

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE DIREITO
GRADUAÇÃO EM DIREITO

Maria Teresa Torido Souto

Inteligência artificial e racismo algorítmico: análise do PL 2338/2023

Juiz de Fora
2025

Maria Teresa Torido Souto

Inteligência artificial e racismo algorítmico: análise do PL 2338/2023

Trabalho de conclusão de curso apresentado ao
Curso de Direito da Universidade Federal de
Juiz de Fora como requisito parcial à obtenção
do título de Bacharel em Direito.

Orientador: Prof. Dr. Sérgio Marcos Carvalho de Ávila Negri

Juiz de Fora
2025

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Torido Souto, Maria Teresa.

Inteligência artificial e racismo algorítmico: análise do PL 2338/2023 / Maria Teresa Torido Souto. -- 2025.
48 f.

Orientador: Sérgio Marcos Carvalho de Ávila Negri
Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Juiz de Fora, Faculdade de Direito, 2025.

1. Discriminação. 2. Inteligência artificial . 3. Projeto de Lei nº 2338/2023. 4. Racismo algorítmico. 5. Racismo estrutural. I. Carvalho de Ávila Negri, Sérgio Marcos, orient. II. Título.

Maria Teresa Torido Souto

Inteligência artificial e racismo algorítmico: análise do PL 2338/2023

Trabalho de Conclusão de curso apresentado
ao Curso de Direito da Universidade Federal
de Juiz de Fora como requisito parcial à
obtenção do título de Bacharel em Direito

Aprovada em 21 de fevereiro de 2025.

BANCA EXAMINADORA

Prof. Dr. Sérgio Marcos Carvalho de Ávila Negri - Orientador
Universidade Federal de Juiz de Fora

Carolina Braga Alvares
PPGD-UFJF

Natália Aparecida de Oliveira
PPGD-UFJF

Dedico este trabalho aos meus pais, cuja força e resiliência sempre me inspiraram a aproveitar as oportunidades que a vida não lhes proporcionou.

Pancrácio aceitou tudo; aceitou até um peteleco que lhe dei no dia seguinte, por me não escovar bem as botas; efeitos da liberdade. Mas eu expliquei-lhe que o peteleco, sendo um impulso natural, não podia anular o direito civil adquirido por um título que lhe dei. Ele continuava livre, eu de mau humor; eram dois estados naturais, quase divinos.

Crônica de Machado de Assis publicada no jornal Gazeta de Notícias, em 19 de maio de 1888.

RESUMO

O presente estudo teve como objetivo principal investigar a seguinte questão: O Projeto de Lei nº 2338, de 2023, é efetivo para a prevenção de possíveis consequências de racismo algorítmico desencadeado pelo racismo estrutural no Brasil? Para isso, foi adotada uma abordagem exploratória, utilizando técnicas como revisão bibliográfica, análise documental e estudo de caso. A pesquisa explorou inicialmente a relação entre racismo estrutural e o uso de inteligência artificial (IA) no Brasil, com ênfase no racismo algorítmico. Referenciais teóricos, como Silvio Almeida, foram empregados para contextualizar o racismo estrutural, enquanto estudos de Tarcízio Silva tratam do racismo algorítmico no contexto brasileiro. Foi realizada uma análise detalhada do Projeto de Lei nº 2338/2023, com o objetivo de avaliar como ele aborda as questões éticas, sociais e legais relacionadas ao uso da inteligência artificial, considerando o histórico de desigualdades estruturais no Brasil e identificando possíveis lacunas na proposta legislativa, especialmente no que tange à proteção contra vieses algorítmicos. Além disso, foi conduzido um estudo para verificar a efetividade do projeto na prevenção de práticas racistas e discriminatórias.

Palavras-chave: discriminação; inteligência artificial; Projeto de Lei nº 2338/2023; racismo algorítmico; racismo estrutural.

ABSTRACT

The primary objective of this study was to investigate the following question: Is Bill nº 2338, of 2023, effective in preventing potential consequences of algorithmic racism triggered by structural racism in Brazil? To this end, an exploratory approach was adopted, employing techniques such as bibliographic review, document analysis, and case study. The research initially explored the relationship between structural racism and the use of artificial intelligence (AI) in Brazil, with a focus on algorithmic racism. Theoretical frameworks, such as those by Silvio Almeida, were utilized to contextualize structural racism, while studies by Tarcízio Silva address algorithm racism in the Brazilian context. A detailed analysis of Bill nº 2338/2023 was conducted to evaluate how it addresses ethical, social, and legal issues related to the use of artificial intelligence, taking into account Brazil's history of structural inequalities and identifying potential gaps in the legislative proposal, particularly concerning protection against algorithmic biases. Furthermore, a study was carried out to assess the effectiveness of the bill in preventing racist and discriminatory practices.

Keywords: algorithmic racism; artificial intelligence; Bill n. 2338/2023; discrimination; structural racism.

LISTA DE ABREVIATURAS E SIGLAS

ABL	Academia Brasileira de Letras
CDC	Código de Defesa do Consumidor
IA	Inteligência Artificial
IBGE	Instituto Brasileiro de Geografia e Estatística
LGPD	Lei Geral de Proteção de Dados
NIST	National Institute of Standards and Technology
PL	Projeto de Lei

SUMÁRIO

1	INTRODUÇÃO	10
2	O USO DE DADOS PARA SISTEMAS DE INTELIGÊNCIA ARTIFICIAL.....	13
3	BREVE CONTEXTO DO RACISMO ESTRUTURAL NO BRASIL.....	16
3.1	RACISMO E DIREITO FUNDAMENTAL À IGUALDADE.....	17
4	ANÁLISE SOBRE O RACISMO ALGORÍTMICO.....	20
5	REGULAMENTAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO BRASIL NO PL 2338/2023.....	24
5.1	AVANÇOS, RISCOS E PARTICIPAÇÃO.....	24
5.2	DISCUSSÕES SOBRE RAÇA E DISCRIMINAÇÃO.....	27
5.3	RESPONSABILIDADE CIVIL: DIRETRIZES E IMPLICAÇÕES JURÍDICAS.....	32
6	CONCLUSÃO	38
	REFERÊNCIAS	44

1 INTRODUÇÃO

O episódio “Queda Livre”, da série *Black Mirror*, aborda uma sociedade distópica onde a vida das pessoas é controlada por um sistema de classificação social alimentado por algoritmos. Nessa realidade, cada interação com outros indivíduos é avaliada, sendo que essas notas atribuídas determinam o acesso a oportunidades e a privilégios. No episódio, Lacie, a personagem principal, tenta, a todo custo, aumentar sua nota para ter acesso a uma melhor moradia, mas diversos eventos ruins acontecem e sua nota acaba, na verdade, diminuindo de forma drástica, fazendo com que ela termine excluída e marginalizada.

Como ocorre no episódio, os algoritmos já são usados para, por exemplo, avaliar a concessão de crédito, para contratar pessoas e para o policiamento nas cidades e, por isso, a história da série, apesar de não se dar exatamente nos mesmos moldes na sociedade atual, retrata como sistemas de inteligência artificial podem reforçar padrões discriminatórios existentes. A narrativa traz a reflexão acerca do problema de usar sistemas de IA (inteligência artificial) de forma irrestrita e expõe o perigo de depender de tecnologias, aparentemente neutras, que carregam vieses a depender dos criadores e dos dados com os quais são alimentadas, resultando em injustiças sistêmicas.

As tecnologias emergentes têm influenciado profundamente as dinâmicas sociais e as interações entre grupos, sendo um fator crucial na configuração do cenário global atual. Exemplos disso incluem o uso de assistentes virtuais, como Siri e Alexa, para realizar tarefas cotidianas e para controlar dispositivos domésticos inteligentes, além dos algoritmos de recomendação, como os usados pela Netflix, que sugerem conteúdos com base no comportamento e nas preferências dos usuários. Além do mais, a inteligência artificial tem sido aplicada em áreas diversas, como para diagnósticos médicos e para uso de reconhecimento facial e de sistemas de predição.

A utilização de sistemas de IA na vida cotidiana, em uma primeira análise, que pode parecer trazer apenas vantagens por auxiliar em tarefas, ao ser visto de forma mais cuidadosa, revela uma série de problemas que surgem de conflitos já existentes nas relações humanas. Tendo em vista as tecnologias que ampliaram o fluxo de informações e as possibilidades de liberdade de expressão, bem como a inclusão de ferramentas que fortaleceram a privacidade por meio do tratamento de dados pessoais, há uma intersecção latente do diálogo entre inovação tecnológica e proteção da personalidade, que se tornou tema central no campo jurídico.

No passado, os sistemas de inteligência artificial eram analisados de forma isolada, desvinculados das pessoas e de seus direitos, sendo compreendidos apenas como ferramentas quantitativas, já que poderiam executar atividades para além da capacidade humana, por meio da rápida e da eficiente execução de cálculos (Doneda *et al.*, 2018). No entanto, tornou-se evidente que essa abordagem é insuficiente diante das complexidades da sociedade contemporânea, especialmente quando se consideram os dilemas éticos e os desafios que o tratamento de dados pode gerar em uma sociedade historicamente marcada por desigualdades.

Diante disso, o fato de sistemas de IA operarem com base em algoritmos complexos e frequentemente incompreensíveis para os consumidores em geral, traz desafios para o Direito. Isso é um problema ainda maior quando esses sistemas aplicam-se em áreas que fazem uso de dados sensíveis, como, por exemplo, raça e etnia. A ausência de mecanismos claros de fiscalização e de controle social, bem como de diversidade no setor tecnológico aprofunda a primazia de determinado grupo em detrimento de uma minoria, sempre sub-representada, o que compromete, flagrantemente, a equidade e a justiça das soluções baseadas em IA.

Levando-se em consideração essas questões, está em processo de tramitação o Projeto de Lei nº 2338, de 2023 - Marco Legal da Inteligência Artificial - o qual busca proteger direitos fundamentais a fim de que a IA seja usada de forma mais segura e que seja criado um ambiente propício à inovação. Isto posto, o presente trabalho¹ objetiva responder o seguinte questionamento: O Projeto de Lei nº 2338, de 2023, é efetivo para a prevenção de possíveis consequências de racismo algorítmico desencadeado pelo racismo estrutural no Brasil? A priori, a hipótese é de que o projeto não fez uma análise criteriosa acerca dessa temática, o que poderá gerar conflitos a serem regulados a partir de casos concretos. Justifica-se, portanto, a elaboração deste trabalho com o propósito de refletir sobre as precauções do Marco Legal da IA, que, embora bem-intencionado, aparenta ser insuficiente para abordar adequadamente as questões raciais relacionadas ao aprendizado de máquina. Assim, destaca-se a necessidade de manter o debate sobre esse tema em evidência.

Dito isso, a metodologia adotada neste estudo é de natureza exploratória, refletindo a dinâmica dos conceitos em constante disputa. Para embasar a pesquisa e para desenvolver as inferências necessárias para confirmar ou para refutar a hipótese apresentada, adotou-se o método técnico de análise textual-discursiva. Além disso, foram empregadas técnicas de revisão bibliográfica, análise documental e estudo da legislação, estruturando o trabalho com

¹ Este trabalho de conclusão de curso foi feito por uma aluna bolsista do CNPq na pesquisa "Inovação e Direito na Inteligência Artificial: mapeamento normativo e análise de impacto para exercício de direitos fundamentais", coordenado pelo Prof. Dr. Sérgio Marcos Carvalho de Ávila Negri.

uma metodologia jurídico-dogmática de abordagem dedutiva. A revisão bibliográfica envolveu a análise de produções acadêmicas recentes sobre racismo algorítmico, discriminação e o Projeto de Lei nº 2338/2023, enquanto a análise documental centrou-se no exame detalhado do próprio texto do PL.

O estudo é composto por um tópico sobre o uso de dados no sistemas de inteligência artificial em geral, devido à presença de dados relacionados a questões raciais. Desse modo, busca-se, também, explorar as complexidades éticas e legais que surgem quando a inteligência artificial absorve, gera e perpetua vieses discriminatórios e minora a equidade racial. Posteriormente, propõe-se uma breve análise do racismo estrutural no Brasil, tendo como referencial teórico principal o estudioso Silvio Almeida. Em seguida, foi analisado o conceito de racismo algorítmico e suas implicações no desenvolvimento tecnológico contemporâneo. Para embasar essa análise, foram apresentados exemplos retirados da bibliografia estrangeira, porém com ênfase no contexto brasileiro, fundamentados nos estudos de Tarcízio Silva.

Na etapa seguinte, realizou-se uma análise detalhada do Projeto de Lei nº 2338, de 2023, com o objetivo de examinar até que ponto seu texto aborda questões relacionadas ao racismo e aos vieses presentes nos sistemas de inteligência artificial. Essa análise considera o contexto histórico e social do Brasil, marcado por profundas desigualdades estruturais e, por essa razão, as discussões raciais, embora frequentemente presentes na formulação de políticas públicas, nem sempre recebem a devida atenção em regulamentações tecnológicas, o que reforça a necessidade de um olhar crítico sobre o Marco Legal da IA. Dito isso, a análise busca identificar possíveis lacunas no projeto de lei, bem como propor reflexões sobre como a regulação pode ser aprimorada para enfrentar as implicações éticas e sociais dos vieses algorítmicos em um país com um histórico de discriminação sistêmica.

Por fim, o trabalho é concluído com a apresentação das considerações finais, que sintetizam as principais reflexões obtidas ao longo da pesquisa. Essa seção busca consolidar os argumentos desenvolvidos, reforçando a relevância do tema abordado e destacando os desafios ainda existentes no contexto da regulamentação da inteligência artificial e do racismo algorítmico. Em seguida, são apresentadas as referências bibliográficas, que incluem obras, artigos, vídeos e documentos consultados, os quais serviram como base teórica e prática para fundamentar as análises realizadas neste estudo.

2 O USO DE DADOS PARA SISTEMAS DE INTELIGÊNCIA ARTIFICIAL

A inteligência artificial não é recente e já foi constatada em diversos momentos históricos. A ideia de que os seres humanos seriam capazes de desenvolver máquinas capazes de pensar por elas próprias sempre esteve presente em filmes e em livros e, a partir disso, o desenvolvimento tecnológico avançou no sentido de criar sistemas cada vez mais sofisticados, possíveis de processar informações, de aprender com os dados e de executar tarefas de forma autônoma, transformando significativamente diversos setores da sociedade.

Sob essa perspectiva, cita-se o teste de Alan Turing que, proposto na década de 1950, buscava avaliar a capacidade de uma máquina de enganar uma pessoa ao se passar por um ser humano durante uma interação textual. Apesar de hoje existirem novas abordagens mais modernas de avaliação, o Teste de Turing ainda é usado, inclusive em competições, como em 2014, ano no qual um *chatbot* enganou o juiz em 30% das vezes durante cinco minutos de conversação pelo teclado (Onody, 2021).

Alguns anos após a criação do Teste de Turing, o termo “inteligência artificial” foi concebido como técnica que possibilita que computadores imitem a inteligência humana e a aprendizagem de máquina como um método para permitir que computadores melhorem a execução de tarefas quando adquirem mais experiência. Nesse cenário, vale analisar como o *big data* e o aprendizado de máquina funcionam e como essa tecnologia está presente nos diversos sistemas, bem como seus impactos.

O aprendizado de máquina permite que o sistema execute a mesma tarefa de forma progressivamente mais eficiente a cada tentativa, o que resulta em uma melhoria contínua e automática de sua performance. Há diversos tipos de aprendizado de máquina e, entre eles, destaca-se o *deep learning*, um método que permite que o *software* treine a si mesmo para realizar tarefas, como a classificação de imagens e o reconhecimento da fala, por meio das redes neurais, as quais são expostas a uma grande quantidade de dados, com o propósito de simular o poder de tomada de decisão do cérebro humano (Raub, 2018).

O *software* de redes neurais é modelado com base no funcionamento das redes de neurônios do cérebro humano, em vez de depender de um conjunto inflexível de instruções criadas anteriormente, como são moldados outros sistemas. Essas redes neurais utilizam *big data* — grandes conjuntos de dados coletados — para analisar e para revelar padrões, sendo certo que a grande quantidade de dados disponibilizada devido aos avanços da internet e ao

poder computacional é essencial para alimentar essas redes, permitindo o processo de *deep learning*.

Levando-se em consideração o exposto, para que o aprendizado de máquina mencionado ocorra, existe o processo conhecido como mineração de dados, o qual analisa dados gerados na *internet* e identifica padrões interpretáveis (Raub, 2018), criando modelos, sendo esse procedimento responsável por recomendações personalizadas nos diversos aplicativos, por exemplo. Nesse aspecto, a escolha de quais dados serão responsáveis por “alimentar” os sistemas traz consequências relevantes, afinal, a inteligência artificial é impossível sem os algoritmos, que podem ser entendidos como opiniões dentro de um código (Mann; O’neil, 2016).

Para desenvolver o modelo, os algoritmos são treinados para reagir de forma específica aos dados que lhe são fornecidos e, nesse processo, as decisões subjetivas acabam sendo incorporadas ao algoritmo por meio dos dados utilizados e dos rótulos atribuídos a determinadas características. Nesse processo, os rótulos subjetivos criados a partir de dados empregados influenciam os resultados e, portanto, geram vieses. Por essa razão, argumenta-se que algoritmos e sistemas de inteligência artificial podem perpetuar preconceitos já existentes e reproduzir padrões discriminatórios contemporâneos (Barocas; Selbst, 2016).

É certo que o uso de IA pode trazer diversas vantagens para o cotidiano, principalmente no caso da automatização de tarefas repetitivas ou até mesmo para recomendar músicas do interesse do ouvinte. Porém, todo esse avanço não está isento de consequências negativas quando esses sistemas são usados para resolver problemas sociais, pois o descuido na seleção de dados, bem como falhas do algoritmo podem desencadear em resultados discriminatórios, perpetuando divisões socioeconômicas pré-existentes (Rainie; Anderson, 2017).

Um exemplo que retrata essa discussão são os bancos de dados de reconhecimento facial, como o *Special Database 32*, criado pelo National Institute of Standards and Technology (NIST). Esse banco de dados reúne imagens de fichas criminais para desenvolver e para testar algoritmos de IA usados, não só em reconhecimento facial, como também em outras aplicações (Crawford, 2021). Imagens originalmente coletadas por sistemas policiais, destinadas à identificação de suspeitos, são usadas com fins comerciais e acadêmicos e, portanto, as pessoas ou suas famílias não têm qualquer controle sobre o uso de suas imagens. Essa análise demonstra que, no setor tecnológico, qualquer dado pode, aparentemente, ser considerado acessível, sem levar em conta sua origem e desconsiderando as implicações éticas e políticas desse processo (Crawford, 2021).

Em uma análise rápida, pode-se pensar que o mero uso dessas imagens não causaria danos aos indivíduos, pois somente estariam sendo utilizadas para treinar sistemas. No entanto, uma observação crítica revela clara violação ética e jurídica aos direitos das pessoas envolvidas. No exemplo citado, ao serem retiradas de fichas criminais, as imagens retratam indivíduos em situações de extrema vulnerabilidade, onde sua privacidade e dignidade já se encontram fragilizadas e, nessa ótica, é razoável presumir que essas pessoas, submetidas a um contexto já estigmatizante, não consentiram e provavelmente não gostariam que suas imagens fossem expostas ou reutilizadas em qualquer meio, especialmente sem transparência ou controle. Tal prática, ao ignorar o princípio ético do respeito à autonomia e à dignidade humana, acaba perpetuando um ciclo de invisibilidade e desproteção daqueles que já se encontram em posições mais vulneráveis socialmente.

Não obstante interpretar imagens seja uma tarefa bastante difícil, já que fotos podem ter diversos significados e contradições, além de que os sistemas não têm capacidade para analisar fatores mais sofisticados, como o impacto social de uma foto, é prática comum coletar imagens da internet, classificá-las e usá-las como base em um sistema de IA. Essas vastas coleções são usadas para definir conjuntos de dados que serão definidos como a “verdade”, deixando de ser uma representação objetiva ou uma realidade consensual e tornando-se uma combinação de imagens extraídas de diversas fontes disponíveis na internet (Crawford, 2021). Com base nisso, torna-se fácil perceber a importância de contar com uma base de dados heterogênea, que objetive refletir a realidade de forma mais fiel possível, considerando as diferenças entre os indivíduos.

O uso de programas de aprendizado de máquina aparentemente autônomos em seus respectivos domínios de aplicação é um aspecto central para compreender a disseminação de sistemas algorítmicos, seus riscos e as estratégias utilizadas pelas organizações para evitar a responsabilização pelos impactos causados por tal tecnologia. É comum empresas alegarem que os *softwares* são segredos de negócio com o intuito de inviabilizar sua auditoria ou análise crítica em relação aos impactos gerados. Também, é alegado frequentemente que a complexidade técnica dos algoritmos torna-os praticamente indecifráveis para a sociedade, especialmente considerando a dificuldade de interpretar individualmente milhões de linhas de código, os quais interagem com diversos pacotes, módulos e bases de dados (Silva, 2022).

3 BREVE CONTEXTO DO RACISMO ESTRUTURAL NO BRASIL

A escravidão foi elemento central na estrutura social e econômica do Brasil colonial e não apenas supriu a mão de obra necessária para o desenvolvimento das atividades agrárias, como também moldou a mentalidade coletiva em relação ao trabalho. A escravização de pessoas com fundamento em questões raciais criou obstáculos, perpetuados até a atualidade, para o desenvolvimento de uma identidade coletiva e de instituições sociais mais justas (Holanda, 1995).

A relação entre escravidão e racismo pode ser explicada de duas formas principais. A primeira considera que o racismo é um legado da escravidão e do colonialismo, com sociedades contemporâneas mantendo padrões mentais e institucionais racistas, autoritários e violentos, sendo, portanto, um entrave à modernização econômica e à democracia, sobretudo em países como o Brasil. A segunda corrente argumenta que o racismo não é um vestígio do passado, mas, sim, um componente intrínseco do capitalismo e da modernidade, ambos forjados pela escravidão (Almeida, 2019).

Dessa forma, a desigualdade racial estaria integrada às estruturas mercantis e de classe, adaptando-se às novas demandas da acumulação capitalista e, nessa ótica, o racismo seria renovado pelo capitalismo, funcionando como uma forma de controle social, com o propósito de justificar salários mais baixos para minorias e dissuadindo trabalhadores brancos de reivindicarem melhores condições, sob a ameaça de serem substituídos por uma mão de obra mais barata e disponível. Nessa perspectiva, o desenvolvimento de técnicas de exploração econômica é acompanhado pela evolução de técnicas de violência e de opressão, como o racismo (Almeida, 2019).

Esse contexto histórico revela como indivíduos foram categorizados e organizados em grupos específicos com o objetivo de estabelecer mecanismos lucrativos de controle social. Nessa toada, legitimado pela Igreja Católica (Vasconcelos, 2013), o sistema escravocrata foi extremamente vantajoso para o enriquecimento de colonizadores e de ricos senhores de terras, pois é evidente o lucro obtido não só com a compra e venda de seres humanos, mas também com o fruto de seu penoso trabalho. Porém, no Brasil, como o processo de miscigenação é latente, o mito da democracia racial perdurou-se por muitos anos, não obstante a clara marginalização de pretos e de pardos após o processo de abolição, facilmente visualizado por

diversos fatores, como, por exemplo, o surgimento de comunidades sem estrutura de moradia e a falta de representatividade nas instâncias de poder.

Com o passar dos anos após a abolição, não obstante a miscigenação ter continuado, o preconceito não foi erradicado e até hoje é possível encontrar dados alarmantes, como o fato de que, mesmo a maior parte da população sendo composta por negros e pardos (Belandi; Gomes, 2022), essas pessoas ainda são sub-representadas na política e no ensino superior, por exemplo. A partir da clara constatação de que o Brasil consolidou-se como um país racista desde sua fundação como Estado Nação devido à política da metrópole colonial, um projeto nacional a fim de desenvolver a sociedade deve enfrentar o racismo no campo prático e simbólico (Almeida, 2019).

Entender que o racismo transcende o âmbito individual e desencadeia no funcionamento das instituições, sendo tratado como um elemento constitutivo da sociedade, é primordial para a análise de que o poder estabelece as relações raciais (Almeida, 2019). As instituições sociais, não obstante não serem responsáveis por criar o racismo, privilegiam certos grupos em detrimento de outros, pois o racismo está intrínseco na sociedade e, assim, as relações reproduzem a violência, seja explícita ou implícita, ao transmitir privilégios e desigualdades, que desencadeiam em microagressões² (Silva, 2022). Cita-se como exemplo a suposição de que indivíduos racializados entraram em alguma universidade ou cargo somente devido às cotas ou, também, a ideia de que essas pessoas seriam mais suscetíveis a serem perigosas a depender de sua cor (Silva, 2022).

Em um contexto extremamente racista, seria impossível que toda essa problemática não se fizesse evidente no meio tecnológico, ambiente no qual as violações a direitos humanos podem ser muito graves, já que as microagressões são reproduzidas de forma algorítmica em bases de dados, refletindo os vieses sociais existentes. Tendo em vista que o racismo está intrínseco na organização social, o combate a essa estrutura para erradicar desigualdades exige uma postura ativa na implementação de práticas antirracistas, implicando no desenvolvimento de políticas que promovam a diversidade, o acolhimento e a revisão constante de medidas, com base em debates que visem mudanças estruturais efetivas (Silva, 2019).

3.1 RACISMO E DIREITO FUNDAMENTAL À IGUALDADE

² Para Tarcizio Silva (2022), o uso do termo “micro” não se refere ao grau de violência, mas, sim, à perversidade e ao fato de que a agressão incide em um nível individual e privado, que permitam anonimato ou a evasão com o pretexto de que seria uma piada.

A igualdade é elemento extremamente caro aos princípios democráticos e, por isso, é resguardada em incontáveis dispositivos no texto constitucional, como, por exemplo, no art. 5º, incisos I, XLI e XLII³, sendo que a doutrina mais moderna considera que existem três dimensões da igualdade. A igualdade formal serve para proteger pessoas contra privilégios e discriminação, enquanto a igualdade material, refere-se à redistribuição de poder e de oportunidade e, por fim, a igualdade como reconhecimento implica no respeito às minorias e às suas particularidades, como a raça (Barroso, 2023).

Considerando o contexto envolvendo a discriminação racial sofrida por pessoas negras com a colonização e mencionando o fato de o Brasil ter sido o país ocidental que mais recebeu pessoas escravizadas, bem como o último a acabar com a escravidão no continente americano (Barroso, 2023), os reflexos disso não poderiam ser questões simples de lidar. A escravização de pessoas com base em características físicas gerou uma realidade de intensa exclusão social que desencadeou um processo desumanizante perpetuado em estruturas sociais que ensejaram um modelo de sociedade excludente, no qual o racismo estrutural, resultado dessa histórica exclusão, continua a operar nas instituições, reproduzindo e ampliando as disparidades sociais, o que impede a construção de uma sociedade verdadeiramente igualitária (Almeida, 2019).

Nesse contexto, não obstante, inicialmente, o princípio da igualdade ser fundamentado na ideia de que todos são iguais, e, por isso, acreditava-se que qualquer indivíduo deveria ser tratado da mesma forma pela lei independentemente de suas circunstâncias pessoais, como ficou nítido que a igualdade formal não afastava injustiças sociais, atribuiu-se um sentido material à igualdade (Scarlet; Marinoni; Mitidiero, 2024), entendendo-se que não bastava eliminar os privilégios, era necessário agir de maneira efetiva para combater a desigualdade econômica e promover a superação de assimetrias (Barroso, 2023).

Como a desigualdade é latente, do ponto de vista jurídico, surgiram as políticas afirmativas para oferecer vantagens competitivas a grupos sociais que foram historicamente privados de oportunidades devido ao racismo estrutural. Assim, as cotas raciais para ingresso em universidades e o incentivo a processos seletivos que prezam por um time diverso em

³ Art. 5º Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade, nos termos seguintes:

I - homens e mulheres são iguais em direitos e obrigações, nos termos desta Constituição;

XLI - **a lei punirá qualquer discriminação atentatória dos direitos e liberdades fundamentais;**

XLII - **a prática do racismo constitui crime inafiançável e imprescritível**, sujeito à pena de reclusão, nos termos da lei (Brasil, 1988, grifo próprio).

empresas privadas são medidas para combater o racismo estrutural que permeia as instituições.

Pensar em igualdade é essencial para iniciar um debate acerca do racismo algorítmico, pois a aplicação de IA pode amplificar e perpetuar desigualdades da sociedade. Caso os avanços tecnológicos e o uso de dados não sejam regulados cuidadosamente, grupos marginalizados podem sofrer ainda mais com as mazelas de um preconceito infundado. Portanto, um debate eficaz sobre racismo algorítmico deve estar centrado na promoção de estratégias que proporcionem que tecnologias não se tornem instrumentos de opressão, mas, sim, ferramentas em consonância com a igualdade material, valor objetivado pela Constituição Federal de 1988.

4 ANÁLISE SOBRE O RACISMO ALGORÍTMICO

Como analisado, o racismo não se trata apenas de um ato discriminatório, mas, de um processo no qual privilégios e desvantagens permeiam grupos em diversos âmbitos sociais, como, na política, na economia e nas relações cotidianas por meio de microagressões, estabelecendo-se, também, nos diversos avanços tecnológicos. Sistemas de IA dependem de grandes volumes de dados para operar, mas esses dados frequentemente refletem desigualdades preexistentes na sociedade e, então, algoritmos treinados com informações já enviesadas tendem a reforçar discriminações.

De acordo com o que foi explorado, as imagens e os dados usados são classificados com o objetivo de serem organizados e analisados pelos sistemas de IA, em uma prática de classificar pessoas que remonta à era dos impérios. Durante as conquistas, os indivíduos eram categorizados e posteriormente organizados em “tipos de povo” pelos ditos especialistas (Hacking, 2007), em atos de nomeação que representavam demonstrações de poder e mecanismos de controle colonial, cujos impactos negativos persistiram em diversos momentos históricos. Essas classificações funcionam como ferramentas que moldam e limitam as formas de conhecimento e estão incorporadas às estruturas da IA (Crawford, 2021). Além disso, o racismo, que sempre serviu como ferramenta a serviço do capitalismo, reflete-se na lógica dos dados, que agora também assumem o papel de capital. Como as atividades humanas podem ser representadas por rastros digitais e, conseqüentemente, classificadas, esses dados tornam-se um meio estratégico de extração de valor, reforçando dinâmicas econômicas e sociais existentes (Crawford, 2021).

Como os algoritmos estão cada vez mais avançados e são aplicados em atividades que demandam a interpretação de universos complexos, como significados, motivações, aspirações, crenças, valores e atitudes, os vieses podem ter impactos sociais relevantes. De forma direta e simplificada, algoritmos são sequências estruturadas de procedimentos organizados de maneira lógica, destinadas a executar tarefas específicas em um ambiente computacional. Os sistemas automatizados expandiram por diversas áreas da vida humana e incorporaram lógicas algorítmicas ao utilizar inteligência artificial em processos que já estavam em transformação devido à digitalização, como no caso da segurança pública, e, assim, vários dados podem ser matéria-prima no capitalismo informacional (Silva, 2022).

A gestão algorítmica de decisões envolvendo indivíduos e grupos tem se fortalecido à medida que os sistemas computacionais são utilizados para maximizar lucros para empresas de tecnologia. Esses sistemas tomam decisões tanto para e sobre os indivíduos e seus impactos podem variar, abrangendo efeitos capazes de moldar comportamentos e atitudes de forma quase imperceptível e, por conseguinte, reproduzindo e consolidando estruturas de poder e opressão (Silva, 2022). Esse contexto revela um dos principais desafios associados à lógica do aprendizado de máquina: a reprodução de injustiças sociais previamente estruturadas, o que desencadeia o que hoje é entendido como racismo algorítmico.

O aprendizado de máquina permite que sistemas computacionais melhorem continuamente suas tarefas com base em grandes volumes de dados, pois redes neurais do aprendizado de máquina simulam o funcionamento do cérebro humano. Entretanto, essas redes refletem os vieses presentes nos dados utilizados para treiná-las, o que pode evidenciar o contexto discriminatório existente. Um caso famoso sobre a temática é o robô Tay, criado pela Microsoft e lançado no Twitter em março de 2016 como um chatbot. Tendo sido projetado para aprender e se comunicar de forma natural com os usuários da plataforma, a ideia era que Tay evoluísse com base nas interações, porém, em menos de um dia, o robô começou a postar mensagens ofensivas e racistas, pois o sistema foi alvo de usuários que enviaram mensagens preconceituosas. Ele incorporou rapidamente esse conteúdo e passou a reproduzi-lo, expondo a falta de salvaguardas no sistema e, diante das repercussões negativas, a empresa tirou o robô do ar.

Esse robô evidencia como um sistema pode apresentar falhas e, rapidamente, produzir vieses discriminatórios, que podem ser concebidos de forma acrítica por quem produz ou treina os sistemas. Isso porque locais onde estão os mais famosos responsáveis por criarem sistemas de IA, como o Vale do Silício, são flagrantemente marcados pela branquitude, o que revela desigualdades estruturais nos mercados e ambientes tecnológicos de ponta. Esses espaços homogêneos geram impactos materiais e simbólicos nos sistemas que influenciam a vida de grande parte da população mundial (Silva, 2022), desconsiderando as peculiaridades e as características da população de cada país, principalmente uma população diversificada e marcada por anos de desigualdade, como é a brasileira.

Sob essa perspectiva, torna-se evidente o fato de que a tecnologia não é neutra. Prova disso são dados de 2020 que revelaram a baixa representatividade de pessoas negras nas principais empresas do setor: apenas 2,5% dos colaboradores do Google, 4% dos profissionais do Facebook e da Microsoft, e 6% dos funcionários do Twitter eram negros. Na Apple, 9% da

força de trabalho era composta por pessoas negras, considerando tanto os setores tecnológicos quanto o setor de varejo. Já na Amazon, esse percentual foi de 26,5% (Vitório, 2020).

Sendo que os dados usados já têm um recorte racial ou discriminatório devido aos próprios preconceitos arraigados na sociedade, não obstante o sistema não ter sido planejado para agir de forma discriminatória, é bem provável que replique preconceitos das próprias bases de dados. A título de exemplo, cita-se o *Special Database 32* - já mencionado - cujo objetivo foi treinar algoritmos de reconhecimento facial, mas possuía uma representação desproporcional de indivíduos brancos em comparação com pessoas negras, o que resultou em falhas.

Essa disparidade tem consequências graves, pois algoritmos treinados em bases de dados predominantemente compostas por rostos brancos apresentam taxas significativamente mais altas de erros ao identificar ou classificar rostos de pessoas negras. A falta de diversidade nos dados utilizados faz com que o código seja enviesado e, desse modo, o racismo estrutural da sociedade perpassa camadas e transforma-se em racismo algorítmico.

O racismo algorítmico pode ser compreendido como uma extensão do racismo sistêmico, em que as tecnologias e os imaginários sociotécnicos, em um mundo marcado pela supremacia branca, realizam classificações racializadas. Essas práticas afetam negativamente grupos minorizados, exacerbando a distribuição desigual de recursos e a violência, bem como moldando dinâmicas de poder futuras, dificultando ainda mais a identificação e o enfrentamento das opressões globais, e, por conseguinte, aprofundando as heranças coloniais de exploração e de injustiça (SILVA, 2022).

Recentemente, a Academia Brasileira de Letras (ABL) divulgou um avatar do escritor Machado de Assis feito por IA para recriar esse importante personagem da história do Brasil e permitir que haja interação com os visitantes. Não obstante tenha sido uma proposta interessante e moderna, o totem embranqueceu Machado de Assis e, então, foi alvo de críticas (Villarinho, 2024). Isso retrata uma triste herança colonial e uma clara tentativa de apagamento da identidade do autor, uma vez que ele foi um homem não branco, filho de escravo, cuja trajetória se deu no enfrentamento das estruturas sociais discriminatórias. Essa reinterpretação digital reforça um imaginário branco e homogêneo, perpetuando a distorção da história e a invisibilidade da contribuição da população negra para a cultura e para a história brasileira.

Há, também, casos em que *softwares* de reconhecimento falharam em detectar rostos de pessoas negras devido à ausência de diversidade nos conjuntos de dados usados para treiná-lo, que eram compostos majoritariamente por rostos brancos (How [...], 2017). Apesar

de ser flagrantemente perigoso para a segurança pública, podendo gerar prisões injustas, qualquer tipo de preconceito é extremamente nocivo ao Estado Democrático de Direito, pois compromete princípios essenciais como igualdade, dignidade e não discriminação, afetando o acesso das pessoas vítimas de racismo algorítmico a direitos básicos, como educação, saúde, emprego e justiça, como já mencionado.

A partir dessa análise, destaca-se que o uso de algoritmos não é revestido de neutralidade⁴, como alegado, pois eles incorporam vieses e preconceitos contidos nas bases de dados. Caso a sociedade desenvolva uma confiança excessiva nos algoritmos ou adote uma postura acrítica em relação a seu uso, as decisões discriminatórias desses sistemas podem ser percebidas como mais legítimas e justificadas do que aquelas feitas por seres humanos. Por isso, destaca-se a relevância do Projeto de Lei nº 2338/2023, o qual visa definir diretrizes para o uso ético e responsável da inteligência artificial, estabelecendo princípios destinados a garantir que os sistemas sejam desenvolvidos e aplicados com respeito aos direitos fundamentais, devido ao seu objetivo de prevenir decisões discriminatórias e de promover maior responsabilidade e equidade nos processos tecnológicos.

⁴ Silvio Almeida na Comissão de Juristas: “Quando se fala de racismo, quando se fala de discriminação em geral, é impossível separar o que chamamos de racismo ou de discriminação de tecnologia” (Relatório [...], 2022, p. 417).

5 REGULAMENTAÇÃO DA INTELIGÊNCIA ARTIFICIAL NO BRASIL NO PL 2338/2023

As discussões para implementação de um projeto de lei que regulamenta o uso de inteligência artificial não são recentes. Em 2020, criou-se o Projeto de Lei nº 21/2020 e, devido ao seu caráter urgente, no mesmo ano houve a discussão e a aprovação pela Câmara dos Deputados. No entanto, o projeto sofreu diversas críticas e, em 2022, foi instaurada uma Comissão de Juristas para a realização de audiências públicas para incentivar o debate acerca do tema, o que ensejou o anteprojeto que, posteriormente, transformou-se no Projeto de Lei nº 2338/2023, cujo objetivo é estabelecer normas gerais para o uso responsável de sistemas de IA.

No relatório final da Comissão de Juristas responsável por subsidiar elaboração de substitutivo sobre a IA no Brasil, disponibilizado pela Coordenação de Comissões Especiais, Temporárias e Parlamentares de Inquérito, é perceptível que o tema da discriminação foi pauta constante trazida por todos os estudiosos. Bojana Bellamy (Relatório [...], 2022, p. 185), por exemplo, declarou que com o intuito de se ter um algoritmo não enviesado, durante o processo, é necessário o tratamento de dados sensíveis⁵, pois esses estão diretamente relacionados às situações discriminatórias (Relatório [...], 2022, p. 629).

A proposta de legislação evidencia intensa preocupação com os dados e, apesar de estar ainda em um patamar de discussões, já traz questões importantes. Como o treinamento de sistemas de IA é feito a partir de dados, os responsáveis pela preparação do sistema de IA devem resguardar princípios da LGPD, pois, caso a base de dados não esteja completamente alinhado com a referida lei (IA no Direito [...], 2024), os sistemas podem apresentar vieses discriminatórios.

5.1 AVANÇOS, RISCOS E PARTICIPAÇÃO

⁵ LGPD, art. 5º Para os fins desta Lei, considera-se:

I - dado pessoal: informação relacionada a pessoa natural identificada ou identificável;

II - dado pessoal sensível: dado pessoal sobre origem racial ou étnica, convicção religiosa, opinião política, filiação a sindicato ou a organização de caráter religioso, filosófico ou político, dado referente à saúde ou à vida sexual, dado genético ou biométrico, quando vinculado a uma pessoa natural; III - dado anonimizado: dado relativo a titular que não possa ser identificado, considerando a utilização de meios técnicos razoáveis e disponíveis na ocasião de seu tratamento [...].

A norma proposta explicitou que pessoas afetadas por sistemas de IA têm “direito à privacidade e à proteção de dados pessoais” (art. 5º, VI), bem como têm direito de receber informações claras e adequadas sobre as “categorias de dados pessoais utilizados no contexto do funcionamento do sistema de inteligência artificial” (art. 7º, V). No que tange aos dados pessoais sensíveis, menciona-se, no art. 12, I, que pessoas afetadas por decisões, previsões ou recomendações de sistemas de IA têm direito a tratamento justo e isonômico, sendo vedado o uso de IA que possa acarretar discriminação, inclusive “em decorrência do uso de dados pessoais sensíveis ou de impactos desproporcionais em razão de características pessoais como origem geográfica, raça, cor ou etnia, gênero, orientação sexual, classe socioeconômica, idade, deficiência, religião ou opiniões políticas”. Assim, todo o escopo do projeto é pautado pela proteção à pessoa e aos dados que podem causar questões discriminatórias, inclusive no art. 42, §3º, há previsão expressa de que a atividade de mineração de textos e dados que envolva dados pessoais está sujeita à LGPD.

Em análise do relatório final feita pela Comissão⁶, em 5 de dezembro de 2024, foram destacados os princípios da integridade da informação e a proteção aos direitos autorais pelo Senador Randolfe Rodrigues pois, apesar de serem constitucionalmente resguardados, estão sendo negligenciados pelos avanços da IA, pois é necessário resguardar que as informações disponibilizadas pelos sistemas sejam íntegras. Ademais, a senadora Zenaide Maia ressaltou a necessidade de centralidade da pessoa humana, pois muitas pessoas já são vítimas de recursos de IA, inclusive relata uma experiência própria de *deepfake*⁷, na qual foi usado seu rosto e sua fala em um vídeo falso para fazer pedidos a prefeitos. Após breves comentários, houve uma votação simbólica positiva para envio do relatório para o Plenário.

O PL 2338/2023 enfrenta o tema da IA sob a ótica dos direitos e do risco, propondo a proteção da pessoa natural impactada por sistemas de IA e a regulamentação da governança, o que foi evidenciado em várias partes do documento. Isso foi feito no intuito de equilibrar a proteção aos direitos e às liberdades fundamentais, a valorização do trabalho e a dignidade humana com o fortalecimento da ordem econômica (Bioni; Garrote; Guedes, 2023).

Para isso, foi adotado um modelo baseado em risco a fim de possibilitar a imputação dos responsáveis, pois, embora haja direitos aplicáveis a quaisquer casos de interação entre

⁶ Comissão (...), 2024.

⁷ O termo *deepfake* denomina uma técnica que consiste na criação de conteúdos sintéticos (não reais), que podem ser áudios e imagens, produzidos com auxílio de inteligência artificial (IA). Na prática, são mídias artificiais geradas a partir de uma grande quantidade de arquivos reais de determinada pessoa e com uso de um algoritmo de aprendizado de máquina - *machine learning* (Projeto Comprova, 2023).

pessoa e máquina, há também uma delimitação específica para casos de diferentes níveis de risco, o que justifica a criação de medidas de governança adaptadas a cada situação. Essa abordagem regulatória resultou na designação de uma autoridade competente (art. 32, caput), que deve colaborar com outras entidades de competências relacionadas para analisar e controlar os riscos (art. 32, incisos V, VII e VIII) (Bioni; Garrote; Guedes, 2023).

A abordagem permite uma regulação mais eficiente, concentrando esforços onde há maior potencial de danos, bem como reforçando a importância de que usuários tenham os meios necessários para compreender e para proteger seus direitos de forma prática e acessível. Também, a criação de uma autoridade reguladora especializada é especialmente relevante, considerando o aumento exponencial no uso de sistemas de IA em diversas áreas, que podem gerar afronta direta aos direitos fundamentais. Essa autoridade poderá centralizar a análise e a gestão de riscos, atuando como um ponto de referência para solucionar conflitos e estabelecer políticas claras de governança.

No PL, categoriza-se os riscos como excessivo (artigo 14), alto risco (arts. 17 e 18) e moderado/baixo, sendo que este último está implícito na legislação (Bioni; Garrote; Guedes, 2023). Dessa forma, a classificação de riscos determina diferentes níveis de exigências regulatórias, ajustados ao grau de risco identificado para aqueles que desenvolvem a tecnologia e a disponibilizam no mercado ou a colocam em uso, sendo que as medidas de regulação aplicadas serão maiores quanto maiores forem os riscos associados ao uso dos sistemas.

Conforme dispõe o art. 13, a categorização dos riscos se dá por uma avaliação preliminar de impacto algorítmico conduzida pelo órgão competente, possivelmente representado pelo Conselho Nacional de Justiça (Mendonça Júnior; Nunes, 2023), antes de ser colocado no mercado, a qual definirá em qual tipo grupo será encaixado - sempre podendo ser reclassificado (art.13, §3º). No projeto, como em outras legislações presentes em diversos ordenamentos jurídicos, a classificação é realizada por meio de uma lista exemplificativa de sistemas de IA e tem como objetivo principal proteger a pessoa natural, regulamentar a governança desses sistemas e, simultaneamente, assegurar o avanço do desenvolvimento tecnológico.

A avaliação de impacto algorítmico tem como objetivo identificar e analisar os benefícios, as preocupações e os riscos relacionados aos sistemas de IA, além de implementar estratégias para prevenir, mitigar e reparar possíveis danos, bem como monitorar os efeitos ao longo do tempo (Bioni; Garrote; Guedes, 2023). Tendo em vista a potencialidade de sistemas

de IA gerarem restrições a direitos humanos, avaliar o impacto é uma tarefa extremamente importante e não deve estar concentrada a um seletivo grupo.

A possibilidade de as partes interessadas participarem na avaliação de impacto das tecnologias, evidenciada no PL no art. 24, §3º e no art. 25, §2º, oferece diversos benefícios tanto para o processo de avaliação quanto para seus resultados, como, por exemplo, contribui para aprimorar a qualidade, a credibilidade e a legitimidade das análises realizadas. Essa abordagem evita que as decisões se baseiem exclusivamente em análises restritas a um grupo restrito, especialmente internas às organizações, reduzindo o risco de enviesamentos ou discriminações. Em vez disso, promove-se a inclusão de agentes diversos, particularmente daqueles potencialmente impactados pela adoção da tecnologia, garantindo uma perspectiva mais ampla e justa (Bioni; Garrote; Guedes, 2023).

A participação de pessoas diretamente impactadas por esses sistemas faz-se ainda mais evidente no caso de sistemas de inteligência artificial classificados como de alto risco, especialmente aqueles com potencial para gerar situações discriminatórias, os quais devem passar por um processo mais rigoroso e contínuo de avaliação. Esse modelo participativo, além de fortalecer a legitimidade das normas, contribui para uma compreensão mais profunda dos efeitos desses sistemas na prática, promovendo uma regulação inclusiva.

Nessa toada, inspirando-se em propostas de outros países, o PL 2338 também prevê a criação de uma base de dados brasileira⁸, reunindo tanto a documentação de autoavaliação dos sistemas quanto as avaliações de impacto relacionadas aos sistemas de alto risco. Essa prática tem o objetivo de promover a transparência e de reduzir a assimetria de informações para os indivíduos potencialmente impactados pelos sistemas de IA, bem como propõe ajudar os próprios fornecedores, os quais podem utilizar a base como uma referência para identificar as melhores práticas na elaboração de relatórios de impacto (Bioni; Garrote; Guedes, 2023).

5.2 DISCUSSÕES SOBRE RAÇA E DISCRIMINAÇÃO

Com o uso de sistemas de IA, há preocupações sobre a exacerbação da desigualdade e do fortalecimento de estigmas e de discursos discriminatórios historicamente vividos por grupos marginalizados (Negri *et al.*, 2023). Isso ocorre, pois, além de as tecnologias serem criadas com base em estruturas de poder estabelecidas, o funcionamento dos algoritmos

⁸ Art. 43. Cabe à autoridade competente a criação e manutenção de base de dados de inteligência artificial de alto risco, acessível ao público, que contenha os documentos públicos das avaliações de impacto, respeitados os segredos comercial e industrial, nos termos do regulamento (Brasil, 2023).

podem facilitar manifestações de racismo, tendo em vista que a circulação e o engajamento de conteúdo discriminatório desencadeiam em faturamento para empresas donas de plataforma (Silva, 2022).

Uma questão importante para ser analisada é que grande parte das ferramentas utilizadas em países como o Brasil foram desenvolvidas por empresas da Europa e dos Estados Unidos e, por esse motivo, aspectos culturais específicos do Brasil não são considerados no desenvolvimento da tecnologia (Cortiz, 2022). Nesse contexto, o uso de sistemas de IA deve ser pautado pelos direitos humanos e deve refletir as diferenças sociais na sociedade brasileira para evitar que grupos marginalizados sofram mais discriminação, com aprofundamento do contexto de vulnerabilidade e de violência institucional ao qual são submetidos (Negri *et al.*, 2023).

O PL é centrado na pessoa e no respeito aos direitos humanos e, ante o exposto acerca da perpetuação do racismo estrutural e de suas implicações nos sistemas de IA, apesar de o projeto de lei não fazer expressa menção ao racismo algorítmico, busca uma estratégia que combina a avaliação de riscos com uma estrutura regulatória fundamentada em direitos (Negri *et al.*, 2023). Assim, o princípio da proteção, com foco nos direitos fundamentais, está explicitamente exposto em diversas partes no projeto de legislação.

Além da clara intenção de proteção aos direitos humanos, conceitos de discriminação, raça e etnia estão presentes em diversos artigos do PL e, logo no art. 2º, já explicita que a igualdade e a não discriminação são fundamentos do desenvolvimento, da implementação e do uso de sistemas de inteligência artificial no Brasil. Posteriormente, no art. 4º é feita a conceituação do significado de discriminação - citando raça⁹ - e de discriminação indireta¹⁰.

Em seguida, no artigo 5º, inciso V, é dito que as pessoas afetadas por sistemas de inteligência artificial têm “direito à não-discriminação e à correção de vieses discriminatórios diretos, indiretos, ilegais ou abusivos” (Brasil, 2023). Ademais, dispõe o art. 7º, VI, que pessoas afetadas por IA têm o direito de receber, previamente à contratação ou à utilização do sistema, informações claras e adequadas quanto às “medidas de segurança, de

⁹ Art. 4º Para as finalidades desta Lei, adotam-se as seguintes definições: VI – **discriminação**: qualquer **distinção, exclusão, restrição ou preferência**, em qualquer área da vida pública ou privada, cujo propósito ou efeito seja **anular ou restringir o reconhecimento**, gozo ou exercício, em condições de igualdade, de um ou mais direitos ou liberdades previstos no ordenamento jurídico, em razão de características pessoais como origem geográfica, **raça, cor ou etnia**, gênero, orientação sexual, classe socioeconômica, idade, deficiência, religião ou opiniões políticas (Brasil, 2023, grifo próprio).

¹⁰ Art. 4º Para as finalidades desta Lei, adotam-se as seguintes definições: VII – **discriminação indireta**: discriminação que ocorre quando normativa, prática ou critério **aparentemente neutro** tem a capacidade de **acarretar desvantagem para pessoas pertencentes a grupo específico**, ou as coloquem em desvantagem, a menos que essa normativa, prática ou critério tenha algum objetivo ou justificativa razoável e legítima à luz do **direito à igualdade e dos demais direitos fundamentais** (Brasil, 2023, grifo próprio).

não-discriminação e de confiabilidade adotadas, incluindo acurácia, precisão e cobertura” (Brasil, 2023).

Também, o art. 12 propõe inibir, diretamente, a implementação e o uso de sistemas de IA que possam acarretar discriminação em seus diversos tipos¹¹ e explicita - no parágrafo único desse mesmo artigo - que, não obstante seja vedada a implementação e o uso de sistemas de inteligência artificial que possam acarretar discriminação direta, indireta, ilegal ou abusiva, são permitidos critérios de diferenciação fundamentados em justificativas legítimas, como ações afirmativas, como as cotas raciais, as quais referem-se à ideia de igualdade material¹², conforme previsão constitucional.

Outrossim, o art. 8º do PL¹³ visa assegurar que indivíduos afetados por atos praticados pela IA solicitem, gratuitamente, explicações acerca da decisão, da previsão ou da recomendação feita pelo sistema, a fim de descobrir procedimentos e critérios adotados na tomada de decisão. Ainda, os artigos 9º e 10º complementam essas garantias ao permitir que o usuário conteste decisões tomadas pela IA e, nos casos em que as ações do sistema resultem em efeitos jurídicos significativos ou impactem de forma relevante seus interesses, podem solicitar intervenção ou revisão por um ser humano.

¹¹ Art. 12. As pessoas afetadas por decisões, previsões ou recomendações de sistemas de inteligência artificial têm direito a tratamento justo e isonômico, **sendo vedadas a implementação e o uso de sistemas de inteligência artificial que possam acarretar discriminação direta, indireta, ilegal ou abusiva**, inclusive:

I – em decorrência do uso de dados pessoais sensíveis ou de impactos desproporcionais em razão de características pessoais como origem geográfica, **raça, cor ou etnia**, gênero, orientação sexual, classe socioeconômica, idade, deficiência, religião ou opiniões políticas; ou

II – em função do estabelecimento de **desvantagens ou agravamento da situação de vulnerabilidade** de pessoas pertencentes a um **grupo específico**, ainda que se utilizem critérios aparentemente neutros.

Parágrafo único. A vedação prevista no caput não impede a adoção de critérios de diferenciação entre indivíduos ou grupos quando tal diferenciação se dê em função de objetivos ou justificativas demonstradas, razoáveis e legítimas **à luz do direito à igualdade e dos demais direitos fundamentais** (Brasil, 2023, grifo próprio).

¹² Art. 5º Todos são iguais perante a lei, sem distinção de qualquer natureza, garantindo-se aos brasileiros e aos estrangeiros residentes no País a inviolabilidade do direito à vida, à liberdade, à igualdade, à segurança e à propriedade (Brasil, 1988).

¹³ Art. 8º **A pessoa afetada por sistema de inteligência artificial poderá solicitar explicação sobre a decisão, previsão ou recomendação, com informações a respeito dos critérios e dos procedimentos utilizados**, assim como sobre os principais fatores que afetam tal previsão ou decisão específica, incluindo informações sobre:

I – a racionalidade e a lógica do sistema, o significado e as consequências previstas de tal decisão para a pessoa afetada;

II – o grau e o nível de contribuição do sistema de inteligência artificial para a tomada de decisões;

III – os dados processados e a sua fonte, **os critérios para a tomada de decisão** e, quando apropriado, a sua ponderação, aplicados à situação da pessoa afetada; IV – os mecanismos por meio dos quais a pessoa pode contestar a decisão; e

V – a possibilidade de solicitar intervenção humana, nos termos desta Lei.

Parágrafo único. As informações mencionadas no caput serão fornecidas por **procedimento gratuito e facilitado**, em linguagem que permita que a pessoa compreenda o resultado da decisão ou previsão em questão, no prazo de até quinze dias a contar da solicitação, permitida a prorrogação, uma vez, por igual período, a depender da complexidade do caso (Brasil, 2023, grifo próprio).

Percebe-se, pois, que o Projeto de Lei nº 2338/2023 configura um bom início das iniciativas de regulamentação e de controle da implementação de IA no Brasil. No entanto, apesar de abordar explicitamente e de forma reiterada a discriminação algorítmica em função de cor, raça e etnia, é nítido que a proposta legal ainda apresenta lacunas, especialmente no que se refere à garantia e fiscalização de suas determinações, bem como às consequências jurídicas em caso de violações.

O PL 2338/2023, devido às discussões durante a Comissão de Juristas, trouxe avanços significativos acerca da temática e deixou explícito a existência de desigualdades e de assimetrias no Brasil. Ao reconhecer as desigualdades estruturais no Brasil pela incorporação das definições de discriminação direta e indireta da Convenção Interamericana contra o Racismo, a qual tem força constitucional desde 2022, é nítido que o PL fortalece a proteção contra discriminações, o que é almejado por meio de diferentes mecanismos, como o direito à informação e à compreensão, o direito à contestação, bem como o direito específico de correção de vieses discriminatórios, além da implementação de medidas preventivas de governança.

Ainda, o texto também enfatiza a proteção de grupos vulneráveis, considerando a discriminação como critério para classificar sistemas de IA de alto risco (art. 18, c) e para avaliar o impacto de sistemas que possam afetar esses grupos (art. 24, §1º, f), o que mostra que é mais afirmativo e protetivo de direitos fundamentais (Bioni; Garrote; Guedes, 2023) caros à sociedade brasileira, sociedade essa marcada por desigualdades. Também, no art. 21¹⁴,

¹⁴ Art. 21. Adicionalmente às medidas de governança estabelecidas neste capítulo, órgãos e entidades do poder público da União, Estados, Distrito Federal e Municípios, ao contratar, desenvolver ou utilizar **sistemas de inteligência artificial considerados de alto risco, adotarão as seguintes medidas:**

I – **realização de consulta e audiência públicas prévias** sobre a utilização planejada dos sistemas de inteligência artificial, com informações sobre os dados a serem utilizados, a lógica geral de funcionamento e resultados de testes realizados;

II – definição de protocolos de acesso e de utilização do sistema que permitam o registro de quem o utilizou, **para qual situação concreta, e com qual finalidade;**

III – utilização de dados provenientes de fontes seguras, que sejam exatas, relevantes, atualizadas e representativas das populações afetadas e testadas contra vieses discriminatórios, em conformidade com a Lei no 13.709, de 14 de agosto de 2018, e seus atos regulamentares;

IV – garantia facilitada e efetiva ao cidadão, perante o poder público, de **direito à explicação e revisão humanas de decisão por sistemas de inteligência artificial que gerem efeitos jurídicos relevantes ou que impactem significativamente os interesses do afetado**, a ser realizada pelo agente público competente;

V – utilização de interface de programação de aplicativos que permita sua utilização por outros sistemas para fins de interoperabilidade, na forma da regulamentação; e

VI – publicização em veículos de fácil acesso, preferencialmente em seus sítios eletrônicos, das avaliações preliminares dos sistemas de inteligência artificial desenvolvidos, implementados ou utilizados pelo poder público da União, Estados, Distrito Federal e Municípios, independentemente do grau de risco, sem prejuízo do disposto no art. 43.

§1º A utilização de sistemas biométricos pelo poder público da União, Estados, Distrito Federal e Municípios será precedida da edição de ato normativo que estabeleça garantias para o exercício dos direitos da pessoa afetada e proteção **contra a discriminação direta, indireta, ilegal ou abusiva**, vedado o tratamento de dados de raça, cor ou etnia, salvo previsão expressa em Lei.

propõe-se medidas adicionais de governança para órgãos públicos que utilizam ou desenvolvem sistemas de IA de alto risco com o intuito de reduzir desigualdades estruturais e de promover maior transparência e equilíbrio entre Estado e sociedade.

Para a efetiva aplicação das normas previstas e para formulação de futuras regulamentações que tratam da questão com maior detalhamento, é essencial considerar pluralidade de raças em todos os aspectos nos processos tecnológicos para superar obstáculos criados pelo racismo estrutural. Com a intenção de que uma abordagem pautada por práticas de transparência funcione, riscos altamente graves, como os que se relacionam com o racismo algorítmico, devem ser prontamente eliminados em qualquer projeto (Negri *et al.*, 2023).

À luz disso, a análise de fatores sociais específicos da realidade brasileira é primordial, uma vez que, como discutido anteriormente, certas aplicações de inteligência artificial podem afetar de forma mais intensa grupos e comunidades específicas. Diante desse contexto, devem ser incorporados os interesses coletivos na avaliação de impacto desses sistemas, tendo em vista que muitos modelos de negócios utilizam dados pessoais para classificação e para o monitoramento de determinados grupos, o que evidencia a tensão constante entre os direitos individuais e as dinâmicas do mercado (Negri *et al.*, 2023).

O PL 2338/23 é, claro, um avanço no que tange à proteção a direitos e ao combate à discriminação algorítmica, mas o texto ainda mantém uma abordagem defensiva ao introduzir mecanismos de governança destinados a proteger contra possíveis resultados ilegítimos ou ilegais que possam ser gerados por sistemas de IA (Bioni; Garrote; Guedes, 2023). O projeto não dispõe explicitamente sobre propostas como, por exemplo, a produção de bancos de dados “éticos, diversos, abertos e multidisciplinares” com o propósito de atingir a “decolonização no design e aplicação de tecnologias de IA”, incluindo todos os níveis de educação, desenvolvendo currículos diversos com a participação de atores diversos para respeitar a herança cultural de diversos povos originários (Silva, 2023). O Brasil, por ser um país extremamente diverso, deve refletir isso no desenvolvimento e no uso de seus sistemas e, sob essa perspectiva, as leis ainda têm muito o que avançar para se adequar ao cenário social, econômico e político brasileiro, tendo em vista ser um país que sofreu bastante com o processo de colonização e ainda sobre com os efeitos do racismo estrutural.

É essencial que o Brasil desenvolva uma regulamentação de IA que esteja alinhada com os modelos internacionais, mas que leve em conta as especificidades nacionais. O PL

§2º Na impossibilidade de eliminação ou mitigação substantiva dos riscos associados ao sistema de inteligência artificial identificados na avaliação de impacto algorítmico prevista no artigo 22 desta Lei, **sua utilização será descontinuada** (Brasil, 2023, grifo próprio).

2338/2023 iniciou essa tutela para proteção dos direitos a partir de princípios como a igualdade, a não-discriminação, a justiça, a inclusão e a explicabilidade, porém não menciona o racismo algorítmico, apesar de possuir caráter protetivo e tentar considerar características brasileiras, como o racismo estrutural, para evitar que a IA intensifique desigualdades e contribua para o progresso social (Bioni; Garrote; Guedes, 2023).

5.3 RESPONSABILIDADE CIVIL: DIRETRIZES E IMPLICAÇÕES JURÍDICAS NO PL 2338/2023

A temática da responsabilidade é abordada em todas as legislações que regulamentam o uso da inteligência artificial e, como foi bastante discutida na Comissão de Juristas, naturalmente, o Projeto de Lei nº 2338/2023 também precisava contemplar esse aspecto. No art. 3º, IX, do PL, é elucidado que um dos princípios do desenvolvimento, da implementação e do uso de sistemas de IA é a rastreabilidade das decisões em todo o ciclo de vida dos sistemas como meio de viabilizar a prestação de contas e a atribuição de responsabilidade a uma pessoa natural ou jurídica.

Posteriormente, o capítulo V é inteiramente dedicado à responsabilidade civil e, com a leitura do art. 27, entende-se que o fornecedor ou o operador de sistema de IA que causar dano patrimonial, moral, individual ou coletivo é obrigado a repará-lo integralmente, não importando o grau de autonomia do sistema. No parágrafo primeiro do mesmo artigo, consta que no caso de ser um sistema de alto risco ou de risco excessivo, o fornecedor ou o operador responde objetivamente pelos danos causados, na medida de sua participação.

Em relação aos sistemas de alto risco, a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima. No art. 29, é dito que as hipóteses de responsabilização civil decorrentes dos danos causados por essas tecnologias no âmbito das relações de consumo estarão, inclusive, sujeitas às regras do Código de Defesa do Consumidor (CDC), o qual institui a responsabilidade objetiva do fornecedor de produtos ou serviços e, assim, cria um sistema de responsabilização sem que se considere o fator subjetivo da culpa (Schreiber, 2024).

Durante as discussões na Comissão, foi bastante discutido sobre qual deveria ser o regime de responsabilidades adotado e a responsabilidade subjetiva adotada no PL nº 21, de 2020 foi bastante criticada por estudiosos¹⁵. Assim, as respostas foram no sentido de adotar a

¹⁵ Nelson Roselvald: “(...) a expressão responsabilidade subjetiva, apesar de ser tradicional, é indesejável. Ela traz uma insegurança jurídica, porque qual responsabilidade subjetiva é essa de que estamos falando?” (Relatório [...], 2022, p. 95)

responsabilidade civil objetiva, conforme consta acima.

No Brasil, a regra é a de que, quem, por ato ilícito, for responsável pelo dano, fica obrigado a repará-lo (Brasil, 2002), sendo que, no parágrafo único do art. 927, consta a cláusula geral de responsabilidade objetiva por atividades de risco, que propõe “a obrigação de reparar o dano, independentemente de culpa, nos casos especificados em lei, ou quando a atividade normalmente desenvolvida pelo autor do dano implicar, por sua natureza, risco para os direitos de outrem” (Brasil, 2002), percebendo-se, portanto, que o legislador tinha a intenção de dispor sobre atividades que tragam algum risco elevado, provável, com perigo de dano (Schreiber, 2024).

Considerando-se que um modelo de responsabilidade subjetiva não é apropriado para lidar com as nuances do desenvolvimento de IA e que o seu uso pode resultar em riscos imensuráveis, é certo que também não é suficiente para contemplar a responsabilização em questões relacionadas à discriminação algorítmica. Nessa linha de pensamento, alguns estudiosos apontam a responsabilidade pela guarda de animais ou objetos¹⁶ (Brasil, 2002) como base para a aplicação da responsabilidade objetiva à inteligência artificial devido ao fato de que os programas, assim como os animais, são imprevisíveis, bem como o fato de as tecnologias estarem sob custódia de algum indivíduo, apesar de existirem críticas a essa ideia por causa da existência de robôs bastante autônomos (Costa, 2023).

Assim, a teoria mais aceita é a responsabilidade objetiva em razão da atividade, sendo

Também Nelson Roselvald: “(...) o PL nº 21, de 2020, no art. 6º, inciso VI, quando traz a responsabilidade subjetiva para os agentes que atuam nessa cadeia, infelizmente trouxe uma grande simplificação que contradiz a própria complexidade do que se quer regular.” (Relatório [...], 2022, p. 405)

Caitlin Mulholland: “se nós pensarmos que a responsabilidade subjetiva seria a preferencial, como previsto no projeto de lei de forma bastante equivocada, nós **traríamos um problema evidente na responsabilidade civil, que é o fato de que a vítima do dano seria obrigada a identificar a conduta culposa, ainda que objetivamente considerada, do agente desenvolvedor da IA, algo que seria demasiadamente gravoso ou pesado para essa vítima (...)**” (Relatório [...], 2022, p. 396, grifos próprios).

Anderson Schreiber: “(...) **não faz sentido que um projeto de lei sobre IA** contenha um dispositivo que prevê que as normas sobre responsabilidade civil dos agentes que atuam no desenvolvimento e operação dos sistemas de IA **devam se pautar pela responsabilidade civil subjetiva.**” (Relatório [...], 2022, p. 392, grifos próprios).

Iris: “**As previsões no PL nº 21, de 2020, sobre o regime de responsabilidade são, a um só tempo, inconsistentes e nocivas.** Traçam diretrizes para outras normas legais futuras, preconizando regra própria de responsabilidade subjetiva e direta, essa mesma a ser mitigada pela eventual adoção de “esforços razoáveis” e “melhores práticas de mercado **O ordenamento já prevê a atribuição de responsabilidade objetiva,** a fim de garantir uma pronta indenização compensatória à vítima, com posterior faculdade de exercício de direito de regresso, quando se pode avaliar a culpa e até mesmo a redistribuição do ônus financeiro aos integrantes da cadeia de produção de produtos ou prestação de serviços. **Essa lógica poderia facilmente ser adaptada à complexidade do mundo da inteligência artificial.**” (Relatório [...], 2022, p. 729).

Anderson Schreiber explicitou que “a legislação consumerista também adota o regime de responsabilidade objetiva, embora com algumas peculiaridades, de maneira que a **introdução de uma nova tecnologia, como é a IA, sem dúvida alguma, se enquadraria** dentro desse conjunto de hipóteses e atrairia, dentro do nosso sistema jurídico, **a responsabilidade civil objetiva.**” (Relatório [...], 2022, p. 93, grifos próprios).

¹⁶ Código Civil, art. 936. O dono, ou detentor, do animal ressarcirá o dano por este causado, se não provar culpa da vítima ou força maior (Brasil, 2002).

que o PL explicitou claramente isso, quando, no art. 28, propôs hipóteses nas quais os agentes não serão responsabilizados, quais sejam, quando comprovarem que não colocaram em circulação, empregaram ou tiraram proveito do sistema ou que o dano é decorrente de fato exclusivo da vítima ou de terceiro, assim como de caso fortuito externo. Ou seja, tanto a empresa que desenvolveu o *software* quanto a corporação que o utiliza seriam responsáveis por eventuais danos, justamente devido ao risco trazido pela atividade.

Sendo o dano conceituado “como a lesão a um interesse juridicamente protegido, a abranger tanto o dano patrimonial quanto o dano moral” (Schreiber, 2024. p.629), uma possível consequência da discriminação algorítmica pode desencadear prejuízos na esfera patrimonial dos indivíduos quando sistemas automatizados discriminam grupos racializados em decisões que envolvem o aspecto financeiro, por exemplo, na concessão de crédito com taxas abusivas ou na filtragem de candidatos em processos seletivos. Já no âmbito moral, é nítido que há afronta à dignidade, causando sofrimento emocional e violando direitos fundamentais.

Diante dessas considerações, o PL está de acordo com a teoria do risco objetivo e deixou de forma explícita que a obrigação de reparar independe do grau de autonomia da IA. Essa proposta de regime de responsabilidade foi acertada principalmente no caso de tecnologias de alto risco ou de risco excessivo - que é o caso de programas que podem causar discriminação - cuja responsabilização se dará de forma objetiva na medida de sua participação no dano, apesar de não haver menção expressa de possível responsabilidade solidária, tendo em vista a presença de diversos atores no uso e IA.

Tendo isso em vista, além das soluções voltadas à conscientização dos atores sociais envolvidos na criação dos sistemas, é essencial a implementação de ferramentas que atuem diretamente nos dados utilizados por esses sistemas. Esse processo é viabilizado por meio da *accountability*, que se configura como um aspecto fundamental para a prevenção da discriminação algorítmica e está alinhado com mecanismos de governança dos algoritmos (Costa, 2023), os quais podem ser definidos como um conjunto de estratégias e ferramentas destinadas a minimizar os impactos negativos da inteligência artificial, garantindo sua eficiência e mitigando possíveis consequências indesejadas (Doneda; Almeida, 2016).

O conceito de *accountability* não foi explicitamente mencionado no PL, mas, sendo uma expressão que remete à responsabilidade com ética, transparência e, claro, à prestação de contas (Costa, 2023), foi tratado em várias passagens do projeto e está condizente com o propósito de não-discriminação. Para que a responsabilização efetivamente ocorra, aliado à

revisão que deve ser feita pela autoridade competente¹⁷, o desenvolvimento econômico da iniciativa privada precisa ser guiada por padrões que considerem o interesse público e a governança (Doneda; Almeida, 2016).

Sendo o propósito da governança a redução dos problemas causados pelos algoritmos, busca-se a preservação da sua eficácia com a redução de resultados indesejáveis (Doneda; Almeida, 2016), como o racismo. No art. 19 do PL nº 2338/2023, é explicitado que os agentes de IA estabelecerão estruturas de governança e processos que serão aplicados a todo o ciclo de vida da tecnologia, bem como devem ser aptos a resguardar os direitos das pessoas afetadas.

No caso de sistemas de alto risco, a lei é ainda mais explícita em seu art. 20¹⁸, dispondo sobre a necessidade de se ter documentação acerca do processo de desenvolvimento e, também, ressaltando medidas de gestão de dados para mitigar e para prevenir vieses discriminatórios por meio da composição de equipe diversa e da avaliação dos dados com foco em evitar a criação de vieses decorrentes de erros ou da ausência de dados, além de possíveis lacunas no que tange à representatividade. Ademais, é mencionada a adoção de medidas para impedir que preconceitos sociais estruturais sejam incorporados ou ampliados pelas tecnologias.

No parágrafo único do mesmo artigo, é tratado que a supervisão humana de sistemas

¹⁷ Art. 4º Para as finalidades desta Lei, adotam-se as seguintes definições: (...)

V – autoridade competente: órgão ou entidade da Administração Pública Federal responsável por zelar, implementar e fiscalizar o cumprimento desta Lei em todo o território nacional (Brasil, 2023).

¹⁸ Art. 20. Além das medidas indicadas no art. 19, os agentes de inteligência artificial que forneçam ou operem sistemas de alto risco adotarão as seguintes medidas de governança e processos internos:

I – documentação, no formato adequado ao processo de desenvolvimento e à tecnologia usada, a respeito do funcionamento do sistema e das decisões envolvidas em sua construção, implementação e uso, considerando todas as etapas relevantes no ciclo de vida do sistema, tais como estágio de design, de desenvolvimento, de avaliação, de operação e de descontinuação do sistema;

II – uso de ferramentas de registro automático da operação do sistema, de modo a permitir a avaliação de sua acurácia e robustez e a apurar potenciais discriminatórios, e implementação das medidas de mitigação de riscos adotadas, com especial atenção para efeitos adversos;

III – realização de testes para avaliação de níveis apropriados de confiabilidade, conforme o setor e o tipo de aplicação do sistema de inteligência artificial, incluindo testes de robustez, acurácia, precisão e cobertura;

IV – medidas de gestão de dados para mitigar e prevenir **vieses discriminatórios**, incluindo: a) avaliação dos dados com medidas apropriadas de controle de vieses cognitivos humanos que possam afetar a coleta e organização dos dados e para evitar a geração de vieses por problemas na classificação, falhas ou falta de informação em relação a grupos afetados, **falta de cobertura ou distorções em representatividade**, conforme a aplicação pretendida, bem como medidas corretivas para evitar a incorporação de **vieses sociais estruturais que possam ser perpetuados e ampliados pela tecnologia**; e b) **composição de equipe inclusiva responsável pela concepção e desenvolvimento do sistema, orientada pela busca da diversidade**.

V – adoção de medidas técnicas para viabilizar a explicabilidade dos resultados dos sistemas de inteligência artificial e de medidas para disponibilizar aos operadores e potenciais impactados informações gerais sobre o funcionamento do modelo de inteligência artificial empregado, explicitando a lógica e os critérios relevantes para a produção de resultados, bem como, mediante requisição do interessado, disponibilizar informações adequadas que permitam a interpretação dos resultados concretamente produzidos, respeitado o sigilo industrial e comercial (Brasil, 2023, grifos próprios).

de alto risco buscará prevenir ou minimizar os riscos para direitos e liberdades das pessoas, viabilizando que sejam identificadas e corrigidas eventuais falhas, a fim de garantir a transparência e um funcionamento alinhado com princípios éticos e jurídicos, evitando discriminações e outros danos decorrentes de vieses algorítmicos. Nessa disposição, está bem claro que a supervisão humana deve ser capaz de compreender as capacidades e as limitações do sistema para, caso haja necessidade, intervir em seu funcionamento e até decidir por não utilizá-lo.

Posteriormente, no art. 21¹⁹, há disposições acerca das medidas de governança a serem adotadas por órgãos e por entidades do poder público quando controlarem ou utilizarem programas de IA. Percebe-se que o projeto priorizou a possibilidade de escutar a sociedade (inciso I), de que os dados usados sejam de fontes seguras e representativos de populações afetadas com vieses discriminatórios (III), sempre pautando-se pela explicabilidade (inciso IV), transparência e publicidade (inciso VI) acerca das tecnologias usadas. Esse capítulo retrata a atenção do PL às particularidades brasileiras, tendo em vista o fato de que o Brasil é um país que precisa de uma maior assistência e de um Estado de Bem-Estar Social para combater desigualdades e para promover a igualdade material (Bioni; Garrote; Guedes, 2023).

Tendo em vista essa análise, é nítido a essencialidade de se pensar em

¹⁹ Art. 21. Adicionalmente às medidas de governança estabelecidas neste capítulo, órgãos e entidades do poder público da União, Estados, Distrito Federal e Municípios, ao contratar, desenvolver ou utilizar sistemas de inteligência artificial considerados de alto risco, adotarão as seguintes medidas:

I – realização de consulta e **audiência públicas** prévias sobre a utilização planejada dos sistemas de inteligência artificial, com informações sobre os dados a serem utilizados, a lógica geral de funcionamento e resultados de testes realizados;

II – definição de protocolos de acesso e de utilização do sistema que permitam o **registro de quem o utilizou, para qual situação concreta, e com qual finalidade**;

III – utilização de dados provenientes de fontes seguras, que sejam exatas, relevantes, atualizadas e **representativas das populações afetadas e testadas contra vieses discriminatórios**, em conformidade com a Lei no 13.709, de 14 de agosto de 2018, e seus atos regulamentares;

IV – garantia facilitada e efetiva ao cidadão, perante o poder público, de **direito à explicação e revisão humanas de decisão por sistemas de inteligência artificial** que gerem efeitos jurídicos relevantes ou que impactem significativamente os interesses do afetado, a ser realizada pelo agente público competente;

V – utilização de interface de programação de aplicativos que permita sua utilização por outros sistemas para fins de interoperabilidade, na forma da regulamentação; e

VI – **publicização em veículos de fácil acesso**, preferencialmente em seus sítios eletrônicos, das avaliações preliminares dos sistemas de inteligência artificial desenvolvidos, implementados ou utilizados pelo poder público da União, Estados, Distrito Federal e Municípios, independentemente do grau de risco, sem prejuízo do disposto no art. 43.

§1º A utilização de sistemas biométricos pelo poder público da União, Estados, Distrito Federal e Municípios será precedida da edição de ato normativo que estabeleça garantias para o **exercício dos direitos da pessoa** afetada e proteção contra a discriminação direta, indireta, ilegal ou abusiva, vedado o tratamento de dados de raça, cor ou etnia, salvo previsão expressa em lei.

§2º **Na impossibilidade de eliminação ou mitigação substantiva dos riscos associados ao sistema de inteligência artificial identificados na avaliação de impacto algorítmico** prevista no artigo 22 desta Lei, **sua utilização será descontinuada** (grifos próprios).

responsabilização adequada para casos de violação de direitos causados por máquinas e, aqui, com foco na discriminação e no racismo. O conceito de máquinas inteligentes não pode ser usado para que encarregados pelo desenvolvimento tecnológico consigam esquivar-se de responsabilização. Como a máquina não pode ser responsabilizada, o PL dispõe corretamente sobre medidas a serem adotadas pelos programadores e pelos controladores para enfrentar a discriminação em todas as etapas de desenvolvimento e de uso desses sistemas, apresentando-se como um primeiro passo importante para uma regulação com foco na pessoa humana que vise, realmente, garantir direitos.

6 CONCLUSÃO

Como estudado no decorrer do trabalho, embora a crescente utilização da inteligência artificial na tomada de decisões se revele estratégica em diversas situações, como na resolução de problemas complexos diante das limitações cognitivas humanas, é fundamental assegurar que seu desenvolvimento ocorra de maneira democrática e antirracista. Caso contrário, corre-se o risco de comprometer uma longa trajetória de conquistas históricas na luta pela igualdade material e pela garantia dos direitos fundamentais da população afro-brasileira.

Diante do exposto neste estudo, para além de se pensar que os algoritmos são racistas, a questão imposta é a “algoritmização do racismo” a partir da ampliação automatizada das desigualdades econômicas, políticas e culturais e intensificação da extração colonial e racializada de dados, bem como o foco no eixo Sul-Norte global (Silva, 2022). Em função disso, enfatiza-se que o racismo algorítmico é uma extensão do racismo estrutural, profundamente enraizado na sociedade brasileira.

Ideias racistas se manifestam quando sistemas de IA reproduzem preconceitos e desigualdades históricas constantes nos dados usados para seu treinamento e algoritmos aparentemente “neutros” acabam perpetuando ou intensificando práticas discriminatórias. Isto posto, é essencial que as medidas impostas para regulamentar e desenvolver tecnologias de IA sejam direcionadas a corrigir vieses nos dados e nos processos algorítmicos, garantindo que os sistemas respeitem os direitos fundamentais. Sem essa abordagem, as tecnologias continuarão reforçando desigualdades estruturais, agravando injustiças já vivenciadas por populações marginalizadas no Brasil.

É fundamental questionar e desconstruir as aparentes neutralidades das tecnologias, bases de dados e representações, como parte de um esforço contínuo para enfrentar opressões e superar limitações impostas por essas ferramentas (Silva, 2022). No contexto dos bancos de imagens, iniciativas como a Nappy e a Young, Gifted & Black atuam para enfrentar representações negativas e suprir a ausência de representatividade, disponibilizando acervos fotográficos que valorizam a presença de pessoas negras. Esse tipo de projeto busca

contrapor-se aos bancos de imagens convencionais que frequentemente reforçam visibilidades seletivas, invisibilidades estruturais e estereótipos (Silva, 2022).

A analogia entre dados e petróleo, a qual trata que "*data is the new oil*" (dados são o novo petróleo) transformou a percepção dos dados, com o intuito de afastá-los de algo pessoal ou sujeito a controle individual para representá-los como um recurso inerte, como algo pronto para exploração, sendo uma metáfora que não os impactos negativos associados à extração de petróleo, como conflitos geopolíticos e impactos ambientais. Como resultado, os dados passaram a ser vistos como algo abstrato e imaterial, o que os coloca fora do escopo de cuidados, consentimento ou responsabilidade ética. Os dados passaram, então, a serem vistos como capital, alinhados à visão neoliberal de mercados como organizadores de valor, onde rastros digitais de atividades humanas são convertidos em métricas que extraem valor (Crawford, 2021).

Assim, ao transformar dados em um recurso abstrato e imaterial, ocultando as suas origens e impactos, reflete-se a adaptação das técnicas de exploração do capitalismo. Como o petróleo, os dados são vistos como um bem a ser extraído sem considerar os custos sociais e éticos envolvidos. Nesse cenário, os dados são convertidos em capital, funcionando dentro da lógica de mercados que organizam valor e extraem riquezas a partir de atividades humanas, reforçando dinâmicas de desigualdade e opressão (Crawford, 2021).

Como o racismo serve como mecanismo de controle social dentro da dinâmica capitalista (Almeida, 2019), ele cumpre um papel central na perpetuação das desigualdades estruturais que sustentam o sistema. O racismo não apenas legitima a exploração e a desumanização de determinados grupos, mas também consolida hierarquias econômicas e sociais que garantem a concentração de riqueza. Nesse enquadramento, a analogia entre dados e petróleo pode ser aplicada para ampliar a discussão do racismo, pois ambos impulsionam o capitalismo contemporâneo, muitas vezes, às custas de direitos. Semelhantemente à forma como o racismo fragmenta a classe trabalhadora, impondo a alguns grupos condições de trabalho precarizadas e salários baixos, a exploração de dados opera de maneira similar, beneficiando apenas grandes corporações enquanto mantém desigualdades no acesso a oportunidades. Dessa forma, tanto o racismo quanto o uso irrestrito de dados cumprem a função de reforçar e de justificar a opressão e o controle, mascarando-os sob narrativas que atendem aos interesses do capital.

Ter em mente que existe certa seletividade no campo jurídico para o tratamento de questões raciais por causa de uma herança colonial, aristocrática, burguesa e patriarcal (Corrêa, 2021), é interessante para direcionar um olhar crítico ao PL e a quaisquer iniciativas

que tentem regular os sistemas de IA. Aliás, uma das formas mais claras de mitigar os danos causados por discriminações algorítmicas é, sempre, promover a diversidade demográfica entre os profissionais responsáveis pela criação das tecnologias em detrimento ao oligopólio das *big techs*, que dominam o desenvolvimento tecnológico global (Silva, 2022) e cujo quadro de empregados é composto majoritariamente por homens brancos.

Por isso, iniciativas que focam em olhares diferentes para o aprendizado de máquina devem ser incentivadas. No Brasil, iniciativas lideradas por mulheres negras, como o portal Geledés Instituto da Mulher Negra (2009)²⁰ e o grupo Blogueiras Negras (20--)²¹ têm desempenhado um papel crucial na promoção de práticas de cuidados digitais. Soma-se a isso um levantamento sobre a presença de mulheres negras e indígenas nas áreas mais privilegiadas da inovação e da tecnologia (Pretalab, 2022) o qual defendeu que a falta de representatividade não é um problema apenas para a inovação, mas também para os direitos humanos e para a liberdade de expressão (Silva, 2022).

A promoção de discussões diversas e a formação de equipes inclusivas são elementos essenciais para o desenvolvimento de sistemas de IA que sejam verdadeiramente representativos e equitativos. Essa necessidade está explicitada no art. 20, IV, b, do Projeto de Lei nº 2338/2023, que a reconhece como uma medida fundamental de governança, aplicável especificamente a sistemas de alto risco. Embora essa ação represente um avanço considerável, é crucial destacar que a promoção da diversidade não deve ser encarada como uma exigência restrita apenas aos sistemas de IA de maior impacto, mas sim como uma responsabilidade aplicável a todas as instâncias da sociedade. O combate ao racismo e à discriminação racial deve ser um princípio norteador em todas as fases do desenvolvimento e da implementação tecnológica, independentemente do grau de risco ou da área de aplicação, a fim de evitar a sub-representação em dados.

O PL foi um pontapé importante, mas ainda é preciso que exista uma legislação mais rígida para que as empresas realmente se comprometam a investir em diversidade, pois muitas decisões corporativas vão na contramão desse objetivo. Como é o caso da Microsoft, que apesar de ter sido uma companhia que pregou o discurso de que iria aumentar a

²⁰ O Geledés Instituto da Mulher Negra é uma organização brasileira fundada em 1988 com o objetivo de promover os direitos das mulheres negras, combatendo o racismo e o sexismo. Atuando em várias frentes, a organização busca a inclusão social, o fortalecimento do movimento negro feminino e a visibilidade das questões que envolvem as mulheres negras. A organização realiza ações educativas, culturais e políticas para promover a inclusão social, a igualdade de oportunidades e o fortalecimento do movimento negro feminino (Geledés, 2009).

²¹ O Blogueiras Negras é um coletivo formado por mulheres negras que buscam dar visibilidade às suas experiências e reflexões por meio da criação de conteúdo digital. O grupo se dedica à construção de narrativas que abordam questões de gênero, raça, identidade e empoderamento. Por meio da plataforma, discute-se temas como racismo, feminismo, cultura negra e outros assuntos relevantes com foco na perspectiva das mulheres negras, se posicionando como um espaço de resistência e promoção da equidade. Blogueiras Negras (20--).

representatividade de líderes negros, em 2024, fez cortes nos investimentos em diversidade sob a alegação de que iniciativas para inclusão, equidade e diversidade não são mais assuntos críticos para os negócios, como evidenciado no Podcast Techish²².

Também em janeiro de 2025, Mark Zuckerberg, CEO da Meta, anunciou mudanças significativas nas políticas de moderação de conteúdo do Facebook e do Instagram com o pretexto de melhorar a liberdade de expressão e lutar contra a censura. A empresa decidiu encerrar o uso de checagem de fatos por terceiros, passando a usar um sistema de "notas da comunidade", o que gerou preocupações de que conteúdos com vieses racistas ou discriminatórios possam permanecer nas plataformas, uma vez que a moderação dependeria da comunidade de usuários. Contrariamente às propostas que combatem o racismo algorítmico, a ausência de checagem profissional facilita a disseminação de desinformação e discurso de ódio.

Dessa forma, debates sobre racismo e responsabilização no caso de discriminação algorítmica devem sempre existir a fim de que sejam desenvolvidos programas de IA alinhadas com a criação de ações antidiscriminatórias, tanto individuais quanto coletivas e difusas, como estabelecido pelo Estatuto da Igualdade Racial (Silva; Barbosa, 2024), lei que define discriminação racial, desigualdade racial, políticas públicas, ações afirmativas e estabelece a população negra como sendo pessoas que se autodeclaram pretas ou pardas segundo o IBGE²³.

Dito isso, deve-se manter vigilância sobre regulações nacionais e internacionais, pois podem ocultar formas contemporâneas de colonialismo mais sutis e complexas. Esses

²²People (...), 2023.

²³ Art. 1º Esta Lei institui o Estatuto da Igualdade Racial, destinado a garantir à população negra a efetivação da igualdade de oportunidades, a defesa dos direitos étnicos individuais, coletivos e difusos e o **combate à discriminação** e às demais formas de intolerância étnica.

Parágrafo único. Para efeito deste Estatuto, considera-se:

I - **discriminação racial ou étnico-racial**: toda distinção, exclusão, restrição ou preferência baseada em raça, cor, descendência ou origem nacional ou étnica que tenha por objeto **anular ou restringir o reconhecimento, gozo ou exercício, em igualdade de condições, de direitos humanos e liberdades fundamentais** nos campos político, econômico, social, cultural ou em qualquer outro campo da vida pública ou privada;

II - **desigualdade racial**: toda situação injustificada de diferenciação de acesso e fruição de bens, serviços e oportunidades, nas esferas pública e privada, em virtude de raça, cor, descendência ou origem nacional ou étnica;

III - desigualdade de gênero e raça: assimetria existente no âmbito da sociedade que acentua a distância social entre mulheres negras e os demais segmentos sociais;

IV - população negra: o conjunto de pessoas que se autodeclaram pretas e pardas, conforme o quesito cor ou raça usado pela Fundação Instituto Brasileiro de Geografia e Estatística (IBGE), ou que adotam autodefinição análoga;

V - políticas públicas: as ações, iniciativas e programas adotados pelo Estado no cumprimento de suas atribuições institucionais;

VI - ações afirmativas: os programas e medidas especiais adotados pelo Estado e pela iniciativa privada para a correção das desigualdades raciais e para a promoção da igualdade de oportunidades (Brasil, 2010, grifos próprios).

mecanismos, ao mascarar a desigualdade de poder sob uma narrativa genérica de regulação baseada em riscos, ameaçam esvaziar direitos e prejudicar a supervisão democrática e, com a finalidade de evitar isso, é crucial implementar práticas robustas de responsabilização que combatam a assimetria informacional e de poder, garantindo que os princípios de transparência e justiça prevaleçam no uso e regulação de tecnologias emergentes (Bioni; Garrote; Guedes, 2023).

Destarte, o Projeto de Lei nº 2338, de 2023, PL estabelece normas gerais para o uso da IA, reconhece as desigualdades e assimetrias estruturais do contexto brasileiro e é mais protetivo de direitos, alinhando-se aos valores constitucionais, assim como a LGPD, o CDC e o Estatuto da Igualdade Racial (Bioni; Garrote; Guedes, 2023). Entretanto, não aborda de forma abrangente a complexidade do racismo estrutural, nem prevê mecanismos eficazes para identificar e para mitigar vieses raciais em sua origem no Brasil²⁴, posto que é uma questão bastante ampla e com acontecimentos históricos, os quais não vão ter seus efeitos suprimidos de forma rápida e sem uma atuação ativa de todas as camadas sociais.

Embora apresente avanços importantes, como responsabilização objetiva, reparação integral, medidas de transparência e de governança e a preocupação com o impacto e o racismo algorítmico, o projeto ainda não avançou com medidas corretivas para incentivar banco de dados diversos (Bioni; Garrote; Guedes, 2023), por exemplo. O Projeto de Lei nº 2338, de 2023 é, então, marco importante para o Brasil, país localizado no sul global, mas não é completamente efetivo para a prevenção de possíveis consequências de racismo algorítmico desencadeado pelo racismo estrutural no Brasil.

²⁴ Considerações relevantes sobre a temática no Relatório Final da Comissão de Juristas:

Nina da Hora: “Eu não acho que, com as possíveis regulações, nós vamos conseguir acabar com o viés algorítmico, o viés racial, com o racismo algorítmico. Não. **Nós vamos conseguir mitigar, porque o viés racial e o viés de gênero são problemas sociais e eles estão sendo reproduzidos em larga escala por tecnologias de IA.**” (Relatório [...], 2022, p. 414)

Iris: “No contexto de combate à discriminação, **o PL também ignora o debate latente sobre racismo algorítmico, que tem ganhado espaço a cada novo caso “isolado” que surge na mídia, envolvendo a discriminação de pessoas negras por meio de novas tecnologias.** Em um país miscigenado, onde o racismo atravessa diversos setores da sociedade, **é essencial a adoção de medidas voltadas a efetivamente combater essa forma de preconceito.** Assim, a inclusão de princípios que abordem, especificamente, o **antirracismo** é fundamental para assegurar que os agentes que atuam no desenvolvimento e uso de sistemas de inteligência artificial comprometam-se, ativamente, contra a existência de vieses discriminatórios.” (Relatório [...], 2022, p. 761).

Juristas Negras: “Apesar da superficial preocupação com a observância do princípio da não discriminação, **o projeto não faz nenhuma menção ao racismo algorítmico e às diversas manifestações reprodutoras do racismo que têm se concretizado a partir do uso da inteligência artificial.**” (Relatório [...], 2022, p. p. 762).

REFERÊNCIAS

ALMEIDA, Silvio Luiz de. **Racismo estrutural**. São Paulo: Pólen, 2019. (Coleção Feminismos Plurais).

ASSIS, Machado de. **Obra Completa**, Vol III. 3ª edição. José Aguilar, Rio de Janeiro. 1973. p. 489 – 491.

BARROSO, Luís R. **Curso de direito constitucional contemporâneo: os conceitos fundamentais e a construção do novo modelo**. 11. ed. Rio de Janeiro: Saraiva Jur, 2023. E-book. p.228. ISBN 9786553624788. Disponível em: <https://integrada.minhabiblioteca.com.br/reader/books/9786553624788/>. Acesso em: 09 fev. 2025.

BAROCAS, Solon; SELBST, Andrew D. Big data's disparate impact. **California Law Review**, v. 104, p. 671-677, 2016. Disponível em: <https://doi.org/10.15779/Z38BG31>. Acesso em: 14 dez. 2024.

BELANDI, Caio; GOMES, Irene. Censo 2022: pela primeira vez desde 1991, a maior parte da população do Brasil se declara parda. **Agência IBGE Notícias**, [S. l.], 26 jan. 2024. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-noticias/noticias/38719-censo-2022-pela-primeira-vez-desde-1991-a-maior-parte-da-populacao-do-brasil-se-declara-parda>. Acesso em: 13 dez. 2024.

BIONI, Bruno; GARROTE, Marina; GUEDES, Paula. **Temas centrais na Regulação de IA: O local, o regional e o global na busca da interoperabilidade regulatória**. São Paulo: Associação Data Privacy Brasil de Pesquisa, 2023. Disponível em: https://www.dataprivacybr.org/wp-content/uploads/2024/02/nota-tecnica-temas-regulatorios-ia_data.pdf. Acesso em: 15 dez. 2024.

BLACK mirror: queda livre. Criação: Charlie Brooker. [S. l.]: Netflix, 2016. 1 episódio (63 min).

BLOGUEIRAS NEGRAS. **Quem somos**: reinventando a tela. [S. l.], [20--]. Disponível em: <https://blogueirasnegras.org/quem-somos/#:~:text=O%20Blogueiras%20Negras%20%C3%A9%20constru%C3%ADdo,por%20e%20para%20mulheres%20negras>. Acesso em: 14 jan. 2025.

BRASIL. **Constituição (1988)**. Constituição da República Federativa do Brasil. Brasília, DF: Senado Federal, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 7 fev. 2025.

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002**. Institui o Código Civil. Diário Oficial da União: seção 1, Brasília, DF, 11 jan. 2002. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/2002/110406compilada.htm. Acesso em: 14 jan. 2025.

BRASIL. **Lei nº 12.288, de 20 de julho de 2010**. Estatuto da Igualdade Racial. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/lei/l12288.htm. Acesso em: 14 jan. 2025.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet). Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 7 fev. 2025.

BRASIL. **Projeto de Lei nº 2.338, de 2023**. Dispõe sobre o uso da Inteligência Artificial. Brasília, DF: Senado Federal, 2023. Disponível em: <https://legis.senado.leg.br/sdleg-getter/documento?dm=9347622&ts=1720545987618&disposition=inline>. Acesso em: 7 fev. 2025.

Comissão sobre Inteligência Artificial analisa relatório final – 5/12/24. Publicado pelo canal: TV Senado. [S. 1: s. n.], 2024, 1 vídeo (77 minutos). Disponível em: <https://www.youtube.com/watch?v=35BiW6s4xwI>. Acesso em: 27 de janeiro de 2025.

CORRÊA, Bianca Kremer Nogueira. **Direito e Tecnologia em Perspectiva Amefricana: autonomia, algoritmos e vieses raciais**. Tese (Doutorado em Direito) - Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio). Orientação de Maria Celina Bodin de Moraes e Co-orientação de Caitlin Sampaio Mulholland. Rio de Janeiro, março de 2021. Disponível em: <https://www.maxwell.vrac.puc-rio.br/58993/58993>. Acesso em: 26 jan. 2025.

CORTIZ, Diogo. Inteligência Artificial: equidade, justiça e consequências. **Panorama setorial da Internet**, 2020, vol. 12, no. 1. Disponível em: https://cetic.br/media/docs/publicacoes/6/20200626161010/panorama_setorial_ano-xii_n_1_inteligencia_artificial_equidade_justi%C3%A7a.pdf. Acesso em: 27 jan. 2025.

COSTA, João Carneiro. A responsabilidade civil pelos danos causados pela inteligência artificial nas hipóteses de discriminação algorítmica. **Revista Direito UNIFACS – Debate Virtual**, n. 272, 2023. Disponível em: <https://revistas.unifacs.br/index.php/redu/article/view/8065/4756>. Acesso em: 9 fev. 2025.

CRAWFORD, Kate. **Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence**. New Haven: Yale University Press, 2021. Disponível em: <https://www.are.na/block/11669646>. Acesso em: 15 dez. 2024.

DONEDA, Danilo Cesar Maganhoto; MENDES, Laura Schertel; SOUZA, Carlos Affonso Pereira de; ANDRADE, Norberto Nuno Gomes de. Considerações iniciais sobre inteligência artificial, ética e autonomia pessoal. **Pensar**, Fortaleza, v. 23, n. 4, p. 1-17, out./dez. 2018. DOI: 10.5020/2317-2150.2018.8257. Disponível em: <https://ojs.unifor.br/rpen/article/download/8257/pdf/31971>. Acesso em: 13 dez. 2024.

DONEDA.NET. DONEDA, Danilo. ALMEIDA, Virgílio A.F. **O que é a governança de algoritmos?** [S. l.], 2016. Disponível em: <https://doneda.net/o-que-e-a-governanca-de-algoritmos/>. Acesso em: 9 fev. 2025.

GELEDÉS. **O que é o Geledés - Instituto da Mulher Negra**. [S. l.], 2009. Disponível em:

<https://www.geledes.org.br/o-que-e-o-geledes-instituto-da-mulher-negra/>. Acesso em: 14 jan. 2025.

HACKING, Ian. Kinds of People: Moving Targets. **Proceedings of the British Academy**, v. 151, 2007. p. 285–318. Disponível em: [https://joelvelasco.net/teaching/5330\(spring2021\)/hacking-kinds_of_people.pdf](https://joelvelasco.net/teaching/5330(spring2021)/hacking-kinds_of_people.pdf). Acesso em: 15 dez. 2024.

HOLANDA, Sérgio Buarque de. **Raízes do Brasil**. 26. ed. São Paulo: Companhia das Letras, 1995.

HOW I'm fighting bias in algorithms. Publicado pelo canal: TED. Palestrante: Joy Buolamwini. [S. l.: s. n.], 2017. 1 vídeo (8:44 min). Disponível em: https://youtu.be/UG_X_7g63rY. Acesso em: 13 jan. 2025.

IA no Direito: Impactos do PL 2338/23 no jurídico , com Dra. Letícia Zanardo #124 #juriscast. Publicado pelo canal: Projuris Plataforma de Inteligência Artificial. [S. l.: s. n.], 2024, 1 vídeo (41:08 min). Disponível em: <https://www.youtube.com/watch?v=mPRAwpsN8Qo>. Acesso em: 27 jan. 2025.

MANN, Gideon; O'NEIL, Cathy. Hiring algorithms are not neutral. **Harvard Business Review**, Massachusetts, 09 dez. 2016. Disponível em: <https://hbr.org/2016/12/hiring-algorithms-are-not-neutral>. Acesso em: 14 dez. 2024.

MENDONÇA JUNIOR, C. do N.; NUNES, D. J. C. Desafios e oportunidades para a regulação da inteligência artificial: a necessidade de compreensão e mitigação dos riscos da IA. **Revista Contemporânea**, [S. l.], v. 3, n. 07, p. 7753–7785, 2023. DOI: 10.56083/RCV3N7-024. Disponível em: <https://ojs.revistacontemporanea.com/ojs/index.php/home/article/view/1146>. Acesso em: 14 jan. 2025.

NEGRI, Sergio M. C. Ávila; MACHADO, Joana de Souza; GIOVANIN, Carolina Fiorini Ramos; BATISTA, Nathan Pascoalini Ribeiro. Sistemas de inteligência artificial e avaliações de impacto para direitos humanos. **Revista Culturas Jurídicas**, V. 10, n. 26, p. 153-181, 2023. Disponível em: https://periodicos.uff.br/culturasjuridicas/article/view/56809?utm_source=chatgpt.com. Acesso em: 27 jan. 2025.

ONODY, Roberto N. Teste de Turing e Inteligência Artificial. **Instituto de Física da Universidade de São Paulo Notícias**, São Paulo, 28 set. 2021. Disponível em: <https://www2.ifsc.usp.br/portal-ifsc/teste-de-turing-e-inteligencia-artificial/>. Acesso em: 13 dez. 2024.

PRETALAB. **Relatório 2022**. Disponível em: <https://www.pretalab.com/report-2022>. Acesso em: 14 jan. 2025.

PROJETO COMPROVA. Saiba o que é deepfake, técnica de inteligência artificial que foi apropriada para produzir desinformação. **CNN Brasil**, [S. l.], 13 jan. 2023. Disponível em: <https://www.cnnbrasil.com.br/noticias/saiba-o-que-e-deepfake-tecnica-de-inteligencia-artificial-l-que-foi-apropriada-para-produzir-desinformacao/>. Acesso em: 27 jan. 2025.

RAINIE, Lee; ANDERSON, Janna. Code-Dependent: Pros and Cons of the Algorithm Age. **Pew Research Center**, 8 fev. 2017. Disponível em: <http://www.pewinternet.org/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/>. Acesso em: 14 dez. 2024.

RAUB, McKenzie. Bots, Bias and Big Data: Artificial Intelligence, Algorithmic Bias and Disparate Impact Liability in Hiring Practices. **Arkansas Law Review**, v. 71, n. 2, p. 529-554, dez. 2018. Disponível em: <https://scholarworks.uark.edu/cgi/viewcontent.cgi?article=1052&context=alr>. Acesso em: 13 dez. 2024.

RELATÓRIO final: comissão de juristas responsável por subsidiar elaboração de substitutivo sobre inteligência artificial no Brasil. Brasília: Senado Federal, 2022. Disponível em: <https://www.gov.br/anpd/pt-br/assuntos/noticias/2-relatorio-final-versao-completa-cjsubia.pdf>. Acesso em: 27 jan. 2025.

RIBEIRO, Darcy. **O povo brasileiro**: a formação e o sentido do Brasil. São Paulo: Companhia das Letras, 1995.

SARLET, Ingo; MARINONI, Luiz G.; MITIDIERO, Daniel. **Curso de Direito Constitucional**. 13. ed. Rio de Janeiro: Saraiva Jur, 2024. E-book. p.542. ISBN 9788553621163. Disponível em: <https://integrada.minhabiblioteca.com.br/reader/books/9788553621163/>. Acesso em: 09 fev. 2025.

SCHREIBER, Anderson. **Manual de Direito Civil Contemporâneo**. 7. ed. Rio de Janeiro: Saraiva Jur, 2024. E-book. ISBN 9788553620234. Disponível em: <https://integrada.minhabiblioteca.com.br/reader/books/9788553620234/>. Acesso em: 09 fev. 2025.

SILVA, Isabela Maria Soares; BARBOSA, Leticia Mendes. Inov(ação): discriminação algorítmica racial e as inteligências artificiais no Brasil. **Revista do CAAP**, [S. l.], v. 28, n. 2, p. 1–30, 2024. DOI: 10.69881/rcaap.v28i2.49200. Disponível em: <https://periodicos.ufmg.br/index.php/caap/article/view/49200>. Acesso em: 14 jan. 2025.

SILVA, Tarcízio. Mapeamento de Danos e Discriminação Algorítmica. **Desvelar**, 2023. Disponível em: <https://desvelar.org/casos-de-discriminacao-algoritmica/>. Acesso em: 13 dez. 2024.

SILVA, Tarcízio. Regular a inteligência artificial no Brasil pode mitigar o racismo algorítmico? **Desvelar**, 30 de junho de 2023. Disponível em: <https://desvelar.org/2023/06/30/regular-a-inteligencia-artificial-no-brasil-pode-mitigar-o-racismo-algoritmico/>. Acesso em: 14 jan. 2025.

SILVA, Tarcízio. **Racismo algorítmico**: inteligência artificial e discriminação nas redes digitais. Edições Sesc SP, 2022.

TECH'S DEI Gaslighting, AI Performance Reviews, Make Politics Attractive Again. Publicado pelo canal: Techish Podcast. [S. 1: s. n.], 2024, 1 vídeo (22:06 min). Disponível em: <https://www.youtube.com/watch?v=jJLneRI0UgQ>. Acesso em: 14 jan. 2025.

VASCONCELOS, Sergio Sezino Douets. Igreja Católica e a escravidão no Brasil Colônia: uma abordagem cultural. In: SNH: SIMPÓSIO NACIONAL DE HISTÓRIA, 27., 2013, Natal. **Anais** [...]. Disponível em: <https://snh2013.anpuh.org/resources/pe/anais/encontro5/04-rep-sociais/Artigo%20de%20Sergio%20Douets.pdf>. Acesso em: 13 dez. 2024.

VILLARINHO, Juliana. Machado de Assis branco feito por IA causa polêmica na web. **TechTudo**, [S. 1.], 2024. Disponível em: <https://www.techtudo.com.br/noticias/2024/03/machado-de-assis-branco-feito-por-ia-causa-polêmica-na-web-veja-edsoftwares.ghtml>. Acesso em: 14 jan. 2025.

VITORIO, Tamires. A desigualdade em números: brancos ainda são maioria nas big techs. **Exame**, [S. 1.], 20 nov. 2020. Disponível em: <https://exame.com/tecnologia/a-desigualdade-em-numeros-brancos-ainda-sao-maioria-nas-big-techs/>. Acesso em: 13 dez. 2024.