



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS E HUMANAS
DEPARTAMENTO DE CIÊNCIAS SOCIAIS APLICADAS
CURSO DE DIREITO

GABRIEL VITOR MENDES CASTRO

**DESIGUALDADE CODIFICADA: INVESTIGANDO A
DISCRIMINAÇÃO ALGORÍTMICA**

MOSSORÓ/RN

2025

GABRIEL VITOR MENDES CASTRO

**DESIGUALDADE CODIFICADA: INVESTIGANDO A
DISCRIMINAÇÃO ALGORÍTMICA**

Artigo apresentado ao Departamento de Ciências Sociais Aplicadas como parte dos requisitos para a obtenção do grau de Bacharel em Direito, pela Universidade Federal Rural do Semi-Árido.

Orientador: Prof. Dr. Rodrigo Vieira Costa.

MOSSORÓ/RN

2025

GABRIEL VITOR MENDES CASTRO

**DESIGUALDADE CODIFICADA: INVESTIGANDO A
DISCRIMINAÇÃO ALGORÍTMICA**

Artigo apresentado ao Departamento de
Ciências Sociais Aplicadas como parte dos
requisitos para a obtenção do grau de Bacharel
em Direito, pela Universidade Federal Rural
do Semi-Árido.

APROVADO EM 10/03/2025

BANCA EXAMINADORA

Rodrigo Vieira Costa, Prof. Dr. (UFERSA)

Presidente

Rayane Cristina de Andrade Gomes, Profa. Dra. (UFERSA)

Membro Examinador

Irineu Francisco Barreto Júnior, Prof. Dr. (FMU-SP)

Membro Examinador

MOSSORÓ/RN

2025

RESUMO

Este trabalho investiga a discriminação algorítmica, visualizando como os algoritmos de inteligência artificial podem reproduzir desigualdades sociais. A pesquisa tem como objetivo explorar casos que exemplificam o enviesamento algorítmico, bem como identificar regulamentações pertinentes à temática. Para isso, foi realizada uma revisão integrativa da literatura sobre discriminação algorítmica por meio do Portal de Periódicos da CAPES com uso de palavras-chave relacionadas ao assunto, com filtros (revisados por pares, publicados nos últimos 05 anos), resultando em 11 trabalhos relevantes. O primeiro capítulo explorará os fundamentos do viés algorítmico, conceituando sobre aprendizado de máquina e o uso de algoritmos. O capítulo seguinte apresentará exemplos de discriminação algorítmica no Brasil e no mundo. O último capítulo identificará o arcabouço jurídico existente e a necessidade de regulamentações mais eficazes para garantir a transparência e a equidade no uso dessas tecnologias. Conclui-se que o uso de sistemas de inteligência artificial traz benefícios evidentes, todavia, a complexidade dos algoritmos que dificulta a sua auditabilidade é um desafio constante. Desse modo, a sua utilização deve ser fundamentada por transparência e responsabilidade por parte dos agentes envolvidos.

Palavras-chave: Viés; Algoritmo; Discriminação; Aprendizagem de Máquina; Transparência.

ABSTRACT

This work investigates algorithmic discrimination, visualizing how artificial intelligence algorithms can reproduce social inequalities. The research aims to explore cases that exemplify algorithmic bias, as well as identify regulations relevant to the topic. To this end, an integrative review of the literature on algorithmic discrimination was carried out through the CAPES Journal Portal using keywords related to the subject, with filters (peer-reviewed, published in the last 5 years), resulting in 11 relevant works. The first chapter will explore the fundamentals of algorithmic bias, conceptualizing about machine learning and the use of algorithms. The following chapter will present examples of algorithmic discrimination in Brazil and around the world. The last chapter will identify the existing legal framework and the need for more effective regulations to ensure transparency and equity in the use of these technologies. It is concluded that the use of artificial intelligence systems brings clear benefits, however, the complexity of the algorithms, which makes their auditability difficult, is a constant challenge. Therefore, its use must be based on transparency and responsibility on the part of the agents involved.

Keywords: Bias; Algorithm; Discrimination; Machine Learning; Transparency.

Sumário

1 INTRODUÇÃO	07
2 Viés Algorítmico e o seu Potencial como Segregador Social	09
2.1 Conceituando a Discriminação Algorítmica consoante a Literatura	11
3 A Utilização de Algoritmos e o Fomento à Discriminação	16
4 Necessária Intervenção: análise de diretrizes regulatórias	23
5 Considerações Finais	29
REFERÊNCIAS	31
APÊNDICE I – IDENTIFICAÇÃO DOS TRABALHOS INCLUÍDOS PELA REVISÃO DE LITERATURA	34
APÊNDICE II – IDENTIFICAÇÃO DOS TRABALHOS EXCLUÍDOS DA REVISÃO DE LITERATURA	35

1 INTRODUÇÃO

Em um mundo cada vez mais tecnológico, a inteligência artificial (IA) é um artifício fundamental para a sistematização e automação de atividades intelectuais, abrangendo diversas áreas de atuação, desde o desenvolvimento da robótica até o aperfeiçoamento de sistemas de planejamento e logística (Fernandes et al., 2022). Nesse viés, podemos visualizar a IA como um mecanismo em que máquinas simulam as habilidades humanas, de modo que, em alguma medida, imitarão o processo cognitivo humano. Para que essa mimetização ocorra e a inteligência artificial demonstre-se efetiva, faz-se necessário a utilização de algoritmos (Junior et al., 2023).

Os algoritmos, em sua definição geral, são um conjunto de instruções lógicas que fornecerão um passo a passo de comandos que serão seguidos para se obter determinado resultado. Funcionando de maneira sistemática, os algoritmos são utilizados em praticamente todos os programas de computadores, auxiliando o processo de tomadas de decisões objetivas e fundamentadas em dados estatísticos (Leal; Paulo, 2023). Apesar de tantos benefícios, os algoritmos, quando especializados para o campo da aprendizagem de máquina e da inteligência artificial, não estão isentos de falhas, principalmente no que tange a neutralidade. Cada humano possui a sua própria concepção de valores e a sua visão da realidade, de modo que, por ser o algoritmo uma criação humana, eles comumente refletem os valores dos seus criadores (De Freitas, 2023).

No tocante ao campo tecnológico e da inteligência artificial, apesar da sua diversidade de modalidades, o uso de algoritmos de aprendizagem de máquina é um dos principais mecanismos associados às discriminações sociais. Esses sistemas são alimentados por uma amostragem de dados que, a partir da análise dos padrões existentes nesses dados, projetam resultados que muitas vezes não foram planejados por seus criadores, reproduzindo vieses (Simões-Gomes et al., 2020).

O fato de os algoritmos serem alimentados por dados históricos cria a hipótese de que, caso os dados sejam danosos ou incoerentes com o panorama social, os algoritmos reproduzirão os estereótipos indesejáveis. Nesse sentido, esses vieses adquiridos por sua base de dados ou programadores intensifica a discriminação no tocante a questões de gênero, raça e classe social (Leal; Paulo, 2023).

Ressalta-se, todavia, que a inteligência artificial não é um mecanismo doente criado com o intuito de discriminar, ela apenas reflete as patologias do mundo em que

está inserida, afinal, os dados que a alimenta vêm desse globo social (Leal; Paulo, 2023). O desafio se encontra, portanto, no conflito entre amenizar os riscos enquanto se resguarda o ambiente de inovação, de modo que a inteligência artificial e os algoritmos não sejam supervalorizados nem demonizados (Kaufman et al., 2023).

A presente pesquisa justifica-se pela importância de compreender os impactos da discriminação algorítmica e de explorar os aspectos jurídicos encontrados na literatura com o intuito de atenuar as mazelas do óbice em questão. Ante o exposto, o trabalho busca responder ao questionamento: “Como os algoritmos podem gerar discriminação para determinados grupos sociais e quais são as possíveis medidas para mitigar esses efeitos?”

Apesar de existir algumas diretrizes secundárias definidas por diferentes instituições que estabelecem parâmetros mínimos a serem seguidos a respeito da transparência dos dados e dos softwares utilizados, bem como dos níveis de acurácia desses algoritmos, tal situação segue sem uma regulamentação específica que fiscalize a criação e a utilização dos sistemas algorítmicos e de inteligência artificial. Nesse contexto, a necessidade de explorar alternativas que possam superar os óbices provenientes da discriminação algorítmica, promovendo soluções jurídicas mais eficazes e alinhadas às demandas da sociedade é pertinente.

Ademais, por ser um recurso que se relaciona e impacta diretamente o corpo social, os algoritmos não podem permanecer sendo orientados por vieses que reproduzem valores discriminatórios. Nesse panorama, observa-se como necessário a aplicação das técnicas de governança em parceria com uma regulamentação eficiente para que a criação e utilização dos algoritmos estejam intrinsecamente atrelados aos princípios da segurança, transparência, dignidade da pessoa humana, não discriminação, liberdade de escolha, proteção da privacidade e responsabilidade (Rossetti; Angelucci, 2021).

Ao realizar esse estudo, uma das possibilidades viabilizadas é oferecer subsídios teóricos que auxiliem para a amenização da discriminação algorítmica, promovendo uma reflexão crítica sobre as consequências da aplicação de um recurso tecnológico enviesado. A expectativa é que os resultados obtidos possam auxiliar juristas, programadores e demais envolvidos na área da tecnologia e do direito a repensarem metodologias e práticas profissionais que proporcionem uma redução do enviesamento algorítmico contemporâneo.

Nesse viés, a construção do estado do conhecimento sobre a segregação social decorrente da discriminação algorítmica se demonstra como importante e atual, levando-se em consideração que situações discriminatórias devem ser sanadas o mais breve possível. Para tal, o presente trabalho será fundamentado pela pesquisa bibliográfica, de modo que, no dia 20 de agosto de 2024, foi realizada uma revisão integrativa da literatura sobre a discriminação algorítmica por meio do Portal de Periódicos da CAPES, utilizando-se de palavras-chave relacionadas à temática (exatamente, em qualquer parte do texto, deveriam constar os termos “Discriminação” e “Algoritmo”), com filtros (revisados por pares, publicados nos últimos 05 anos), resultando em 11 trabalhos relevantes.

Estruturalmente, o trabalho se desenvolveu em três capítulos. O primeiro capítulo tem como objetivo introduzir conceitos pertinentes acerca do enviesamento algorítmico tomando como base a referida literatura, examinando os fundamentos da injustiça algorítmica e a relação entre o aprendizado de máquina e a tendência discriminatória desses sistemas. O segundo capítulo se propõe a investigar como esse preconceito automatizado acontece na prática, abordando exemplos nacionais e internacionais que demonstrarão que tal assunto não pode ser negligenciado, sendo indispensável uma intervenção para que as situações discriminatórias não se perpetuem. Por fim, o terceiro capítulo abordará os dispositivos jurídicos relevantes disponíveis, salientando a necessidade de uma regulamentação específica e eficiente.

2 Viés Algorítmico e o seu Potencial como Segregador Social

O foco do presente capítulo será expor a demanda por mudanças no que se refere ao uso de algoritmos que possuem vieses discriminatórios em sua constituição, introduzindo o leitor aos conceitos pertinentes à temática. Nessa conjuntura, observa-se a importância da pesquisa, processo específico que favorece a produção e difusão de conhecimento e que pressupõe o emprego de uma revisão bibliográfica para fundamentar o estado de conhecimento sobre determinado objeto de estudo (Rodrigues; Grubba, 2023).

No que tange o campo do Direito, a revisão integrativa, espécie de revisão bibliográfica, permite condensar o conhecimento existente sobre determinado tema, bem

como identificar as lacunas nesse conhecimento, contribuindo para o aperfeiçoamento da prática científica e jurídica (Rodrigues; Grubba, 2023). Este método possibilita, portanto, compilar informações de diversas fontes com o intuito de fomentar o questionamento e desenvolvimento do assunto. Com esse intuito, foi realizada uma revisão integrativa utilizando o Portal de Periódicos da CAPES como fontes de dados, por meio do acesso CAFe¹ disponibilizado pela Universidade Federal Rural do Semi-Árido.

Para a aplicação efetiva da revisão integrativa de literatura, faz-se necessário a utilização de um processo rigoroso e sistemático, em que as etapas serão descritas com detalhes, uma vez que, por ser um método PBE,² o processo tem que ser passível de ser reproduzido por outros pesquisadores (Rodrigues; Grubba, 2023). Desse modo, segue a descrição de como o estudo foi conduzido.

Em 20 de agosto de 2024, realizou-se uma busca avançada por trabalhos contendo exatamente os termos “Discriminação” E (and) “Algoritmo”, obtendo-se 138 resultados majoritariamente em português e subsidiariamente em inglês. Com o intuito de refinar e garantir a pertinência dos trabalhos, aplicou-se o filtro revisado por pares³, resultando em 45 produções científicas. Ademais, objetivando que a pesquisa seja relevante para a contemporaneidade, aplicou-se o filtro para artigos publicados nos últimos cinco anos (2020-2024), de modo a obter-se 18 artigos.

Após a leitura dos títulos e resumos, observou-se que 7 artigos tangenciavam a temática⁴, não abordando os pontos objetivados na pesquisa, de modo que foram excluídos. Portanto, mediante a aplicação dos critérios de inclusão e exclusão, a presente revisão integrativa conterá 11 artigos⁵ revisados por pares, publicados nos últimos 5 anos e que contém, exatamente, em alguma parte do seu corpo textual, os termos “Discriminação” e “Algoritmo”.

¹ O CAFe é um serviço que integra uma base de dados provenientes de instituições de ensino e pesquisa brasileiras. O acesso a publicações científicas é um dos seus maiores benefícios.

² O método PBE identifica que a revisão integrativa obedecerá aos critérios de viabilidade, adequação, pertinência e eficácia (Rodrigues; Grubba, 2023, p. 193).

³ A revisão por pares é um processo por meio do qual uma produção científica será avaliada por dois outros cientistas do mesmo campo e com mesmo título, de modo a avaliar e garantir a qualidade do trabalho.

⁴ Os textos excluídos da revisão integrativa não abordavam sobre discriminação algorítmica. A tabela de identificação dos trabalhos excluídos poderá ser visualizada no **Apêndice II** anexado ao final.

⁵ Os textos selecionados poderão ser consultados ao final, no **Apêndice I**. Os trabalhos foram organizados de acordo com a ordem dos resultados da CAPES.

2.1 Conceituando a Discriminação Algorítmica consoante a Literatura

A utilização de recursos atrelados à aplicabilidade de sistemas algorítmicos se demonstra fundamental para o auxílio das atividades humanas diante de um corpo social cada vez mais globalizado e tecnológico. Todavia, a sua natureza complexa somada à uma enviesada base de dados e uma má execução resultam em uma reprodução de estereótipos discriminatórios. Partindo da concepção de que uma estrutura que discrimina grupos sociais não deve se manter vigente, será explorada a literatura no que diz respeito à utilização de algoritmos enviesados e os seus impactos.

O artigo científico 03⁶ conceitua os algoritmos como modelos escalonados de comando, que funcionam de forma sistemática, a partir de uma sequência de instruções previamente ordenadas para a obtenção de um resultado almejado. Nessa perspectiva, os autores afirmam que os algoritmos tendem a discriminar grupos sociais pois poderão herdar vieses da sua base de dados ou de seus programadores. Em um contexto de discriminação histórica e estrutural, aspectos como gênero, raça e classe social estão arraigados ao funcionamento da sociedade e servem para a classificação de pessoas, demonstrando-se potencial lesivo no que tange a observação desses vieses dentro da linguagem de máquina (Leal; Paulo, 2023).

Continuando, o artigo supracitado ressalta que no nosso sistema jurídico as pessoas devem ser julgadas pelas suas ações e não por serem quem elas são. Logo, é importante ressaltar que a Lei nº 13.709, de 14 de agosto de 2018, que instituiu a Lei Geral de Proteção de Dados (LGPD), disciplina sobre o tratamento de dados pessoais, com o fito de proteger os direitos de liberdade, de privacidade e do livre desenvolvimento da personalidade da pessoa natural (Leal; Paulo, 2023).

Nesse sentido, os autores defendem ser necessário o cuidado com a qualidade dos dados utilizados pelos algoritmos, a transparência do seu funcionamento e com a ética e responsabilidade social das empresas no que tange o resultado de seus produtos (Leal; Paulo, 2023). Nota-se que os três pontos convergem em uma demanda por uma regulamentação efetiva.

⁶ O artigo 03 é “Algoritmos discriminatórios e jurisdição constitucional: os riscos jurídicos e sociais do impacto dos vieses nas plataformas de inteligência artificial de amplo acesso” (Leal; Paulo, 2023).

O artigo científico 04⁷ afirma que um dos principais objetivos dos algoritmos é realizar previsões probabilísticas por meio da análise de uma amostragem de dados. É nessa conjuntura que se observa os dados como a base do mecanismo algorítmico, sendo determinante a quantidade e a qualidade desses dados para a obtenção do melhor resultado, uma vez que a ausência de informações verossímeis com a realidade aumenta as chances de enviesamento e de empobrecimento a longo prazo dos resultados das inteligências artificiais. O artigo pontua que a inteligência artificial busca ensinar atividades humanas para as máquinas, sendo a aprendizagem de máquina (machine learning) uma dessas atividades (De Freitas, 2023).

Nesse viés, a autora cita que a inteligência artificial, em um dos seus campos de atuação, é pautada pela abordagem das redes neurais. Essa abordagem foca em imitar a estrutura cerebral humana, criando neurônios artificiais capazes de receber e transmitir informações. Com a inserção de uma amostragem de dados, essas redes neurais artificiais identificarão padrões, alcançando de forma autônoma uma decisão ou resultado. A evolução do machine learning foi o que permitiu, por exemplo, a criação dos atuais sistemas de reconhecimento facial, análises de crédito e policiamento preditivo (De Freitas, 2023).

Todavia, a autora pontuou também que diante da opacidade algorítmica e da ausência de neutralidade, viabiliza-se recorrentemente a ocorrência de discriminação. Nessa conjuntura, viabiliza-se quatro espécies principais de discriminação algorítmica: a) discriminação por erro estatística, que são erros durante a construção ou calibração dos algoritmos; b) discriminação por generalização, momento em que algumas pessoas são classificadas em agrupamentos dos quais não pertencem; c) discriminação pelo uso de informações sensíveis, quando ocorre a utilização de dados sensíveis baseados em atributos endógenos; d) discriminação limitadora do exercício de direitos, momento em que um direito é afetado em decorrência do resultado produzido por um algoritmo (De Freitas, 2023).

⁷ O artigo 04 é “Policiamento Preditivo: aspectos discriminatórios no uso das novas tecnologias” (De Freitas, 2023).

O artigo científico 01⁸ introduz a conceituação dos sistemas algorítmicos como procedimentos de tratamento de dados que utilizam a matemática para deduzir correlações e padrões probabilísticos (Do Amaral et al., 2021).

Os autores também indagam que, apesar dos benefícios, a opacidade da ferramenta algorítmica dificulta descobrir quais foram os critérios utilizados durante o procedimento e se eles estão ou não atrelados a concepções discriminatórias. Tal fato assimilaria o procedimento algorítmico à uma ‘caixa preta’, haja vista que a sua opacidade torna inacessível aos afetados verificarem os processos pelos quais o algoritmo correlaciona os dados e transforma-os em padrões injustos (Do Amaral et al., 2021).

Nessa conjuntura, os algoritmos poderiam funcionar como ‘armas de destruição matemática’, criando ‘espirais viciosas’ que acarretam a ampliação da desigualdade tomando como ponto de partida interesses privados ou governamentais (O’Neil, 2016 apud Do Amaral et al., 2021). Por fim, os autores desenvolvem essa problematização apresentando exemplos em que a discriminação algorítmica se demonstrou explícita.

O artigo científico 07⁹ afirma que o algoritmo é um conjunto de ações ordenadas com o intuito de solucionar um problema, comparando-o a uma receita culinária humana. Entretanto, a necessidade de algoritmos complexos para o auxílio em atividades se faz frequente em computadores e máquinas, de modo que um subconjunto algorítmico específico passa a ser utilizado. Esse subconjunto baseia-se no aprendizado de máquina (machine learning), que são algoritmos capazes de aprender a partir de dados, de forma que identificará padrões e deduzirá resultados (Chaves Junior et al., 2023).

Após as conceituações iniciais e algumas exemplificações, o artigo viabiliza que, para o combate à ausência de neutralidade, uma solução plausível seria a remoção dos atributos sensíveis presentes nos dados. Características como raça, gênero e religião deixariam, portanto, de serem adicionadas ao treinamento do modelo algorítmico. Nessa conjuntura, os autores afirmam que os conflitos éticos demandam fiscalização e regulamentação, mas não devem promover um repúdio exagerado, uma vez que a automação e a inteligência artificial apresentam diversos benefícios no que tange funções

⁸ O artigo 01 é “Gubernamentalidad algorítmica y nuevas prácticas punitivas” (Do Amaral et al., 2021).

⁹ O artigo 07 é “Segregação racial e vieses algorítmicos: máquinas racistas no âmbito do controle penal” (Chaves Junior et al., 2023).

de auxílio. Citou, inclusive, a inserção de inovações tecnológicas nos Tribunais brasileiros como um poderoso aliado (Chaves Junior et al., 2023).

O artigo científico 05¹⁰ fundamenta-se principalmente na análise ética da discriminação algorítmica, com especial atenção à opacidade e à neutralidade. Os escritores apontam que a opacidade é uma questão ética proveniente da natureza complexa do algoritmo, de modo que as etapas de criação de um mecanismo algorítmico dificilmente estão disponíveis ao usuário ou ao pesquisador, demonstrando-se uma ausência de transparência (Rossetti; Angelucci, 2021).

Segundo o artigo, o argumento dos programadores quanto à opacidade está atrelado ao princípio da privacidade. Uma vez que os dados são privados, torná-los públicos afetariam questões de segurança, bem como possibilitaria uma apropriação indevida das informações pela concorrência. Todavia, o autor ressalta que a transparência algorítmica é indispensável para que se tenha previsibilidade, controle, monitoramento e correção no que tange a compreensão dessas informações (Rossetti; Angelucci, 2021).

Dando prosseguimento, os autores pontuam que a problemática do viés se encontra nas decisões algorítmicas tendenciosas, que muitas vezes estão relacionadas a preferência de determinados valores. Desse modo, mesmo que em essência um algoritmo não seja preconceituoso, ele acaba por refletir atributos valorativos de seus criadores, de modo que a neutralidade passa a não existir e o algoritmo permanece sujeito a um enviesamento social e cultural (Rossetti; Angelucci, 2021).

O artigo científico 08,¹¹ após conceituações iniciais, também focaliza na análise ética e regulatória. O artigo aponta que a ética é uma característica intrínseca da sociedade humana e a inteligência artificial é uma ferramenta tecnológica não determinista, logo, cabe aos próprios humanos a responsabilidade de definir e disciplinar a utilização desse instrumento. Nessa perspectiva, os autores defendem a necessidade de auditoria humana nas situações de automação, citando diversos dispositivos constitucionais que impedem a discriminação, bem como a Declaração Universal dos Direitos Humanos (Kaufman et al., 2023).

¹⁰ O artigo 05 é “Ética Algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação” (Rossetti; Angelucci, 2021).

¹¹ O artigo 08 é “Externalidades negativas da inteligência artificial: conflitos entre limites da técnica e direitos humanos” (Kaufman et al., 2023).

O artigo científico 02,¹² apesar de também abordar conceitos introdutórios sobre a temática, focou-se principalmente em uma análise do arcabouço jurídico brasileiro. A produção textual afirma que, por ser o Brasil um país fundamentado pelo Estado Democrático de Direito, situações marcadas pela discriminação propagam preconceitos que a própria Constituição Federal de 1988 considera inaceitáveis (Soares et al., 2022).

O artigo pontua a inexistência de legislação específica sobre o caso, citando o Projeto de Lei nº 21/2020 como uma potencial regulamentação que disciplinará sobre a transparência de sistemas de inteligência artificial e a possível responsabilização dos seus criadores (Soares et al., 2022).

O artigo científico 14¹³ abordou um registro analítico das diretrizes regulatórias disponíveis em documentos oficiais públicos, contribuindo para a descrição de boas práticas regulatórias no que tange a inteligência artificial. Nesse sentido, boas práticas para a regulamentação dos sistemas de inteligência artificial incluem, por exemplo, utilizar conceitos claros e abrangentes, adotar medidas de proteção da concorrência e exigir transparência, explicabilidade, segurança, equidade e não discriminação (Leitão; Belchior, 2022).

O artigo científico 10¹⁴ ressalta o posicionamento de que o aprimoramento tecnológico tem auxiliado a desenvolver o campo político, econômico e social. Todavia, quando mal utilizado, acarretará retrocessos desastrosos para diversificadas extensões da sociabilidade. Trata-se, portanto, de um instrumento que não é bom ou mal em seu âmago, cabendo ao ser humano utilizá-la levando em consideração os valores que caracterizam o bem social. Para fundamentar a necessidade de uma correta manipulação dos sistemas algorítmicos, desvinculando-os de vieses discriminatórios, o artigo cita a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância (Feliciano et al., 2022).

O artigo científico 06¹⁵ produziu um inventário pautado por um levantamento bibliográfico sobre as principais reflexões de estudos acerca do viés algorítmico. Os

¹² O artigo 02 é “Inteligência artificial e discriminação: um panorama sobre a antagonização entre exclusão e o estado democrático de direito brasileiro à luz dos direitos da personalidade” (Soares et al., 2022).

¹³ O artigo 14 é “Diretrizes Regulatórias para sistemas de inteligência artificial: análise documental das iniciativas dos Estados Unidos e da União Europeia” (Leitão; Belchior, 2022).

¹⁴ O artigo 10 é “O capitalismo de vigilância e seus efeitos: discriminação algorítmica e reificação humana” (Feliciano et al., 2022).

¹⁵ O artigo 06 é “Viés algorítmico – um balanço provisório” (Simões-Gomes et al., 2020).

escritores analisam possíveis origens para o viés algorítmico, de modo a concluir que a busca pela subjetividade atrelada ao sistema é um fato recorrente. Todavia, apesar da subjetividade do desenvolvedor e da subjetividade do implementador serem importantes, pensar nelas como a única causa para a discriminação é insuficiente, de modo que negligenciaria a desigualdade estrutural presente nos próprios dados. Desse modo, os autores defendem que, para além da correção do viés subjetivo, o estudo sobre as lacunas existentes no debate sobre a reprodução das desigualdades sociais é indispensável (Simões-Gomes et al., 2020).

O artigo científico 16¹⁶ analisou a exclusão da produção de dados em relação às pessoas trans. Desta forma, observou-se que existem poucos dados e estudos sobre a discriminação às pessoas trans dentro das redes. Os autores defendem que esse fato se dá mediante um padrão histórico que excluem as pessoas trans, de modo que se faz necessário uma maior transparência dos dados para possibilitar um controle social mais efetivo, retirando os vieses cognitivos dos processos de aprendizagem dos algoritmos, resultando em melhorias para a corpo civil (Bomfim et al., 2023).

A revisão de literatura evidencia que a utilização de algoritmos para o auxílio às atividades humanas é um benefício cada vez mais indispensável para a comunidade. Todavia, diante de um enviesamento proveniente dos dados utilizados ou dos manuseadores da ferramenta, esse recurso benéfico se torna um potencial intensificar de segregações sociais.

Ademais, percebe-se a constante necessidade de criação e aplicação de regulamentações específicas que disciplinem a utilização dos sistemas algorítmicos e que exijam a transparência e a responsabilidade dos envolvidos com os resultados alcançados. Tais fatos precisam ser sanados para amenizar os conflitos éticos provenientes da discriminação algorítmica.

3 A Utilização de Algoritmos e o Fomento à Discriminação

A presente seção abordará a discriminação algorítmica na prática e problematizará a sua manutenção, por meio da exemplificação de casos internacionais e

¹⁶ O artigo 16 é “Abjeção de mercado: uma análise acerca da discriminação algorítmica de pessoas trans” (Bomfim et al., 2023).

nacionais em que se observou uma parcialidade algorítmica que produziu uma consequente disseminação de preconceitos.

Em artigo publicado na Michigan Institute Technology Review, pesquisadores da ciência de dados observaram padrões em inteligências artificiais geradoras de imagens. Por meio da aprendizagem de máquina (machine learning), os sistemas foram alimentados com um banco de imagens e o objetivo era identificar o quão enviesados são esses modelos a partir de solicitações fundamentadas em pequenas descrições como ‘uma mulher’ e ‘um encanador ambicioso’ (Leal, Paulo, 2023).

Com o uso da DALL-E 2 (sistema de inteligência artificial para criação de imagens), identificou-se que os resultados se voltavam à criação de 97% mais homens brancos quando o pedido objetivava a representação de pessoas em posições de autoridade, liderança ou poder. Uma vez que para a utilização da aprendizagem de máquina é necessário a utilização de um banco de dados, percebe-se que o problema não está na ferramenta em si e sim no panorama de desigualdade estrutural que fundamentou esses dados e tornou o mecanismo um potencial discriminador (Leal, Paulo, 2023).

Dando seguimento, em 2018, a Amazon descartou um modelo que utilizava para a escolha de currículos para novas contratações por ter apresentado um viés sexista. O dispositivo foi treinado a partir dos últimos 10 anos de dados sobre os currículos enviados, que em sua maior parte eram de homens, desenvolvendo uma discriminação perante as mulheres, uma vez que desprezava os currículos femininos e atribuía pontuações mais altas aos masculinos (Chaves Junior et al., 2023).

Nesse sentido, o AI INDEX, relatório independente que analisou o setor profissional tecnológico nos Estados Unidos, pontuou que 71% dos candidatos para vagas no campo da inteligência artificial eram homens, de forma que o algoritmo alimentado com essa amostragem de dados passou a identificar os homens como candidatos mais aptos. Diante do fato de que a aprendizagem de máquina se fundamenta em dados do passado, o algoritmo identificou que os homens, além de serem uma maioria, recebiam promoções e aumento salarial com maior frequência. De forma contrária, as mulheres, além de serem uma minoria, recebiam menos aumento salarial, indicando, estatisticamente, que os homens seriam as melhores contratações. Logo, a permanência desse enviesamento conserva e dissemina uma discriminação histórica e sexista (Chaves Junior et al., 2023).

Os exemplos supracitados indicam a insuficiente diversidade dos profissionais na indústria tecnológica e a predominância masculina e branca nesse setor. Dessa forma, o AI Now Institute, instituto de pesquisa sobre as implicações da inteligência artificial, realizou estudo constatando que somente 15% dos pesquisadores de IA do Facebook e 10% do Google são mulheres, fato que reforça a ausência de diversidade que preserva os preconceitos de gênero (Chaves Junior et al., 2023).

Nessa perspectiva, o Lensa, aplicativo que utiliza inteligência artificial para a geração de avatares de usuários mediante as suas fotos, recebeu acusações no que tange a criação de imagens sexualizadas de mulheres sem o consentimento prévio, em especial ao público feminino asiático, latino e negro. Além da discriminação algorítmica inicialmente identificada, essa sexualização possibilita assédio ou difamação, afetando a dignidade da pessoa humana e o direito de intimidade das pessoas afetadas (Leal, Paulo, 2023).

Dando prosseguimento às exemplificações, a Ferramenta de Rastreamento Familiar de Allegheny (Allegheny Family Screening Tool — AFST) previa a proteção do bem-estar infantil no Condado de Allegheny, Pensilvânia, utilizando um sistema estatístico preditivo para o auxílio ao processo de triagem de famílias pobres estadunidenses, gerando uma pontuação automática quando identificasse alerta de risco de abuso ou negligência infantil (Do Amaral et al., 2021).

Levando em consideração que o mecanismo apresentou quase três quartos dos casos como de negligência e que grande parcela dos indicadores de negligência infantil está associado a indicadores de pobreza, viabiliza-se que o sistema possui falhas de design que visa injustamente os pobres, uma vez que uma família desabrigada enfrentando desafios, por exemplo, seria considerada negligente (Do Amaral et al., 2021).

Um agravante observado é que o sistema é baseado somente em dados de famílias que são beneficiadas por serviços sociais públicos. Nesse panorama, famílias que recorrem a recursos públicos serão submetidas a uma ingerência maior da autoridade governamental, permitindo a interpretação de que a utilização de recursos públicos sinalizaria fraqueza, de forma a assumir que pais e mães auxiliados pelo serviço público representam riscos para os seus filhos (Do Amaral et al., 2021).

Prosseguindo, no que tange ao uso de mecanismos algorítmicos para reconhecimento facial, o Instituto Nacional de Padrões e Tecnologia dos Estados Unidos

(National Institute of Standards and Technology — NIST) realizou uma pesquisa em que se constatou taxas expressivas de erros de identificação das faces de afro-americanos, de asiáticos, de mulheres e de pessoas com idade avançada pelos sistemas comerciais de reconhecimento facial que identificavam suspeitos para os departamentos de polícia dos Estados Unidos (Do Amaral et al., 2021).

Em um contexto estadunidense em que os afro-americanos apresentam até 2,5 vezes mais probabilidades de sofrerem com a vigilância policial, bem como pelo menos duas vezes mais chances de serem aprisionados quando comparados a qualquer outra raça (Do Amaral et al., 2021), as falhas de reconhecimento facial apresentam aptidão para intensificar injustiças raciais. Esse potencial lesivo também é exemplificado pelo documentário *Coded Bias*, que demonstrou a dificuldade dos softwares de reconhecimento facial em identificar o rosto negro quando comparado ao rosto branco diante de uma baixa diversidade e má qualidade dos dados apurados (Leal; Paulo, 2023).

Dando continuidade à análise de casos de parcialidade algorítmica, cabe ressaltar a temática do policiamento preditivo, mecanismo que tem como fito a identificação de atividades criminosas em potencial por meio de investigações matemáticas. Diante das inovações tecnológicas, a utilização de algoritmos potencializa os softwares preditivos, permitindo aos sistemas determinarem pessoas e grupos com maiores chances de cometer ilícitos, resultando em uma vigilância policial diversas vezes desnecessária (De Freitas, 2023).

Nesse sentido, a ProPublica realizou uma pesquisa em que examinou mais de dez mil réus criminais no Condado de Broward, Flórida. Baseado pela comparação entre as taxas de reincidência previstas pelo algoritmo com as taxas reais, concluiu-se que os réus negros apresentavam 45% mais chances de serem classificados como de alto risco quando comparados com réus brancos, bem como possuíam duas vezes mais probabilidades de serem categorizados erradamente no que tange o risco de reincidência violenta (Chaves Junior et al., 2023).

O interessante a ser analisado é que os dados não estavam incorretos. Tendo em vista um contexto em que os negros representam somente 13% da população estadunidense, mas ocupam 40% das vagas em presídios, o algoritmo adotou como um padrão os negros serem mais perigosos. Cita-se que, segundo a Associação Americana para Liberdades Cívicas (ACLU), os homens negros recebem sentenças 20% maiores do

que homens brancos em crimes similares (Chaves Junior et al., 2023). As situações descritas reforçam o potencial segmentador do racismo algorítmico, fato este que não pode ser aceito em uma sociedade historicamente segmentadora.

Para a sociedade brasileira, marcada por uma seletividade penal racista, a utilização de modelos preditivos acabará incentivando o encarceramento em massa de grupos específicos. Contando com mais de 700 mil pessoas encarceradas, dentre as quais 64% são negras e 75% não realizaram o ensino médio, o Brasil possui a terceira maior população carcerária mundial (Silva, 2022).

Informações divulgadas no 14º Anuário Brasileiro de Segurança Pública, em outubro de 2020, relatou que a mortalidade entre pessoas negras é 183,2% superior à taxa observada entre brancos no tocante a intervenções policiais. Nesse sentido, a pesquisa também relatou que as políticas de encarceramento brasileiras possuem os negros como alvos preferenciais, de modo que a cada três pessoas encarceradas, duas são negras (Chaves Junior et al., 2023).

O setor de segurança pública do Brasil estreou a utilização de reconhecimento facial oficialmente em 2018. Em um país que possui 56% da população como autodeclarada negra, classe composta por pretos e pardos, a Rede de Observatórios da Segurança realizou levantamento demonstrando que, após um ano de experiências com reconhecimento facial para o monitoramento dos casos de prisões em cinco estados brasileiros, 90,5% dos presos a partir do reconhecimento facial foram negros (Kremer, 2023). Os estados da Bahia, Rio de Janeiro e Santa Catarina foram líderes no uso desse mecanismo (Silva, 2022).

Nesse viés, o Laboratório de Políticas Públicas e Internet (Lapin) realizou um estudo em que foi analisado as diferentes políticas públicas que utilizam do reconhecimento facial em suas atividades. Essa análise foi fundamentada por cinco requisitos básicos que auxiliam a transparência e a proteção dos direitos dos cidadãos (existência de regulação específica, adoção de boas práticas, publicação do número de erros e da acurácia, existência de mecanismos para proteger os direitos do titular e informações sobre as formas de aquisição da tecnologia), de modo que se constatou uma grave falha na implementação e uso dos instrumentos de inteligência artificial para a segurança pública, haja vista que nenhum dos casos analisados cumpriam ao menos dois dos requisitos básicos (Lima et al., 2024).

Em julho de 2019, jornais noticiaram que o mecanismo experimental de reconhecimento facial utilizado pela polícia do Rio de Janeiro indicou erroneamente uma mulher como procurada pela Justiça na orla de Copacabana. Descobriu-se, dias depois, que a verdadeira criminosa já estava presa há quatro anos, demonstrando que o banco de dados utilizado estava desatualizado e a abordagem policial enviesada (Kremer, 2023). A ineficiência dos sistemas de reconhecimento facial também foi visualizada na Micareta de Feira de Santana, na Bahia, momento em que foram coletados 1,3 milhões de rostos para o cumprimento de somente 18 mandados (Silva, 2022).

Diante do exposto, o policiamento preditivo fundamentado pelos algoritmos de reconhecimento facial é um tipo de tecnologia que reproduz vieses discriminatórios vinculados a uma estrutura social racista. Todavia, desde 2021, a utilização desses sistemas vem ganhando forças em solo brasileiro, de modo que o Instituto Igarapé, em mesmo ano, lançou edital para que cinco municípios testassem o seu novo dispositivo chamado CrimeRadar para previsão da ocorrência de delitos. Diante de diversas críticas do corpo cível e de ativistas, o projeto foi suspenso indeterminadamente (Kremer, 2023).

Diante do célere crescimento na implementação dos sistemas de reconhecimento facial, torna-se dificultoso que as regulamentações consigam acompanhar tal avanço e garantir a proteção dos direitos fundamentais, o que permite que os espaços públicos se tornem laboratórios de experimentação para essas tecnologias.

Cabe, portanto, a citação do relatório “O sertão vai virar mar: expansão do reconhecimento facial na Bahia” (2023) que demonstrou como a utilização das tecnologias de reconhecimento facial (TRF) se disseminaram por mais de 70 municípios da Bahia sem que fosse indicado claramente que esses sistemas permitiram a redução da criminalidade. Segundo os dados da Segurança Pública da Bahia, no período entre 2018 e 2022, os índices de criminalidade e violência não apresentaram mudanças significativas, demonstrando que esse avanço abrupto do uso de TRFs não resulta diretamente na redução do crime (Lima et al., 2024).

Ademais, no relatório “Das Planícies ao Planalto: como Goiás influenciou a expansão do reconhecimento facial na segurança pública brasileira” (2023) foi observado que o encorajamento à ‘modernização da segurança’ se avançou sem a devida preocupação com a responsabilização, prestação de contas e transparência no que tange o uso das tecnologias de reconhecimento facial. Em Goiás, itens básicos como quem gere

os dados e como eles são armazenados são questões não respondidas pelas prefeituras e órgãos responsáveis. Tal opacidade é exemplificada pelo fato de que 12 dos 37 municípios goianos que realizaram a aquisição de infraestrutura para o reconhecimento facial não possuíam em seus Portais de Transparência os contratos atrelados a essas atividades (Lima et al., 2024).

Viabilizando que a transparência e a visibilidade dos resultados do uso de TRF permitem a mensuração da eficiência da tecnologia e da política pública de segurança, a manutenção da negligência quanto à disponibilização de relatórios e indicadores fazem com que a população não consiga verificar se as políticas instauradas em seu estado ou município estão sendo efetivas, permite que a sociedade não averigue se o dinheiro público está sendo bem utilizado. Tais situações explicitam que, apesar de a Lei de Acesso à Informação disponibilizar meios para que seja aplicado o princípio da publicidade, o uso dos sistemas de inteligência artificial, especialmente os de reconhecimento facial, seguem como gargalos que demonstram uma falta de transparência deliberada (Lima et al., 2024).

Dados limitantes sobre raça e gênero para o desenvolvimento dos dispositivos implica em uma cegueira do algoritmo de reconhecimento facial. Não existe, pois, representatividade estatística. Além da evidente problemática de diversificação, não há parâmetros que disciplinem os softwares, de forma que eles não são auditados, fiscalizados nem revisados com o intuito de investigar a existência de enviesamentos incrustados. Não existe, portanto, regulamentação estabelecendo critérios sobre o nível de acurácia, a transparência sobre o uso e a proteção dos dados coletados (Chaves Junior et al., 2023).

Iniciativas emergentes no setor tecnológico estão surgindo com o intuito de possibilitar uma formação horizontal, um compartilhamento de saberes e um apoio mútuo, como a Minas Programam, a Conexão Malunga, o Kilombotech, o Perifacode, o Tecnogueto, o Afropython, o Afrotech, o Quebradev e a InspirAda. Na coordenação do Perifacode, a engenheira de software Carla Vieira defende o combate aos vieses discriminatórios, ressaltando a importância da diversidade na tecnologia com o intuito de ‘representar o mundo como ele é’ (Silva, 2022).

Por sua vez, Silvana Bahia, diretora do Olabi e idealizadora do PretaLab, liderou um levantamento de dados em que analisou a presença de mulheres negras e indígenas

no setor tecnológico, identificando que “a falta de representatividade é um problema não só para o ecossistema de tecnologia e inovação, mas para os direitos humanos e a liberdade de expressão” (Silva, 2022).

A parcialidade algorítmica que permite e ocasiona a criação de preconceitos automatizados em diferentes ambientes é evidente mediante os exemplos supracitados. Conclui-se, portanto, que o surgimento de instrumentos regulatórios que lidem especificamente com os algoritmos se faz indispensável e irremediável, uma vez que a permanência de discriminações de raça, gênero e afins provenientes de injustiças algorítmicas intensificam a opressão perante pessoas e grupos que historicamente são marginalizados.

4 Necessária Intervenção: análise de diretrizes regulatórias

O desenvolvimento tecnológico implica em diversos benefícios para a vida em sociedade, sendo um grande aliado para os Tribunais brasileiros, por exemplo, melhorando o ambiente jurídico uma vez que permite maior celeridade, efetividade e segurança jurídica. A sua utilização pode, inclusive, reduzir vieses cognitivos humanos atrelados a elementos externos que interferem na tomada de decisão judicial, como a fome, a fadiga e o humor (Chaves Junior et al., 2023).

Ressalta-se que os algoritmos utilizados na justiça brasileiro atuam, em relação ao auxílio à tomada de decisão, de modo sugestivo e não decisório. A supervisão humana para o uso de inteligência artificial no judiciário é um requisito exigido pelo Conselho Nacional de Justiça (Chaves Junior et al., 2023) e que foi implementado inicialmente pela Resolução n. 332/2020, exigindo-se, em caso de decisão automatizada, explicação satisfatória e passível de auditoria humana (Kaufman et al., 2023). Recentemente, o Conselho Nacional de Justiça aprovou o Ato Normativo 0000563-47.2025.2.00.0000, atualizando a Resolução n. 332/2020 e incorporando ao Poder Judiciário um novo conjunto de normas com o objetivo de acompanhar a célere evolução tecnológica.

Todavia, há situações em que a utilização de algoritmos reforçará um modelo dominante e desviado de pensamento (Feliciano et al., 2022), de modo que, por interagirem com a sociedade, sistemas algorítmicos enviesados que produzem valores sociais discriminatórios não podem ser mantidos. Há, portanto, a necessidade de pelo

menos o mínimo de governança associada com regulamentação que garantam os princípios da segurança, da transparência, da dignidade da pessoa humana, da não discriminação, da liberdade de escolha, da proteção da privacidade e da responsabilidade (Rossetti; Angelucci, 2021).

Levando-se em consideração que o processo de tomada de decisão dos algoritmos é difícil de ser analisado, não sendo possível verificar se elas são justas e imparciais, uma das principais problemáticas no que tange a discriminação algorítmica é a ausência de transparência do código-fonte e da base de dados utilizada para o treinamento do modelo (Leal; Paulo, 2023). A descrição de motivos do Anteprojeto de Lei de Proteção de Dados para Segurança Pública e Persecução Penal exemplifica tal situação:

[...] há um enorme déficit de proteção dos cidadãos, visto que não há regulação geral sobre a licitude, a transparência ou a segurança do tratamento de dados em matéria penal, tampouco direitos estabelecidos ou requisitos para utilização de novas tecnologias que possibilitam um grau de vigilância e monitoramento impensável há alguns anos. Apesar do crescimento vertiginoso de novas técnicas de vigilância e de investigação, a ausência de regulamentação sobre o tema gera uma assimetria de poder muito grande entre os atores envolvidos (Estado e cidadão). Nesse contexto, o titular dos dados é deixado sem garantias normativas mínimas e mecanismos institucionais aplicáveis para resguardar seus direitos de personalidade, suas liberdades individuais e até a observância do devido processo legal (Chaves Junior et al., 2023).

Nesse sentido, o Código de Defesa do Consumidor delineou um dos primeiros dispositivos jurídicos para a proteção de dados. Em seu artigo 43 observa-se a garantia de direitos relacionados aos elementos pessoais armazenados em bancos de dados, como (i) o direito de acesso a essas informações; (ii) o princípio da qualidade dos dados e (iii) o direito de retificação de dados imprecisos (Oliveira Paranhos, 2022).

A Lei Federal n.º 12.414/2011, ou Lei do Cadastro Positivo, também disciplina sobre a formação e consulta a banco de dados, estabelecendo os dados que poderão ser guardados (artigo 3º, §§s 2 e 3), os direitos do titular das informações (artigo 5º), as obrigações dos administradores dos bancos de dados (artigo 6º), bem como os responsáveis para as situações de danos (artigo 16) (Oliveira Paranhos, 2022).

Por sua vez, a Lei Federal n.º 12.527/2011, ou Lei de Acesso à Informação, disciplina o direito de receber informações de interesse particular ou coletivo dos órgãos públicos, assegurado pelo artigo 5º, inciso XXXIII, da Constituição Federal. Essa lei, semelhante às anteriores, exemplifica mais um mecanismo jurídico defensor do dever da transparência no tocante à atuação do poder público (Oliveira Paranhos, 2022).

Diante do exposto, apesar de ainda não existir uma legislação específica sobre inteligência artificial, observa-se a tendência do Estado em contestar essas situações discriminadoras que violam princípios constitucionais e que devem ser identificadas e combatidas, haja vista o histórico de violência contra grupos marginalizados (Soares et al., 2022). Cabe, portanto, analisar alguns dispositivos constitucionais que reprovam esse preconceito automatizado.

Inicialmente, no artigo 1º, inciso III, já se destaca a dignidade da pessoa humana como um dispositivo basilar para a República Federativa do Brasil. Logo após, no artigo 3º, inciso IV, é disciplinado que as pessoas serão tratadas sem preconceitos fundamentados em origem, raça, sexo, cor, idade e afins, em conformidade com o artigo 4º, inciso VIII, que disciplina o repúdio ao racismo (Soares et al., 2022).

Ademais, outras leis nacionais importantes contra a discriminação podem ser citadas: Lei do Racismo (Lei n.º 7.716/1989); Estatuto da Criança e do Adolescente (Lei n.º 8.069/1990); Estatuto do Idoso (Lei n.º 10.741/2003); Lei Maria da Penha (Lei n.º 11.340/2006); Estatuto da Pessoa com Deficiência (Lei n.º 13.146/2015) (Soares et al., 2022). Internacionalmente, pode-se mencionar que o artigo VII da Declaração Universal dos Direitos Humanos defende a proteção à igualdade e à não discriminação, existindo também outras declarações pertinentes: Declaração das Nações Unidas sobre a Eliminação de Todas as Formas de Discriminação Racial; Convenção sobre a Eliminação de todas as Formas de Discriminação contra a Mulher (1979); Convenção n. 111, da Organização Internacional do Trabalho (OIT), que dispõem sobre discriminação em matérias de emprego e ocupação (Kaufman et al., 2023).

A República Federativa do Brasil, inclusive, assinou a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância e a internalizou por meio do Decreto n.º 10.932, de 10 de janeiro de 2022. A Convenção anteriormente citada, em seu artigo 1º, afirma que a discriminação racial é:

[...] 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 humanos e liberdades fundamentais consagrados nos instrumentos internacionais aplicáveis aos Estados Partes (Feliciano et al., 2022).

Diante do fato de que o nosso arcabouço jurídico rejeita as formas de discriminação e vem criando dispositivos que combatam essas situações, salienta-se que o mundo da tecnologia não pode existir sem limites (Soares et al., 2022). Mesmo que por

um lado seja necessário proteger a liberdade dos agentes privados de promover o desenvolvimento da inovação tecnológica, essa liberdade deve ser utilizada nos limites da sua função social, observando-se uma sociedade justa (Oliveira Paranhos, 2022).

A aplicação de algoritmos de aprendizagem de máquina de forma contínua e aparentemente autônoma é um fato importante para a análise dos seus riscos e das estratégias que permitem a evasão de responsabilidade das empresas. A partir das ideias de que os algoritmos são segredos de negócio e apresentam um complexo processo de desenvolvimento que fomenta a sua tomada de decisões, as empresas argumentam que, portanto, os mecanismos algorítmicos não poderiam ser auditáveis ou terem seus impactos criticados (Silva, 2022).

Todavia, uma vez que as tecnologias interferem diretamente nos direitos sociais, políticos, civis, culturais e econômicos, o Estado Democrático de Direito brasileiro deverá disseminar o compromisso normativo com a proteção dos direitos humanos (Fernandes et al., 2022). Nesse panorama, pontua-se que a legislação do Brasil agiu em benefício da proteção dos dados pessoais quando, por meio da Emenda Constitucional nº 115/2022, elevou a proteção de dados pessoais ao status de direito fundamental, bem como, por intermédio da Lei Geral de Proteção de Dados (LGPD), criou instrumentos acerca do tratamento de dados pessoais, garantindo o direito de privacidade e o livre desenvolvimento da personalidade da pessoa natural (Leal; Paulo, 2023).

A LGPD também traz, em seu artigo 6º, o princípio da não discriminação, reforçando que os dados pessoais não poderão ser utilizados para fomentar a discriminação (Leal; Moreschi Paulo, 2023). O artigo 20, por sua vez, criou norma inédita no ordenamento jurídico brasileiro e pertinente ao uso dos algoritmos, afirmando que, em caso de tratamento automatizado de dados pessoais, o titular poderá requisitar a revisão de decisões quando estas afetarem os seus interesses. Além disso, o artigo 20, em seus parágrafos únicos, exige o fornecimento de informações claras e adequadas no tocante à escolha dos parâmetros e métodos usados para a decisão algorítmica, o que demonstra uma real preocupação do legislador com as injustiças provenientes do enviesamento dos algoritmos (Oliveira Paranhos, 2022).

Nesse contexto, o Projeto de Lei n.º 21/2020, proposto pela Câmara dos Deputados, disciplina sobre os princípios, direitos e deveres associados à utilização de inteligência artificial no Brasil. O dispositivo prevê regulamentações importantes sobre o

conflito existente entre empresas e corpo civil, viabilizando ‘a relevância da inteligência artificial para a inovação, o aumento da competitividade, o crescimento econômico sustentável e inclusivo e a promoção do desenvolvimento humano e social’ (art. 3º), com foco no fomento tecnológico, inovação e livre iniciativa (art. 4º, incisos I e II), mas sem deixar de lado o ‘respeito aos direitos humanos e aos valores democráticos’ (art. 4º, inciso III) e ‘a igualdade, a não discriminação, a pluralidade e o respeito aos direitos trabalhistas’ (art. 4º, inciso IV) (Oliveira Paranhos, 2022).

Além de normas principiológicas, o referido projeto também abordou regras pertinentes relativas aos princípios da governança: i) finalidade, de modo que a IA deve almejar resultados benéficos para a sociedade; ii) centralidade no ser humano, respeitando a dignidade humana, a privacidade e à proteção de dados; iii) não discriminação; iv) transparência e explicabilidade; v) segurança e vi) responsabilização e prestação de contas (Oliveira Paranhos, 2022). Portanto, essas diretrizes representam novamente a preocupação do legislador quanto à temática e são indispensáveis para a criação de um ambiente que permite o desenvolvimento tecnológico ao mesmo tempo que respeita os direitos humanos.

Todavia, o PL 21/2020 gerou diversas discussões pois estabelecia uma responsabilidade civil subjetiva para os agentes atuantes no desenvolvimento e operação dos sistemas de IA, o que exigiria a necessária comprovação de culpa ou dolo por parte do responsável pelo dano. Nesse sentido, em nota técnica, o Laboratório de Políticas Públicas e Internet ressaltou a dificuldade, quase impossibilidade, que a vítima teria para comprovar um eventual dano proveniente de sistemas de inteligência artificial, uma vez que eles apresentam comportamentos imprevisíveis. Desse modo, a complexidade algorítmica, que tornaria a comprovação de um dano quase impossível, somado a um regime pautado por uma responsabilidade subjetiva, acarretaria uma situação de desamparo para as vítimas (Melo, 2024).

Diante disso, em 07 de fevereiro de 2022, foi emitido o Ato do Presidente do Senado Federal (ATS) nº 4, de 2022, que estabeleceu uma Comissão de Juristas (CJSUBIA) responsáveis pela elaboração de uma minuta de substitutivo para os Projetos de Lei nºs 5.051/2019, 21/2020 e 872/2021. Essa comissão tinha como objetivo convocar diversos especialistas e juristas tanto nacionais quanto internacionais para discutir a temática da IA. Destaca-se que, nas audiências, a responsabilização subjetiva prevista no PL 21/2020 foi alvo de muitas críticas (Manso; Junqueira Calixto, 2024).

Nessa conjuntura, criou-se, então, o Projeto de Lei 2338/2023, que além de disciplinar sobre o uso das inteligências artificiais, apresentou novidades no que tange a responsabilização civil por meio dos seus artigos 27 a 29. O caput do artigo 27 prevê que o fornecedor ou operador do sistema de IA que ocasionar dano é obrigado a repará-lo integralmente, de modo que o parágrafo primeiro do referido artigo salienta a responsabilidade objetiva do agente quando se tratar de sistema de alto risco ou risco excessivo e o parágrafo segundo salienta que nas demais ocasiões a culpa do agente causador do dano será presumida, aplicando-se a inversão do ônus da prova em favor da vítima (Melo, 2024).

Cabe também ressaltar a importância de se efetivar os princípios dispostos no artigo 3º do PL 2338/2023, como a transparência, a explicabilidade, a inteligibilidade, a auditabilidade, a rastreabilidade das decisões e a prestação de contas, que auxiliam para a compreensão da atividade desenvolvida por sistemas de inteligência artificial e a individualização da responsabilidade de cada um dos envolvidos nesse processo (Manso; Junqueira Calixto, 2024).

Ante o exposto, é possível enumerar três pontos fundamentais para garantir decisões algorítmicas mais justas e imparciais, resultando na redução de discriminações. O primeiro deles diz respeito à qualidade dos dados empregados no modelo algorítmico, ou seja, os dados precisam ser imparciais, relevantes e precisos, com o fito de representarem a diversidade da sociedade e não perpetuarem estereótipos (Leal; Paulo, 2023).

O segundo ponto diz respeito à necessária transparência sobre como operam os algoritmos, visto que os seus processos ainda são uma caixa-preta, exigindo-se uma atuação regulatória a partir do conflito existente entre o direito à propriedade intelectual da empresa desenvolvedora e o direito à transparência, à fundamentação e à não discriminação dos algoritmos. O terceiro ponto crucial é o cuidado com a ética das empresas de desenvolvimento, de modo que elas sejam responsabilizadas pelos resultados dos seus produtos, aplicando parâmetros de ESG (Environmental, Social and Governance) para incentivar a diversidade às suas equipes (Leal; Paulo, 2023).

Esses três pontos ressaltam a importância de uma regulamentação eficaz e que permita que os processos sejam auditáveis, buscando o equilíbrio entre os interesses

comerciais e a valorização do ser humano enquanto utilizador e consumidor final dos algoritmos (Leal; Paulo, 2023).

Conclui-se, portanto, que para além do combate à discriminação algorítmica e da exigência de que empresas desenvolvam culturas que possibilitem a diversidade em suas equipes, faz-se imprescindível a adoção de leis e regulamentos sobre a criação e utilização de algoritmos, de modo a reduzir o enviesamento desses recursos tecnológicos e garantir a transparência, a responsabilidade e a justiça. Salienta-se, por fim, que a discriminação é um mal humano que pode ser reproduzida para os algoritmos, mas as raízes desses preconceitos não estão em softwares ou hardwares e sim na visão obsoleta que não enxerga a pluralidade e a diversidade como a essência do que nos faz humanos. Deve-se, portanto, combater quaisquer formas de discriminação, almejando uma comunidade cada vez mais unida, plural e livre de injustiças (Leal; Paulo, 2023).

5 Considerações Finais

A pesquisa desenvolvida ao longo deste trabalho evidenciou que a inteligência artificial e os algoritmos, apesar do seu benéfico uso para a automação de processos e otimização de atividades humanas, são potenciais instrumentos para a reprodução e amplificação de desigualdades sociais. Uma vez que o treinamento dos modelos é alimentado por dados que podem refletir vieses históricos, a discriminação algorítmica possibilita que aspectos discriminatórios sobre grupos vulneráveis sejam reforçados, perpetuando a manutenção de estereótipos e preconceitos.

Os casos citados demonstram que a ausência de transparência e de regulamentação, bem como a dificuldade de fiscalizar os sistemas algorítmicos contribuem para a conservação de um cenário discriminatório. Nesse sentido, o enviesamento algorítmico pode ser facilmente encontrado em processos seletivos, no uso do reconhecimento facial e no policiamento preditivo, o que demonstra como a utilização de sistemas de inteligência artificial sem uma supervisão efetiva pode resultar em injustiças sociais e em violações de direitos fundamentais.

Nessa conjuntura, faz-se notória a necessidade de desenvolver normativas cada vez mais eficazes para combater a discriminação algorítmica, garantindo que o

desenvolvimento e a operação dos algoritmos sejam auditáveis e alinhados aos princípios da transparência, da equidade e da não discriminação.

Viabiliza-se no corpo normativo brasileiro avanços na regulamentação da temática, situação exemplificada pelos Projetos de Lei 21/2020 e 2338/2023, todavia, o incentivo constante e crescente para a adoção de práticas éticas e a promoção da diversidade tanto na base de dados quanto nas equipes que desenvolverão e utilizarão os algoritmos são medidas indispensáveis para amenizar os efeitos negativos do uso indevido da inteligência artificial na sociedade.

Ademais, é necessário um esforço conjunto entre juristas, legisladores, cientistas da computação e demais especialistas com o fito de estabelecer diretrizes que equilibrem a inovação tecnológica e a proteção dos direitos fundamentais. O avanço no uso da inteligência artificial e dos sistemas algorítmicos deve ser permeado por responsabilidade, assegurando que seu impacto seja positivo e promova um corpo social mais justo e inclusivo.

Ressalta-se, por fim, que a discriminação não é um aspecto inerente aos algoritmos, mas sim um reflexo das falhas humanas presentes nos dados que os alimentam. Nesse sentido, o combate ao enviesamento algorítmico transcende o aprimoramento técnico dos instrumentos tecnológicos, exigindo uma transformação cultural que promova a equidade e a justiça para todos os indivíduos.

REFERÊNCIAS

- BOMFIM, R.; JANUZZI, G.; MÁXIMOTEODORO, M. C. *Abjeção de mercado: uma análise acerca da discriminação algorítmica de pessoas trans*. **Revista Videre**, [S. l.], v. 15, n. 32, p. 285–303, 2023. DOI: 10.30612/videre.v15i32.17130. Disponível em: <https://ojs.ufgd.edu.br/videre/article/view/17130>. Acesso em: 20 ago. 2024.
- CHAVES JUNIOR, A.; GUASQUE, B.; DE PADUA, T. S. A. *Segregação racial e vieses algorítmicos: máquinas racistas no âmbito do controle penal*. **Revista Brasileira de Direito**, Passo Fundo, v. 19, n. 2, p. e4768, set. 2023. DOI: <https://doi.org/10.18256/2238-0604.2023.v19i2.4768>. Disponível em: <https://seer.atitus.edu.br/index.php/revistadedireito/article/view/4768>. Acesso em: 20 ago. 2024.
- DE FREITAS, I. B. *Policiamento Preditivo: aspectos discriminatórios no uso das novas tecnologias*. **Revista de Direito, Governança e Novas Tecnologias**, [S. l.], v. 9, n. 1, p. 113-129, 2023. DOI: <https://doi.org/10.26668/IndexLawJournals/2526-0049/2023.v9i1.9732>. Disponível em: <https://indexlaw.org/index.php/revistadgnt/article/view/9732/pdf>. Acesso em: 20 ago. 2024.
- DO AMARAL, A. J.; SANTOS ELESBÃO, A. C.; DA VEIGA DIAS, F. *Gubernamentalidad algorítmica y nuevas prácticas punitivas*. **Derechos en Acción**, [S. l.], v. 20, n. 20, p. 549, 2021. DOI: 10.24215/25251678e549. Disponível em: <https://revistas.unlp.edu.ar/ReDeA/article/view/12557>. Acesso em: 20 ago. 2024.
- FELICIANO, G. G.; NASPOLINI, S. H. D. F.; FOGAROLLI FILHO, P. R.; DE AMORIM SOUZA, D. *O capitalismo de vigilância e seus efeitos: discriminação algorítmica e reificação humana*. **Revista de Direito Brasileira**, [S. l.], v. 33, n. 12, p. 309-332, 2022. DOI: <https://doi.org/10.26668/IndexLawJournals/2358-1352/2022.v33i12.9167>. Disponível em: <https://www.indexlaw.org/index.php/rdb/article/view/9167/6784>. Acesso em: 20 ago. 2024.
- FERNANDES, A. L.; BRANCO, C.; MENDES, C.; PEREIRA, L.; SANTANA, L.; SARAIVA, R.; ALEXANDRE, R. Contribuição à consulta pública da Comissão de Juristas do Senado responsável por subsidiar a elaboração de substitutivo sobre inteligência artificial no Brasil. **Instituto de Pesquisa em Direito e Tecnologia do Recife – Ip.rec**, 2022. Disponível em: <https://ip.rec.br/publicacoes/contribuicao-a-consulta-publica-da-comissao-de-juristas-do-senado-responsavel-por-subsidiar-a-elaboracao-de-substitutivo-sobre-inteligencia-artificial-no-brasil-cjsubia/>. Acesso em: 10 ago. 2024.

KAUFMAN, D.; JUNQUILHO, T.; REIS, P. *Externalidades negativas da inteligência artificial: conflitos entre limites da técnica e direitos humanos*. **Revista de Direitos e Garantias Fundamentais**, [S. l.], v. 24, n. 3, p. 43–71, 2023. DOI: 10.18759/rdgf.v24i3.2198. Disponível em: <https://sisbib.emnuvens.com.br/direitosegarantias/article/view/2198>. Acesso em: 20 ago. 2024.

KREMER, B. **Racismo Algorítmico**. Rio de Janeiro: CESeC, 2023. Disponível em: <https://www.opanoptico.com.br/#publicacoes>. Acesso em: 10 ago. 2024

LEAL, M. C. H.; PAULO, L. M. *Algoritmos discriminatórios e jurisdição constitucional: os riscos jurídicos e sociais do impacto dos vieses nas plataformas de inteligência artificial de amplo acesso*. **Revista de Direitos e Garantias Fundamentais**, [S. l.], v. 24, n. 3, p. 165–187, 2023. DOI: 10.18759/rdgf.v24i3.2311. Disponível em: <https://sisbib.emnuvens.com.br/direitosegarantias/article/view/2311>. Acesso em: 20 ago. 2024.

LEITÃO, R. G.; BELCHIOR, W. S. *Diretrizes Regulatórias para sistemas de inteligência artificial: análise documental das iniciativas dos Estados Unidos e da União Europeia*. **Revista Eletrônica Direito e Sociedade – REDES**, [S. l.], v. 10, n. 3, p. 187–204, 2022. DOI: <https://doi.org/10.18316/redes.v10i3.9097>. Disponível em: <https://revistas.unilasalle.edu.br/index.php/redes/article/view/9097>. Acesso em: 20 ago. 2024.

LIMA, T. et al. *Vigilância por lentes opacas: mapeamento da transparência e responsabilização nos projetos de reconhecimento facial no Brasil*. Rio de Janeiro: CESeC, 2024. Disponível em: <https://www.opanoptico.com.br/#publicacoes>. Acesso em: 15 fev. 2025.

MANSO, A. M.; JUNQUEIRA CALIXTO, M. Os regimes de responsabilização civil no novo Marco Legal para Inteligência Artificial: o Projeto de Lei nº 2.338/23: The civil liability regimes in the new Legal Framework for Artificial Intelligence: Bill No. 2,338/23. **International Journal of Digital Law**, Belo Horizonte, v. 4, n. 3, p. 111–129, 2024. DOI: 10.47975/digital.law.vol.4.n.3.manso. Disponível em: <https://journal.nuped.com.br/index.php/revista/article/view/1261>. Acesso em: 15 fev. 2025.

MELO, G. da S. *Inteligência Artificial e responsabilidade civil: uma análise do anteprojeto do Marco Legal da Inteligência Artificial e do Projeto de Lei 2338/2023*. **Revista IBERC**, Belo Horizonte, v. 7, n. 1, p. 49–65, 2024. DOI: 10.37963/iberc.v7i1.271. Disponível em: <https://revistaiberc.responsabilidadecivil.org/iberc/article/view/271>. Acesso em: 15 fev. 2025.

OLIVEIRA PARANHOS, M. C. **Viés Algorítmico: uma análise sobre discriminações automatizadas**. Rio de Janeiro: Lumen Juris, 2022.

RODRIGUES, H. W.; GRUBBA, L. S. **Pesquisa Jurídica Aplicada**. Habitus Editora, 2023. Disponível em: <https://www.habituseditora.com.br/index.php?q=ed23>. Acesso em 10 ago. 2024.

ROSSETI, R.; ANGELUCI, A. *Ética Algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação*. **Galáxia (São Paulo)**, [S. l.], n. 46, p. e50301, 2021. DOI: <https://doi.org/10.1590/1982-2553202150301>. Disponível em: <https://www.scielo.br/j/gal/a/R9F45HyqFZMpQp9BGTfZnyr/?format=pdf&lang=pt>. Acesso em: 20 ago. 2024.

SILVA, T. **Racismo algorítmico: inteligência artificial e discriminação nas redes digitais**. São Paulo: Edições Sesc São Paulo, 2022.

SIMÕES-GOMES, L.; ROBERTO, E.; MENDONÇA, J. *Viés algorítmico – um balanço provisório*. **Estudos de Sociologia**, Araraquara, v. 25, n. 48, 2020. DOI: 10.52780/res.13402. Disponível em: <https://periodicos.fclar.unesp.br/estudos/article/view/13402>. Acesso em: 20 ago. 2024.

SOARES, M. N.; CENTURIÃO, L. F. .; TOKUMI, C. A. L. *Inteligência artificial e discriminação: um panorama sobre a antagonização entre exclusão e o estado democrático de direito brasileiro à luz dos direitos da personalidade*. **Revista Direitos Sociais e Políticas Públicas (UNIFAFIBE)**, [S. l.], v. 10, n. 2, p. 567–597, 2022. DOI: 10.25245/rdspp.v10i2.1311. Disponível em: <https://portal.unifafibe.com.br:443/revista/index.php/direitos-sociais-politicas-pub/article/view/1311>. Acesso em: 20 ago. 2024.

APÊNDICE I – IDENTIFICAÇÃO DOS TRABALHOS INCLUÍDOS PELA REVISÃO DE LITERATURA

Nº	TÍTULO	ANO	AUTORIA
1.	“Governamentalidade algorítmica e novas práticas punitivas”	2021	Do Amaral, A. J.; Santos Elesbão, A. C.; Da Veiga Dias, F.
2.	“Inteligência Artificial e discriminação: um panorama sobre a antagonização entre exclusão e o estado democrático de direito brasileiro à luz dos direitos de personalidade”	2022	Soares, M. N.; Centurião, L. F.; Tokumi, C. A. L.
3.	“Algoritmos discriminatórios e jurisdição constitucional: os riscos jurídicos e sociais do impacto dos vieses nas plataformas de inteligência artificial de amplo acesso”	2023	Leal, M. C. H.; Paulo, L. M.
4.	“Policiamento preditivo: aspectos discriminatórios no uso das novas tecnologias”	2023	De Freitas, I. B.
5.	“Ética algorítmica: questões e desafios éticos do avanço tecnológico da sociedade da informação”	2021	Rossetti, R.; Angelucci, A.
6.	“Viés algorítmico – um balanço provisório”	2020	Simões-Gomes, L.; Roberto, E.; Mendonça, J.
7.	“Segregação racial e vieses algorítmicos: máquinas racistas no âmbito do controle penal”	2023	Chaves Junior, A.; Guasque, B.; De Padua, T. S. A.
8.	“Externalidades negativas da inteligência artificial: conflitos entre limites da técnica e dos direitos humanos”	2023	Kaufman, D.; Junquillo, T.; Reis, P.
10.	“O capitalismo de vigilância e seus efeitos: discriminação algorítmica e reificação humana”	2022	Feliciano, G. G.; Naspolini, S. H. D. F.; Fogarolli Filho, P. R.; De Amorim Souza, D.
14.	“Diretrizes regulatórias para sistemas de inteligência artificial: análise documental das iniciativas dos Estados Unidos e União Europeia”	2022	Leitão, R. G.; Belchior, W. S.
16.	“Abjeção de mercado: uma análise acerca da discriminação algorítmica de pessoas trans”	2022	Bomfim, R.; Januzzi, G.; Máximoteodoro, M. C.

Desenvolvida pelo autor em 2025

APÊNDICE II – IDENTIFICAÇÃO DOS TRABALHOS EXCLUÍDOS DA REVISÃO DE LITERATURA

Nº	TÍTULO	ANO	AUTORIA
9.	“O estado da arte da aplicação da inteligência artificial no esfregaço de sangue periférico”	2023	De Carvalho et al.
11.	“Efetivação de direitos fundamentais e difusão do conhecimento empírico”	2023	Freire, A. R. S.; Vieira, T.
12.	“Medidas de densidade e lacunaridade na discriminação da forma urbana: o caso das zeis em Campina Grande, PB”	2022	Barros Filho, M. N. M.; Simões, M. B.; Silva, Y. S. A.
13.	“Acurácia do Ecocardiograma Transtorácico como Método de Triagem na Prática Clínica da Investigação da Hipertensão Pulmonar”	2023	Rezende, C. F.; Mancuzo, E. V.; Nunes, M. C. P.; Corrêa, R. A.
15.	“Utilização de imagens SAR na classificação de formações florestais brasileiras”	2021	De Jesus, J. B.; Kuplich, T. M.
17.	“Variações na forma do otólio sagitta de <i>Coryphaena hippurus</i> (Actinopterygii: Coryphaenidae) em uma área de ressurgência na costa sudoeste do Oceano Atlântico”	2020	Almeida, P. R. C.; Monteiro-Neto, C.; Tubino, R. A.; Costa, M. R.
18.	“Discriminação de Classes de Cobertura Vegetal em uma Região de Transição Amazônica/Cerrado no Estado do Mato Grosso por meio de Imagens do Satélite ALOS-2/PALSAR-2”	2020	Silva, V. S.; Sano, E. E.; De Almeida, T.; De Mesquita Júnior, H. N.

Desenvolvida pelo autor em 2025