considerados, os processos e/ou a lógica que levou à previsão, recomendação ou decisão. Também é necessário garantir que as pessoas afetadas por sistemas de IA tenham acesso a meios eficazes para contestar os resultados que lhes sejam desfavoráveis (OCDE, 2019).

Assim, embora a IA possa contribuir significativamente para a eficiência do sistema judiciário, seu uso deve estar incondicionalmente subordinado ao respeito aos princípios constitucionais do devido processo legal. Isso implica a adoção de medidas técnicas e normativas que assegurem não apenas a justiça do resultado, mas também a legitimidade do processo decisório.

VIESES ALGORÍTMICOS E IMPACTOS EM DIREITOS FUNDAMENTAIS

Vieses algorítmicos referem-se a distorções sistemáticas nas saídas de um modelo de IA que resultam em decisões injustas ou prejudiciais para determinados grupos. Esses vieses podem surgir em diferentes etapas do desenvolvimento de um algoritmo — desde a coleta e seleção dos dados até o design e a implementação do modelo.

Serhii Pospielov explica que o viés de aprendizado de máquina ocorre quando um algoritmo produz sistematicamente resultados tendenciosos, em razão de suposições incorretas durante o processo de aprendizado. Tais distorções podem assumir diversas formas, incluindo viés de gênero, preconceito racial, discriminação por idade e tratamento desigual em processos seletivos (Pospielov, 2022).

Os vieses algorítmicos podem ser classificados em várias categorias, entre as quais se destacam:

- Viés de amostragem: ocorre quando os dados utilizados no treinamento do modelo não representam adequadamente a diversidade da população, o que leva a previsões distorcidas para grupos sub-representados (Mehrabi et al., 2022).
- Viés de medição: surge quando as variáveis utilizadas para mensurar determinados atributos não capturam com precisão as características que se deseja avaliar, resultando em discriminação indireta. Um exemplo é o sistema COMPAS, que utilizava variáveis como prisões anteriores ou de amigos e familiares como proxies para o nível de risco de reincidência (Mehrabi et al., 2022).
- Viés de variável omitida: refere-se à ausência de variáveis relevantes no modelo, o que compromete a acurácia e a neutralidade das previsões (Mehrabi et al., 2022).
- Viés de representação: decorre de processos de coleta de dados que deixam de incluir adequadamente certos subgrupos da população, o que compromete a diversidade e a abrangência da base de dados (Mehrabi et al., 2022).

Panch et al. definem o viés algorítmico como a aplicação de um algoritmo que acentua desigualdades pré-existentes — de ordem socioeconômica, racial, étnica, religiosa, de gênero, deficiência ou orientação sexual. Segundo os autores, os algoritmos não apenas refletem desigualdades sociais já estabelecidas, mas também têm o potencial de agravá-las: "Se o mundo tiver determinada aparência, isso se refletirá nos dados, seja diretamente ou por meio de proxies, e, portanto, nas decisões algorítmicas" (Panch et al., 2019, p. 01).

Isso quer dizer que os modelos algorítmicos, como elucida O'Neil, apesar de desfrutarem de uma reputação de imparcialidade, são "opiniões embutidas em matemática". Eles expressam os objetivos e ideologias de seus criadores, e, "seus pontos cegos, refletem as prioridades e o julgamento de quem os alimentou" (O´Neil, 2020, pp. 30-35).

Essa crítica é reforçada por O'Neil, que afirma que, apesar da reputação de imparcialidade, os modelos algorítmicos são, em essência, "opiniões embutidas em matemática". Eles expressam os objetivos, as ideologias e os pontos cegos de seus desenvolvedores, refletindo prioridades e julgamentos implícitos (O'Neil, 2020, p. 30-35).

Como os algoritmos operam com base em dados estatísticos, eles não realizam juízos de valor, mas apenas inferem padrões com base no que ocorreu no passado. Nesse sentido, refletem

a dinâmica social vigente e dificilmente projetam um futuro diferente, o que implica a reprodução — ou até ampliação — de injustiças históricas.

Esses vieses já foram amplamente documentados em diversas aplicações da IA, com impactos significativos sobre os direitos fundamentais. Ainda que constituam ferramentas extremamente promissoras, a inteligência artificial possui limitações importantes e, em muitos casos, sua aplicação acarreta graves consequências jurídicas e sociais.

Embora o uso da IA no Poder Judiciário brasileiro esteja parcialmente regulamentado pela Resolução nº 332/2020 do Conselho Nacional de Justiça, a maior parte das aplicações da tecnologia — em áreas como segurança pública, mercado de trabalho, saúde, crédito e seguros — segue desprovida de normas específicas, o que aumenta os riscos de discriminação, sexismo e racismo algorítmico.

Diversos estudos demonstram a existência de vieses discriminatórios em softwares utilizados para recrutamento de candidatos, análise de crédito, diagnóstico médico, avaliação de risco de reincidência penal, reconhecimento facial e predição de crimes. A análise desses sistemas revela distorções que resultam em discriminações sistemáticas com base em renda, raça e gênero.

Nos Estados Unidos, o já mencionado sistema COMPAS mostrou-se enviesado ao classificar pessoas negras como mais propensas à reincidência do que pessoas brancas em situações equivalentes, perpetuando a desigualdade racial (Larson et al., 2016).

Em 2018, a Amazon descontinuou um algoritmo de recrutamento que apresentava viés sexista, penalizando sistematicamente candidatas do sexo feminino. A discriminação da ferramenta contra candidatas do sexo feminino acontecia, porque os dados utilizados para treinamento e validação do modelo foram currículos enviados para a empresa nos últimos 10 anos, que são em sua imensa maioria de homens, como acontece na maior parte da indústria de tecnologia (Dastin, 2018). Assim, o algoritmo passou a entender que os homens naturalmente eram mais aptos para as vagas, refletindo a histórica desigualdade de gênero no setor de tecnologia.

Outro exemplo é o software Rekognition, da Amazon, amplamente utilizado por órgãos policiais norte-americanos para reconhecimento facial. Um estudo da ACLU (American Civil Liberties Union) revelou que o sistema identificou erroneamente 28 membros do Congresso como criminosos procurados. "As falsas correspondências eram desproporcionalmente de pessoas de cor". Cerca de 40% dos identificados eram negros — embora esse grupo representasse apenas 20% dos congressistas (Snow, 2018).

Joy Buolamwini e Timnit Gebruque no artigo, "*Tons de gênero: disparidades de precisão interseccional em classificação comercial de gênero*", trouxeram à tona inúmeros estudos que foram feitos comprovando disparidades substanciais na precisão do reconhecimento facial. A conclusão foi a de que mulheres de pele mais escura constituem o grupo mais mal classificado, com taxas de erro de até 34,7%. Enquanto isso, a taxa máxima de erro para homens de pele mais clara é de 0,8% (Buolamwini, Gerbruque, 2018, p.08).

Em "AI, não sou uma mulher?, Buolamwini resolveu testar os softwares de reconhecimento facial das maiores gigantes da tecnologia, como Microsoft, Google, Amazon, IBM e Face++. O teste se utilizou de fotografias de mulheres negras icônicas, como Michelle Obama, Serena Williams e Oprah Winfrey. Serena Williams foi rotulada como "homem"; Michelle Obama foi rotulada com o descritor "*um jovem homem vestindo uma camisa preta e com peruca*" e Oprah foi rotulada como "*aparentemente um homem*". Já a foto de Oprah quando era jovem recebeu a rotulagem de "nenhum rosto detectado" (Buolamwini, Gerbruque, 2018).

Nesta mesma linha está o *software PredPol*, utilizado por departamentos de polícia para prever lugares onde crimes futuros podem acontecer. Programas como este vêm sendo amplamente utilizados em departamentos de polícia que dispõem de reduzido orçamento nos Estados Unidos. O *PredPol* é baseado em software sísmico, ou seja, à medida que ele vê um crime acontecer em dado local, ele incorpora em padrões de histórico, a fim de prever os locais aonde novos crimes poderão acontecer futuramente. Diferentemente do COMPAS, este tipo de software não se concentra no indivíduo, levando em conta apenas a localização geográfica. No entanto, ao configurar o sistema *PredPol*, os departamentos de polícia inserem, como dados de treinamento, relatórios policiais de crimes anteriores. E estes relatórios contemplam majoritariamente crimes de perturbação (já que os policiais foram todos treinados na ortodoxia da política da "tolerância zero") (O´Neil, 2020, p. 135-137).

É fato que esses tipos de notificações de práticas delitivas são endêmicos em bairros de baixa renda, o que alimenta o sistema com mais casos, gerando alta pontuação no *PredPol*. A consequência é que o sistema reforça o policiamento em bairros pobres e majoritariamente negros, criando um ciclo contínuo de estigmatização e violência institucional, com base em evidências enviesadas. Logo, o *PredPol* acaba por automatizar os preconceitos e o racismo com ares de evidência matemática e a falsa percepção de legitimidade, pois opera mediante algoritmos "imparciais". (O´Neil, 2020, p. 135-137).

Estudo apresentado por Marda e Narayan (2020) na *Conference on Fairness, Accountability, and Transparency*, demonstrou que software similar utilizado em Nova Deli, na Índia, apresentava tendência discriminatória contra comunidades de migrantes e minorias étnicas, intensificando a vigilância sobre esses grupos (Marda; Narayan, 2020).

Esse enviesamento ocorre porque os sistemas algorítmicos realizam inferências com base em dados históricos, o que os impede de projetar um futuro distinto do passado. Além de refletirem preconceitos estruturais incorporados ao longo da história, esses registros não necessariamente correspondem aos valores contemporâneos e tampouco admitem revisão espontânea. Assim, ao serem convertidos em dados de treinamento para modelos de IA, tornam-se obstáculos à superação do racismo e de outras formas de injustiça histórica.

Na área da saúde, embora ferramentas de inteligência artificial apresentem avanços notáveis na detecção de tumores por meio de exames de imagem e no diagnóstico precoce de doenças, seus efeitos discriminatórios também se fazem presentes. Em artigo publicado na revista *Science*, identificaram viés racial significativo em um algoritmo amplamente utilizado no sistema de saúde dos Estados Unidos. O estudo demonstrou que, mesmo quando o algoritmo atribuía o mesmo nível de risco a pacientes brancos e negros, os pacientes negros apresentavam, na prática, condições de saúde mais graves. Como resultado, menos da metade dos pacientes negros elegíveis recebia cuidados adicionais. A distorção decorre do uso de custos médicos anteriores como variável de treinamento. Historicamente, os gastos com pacientes negros são menores do que com pacientes brancos, em razão do acesso desigual ao sistema de saúde. Assim, o algoritmo interpretava falsamente que os pacientes negros eram mais saudáveis, perpetuando a exclusão desses grupos do cuidado adequado (Obermeyer al, 2019, p.05).

Discriminações semelhantes ocorrem em relação ao gênero. Mulheres têm sido, historicamente, subdiagnosticadas em doenças cardiovasculares e recebem, com frequência, tratamentos inadequados ou insuficientes (Niethammer, 2020).

No setor de seguros, a utilização de dados para a classificação e precificação de riscos é uma prática tradicional e inerente à própria lógica da atividade securitária, sendo essencial para sua

sustentabilidade econômica. Embora o funcionamento estrutural das seguradoras não tenha se transformado radicalmente, o acesso ampliado a volumes massivos de dados — incluindo informações comportamentais, genéticas e de saúde — tem potencializado os efeitos dessa prática, ampliando não apenas sua eficácia, mas também seus impactos sociais.

A adoção de algoritmos de *deep learning* e técnicas de *Big Data Analytics* tem proporcionado ganhos significativos em eficiência, precisão na precificação e capacidade de detecção de fraudes — estas, por sua vez, representam um dos maiores desafios históricos do setor e influenciam diretamente o valor dos prêmios. Nesse contexto, é recorrente o argumento das seguradoras de que o compartilhamento voluntário de dados pelos usuários viabiliza uma análise de risco mais refinada e, consequentemente, a redução dos custos do seguro para determinados perfis.

O problema ocorre, porém, quando algoritmos de aprendizagem profunda (*deep learning*) têm acesso a um conjunto enorme de dados sensíveis ou *proxies* comportamentais, muitas vezes coletados sem consentimento — como localização, dados de redes sociais, histórico de compras e uso de wearables, como os *smart watch*.

Essa quantidade massiva de dados permite a análise de inteligência de mídia social e a extração de dados sensíveis, traçando um perfil da personalidade do indivíduo, locais que frequenta, hábitos de compra, risco de morbidade ou longevidade, nível de agressividade e exposição a risco, histórico familiar, amizades, realização de exames, batimentos cardíacos, nível de atividade física, etc. A utilização deste Big Data pelas companhias de seguro certamente contribui para análise de risco e na precificação precisa, mas também permite a amplificação da discriminação.

Pesquisa do Institute and Faculty of Actuaries, em parceria com a Fair By Design, demonstrou que pessoas em situação de pobreza pagam, anualmente, cerca de quinhentas libras a mais por serviços essenciais, como seguros, energia e crédito, devido à precificação algorítmica baseada em fatores discriminatórios (Fair By Design, 2021). Já o estudo da Citizen Advice mostrou que pessoas negras pagam, em média, duzentos e oitenta libras a mais por seguros de automóveis do que pessoas brancas em condições semelhantes. Isso quer dizer que pessoas negras estão pagando duzentos e treze mihões de libras a mais em seguro de carro todos os anos, simplesmente por residirem em áreas com grandes comunidades de cor (Citizen Advice, 2022, p.41-42).

Dados históricos utilizados no treinamento de algoritmos de aprendizado de máquina podem conduzir à formulação de premissas equivocadas, como a suposição de que pessoas em situação de pobreza representam maior risco e, portanto, devem arcar com prêmios securitários mais elevados. Esse tipo de inferência tende a gerar um ciclo de retroalimentação nocivo, que não apenas perpetua como também intensifica desigualdades sociais preexistentes.

A discriminação algorítmica pode manifestar-se, por exemplo, por meio da aplicação de prêmios mais altos a determinados grupos sociais, raciais ou étnicos, com base em padrões históricos de mortalidade ou morbidade. Se tais grupos apresentaram, historicamente, taxas de mortalidade mais elevadas — frequentemente em decorrência de fatores estruturais como exclusão social, condições precárias de habitação, violência urbana ou ausência de acesso adequado à saúde — o algoritmo pode, de forma equivocada, associá-los a um risco intrínseco mais alto, sem considerar os determinantes sociais subjacentes. Dessa forma, mesmo quando não há justificativa objetiva para essa classificação, o sistema reproduz e legitima desigualdades sob a aparência de neutralidade técnica.

Além da majoração de prêmios, a parcialidade dos dados pode resultar na recusa de cobertura para indivíduos justamente pertencentes aos grupos mais vulneráveis e mais necessitados de proteção securitária.

Embora a discriminação direta seja juridicamente vedada, a discriminação indireta — mediada por *proxies* estatísticos ou por algoritmos complexos e opacos — ainda carece de regulamentação específica, criando um vazio normativo que permite a reprodução de injustiças sob a aparência de objetividade matemática.

Como agravante, está o fato de que estes sistemas são invisíveis. São utilizados por empresas privadas sem a ciência dos consumidores. As empresas não informam de que maneira a decisão foi tomada.

Outro problema relevante diz respeito à forma como seguradoras têm explorado a proliferação de dados pessoais para estabelecer correlações e inferências que, além de questionáveis do ponto de vista técnico, são frequentemente ocultadas dos indivíduos afetados. Por exemplo, ao analisar dados de programas de fidelidade de supermercados, o sistema pode concluir que determinado consumidor adota hábitos alimentares prejudiciais à saúde, elevando seu risco de desenvolver doenças. Por outro lado, o simples fato de uma pessoa estar vinculada a uma academia pode ser interpretado como indicativo de um estilo de vida saudável, mesmo sem evidências de frequência ou prática efetiva de atividade física. Essas inferências, embora revestidas de aparente racionalidade estatística, podem ser imprecisas ou absolutamente equivocadas, resultando em práticas discriminatórias não apenas ilegais, mas também impossíveis de serem contestadas — justamente porque não são comunicadas ao titular dos dados.

Embora a inteligência artificial tenha potencial para ampliar o acesso ao seguro para certos grupos, com redução do valor das apólices e flexibilização de critérios, há também o risco de exclusão de pessoas consideradas excessivamente arriscadas ou economicamente inviáveis. Paradoxalmente, são justamente os indivíduos em situação de maior vulnerabilidade — e, portanto, os que mais necessitam de proteção securitária — que tendem a ser negativamente impactados por essa lógica de seleção algorítmica.

Apesar do discurso de que o uso da análise de dados e da inteligência artificial pode proporcionar ganhos de eficiência e permitir a personalização do valor do prêmio, não se pode ignorar que essa prática também pode catalisar preconceitos, além de afrontar diretamente direitos fundamentais como a privacidade, a autonomia e a dignidade da pessoa humana.

Ainda que a atividade securitária se fundamente na aferição e na precificação de riscos, é legítimo — do ponto de vista ético e jurídico — que uma seguradora utilize dados sensíveis, alternativos — muitas vezes obtidos sem consentimento, de forma opaca e silenciosa — para traçar perfis genéticos, hereditários e comportamentais, e a partir deles, aplicar práticas excludentes? A ausência de regulação específica para essas situações deixa o indivíduo vulnerável a decisões automatizadas que o rotulam silenciosamente, sem possibilidade de defesa ou correção.

Contudo, esses sistemas estão alcançando uma adoção rápida e generalizada, prometendo acelerar a tomada de decisões, eliminar atrasos, fazer avaliações mais assertivas e reduzir custos. O Brasil tem incorporado, de forma crescente, o uso de sistemas de inteligência artificial em múltiplos setores, como a segurança pública, o mercado de seguros, a análise de crédito e a triagem automatizada de currículos para processos seletivos.

No Brasil, já há uso concreto e documentado de sistemas com potencial discriminatório semelhante. Ferramentas de reconhecimento facial, como as utilizadas por forças policiais em estados como Bahia, Rio de Janeiro, São Paulo e Santa Catarina, têm apresentado altas taxas de erro, especialmente contra pessoas negras. Tais sistemas foram responsáveis por prisões indevidas, predominantemente de pessoas negras (Alencar, 2023).

Consoante dados levantados pela Folha, por meio das secretarias estaduais de Segurança e das polícias Civil e Militar, vinte estados brasileiros utilizam ou estão implementando a tecnologia de reconhecimento facial na segurança pública local. Outros três estudam sua implementação e apenas quatro estados não utilizam, não tiveram contato com o sistema ou planejam utilizar (Damasceno; Fernandes, 2021).

Merece destaque o software CrimeRadar, desenvolvido pelo Instituto Igarapé e utilizado no Brasil como ferramenta de apoio à atuação policial. Inspirado no modelo norte-americano PredPol, o sistema utiliza algoritmos de aprendizado profundo para mapear e prever a ocorrência de crimes em diferentes bairros e horários. De acordo com seus desenvolvedores, a plataforma seria capaz de contribuir para a redução de homicídios em até 10%, diminuir registros criminais em até 40% e encurtar significativamente o tempo de resposta a ocorrências (Instituto Igarapé, 2023).

Apesar dessas promessas, não há, atualmente, regulamentação que estabeleça critérios mínimos para o funcionamento de tais sistemas, tampouco mecanismos de controle sobre aspectos como a origem e a qualidade dos dados utilizados para treinamento, os protocolos de auditoria externa, o nível de acurácia, ou a transparência e proteção dos dados coletados. Isso se agrava diante do disposto no art. 4º da Lei Geral de Proteção de Dados (LGPD), que exige a edição de legislação específica para o tratamento de dados pessoais em atividades de segurança pública e investigação criminal — norma que, até o presente momento, não foi implementada.

Além do campo penal e da segurança pública, o Brasil tem adotado ferramentas de inteligência artificial em outros setores igualmente sensíveis — como seguros, análise de crédito e processos seletivos — cujos impactos podem ser profundos. Sem a devida regulamentação, publicidade, transparência e supervisão ética, essas tecnologias têm o potencial de ampliar desigualdades e violar direitos fundamentais, como a privacidade, a autonomia e a dignidade da pessoa humana.

A plataforma MetLife Xcelerator, desenvolvida em parceria com a insurtech Klimber, oferece soluções baseadas em inteligência artificial e análise de dados voltadas ao mercado de seguros no Brasil e na América Latina (MetLife, 2023). Já a empresa catarinense Neoway, especializada em Big Data e IA, promove a chamada "precificação cirúrgica", prometendo definir valores com maior precisão a partir de múltiplas fontes de informação. Segundo a própria empresa, "as informações que a organização **coleta diariamente** revelam fatos interessantes sobre o comportamento do consumidor e podem indicar caminhos mais precisos para precificar" (Neoway, 2022).

Outro exemplo é a Neurotech, empresa brasileira que desenvolve soluções baseadas em inteligência artificial, machine learning e Big Data aplicadas aos setores de crédito, seguros e saúde. Sua plataforma "Neurotech Saúde" já está em operação no país, oferecendo análise automatizada de risco para diferentes perfis populacionais.

Essas tendências foram discutidas no Insurtech Latam Fórum 2023, um dos principais eventos latino-americanos sobre inovação no setor segurador, realizado em agosto. O fórum reuniu empresas líderes em tecnologia e os principais bancos e seguradoras do Brasil, evidenciando que a incorporação da IA ao mercado securitário nacional tende a se expandir significativamente nos próximos anos (Universo do Seguro, 2023).

A Neoway, que se apresenta como a maior empresa da América Latina em Big Data Analytics e inteligência artificial para negócios, já anuncia a aplicação de seus serviços na gestão da saúde pública brasileira. Segundo declarações da empresa, "a ciência de dados, quando atrelada a uma

série de tecnologias, aprimora a gestão da saúde pública e favorece a tomada de decisões em municípios e hospitais" (Neoway, 2023).

No mercado de crédito, o uso de IA e análise de dados também avança rapidamente. A empresa americana FIS, em parceria com a startup *Neener Analytics* — ambas com atuação no Brasil —, desenvolveu um sistema que utiliza **informações extraídas de redes sociais** para traçar perfis comportamentais e balizar a concessão de empréstimos (Malar, 2023). Conforme divulgado pela *Neener Analytics*, a ferramenta é capaz de estimar, com quase 80% de precisão, características como propensão à inadimplência, atitude de risco, paciência, impaciência e veracidade das informações prestadas pelo cliente. Além disso, o sistema classifica os usuários como *Transactor*, *Revolver* ou *Dormant1*, identificando padrões de comportamento financeiro e permitindo previsões sobre a forma como o empréstimo será quitado (Neener Analytics, 2023).

Também no setor de crédito, destaca-se a plataforma Riskpack, desenvolvida pela Neurotech, que utiliza dados de geolocalização e perfis comportamentais para avaliar o risco de inadimplência e conceder crédito de forma automatizada. A empresa afirma ter acesso a mais de mil fontes externas para coleta de dados, com as quais é possível identificar padrões de comportamento financeiro, prever capacidades de pagamento e classificar os clientes segundo suas rotinas de consumo (Neurotech, 2023).

A crescente utilização de dados pessoais e comportamentais — inclusive provenientes de fontes não tradicionais — para análise e precificação de riscos no mercado de seguros e crédito tem suscitado sérias preocupações. Embora o argumento da eficiência e da personalização dos serviços seja recorrente, essas práticas também podem ferir direitos fundamentais como a igualdade, a privacidade, a autonomia e a dignidade da pessoa humana. O uso de informações sensíveis, como dados de localização, hábitos de consumo, registros em redes sociais e até mesmo métricas obtidas por dispositivos vestíveis (*wearables*), muitas vezes ocorre sem o consentimento informado dos indivíduos, sendo utilizados em seu prejuízo.

Embora não se deva desconsiderar os inúmeros e relevantes benefícios proporcionados pela utilização da inteligência artificial, é inegável que seu uso também impõe desafios inéditos e substanciais. A crescente adoção de conjuntos de dados massivos — muitas vezes extraídos de fontes não tradicionais e sem o consentimento informado dos titulares — tem alimentado sistemas cuja lógica decisória é opaca, de difícil explicação e invisível ao cidadão comum. Esse cenário amplia o risco de violações a direitos fundamentais como a liberdade, a igualdade, a autonomia, o devido processo legal e a privacidade.

Diante disso, torna-se imperativo compreender a complexidade desses riscos e avaliar, com rigor, se a proposta regulatória em elaboração é capaz de prever mecanismos eficazes de mitigação, proteção e fiscalização.

PROJETO DE LEI Nº 2338/2023 E A REGULAMENTAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO BRASIL

O Brasil, como muitas outras nações, está enfrentando o desafio de regular o rápido avanço das tecnologias de IA. O Projeto de Lei nº 2338/2023, atualmente em tramitação, é a iniciativa mais avançada do país para criar um marco regulatório abrangente para a IA. Este projeto visa

[1] Transactores são aqueles que pagam a fatura integralmente a cada mês, evitando juros. Revolvers são aqueles que não pagam a fatura integralmente e carregam saldos mensais, pagando juros. (Consolidated Credit, 2024)

estabelecer normas gerais para o desenvolvimento, implementação e uso de sistemas de IA, com foco na proteção de direitos fundamentais e na promoção de inovação responsável (Senado Federal, 2023).

O Projeto de Lei nº 2338/2023 representa uma das iniciativas mais relevantes no cenário legislativo brasileiro para disciplinar o uso da inteligência artificial (IA). Estruturado a partir de uma abordagem orientada por riscos, o PL busca estabelecer fundamentos jurídicos, éticos e técnicos para assegurar que o desenvolvimento e a aplicação da IA no país ocorram em conformidade com os direitos fundamentais e os valores democráticos.

Entre os principais pontos do projeto, destacam-se:

a) Princípios éticos: O projeto estabelece um conjunto de princípios que devem nortear todo o ciclo de vida dos sistemas de IA. Entre eles estão a transparência, a equidade, a não discriminação, a privacidade, a responsabilidade, o respeito ao devido processo legal, o contraditório e a contestabilidade. Esses princípios buscam garantir que a IA promova a inovação tecnológica sem comprometer a dignidade humana nem acentuar desigualdades sociais (Wachowicz, 2024).

b) Definições e escopo: O texto legal define com precisão conceitos como "sistema de inteligência artificial", "fornecedor" e "operador" de IA, além de "discriminação direta" e "indireta", e "mineração de textos e dados". Essa taxonomia é relevante para delimitar obrigações e responsabilidades, bem como os direitos dos afetados.

c) Direitos dos afetados: O PL garante um conjunto de direitos fundamentais às pessoas naturais impactadas por sistemas de IA, incluindo:

- direito à informação prévia e acessível sobre interações com IA;
- direito à explicação das decisões, previsões ou recomendações automatizadas;
- direito à contestação de decisões automatizadas;
- direito à autodeterminação e supervisão humana;
- direito à privacidade e à proteção de dados pessoais;
- direito à não discriminação e à correção de vieses abusivos, ilegais ou indiretos.

Todos os sistemas que apresentem alguma camada de inteligência artificial estarão sujeitos a esses direitos, o que representa avanço normativo relevante (Unzelte, 2024).

d) Proibições expressas: O PL veda, de forma categórica, determinados usos considerados incompatíveis com os direitos fundamentais, como:

- técnicas subliminares que induzam comportamentos prejudiciais à saúde ou segurança;
- exploração de vulnerabilidades de grupos específicos (ex.: crianças, idosos ou pessoas com deficiência);
- uso, pelo poder público, de sistemas de ranqueamento social que classifiquem cidadãos com base em comportamento ou traços de personalidade, em contextos desproporcionais ou ilegítimos.

No campo da segurança pública, a identificação biométrica à distância em espaços públicos só será admitida em situações excepcionais e específicas: crimes com pena superior a dois anos, flagrantes, ou busca por vítimas e desaparecidos (Senado Federal, 2023).

e) Avaliação de impacto: O projeto exige a realização obrigatória de avaliações de impacto para sistemas classificados como de "alto risco". Essas avaliações devem anteceder a implementação

e analisar os riscos aos direitos fundamentais, bem como prever medidas técnicas, organizacionais e jurídicas para mitigação (Senado Federal, 2023).

f) Transparência e explicabilidade: Sistemas que impactem diretamente os direitos dos cidadãos deverão ser compreensíveis, auditáveis e acompanhados de informações claras sobre sua lógica decisória. Isso visa combater a opacidade dos algoritmos e possibilitar o exercício efetivo do contraditório.

g) Governança algorítmica: Desenvolvedores e operadores deverão adotar práticas que assegurem a auditabilidade e a explicabilidade dos sistemas. Isso inclui testes de robustez, acurácia, precisão e cobertura, além de estratégias para a gestão de dados e mitigação de vieses discriminatórios.

h) Supervisão regulatória: O PL cria o Sistema Nacional de Regulação e Governança de Inteligência Artificial (SIA), coordenado pela Autoridade Nacional de Proteção de Dados (ANPD), designada pelo Congresso. À ANPD caberá regulamentar, fiscalizar e aplicar sanções, bem como atualizar continuamente a lista de sistemas classificados como de alto ou risco excessivo. Sistemas de alto risco deverão ser submetidos a testes e supervisão humana. Já os sistemas de risco excessivo estarão proibidos de operar no território nacional (Senado Notícias, 2024).

i) *Sandboxes* regulatórios: O PL prevê a criação de *sandboxes* — ambientes experimentais controlados em que tecnologias podem ser testadas sob monitoramento da autoridade reguladora. Essa abordagem permite fomentar a inovação com segurança, reduzindo o risco de impactos adversos antes da adoção em larga escala.

j) Direitos autorais e mineração de dados: Um avanço importante é a previsão de que determinadas atividades de mineração de dados, realizadas por instituições de pesquisa, bibliotecas, museus e arquivos, não constituirão violação de direitos autorais — desde que não tenham como finalidade a reprodução ou exibição da obra original, e desde que respeitem critérios de necessidade, proporcionalidade e não concorrência com a exploração econômica da obra (Senado Notícias, 2024). Essa exceção é essencial para garantir a viabilidade de aplicações em IA generativa e sistemas de linguagem.

k) Responsabilidade civil: O projeto adota uma lógica de responsabilidade objetiva para fornecedores e operadores de sistemas de alto risco ou risco excessivo, obrigando-os à reparação integral dos danos patrimoniais, morais, individuais ou coletivos causados. Para sistemas de menor risco, presume-se a culpa, com inversão do ônus da prova em favor da vítima (Almeida; Chang, 2025).

l) Sanções administrativas: A ANPD poderá aplicar sanções que vão desde advertência até multa de até R$ 50 milhões por infração, ou até 2% do faturamento anual da empresa, além de medidas como: suspensão temporária ou definitiva do sistema, exclusão de programas de *sandbox* e publicização da infração confirmada (Senado Federal, 2023).

O Projeto de Lei nº 2338/2023 representa um avanço importante na tentativa de estabelecer um marco normativo para a inteligência artificial no Brasil. No entanto, sua consolidação como instrumento efetivo de regulação ainda enfrenta desafios legislativos, técnicos e institucionais consideráveis.

O primeiro entrave diz respeito à própria tramitação do texto. Desde sua apresentação na Comissão Temporária sobre Inteligência Artificial, o projeto recebeu 145 emendas, teve sua votação adiada mais de três vezes e permanece em fase de audiências públicas. Setores econômicos, como o das big techs e da indústria nacional, vêm pressionando por alterações substanciais no texto

original. Há receios de que o projeto seja progressivamente desidratado diante das alegações de que ele representaria um entrave à inovação e ao desenvolvimento tecnológico.

Segundo relatório da Confederação Nacional da Indústria (CNI), o PL apresenta um modelo regulatório de amplitude e rigor sem precedentes, o que poderia "colocar o país sob o risco de isolamento e atraso tecnológico" (Portal da Indústria, 2024). Entre as críticas centrais está o fato de que a proposta regula o ciclo completo da IA — desde a concepção e desenvolvimento até a adoção e operação dos sistemas —, e não apenas sua implementação com base no risco. Para a CNI, isso impõe barreiras excessivas à pesquisa e ao avanço tecnológico.

Contudo, como demonstrado ao longo deste trabalho, os principais riscos da IA emergem justamente nas fases iniciais do ciclo de vida dos sistemas: na curadoria dos dados, no design algorítmico e nos conjuntos de treinamento. São nesses estágios que se originam os vieses discriminatórios e os riscos à privacidade e à equidade. Regular apenas a etapa final da aplicação equivaleria a ignorar a gênese dos problemas. Assim, enfraquecer essas disposições seria comprometer as salvaguardas mais relevantes do texto.

Outra objeção da CNI refere-se ao que qualifica como "governança excessiva", inclusive para sistemas de baixo risco, além da alegada invasão em processos internos das empresas, afetando segredos comerciais e a livre iniciativa. De fato, o PL exige que todos os sistemas de IA observem padrões mínimos de governança, incluindo: transparência na interação com usuários, medidas adequadas de segurança da informação e gestão de dados para prevenção de vieses discriminatórios.

Essas exigências, porém, são imprescindíveis. A transparência sobre os dados utilizados, os critérios de validação, o desempenho do sistema (como nível de acurácia), e a ciência dos usuários quanto à presença de IA em suas interações são requisitos básicos para garantir contestabilidade, correção de erros e responsabilização. Suprimir tais obrigações seria enfraquecer a proteção de direitos fundamentais, como o devido processo legal, a privacidade e a autodeterminação informativa.

A CNI também contesta a designação da Autoridade Nacional de Proteção de Dados (ANPD) como órgão responsável pela coordenação do Sistema Nacional de Regulação e Governança da Inteligência Artificial (SIA). A crítica se baseia na alegação de que a ANPD ainda está em fase de estruturação e carece de expertise técnica específica sobre IA. De fato, há preocupações legítimas sobre a capacidade institucional da ANPD para exercer um papel regulador tão abrangente, especialmente diante das pressões políticas e da complexidade do setor.

Adicionalmente, questiona-se se a ANPD disporá de recursos humanos, técnicos e financeiros suficientes para cumprir a missão de fiscalizar, categorizar riscos, aplicar sanções e promover a inovação segura. Conforme alertado por Unzelte (2024), "o texto do projeto propõe uma carga de governança extremamente densa, complexa e custosa".

Outra preocupação diz respeito ao tratamento isonômico entre empresas de diferentes portes. A CNI sustenta que pequenas e médias empresas podem ser desproporcionalmente afetadas por exigências rigorosas de governança, como auditorias externas. No entanto, o PL adota justamente uma abordagem baseada no risco das aplicações, e não no tamanho das empresas. Em outras palavras, é o potencial de dano que deve determinar a intensidade da fiscalização e das salvaguardas — afinal, uma pequena empresa também pode desenvolver um sistema de IA com alto risco para os direitos fundamentais.

No campo conceitual, Marcílio (Marcílio, 2024) destaca uma importante lacuna na definição de sistemas de IA apresentada no artigo 4º do PL. A redação atual inclui apenas os sistemas

voltados à produção de previsões, recomendações ou decisões, o que poderia excluir da regulação as aplicações de IA generativa. Considerando o crescimento exponencial dessas tecnologias e sua capacidade de produzir textos, imagens, vídeos e *deepfakes*, trata-se de uma omissão que exige urgente revisão legislativa.

Marcílio também aponta fragilidades na previsão sobre consentimento informado. Embora o projeto garanta o direito à informação prévia, ainda é incerta a forma de sua implementação prática. A simples fixação de placas alertando para a captação de imagens, ou a coleta automática de cookies, dificilmente podem ser consideradas como formas válidas de consentimento informado, já que não asseguram ao cidadão real compreensão das implicações da interação com sistemas automatizados.

Outro desafio relevante será a implementação efetiva do PL e a supervisão contínua de suas disposições. Ainda que o texto preveja a criação de *sandboxes* regulatórios — ambientes controlados para testes de IA —, a eficácia desses instrumentos depende da existência de mecanismos robustos de monitoramento, bem como da capacidade do Estado de mobilizar conhecimento técnico qualificado.

Além disso, a legitimidade do processo regulatório depende da participação contínua e qualificada de diferentes *stakeholders*: setor privado, governo, academia, organizações da sociedade civil e grupos vulnerabilizados. Essa multiplicidade de vozes é fundamental para que o marco regulatório da IA reflita as reais preocupações da sociedade brasileira e contribua para a construção de sistemas mais justos, inclusivos e seguros.

Sem a pretensão de esgotar o debate — até porque o texto ainda está em discussão e sujeito a alterações —, é possível afirmar que o PL nº 2338/2023 constitui uma tentativa ambiciosa, pioneira e necessária de regular a inteligência artificial no Brasil. Trata-se de uma das primeiras propostas legislativas no mundo a adotar um modelo normativo estruturado por risco e com forte orientação garantista. Fora esse projeto, destacam-se apenas a Recomendação da OCDE e o AI Act da União Europeia como referências comparáveis.

Elaborado por uma comissão de especialistas altamente qualificada, o texto demonstra sensibilidade tanto aos potenciais benéficos da IA quanto às suas armadilhas estruturais, especialmente os vieses algorítmicos e as ameaças à privacidade e à dignidade. Seu caráter abrangente é um mérito, não um defeito.

Se aprovado em sua forma atual, o Brasil se posicionará como protagonista na construção de uma IA ética, democrática e voltada ao interesse público na América Latina. Contudo, a efetividade da lei dependerá não apenas da qualidade da norma escrita, mas de sua implementação prática, da estrutura regulatória adequada, e da disponibilidade de recursos e capacitação técnica da autoridade supervisora.

A regulação da IA é um campo dinâmico, que exigirá constante atualização legislativa, mecanismos flexíveis de governança e, acima de tudo, uma escuta ativa da sociedade civil. Apenas com esse esforço conjunto será possível compatibilizar inovação tecnológica com a proteção efetiva dos direitos fundamentais.

CONSIDERAÇÕES FINAIS

O presente artigo buscou analisar criticamente os desafios jurídicos e constitucionais associados ao uso de sistemas de inteligência artificial no Brasil, especialmente no que tange à proteção dos direitos fundamentais. A discussão abordou de forma pormenorizada os riscos

decorrentes dos vieses algorítmicos, a opacidade decisória e os obstáculos impostos ao devido processo legal quando decisões são delegadas a sistemas automatizados sem mecanismos adequados de transparência, contestação ou supervisão.

A centralidade dos dados pessoais no funcionamento da IA exige atenção redobrada à sua proteção. O tratamento massivo e automatizado de dados — muitas vezes sensíveis — acarreta riscos reais à privacidade, à intimidade e ao devido processo legal, pilares do artigo 5º, incisos X e XII da Constituição Federal. A ausência de controles efetivos pode perpetuar discriminações históricas, afetar negativamente grupos vulneráveis e comprometer a confiança dos cidadãos nas instituições públicas e privadas.

A implementação de sistemas de inteligência artificial no Brasil e no mundo apresentam uma dualidade de potenciais benefícios e riscos significativos para os direitos fundamentais. À medida que a tecnologia avança e se torna mais integrada em setores cruciais, como o judiciário, a segurança pública, mercado de seguro e de crédito, dentre outras, é imprescindível que o uso da IA seja acompanhado de uma análise crítica e regulamentações eficazes.

No contexto brasileiro, o Projeto de Lei nº 2338/2023 representa um avanço importante ao propor diretrizes para o desenvolvimento e uso ético da inteligência artificial, com inspiração em marcos regulatórios internacionais como o AI Act da União Europeia. Entre seus méritos, destaca-se o reconhecimento do risco como critério para regulação, a exigência de supervisão humana nos sistemas de alto risco e a previsão de deveres de transparência por parte dos agentes desenvolvedores e operadores de IA.

O Projeto de Lei nº 2338/2023 representa um passo importante em direção à regulamentação da IA no Brasil, mas é evidente que, para sua eficácia, é necessário uma implementação robusta e a colaboração entre diversos *stakeholders*. A inclusão de vozes de diferentes setores, incluindo a sociedade civil, é fundamental para garantir que a regulação reflita uma diversidade de interesses e promova a equidade. Somente através de uma governança inclusiva e da criação de mecanismos de fiscalização adequados será possível mitigar os riscos associados ao uso da inteligência artificial e proteger os direitos fundamentais dos cidadãos.

Portanto, o futuro da inteligência artificial no Brasil deve ser construído sobre os pilares da transparência, responsabilidade e respeito a valores democráticos e às conquistas civilizatórias. Assim, a construção de um marco regulatório efetivo não apenas assegurará a justiça e a equidade na aplicação da IA, mas também fortalecerá a confiança da sociedade nas inovações tecnológicas, permitindo que o Brasil se posicione de maneira ética e responsável em um cenário global cada vez mais desafiador.

Dessa forma, conclui-se que a regulamentação da inteligência artificial no Brasil deve partir de uma perspectiva profundamente comprometida com a tutela dos direitos fundamentais, assegurando que a inovação tecnológica ocorra de forma ética, transparente, segura e inclusiva. A construção de um marco regulatório robusto não deve apenas incentivar o desenvolvimento econômico, mas também garantir que a IA seja uma ferramenta a serviço da dignidade humana, da justiça social e da promoção da igualdade — valores fundantes do Estado Democrático de Direito.

REFERÊNCIAS

AGÊNCIA SENADO. Senado aprova regulamentação da inteligência artificial; texto vai à Câmara, 10. dez. 2024. Disponível em: < https://www12.senado.leg.br/noticias/materias/2024/12/10/senado-aprova-regulamentacao-da-inteligencia-artificial-texto-vai-a-camara > acesso em 10.fev.2025

ALENCAR, Itana. Com mais de mil prisões na BA, sistema de reconhecimento facial é criticado por 'racismo algorítmico'; inocente ficou preso por 26 dias. **G1**. Disponível em: < https://g1.globo.com/ba/bahia/noticia/2023/09/01/com-mais-de-mil-prisoes-na-ba-sistema-de-reconhecimento-facial-e-criticado-por-racismo-algoritmico-inocente-ficou-preso-por-26-dias.ghtml > acesso em 12.mai.2024

ALMEIDA, André C. M. M.; Chang, luiza. Inteligência artificial: uma perspectiva da regulação no Brasil e na China. **JOTA.** 17. Jan 2025. Disponível em: < https://www.jota.info/artigos/inteligencia-artificial-uma-perspectiva-da-regulacao-no-brasil-e-na-china > acesso em 15.fev.2025.

BRASIL. Senado Federal. Projeto de Lei nº 2338, de 2023. Dispõe sobre o uso da inteligência artificial no Brasil. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233. Acesso em: 10 mar. 2025.

Buolamwini, Joy, and Timnit Gebru. **Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification**. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency,* 2018, pp. 77-91.

BUOLAMWINI, Joy. **AI, Ain't I A Woman?** Disponível em: < www.youtube.com/watch?v=QxuyfWoVV98 > acesso em 23.ago.2023.

CB INSIGHTS. 12 Startups Fighting Cancer With Artificial Intelligence. 2016. Disponível em: < https://www.cbinsights.com/research/ai-startups-fighting-cancer/ > Acesso em 14.Nov.2023.

CITIZEN ADVISE. Discriminatory pricing*: Exploring the 'ethnicity penalty' in the insurance market.* Disponível em: < https://www.citizensadvice.org.uk/about-us/our-work/policy/policy-research-topics/consumer-policy-research/consumer-policy-research/discriminatory-pricing-exploring-the-ethnicity-penalty-in-the-insurance-market1/ > acesso em 04.dez.2023

CONSOLIDATE CREDIT. **Revolver vs. Transactor: What Kind of Credit User are You?** Disponível em: < https://www.consolidatedcredit.org/financial-news/revolver-vs-transactor/ > acesso em 15.mar.2025

DA ROSA, Alexandre; GUASQUE, Bruno. O avanço da disrupção nos tribunais brasileiros. In: NUNES, Daniel; LUCON, Paulo Henrique dos Santos; NAVARRO WOLKART, Eduardo (Orgs.). **Inteligência artificial e direito processual:** os impactos da virada tecnológica no Direito Processual. Belo Horizonte: JusPODIVM, 2021. p. 93–121

DAMASCENO, Victoria; FERNANDES, Samuel. **Sob críticas por viés racial, reconhecimento facial chega a 20 estados**. Folha, 2021. Disponível em: < https://www1.folha.uol.com.br/cotidiano/2021/07/sob-criticas-por-vies-racial-reconhecimento-facial-chega-a-20-estados.shtml > acesso em 20.0ut.2022

DASTIN, J. (2018, 10 de outubro). Amazon scraps secret AI recruiting tool that showed bias against women. **Reuters**. Disponívem em: < https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G > acesso em 15.fev.2025.

FAIR BY DESIGN. The hidden risks of being poor: the poverty premium in insurance. Disponível em: < https://fairbydesign.com/wp-content/uploads/2021/09/IFoA_Hidden_Risks_of_Being_Poor_Aug_21_v09.pdf > acesso em 04.dez.2023

FERRAJOLI, Luigi. **Por uma Teoria dos Direitos e dos Bens Fundamentais.** Tradução de Alexandre Salim, Alfredo Copetti Neto, Daniela Cademartori, Hermes Zaneti Júnior, Sérgio Cademartori. Porto Alegre: Livraria do Advogado, 2011. (sem título original no exemplar utilizado), p.49).

FGV CONHECIMENTO– Centro de Inovação, Administração e Pesquisa do Judiciário. In: **Inteligência Artificial:** Tecnologia aplicada à gestão dos conflitos no âmbito do Poder Judiciário Brasileiro.

Disponível em: < https://ciapj.fgv.br/sites/ciapj.fgv.br/files/relatorio_ia_2fase.pdf > acesso em 15.dez.2023.

GESSER, Rafael. **Machine Learning e a precificação de seguros**. Disponível em: < https://oinsurance.com.br/machine-learning-precificacao-de-seguros/ > acesso em 21 Nov. 2023

INSTITUTO IGARAPÉ. **Crime Radar**. Disponível em: < https://igarape.org.br/tech/crimeradar/ > acesso em 29.out.2023.

LARSON, Jeff; MATTU, Surya; KIRCHNER, Lauren; ANGWIN, Julia. How we analyzed the COMPAS recidivism algorithm. **ProPublica,** [s.l.], 23 maio 2016. Disponível em: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm. Acesso em: 15 fev. 2025.

LORIDI, Luciano et al. AI4People – An ethical framework for a good AI society. Minds and Machines, Dordrecht, v. 28, p. 689–707, 2018. Disponível em: https://link.springer.com/article/10.1007/s11023-018-9482-5. Acesso em: 13 maio 2025.

MALAR, João Pedro**. Inteligência artificial usa posts nas redes sociais para fazer análise de crédito**. 2023. Disponível em: < https://exame.com/future-of-money/inteligencia-artificial-posts-redes-sociais-analise-credito/ > acesso em 21. nov. 2023

MARCÍLIO, Thiago. Quatro polêmicas sobre o projeto que regula a inteligência artificial. **Consultor Jurídico**, 29 fev. 2024. Disponível em: https://www.conjur.com.br/2024-fev-29/quatro-polemicas-sobre-o-projeto-que-regula-a-inteligencia-artificial/. Acesso em: 20 set. 2024.

MARDA, Vidushi; NARAYAN, Shivangi. Data in New Delhi's predictive policing system. **ACM Digital Library**. Disponível em: < https://dl.acm.org/doi/abs/10.1145/3351095.3372865 > acesso em 20. Out. 2023.

MEHRABI, Ninareh et al. A Survey on Bias and Fairness in Machine Learning, **arXiv**:1908.09635, v3, 25.jan.2022. Disponível em: < https://arxiv.org/abs/1908.09635 > acesso em 10.mai.2025

METLIFE. Disponível em: < https://www.metlife.com/us/lp/xcelerator-pt/ > acesso em 14.nov. 2023.

Ministério da Justiça e Segurança Pública. ANPD é formalizada como coordenadora do Sistema Nacional de Inteligência Artificial. 20.jun.2024. Disponível em: < https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-e-formalizada-como-coordenadora-do-sistema-nacional-de-inteligencia-artificial > acesso em 25.nov.2024.

NAVARRO, Susana Navas. Derecho e inteligencia artificial desde el diseño: aproximaciones. In: NAVARRO, Susana Navas (Coord.). **Inteligencia artificial:** tecnología, derecho. Valência: Tirant Lo Blanch, 2017. p. 23–72

NEENER ANALYTICS. Products. [s.l.], [s.d.]. Disponível em: https://www.neeneranalytics.com/products.html. Acesso em: 20 nov. 2023.

NEOWAY. Como os dados estruturados melhoram a gestão da saúde pública. Disponível em: < https://blog.neoway.com.br/podcasts/dados-estruturados/?lang=pt-br > acesso em 06.dez.2023.

NEOWAY. Precificação Inteligente: como o Big Data ajuda neste processo? Disponível em: < https://blog.neoway.com.br/precificacao-inteligente/ > acesso em 14. Nov. 2023.

NEUROTECH. Riskpack. [s.l.], [s.d.]. Disponível em: https://www.neurotech.com.br/riskpack/. Acesso em: 20 nov. 2023.

NIETHAMMER, Carmen. *AI Bias Could Put Women's Lives At Risk - A Challenge For Regulators*. **Forbes**, 2020. Disponível em: < https://www.forbes.com/sites/carmenniethammer/2020/03/02/ai-bias-could-put-womens-lives-at-riska-challenge-for-regulators/?sh=e5abaaf534f2 > acesso em 14.nov.2023

NIEVA FENOLL, Jordi. Inteligencia artificial y proceso judicial. Madri: Marcial Pons, 2018.

O'Neil, Cathy. **Algoritmos de destruição em massa**: como o Big Data aumenta a desigualdade e ameaça a democracia (Tradução Rachel Abraham). São Paulo: Editora Rua do Sabão,2020.

O'NEIL, Cathy. The authority of the inscrutable: an interview with Cathy O'Neil. *CCBLab*, [s.l.], 22 jan. 2019. Disponível em: https://lab.cccb.org/en/the-authority-of-the-inscrutable-an-interview-with-cathy-oneil/. Acesso em: 15 fev. 2025.

OBERMEYER, Ziad; POWERS, Brian; VOGELI, Christine; MULLAINATHAN, Sendhil. Dissecting racial bias in an algorithm used to manage the health of populations. **Science,** [s.l.], v. 366, n. 6464, 2019. Disponível em: https://www.science.org/doi/10.1126/science.aax2342. Acesso em: 14 nov. 2023.

OECD. Recommendation of the Council on Artificial Intelligence. 2019. Disponível em: < https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449 > acesso em 13.mai.2025

PANCH, Trishan; MATTIE, Heather; ATUN, Rifat. Artificial intelligence and algorithmic bias: implications for health systems. **Journal of Global Health**, [s.l.], v. 9, n. 2, 2019. Disponível em: https://doi.org/10.7189/jogh.09.020318 . Acesso em: 24 fev. 2025.

PORTAL DA INDÚSTRIA. PL da inteligência artificial prejudica desenvolvimento da tecnologia no Brasil e a inovação no setor produtivo. **Portal da Indústria**, [s.l.], 2024. Disponível em: https://noticias.portaldaindustria.com.br/posicionamentos/pl-da-inteligencia-artificial-prejudica-desenvolvimento-e-uso-da-tecnologia-no-brasil/. Acesso em: 20 set. 2024.

POSPIELOV, Serhii. How to reduce bias in machine learning. **Spiceworks**, [s.l.], 2022. Disponível em: https://www.spiceworks.com/tech/artificial-intelligence/guest-article/what-is-ml-bias-and-where-can-we-see-it/. Acesso em: 24 fev. 2025.

RODRÍGUEZ, Pedro. **Inteligencia artificial:** cómo cambiará el mundo (y tu vida). Barcelona: Ediciones Deusto, 2018.

SNOW, Jacob. Amazon's face recognition falsely matched 28 members of Congress with mugshots. **ACLU,** [s.l.], 26 jul. 2018. Disponível em: https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28. Acesso em: 10 ago. 2023.

UNIVERSO DO SEGURO. Inteligência artificial e análise de dados transformam o mercado de seguros, afirma country manager da Klimber. **Universo do Seguro**, [s.l.], 2023. Disponível em: https://universodoseguro.com.br/inteligencia-artificial-e-analise-de-dados-transformam-o-mercado-de-seguros-afirma-country-manager-da-klimber/. Acesso em: 14 nov. 2023.

UNZELTE, Carolina. Marco legal da IA: entenda os principais pontos do texto preliminar. **JOTA**, [s.l.], 2024. Disponível em: https://www.jota.info/legislativo/marco-legal-da-ia-entenda-os-principais-pontos-do-texto-preliminar. Acesso em: 20 set. 2024.

VEGA IRACELAY, José Javier. Inteligencia artificial y derecho: principios y propuestas para una gobernanza eficaz. **Informática y Derecho: Revista Iberoamericana de Derecho Informático,** [s.l.], n. 5, p. 13–48, 2018. Disponível em: https://dialnet.unirioja.es/servlet/articulo?codigo=6845781. Acesso em: 16 jun. 2024.

WACHOWICZ, Marcos. Regulação da inteligência artificial: garantindo um futuro ético e inclusivo no Brasil. *GEDAI – UFPR,* 2024. Disponível em: https://gedai.ufpr.br/regulacao-da-inteligencia-artificial-garantindo-um-futuro-etico-e-inclusivo-no-brasil/. Acesso em: 20 fev. 2025.