



Racismo Algorítmico e Discriminação Automatizada

Algorithmic Racism and Automated Discrimination

André Felipe Cardoso Cruz Veloso

Mestrando em Função Social do Direito pela Faculdade Autônoma de Direito de São Paulo (2025-Presente). Pós-graduado em Direito Médico e Bioética pela Faculdade de Ciências Médicas da Santa Casa de São Paulo (2024-2025) e especialista em Direito Médico pelo Instituto Paulista de Direito Médico e da Saúde (2022). Graduado em Direito pela Universidade Federal do Amazonas (2024). Integrou o quadro de Advogados do Conselho Regional de Medicina do Estado do Amazonas (2024-2025). Atualmente, exerce a função de Auxiliar de Gabinete da Presidência do Tribunal de Justiça do Amazonas TJAM, com atuação voltada a procedimentos licitatórios (2025- Presente).

Resumo: O presente estudo analisa o fenômeno do racismo algorítmico no contexto dos avanços tecnológicos contemporâneos, com ênfase na expansão da internet e na crescente utilização de sistemas automatizados baseados em algoritmos. Embora tais tecnologias tenham ampliado significativamente o acesso à informação e promovido maior integração social, cultural e acadêmica, persistem desafios relevantes relacionados à desigualdade no acesso aos recursos digitais e à presença de vieses nos sistemas algorítmicos. Esses vieses, frequentemente decorrentes da qualidade dos dados de treinamento e da ausência de diversidade em seu desenvolvimento, resultam em discriminações sistemáticas, sobretudo contra pessoas negras. O estudo evidencia que tais sistemas, longe de serem neutros, podem reproduzir e intensificar desigualdades históricas, afetando áreas sensíveis como recrutamento, concessão de crédito e segurança pública. Diante disso, destaca-se a necessidade de medidas regulatórias, maior transparência e estratégias de mitigação que assegurem o uso ético, inclusivo e responsável das tecnologias digitais.

Palavras-chave: racismo algorítmico; discriminação automatizada; algoritmos; desigualdade racial; tecnologia e sociedade.

Abstract: This study analyzes the phenomenon of algorithmic racism in the context of contemporary technological advancements, with an emphasis on the expansion of the internet and the increasing use of automated systems based on algorithms. Although such technologies have significantly expanded access to information and promoted greater social, cultural, and academic integration, important challenges remain regarding unequal access to digital resources and the presence of biases in algorithmic systems. These biases, often resulting from the quality of training data and the lack of diversity in their development, lead to systematic discrimination, particularly against Black individuals. The study demonstrates that such systems, far from being neutral, can reproduce and intensify historical inequalities, affecting sensitive areas such as recruitment, credit approval, and public security. Therefore, the need for regulatory measures, greater transparency, and mitigation strategies is emphasized to ensure the ethical, inclusive, and responsible use of digital technologies.

Keywords: algorithmic racism; automated discrimination; algorithms; racial inequality; technology and society.

INTRODUÇÃO

Não há como negar que, nos últimos anos, os avanços tecnológicos — com especial destaque para a expansão da internet — têm representado um progresso

significativo em diversos campos, sobretudo na vida em sociedade. Esse cenário se evidencia, principalmente, pela ampliação das possibilidades de integração cultural, acadêmica e artística, muitas vezes de forma gratuita, por meio de sites e plataformas voltadas ao compartilhamento de produções e criações. Tal dinâmica é ainda potencializada pela facilidade de troca instantânea de informações, ideias e referências, contribuindo para a formação de redes colaborativas e para a democratização — ainda que parcial — do conhecimento.

Entretanto, embora tais avanços constituam marcos relevantes no mundo contemporâneo, persistem desafios expressivos, especialmente — mas não exclusivamente — no que se refere ao acesso a essas informações e dados. Isso se deve, em grande medida, à distribuição desigual de recursos tecnológicos, que limita a inclusão digital de parcelas significativas da população. Mais recentemente, soma-se a esse quadro o problema dos vieses algorítmicos, que reproduzem e, por vezes, intensificam padrões e construções distorcidas da realidade, frequentemente apresentando maiores taxas de erro em relação a pessoas negras. Esse fenômeno não apenas contribui para a disseminação de informações imprecisas, mas também reforça estigmas e desigualdades históricas, consolidando o que se convencionou denominar racismo algorítmico.

A relevância dessa problemática torna-se ainda mais evidente ao se considerar que a tecnologia passa a permear, de forma quase integral, todas as esferas da vida social. As decisões automatizadas baseadas em algoritmos tornam-se cada vez mais frequentes, abrangendo desde processos aparentemente simples — como a triagem de currículos em recrutamento e seleção ou a análise para concessão de crédito — até aplicações mais complexas e sensíveis, como sistemas de reconhecimento facial, mecanismos de recomendação e moderação automatizada de conteúdo. Em todos esses casos, há impactos concretos na vida dos indivíduos, muitas vezes sem a devida transparência ou possibilidade de contestação.

Nesse contexto, a utilização de algoritmos matemáticos para o processamento e a análise de dados não constitui novidade, sendo, na verdade, uma prática consolidada e continuamente aprimorada ao longo dos anos. O que se observa atualmente, contudo, é que a eficácia desses sistemas está diretamente condicionada à qualidade dos dados utilizados em seu treinamento e à diversidade das interações com usuários reais. Esse fator influencia decisivamente sua capacidade de tomada de decisão e de geração de respostas, o que explica por que tais algoritmos frequentemente reproduzem e amplificam desigualdades já presentes na sociedade, como racismo, sexismo e misoginia. Longe de serem neutros, esses sistemas refletem contextos sociais e históricos específicos, podendo reforçar estruturas de exclusão quando não submetidos a mecanismos adequados de controle e revisão.

Diante disso, não é exagero afirmar que, caso não sejam implementadas medidas corretivas eficazes — como maior transparência, auditorias independentes e diversidade nos processos de desenvolvimento tecnológico —, tais sistemas poderão contribuir para a perpetuação e o agravamento das desigualdades sociais e raciais existentes, inclusive em escala ampliada.

Assim, nos tópicos seguintes, procede-se à análise mais aprofundada desse fenômeno, abordando seus conceitos, impactos, desafios, bem como as perspectivas de regulação e estratégias de mitigação, com vistas à construção de um uso mais ético, inclusivo e responsável das tecnologias digitais.

RACISMO ALGORÍTMICO: A PERSISTÊNCIA DO RACISMO SOB NOVAS FORMAS TECNOLÓGICAS

De início, antes de examinar propriamente o fenômeno do racismo algorítmico, é fundamental compreender, de maneira prévia, o conceito de racismo em si. Para tanto, impõe-se recorrer aos fundamentos estabelecidos pela Constituição Federal de 1988, especialmente em seu artigo 3º, inciso IV, que elenca como um dos objetivos fundamentais da República a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor, idade ou quaisquer outras formas de discriminação. No mesmo sentido, o artigo 4º, inciso VIII, dispõe que as relações internacionais do Brasil são orientadas, entre outros princípios, pelo repúdio ao terrorismo e ao racismo.

Nesse contexto, Silvio Luiz de Almeida (2018, p. 25) conceitua o racismo da seguinte forma: “Racismo é uma forma sistemática de discriminação que tem a raça como fundamento, e que se manifesta por meio de práticas conscientes ou inconscientes que culminam em desvantagens ou privilégios, a depender do grupo racial ao qual pertençam”.

A discriminação, por sua vez, representa a exteriorização concreta desse processo, caracterizando-se pelo tratamento desigual dirigido a indivíduos ou grupos em razão de sua raça, em manifesta violação aos princípios da igualdade e da dignidade da pessoa humana. Já o preconceito situa-se em uma dimensão subjetiva, consistindo em uma predisposição interna, uma concepção prévia frequentemente alicerçada em estereótipos e generalizações. Embora o preconceito, em si, não produza necessariamente efeitos diretos no plano fático, ele frequentemente atua como substrato das práticas discriminatórias, na medida em que, ao ser externalizado, converte-se em ações que geram exclusão, marginalização e desigualdade.

A distinção entre esses conceitos permite compreender que um indivíduo pode, em tese, nutrir convicções preconceituosas sem que isso se traduza imediatamente em condutas discriminatórias. Todavia, quando tais concepções são efetivamente materializadas em práticas sociais, seus efeitos tornam-se significativamente mais gravosos, repercutindo diretamente no acesso a direitos fundamentais, oportunidades e condições dignas de existência.

Importa destacar que o racismo está intimamente vinculado ao conceito de raça, o qual, embora desprovido de fundamento biológico capaz de sustentar hierarquizações humanas, constitui uma construção histórica, social e cultural. Ao longo do tempo, essa construção foi instrumentalizada como mecanismo de legitimação de relações de poder e dominação, sustentando narrativas fictícias que atribuem características negativas a determinados grupos raciais. Entre tais

construções, destaca-se a persistência — ainda que frequentemente dissimulada — de estereótipos que associam pessoas negras à inferioridade intelectual, à violência ou à incapacidade civilizatória, como se tais atributos fossem inerentes à sua condição racial.

Essas narrativas integram um sistema mais amplo de reprodução de desigualdades, no qual a branquitude é frequentemente tomada como padrão normativo. Nesse cenário, saberes, trajetórias e contribuições de pessoas negras são sistematicamente invisibilizados ou deslegitimados, contribuindo para a manutenção de estruturas excludentes.

Cumprе ressaltar, ademais, que o racismo não constitui fenômeno recente. Ao contrário, tem sido amplamente estudado por diferentes campos do conhecimento, como a sociologia, a história e a psicanálise, que evidenciam sua permanência e adaptação ao longo do tempo. No contexto brasileiro, sua compreensão demanda necessariamente a consideração do legado de mais de três séculos de escravidão, período que moldou profundamente as estruturas sociais, econômicas e simbólicas do país, estabelecendo hierarquias raciais e delimitando espaços de pertencimento e exclusão.

Tal herança histórica mostra-se essencial para a análise contemporânea do tema, sobretudo quando confrontada com dados empíricos, como os produzidos pelo IBGE¹ que indicam que mais da metade da população brasileira se autodeclara negra ou parda. Ainda assim, as desigualdades raciais persistem de forma significativa, revelando a continuidade de práticas discriminatórias, agora também mediadas por novas tecnologias.

É nesse contexto que emerge o racismo algorítmico, entendido como a reprodução e, em muitos casos, a amplificação de vieses raciais por sistemas automatizados. Alimentados por bases de dados historicamente marcadas por desigualdades, tais sistemas tendem a replicar padrões discriminatórios preexistentes, projetando no ambiente digital as mesmas distorções estruturais observadas na sociedade. Dessa forma, embora se apresente como um fenômeno contemporâneo, o racismo algorítmico possui raízes profundas, estando intrinsecamente ligado a processos históricos que continuam a influenciar, de maneira decisiva, as dinâmicas sociais atuais.

O VÉU DA OPACIDADE: DESIGUALDADES E INVISIBILIDADE NAS DECISÕES ALGORÍTMICAS

Não obstante, para além do problema anteriormente destacado — consubstanciado no enviesamento de dados —, impõe-se analisar, com maior

1 De acordo com o último censo demográfico, de 2010, a população brasileira é constituída majoritariamente por negros; são 15 milhões de autodeclarados pretos (7,6%), somados a 82 milhões de pardos (43,1%). Na outra parcela da população são 91 milhões que se classificaram como brancos (47,7%), dois milhões como amarelos (1,1%) e 817 mil indígenas (0,4%). Dados de 2015, da Pesquisa Nacional por Amostra de Domicílio (PNAD), já revelam um incremento do número de autodeclarados negros em relação ao último censo, 53,9% das pessoas se declaravam de cor/raça preta ou parda (IBGE, 2016).

profundidade, a ausência de transparência que permeia as inteligências artificiais e os ambientes digitais contemporâneos. Tal opacidade não constitui um aspecto meramente técnico, mas um obstáculo concreto à efetivação de direitos fundamentais. Na prática, isso significa que indivíduos que eventualmente sofram discriminações algorítmicas encontram enormes dificuldades em reconstruir o percurso decisório que levou ao resultado, sendo, portanto, incapazes de comprovar, de forma clara e objetiva, que foram alvo de tratamento discriminatório, seja ele direto ou indireto.

Essa dificuldade decorre, em grande medida, da distância existente entre a atuação dos programadores e o comportamento efetivamente apresentado pelos sistemas de *machine learning*. Diferentemente de sistemas tradicionais, nos quais há uma correspondência mais evidente entre código e resultado, os modelos de aprendizado de máquina desenvolvem padrões próprios a partir dos dados que lhes são fornecidos. Isso faz com que até mesmo seus desenvolvedores enfrentem limitações relevantes ao tentar descrever, com precisão, os caminhos percorridos pelo algoritmo até a formação de determinada decisão.

Nesse contexto, ainda que vieses possam ser identificados — sobretudo por aqueles que detêm conhecimento técnico mais aprofundado —, persiste uma barreira estrutural que contribui para a reprodução e o agravamento de desigualdades já existentes. Trata-se do fenômeno conhecido como “caixas-pretas” (*black boxes*) decisórias, expressão que designa sistemas cujo funcionamento interno é inacessível ou dificilmente compreensível. Tal opacidade compromete não apenas a auditabilidade dos sistemas, mas também a possibilidade de responsabilização, criando um ambiente propício à naturalização de decisões potencialmente injustas.

Diante desse cenário, a superação da opacidade exige a adoção de medidas que promovam transparência desde a própria concepção dos algoritmos. Não se trata apenas de disponibilizar informações técnicas, mas de assegurar que as decisões automatizadas sejam explicáveis em termos inteligíveis, inclusive para indivíduos não especializados. A explicabilidade, nesse sentido, deve ser compreendida como um requisito essencial de legitimidade, e não como um atributo opcional dos sistemas.

É certo que tal exigência encontra limites na necessidade de proteção da propriedade intelectual e dos segredos industriais das empresas desenvolvedoras. Todavia, a invocação desses interesses não pode justificar a manutenção de uma opacidade absoluta. Nesse ponto, ganha relevância a ideia de modelos intermediários, frequentemente denominados “*gray boxes*”, que buscam conciliar a preservação de aspectos estratégicos dos sistemas com a garantia de um nível mínimo de transparência. Essa abordagem impede que a complexidade técnica seja utilizada como mecanismo de blindagem frente a práticas discriminatórias.

Com efeito, reivindicações relacionadas à propriedade e à privacidade, embora legítimas e juridicamente relevantes, não podem servir como barreiras intransponíveis ao acesso à informação por parte dos indivíduos afetados. É imprescindível que estes disponham de elementos suficientes para compreender, questionar e, se necessário, contestar decisões algorítmicas. Tal prerrogativa não apenas contribui para o aperfeiçoamento dos sistemas, mas também representa

instrumento fundamental de proteção de identidades culturais, históricas e sociais frequentemente marginalizadas.

A construção da confiança pública no uso dessas tecnologias está intrinsecamente vinculada à transparência e ao reconhecimento de suas limitações. A percepção de confiabilidade não decorre da infalibilidade dos sistemas, mas da possibilidade de escrutínio, auditoria e controle de seus processos decisórios. Nesse sentido, garantir mecanismos de supervisão e alinhamento com princípios éticos e normativos torna-se condição indispensável para a legitimidade do uso da inteligência artificial em contextos sensíveis.

É nesse cenário que se insere a necessidade de adoção de mecanismos de Inteligência Artificial Explicável (Explainable Artificial Intelligence – XAI). Mais do que uma inovação técnica, a XAI representa uma resposta às demandas sociais por maior clareza e responsabilização. A literatura especializada², há décadas, vem desenvolvendo diretrizes nesse campo, destacando que as explicações devem ser adequadas ao perfil do público destinatário. Assim, modelos e descrições que podem ser plenamente compreendidos por especialistas — como redes bayesianas ou estruturas probabilísticas complexas — podem se mostrar completamente inacessíveis para usuários leigos.

Por outro lado, explicações excessivamente simplificadas podem comprometer a efetividade de auditorias realizadas por profissionais qualificados, tornando-se insuficientes para a identificação de falhas estruturais. Desse modo, a explicabilidade deve ser calibrada conforme o contexto de uso e o perfil do usuário final, garantindo um equilíbrio entre profundidade técnica e acessibilidade. Trata-se, portanto, de um exercício de proporcionalidade informacional.

Nesse contexto, ferramentas como as SHAP (*SHapley Additive exPlanations*) assumem papel relevante. Tais técnicas permitem atribuir valores específicos à influência de cada variável no resultado produzido pelo algoritmo, possibilitando a construção de um mapa interpretativo do processo decisório. Com isso, torna-se viável identificar quais fatores exercem maior peso nas decisões, o que se revela particularmente útil na detecção de padrões discriminatórios.

Esse aspecto ganha especial importância quando se analisa o fenômeno do racismo algorítmico. Em muitos casos, os sistemas reproduzem — e até amplificam — distorções presentes nas bases de dados utilizadas em seu treinamento, frequentemente marcadas por perspectivas eurocêntricas e excludentes. Assim, mesmo na ausência de uma discriminação explícita, os resultados produzidos podem refletir desigualdades estruturais profundamente enraizadas.

Importa destacar que tais discriminações, em regra, não se manifestam de forma direta ou facilmente identificável. Não há, na maioria dos casos, uma relação linear evidente entre variáveis como raça ou gênero e os resultados obtidos, o que

2 RIBEIRO, Marco Túlio; SINGH Sameer; GUESTRIN, Carlos. “Why should I trust you?”: explaining the predictions of any classifier. arXiv:1602.04938, 16 de fevereiro de 2016. Disponível em: <https://arxiv.org/abs/1602.04938>. Acesso em: 26 set. 2021.

leva muitas empresas a classificarem tais ocorrências como meros “erros pontuais”. No entanto, análises mais abrangentes permitem identificar padrões consistentes de distorção, possibilitando sua correção tanto no plano informacional quanto em aplicações mais sensíveis.

Entre essas aplicações, destacam-se os sistemas de reconhecimento facial, amplamente criticados por apresentarem taxas de erro significativamente mais elevadas para pessoas negras, bem como ferramentas de geração de imagens por inteligência artificial, que frequentemente reproduzem padrões de embranquecimento associados a ideais de beleza, prestígio e credibilidade. Tais exemplos evidenciam que os impactos da opacidade algorítmica não se restringem ao plano abstrato, mas produzem efeitos concretos e, por vezes, graves na vida dos indivíduos.

Diante disso, a explicitação dos processos preditivos — ainda que em nível simplificado — mostra-se essencial para a identificação e correção dessas falhas. O uso de técnicas como as SHAP não apenas permite localizar distorções, mas também viabiliza maior interação entre usuários e sistemas, promovendo um ambiente mais transparente e colaborativo.

A esse respeito, estudo conduzido por pesquisadores da University of Washington (Ribeiro *et al.*, 2016) fornece evidências empíricas relevantes. Ao analisar a percepção de estudantes de programação acerca da confiabilidade de algoritmos de classificação de imagens, constatou-se que, inicialmente, pouco mais de um terço dos participantes demonstrava confiança no sistema. Contudo, após a apresentação de explicações baseadas no método LIME (*Local Interpretable Model-Agnostic Explanations*), que evidenciaram a influência de fatores como o fundo das imagens na classificação, houve uma redução ainda maior desse nível de confiança.

Tal resultado revela que a exposição dos critérios utilizados pelos algoritmos não apenas amplia a compreensão sobre seu funcionamento, mas também contribui para uma avaliação mais crítica de seus resultados. Em última análise, isso evidencia a fragilidade da premissa segundo a qual decisões algorítmicas seriam inerentemente superiores às decisões humanas, reforçando a necessidade de transparência, supervisão e constante aprimoramento dessas tecnologias.

O RACISMO ALGORÍTMICO NO BRASIL

Para fins de melhor elucidação sobre a temática do racismo algorítmico no Brasil, passa-se a seguir, no capítulo seguinte, à análise de casos ocorridos em território nacional. Embora não se trate de episódios isolados ou exclusivos, esses exemplos permitem compreender de maneira mais clara a natureza do problema, evidenciando a forma como os sistemas algorítmicos podem impactar de maneira desproporcional a comunidade negra e reforçar desigualdades históricas e estruturais presentes na sociedade.

a) Vigilância Digital e Reconhecimento Facial

Tecnologias de reconhecimento facial, quando incorporadas de forma rotineira às práticas de segurança pública, têm demonstrado níveis preocupantes de imprecisão. Mais alarmante do que a existência dessas falhas é o fato de que elas não são desconhecidas pelo poder público. Ao contrário, há um volume crescente de relatórios — elaborados tanto por órgãos estatais quanto por instituições independentes — que evidenciam, de maneira consistente, as limitações técnicas, os vieses estruturais e os riscos inerentes à utilização desses sistemas. Ainda assim, tais evidências não têm sido suficientes para frear ou ao menos redirecionar sua expansão.

No Brasil, observa-se um processo de normalização progressiva do uso dessas tecnologias, frequentemente implementadas sem o devido debate público ou a cautela institucional necessária. Um exemplo emblemático dessa tendência ocorreu no estado de São Paulo, em 2020, durante a gestão do ex-governador João Dória, com a inauguração do Laboratório de Identificação Biomédica Facial e Digital.

A iniciativa tinha como propósito central a ampliação da gestão de dados biométricos digitais e a consolidação do uso do reconhecimento facial como ferramenta de segurança pública. À época, o projeto foi amplamente divulgado como um avanço tecnológico significativo, sendo celebrado com declarações que exaltavam sua suposta eficácia, como a afirmação de que o sistema seria capaz de “localizar o bandido antes que ele execute o crime”.

Paralelamente a esse discurso, destaca-se a criação de uma vasta base de dados contendo mais de 30 milhões de fotografias de cidadãos, armazenadas e utilizadas para fins de identificação automatizada. Esse processo de coleta massiva de dados levanta questionamentos relevantes acerca de privacidade, consentimento e proteção de informações pessoais. Ademais, a retórica adotada por autoridades públicas à época reforçava uma lógica de eficiência punitiva, evidenciada em declarações que celebravam o aumento do número de prisões e elogiavam a atuação policial sob a ótica quantitativa, como no caso do reconhecimento ao delegado-geral da Polícia Civil, que “nunca prendeu tanto quanto nestes 13 meses”³.

Quando analisadas sob a perspectiva do racismo algorítmico, tais iniciativas revelam implicações ainda mais profundas e preocupantes. Isso porque grande parte dos sistemas de reconhecimento facial é treinada com bases de dados desproporcionais, compostas majoritariamente por rostos de pessoas brancas. Essa assimetria resulta em maior margem de erro na identificação de indivíduos negros, que acabam sendo mais frequentemente associados, de forma equivocada, a práticas criminosas. Trata-se, portanto, de uma tecnologia que, longe de ser neutra, reproduz e potencializa desigualdades estruturais já existentes na sociedade.

Apesar da gravidade desse cenário, a discussão sobre os vieses raciais presentes nessas ferramentas muitas vezes é relegada a um plano secundário,

³ “Discurso de João Dória na inauguração do Laboratório de Identificação Biométrica em 28 de janeiro”, Governo do Estado de São Paulo, fev. 2020, disponível em: <<https://soundcloud.com/governosp/discurso-de-joao-504599548>>, acesso em: out. 2021.

especialmente diante da pressão política por medidas de endurecimento da segurança pública. Discursos pautados na ampliação do controle social e no agravamento de penalidades tendem a se sobrepor a debates mais complexos sobre direitos fundamentais, discriminação e limites éticos da tecnologia.

Os impactos concretos dessa realidade podem ser observados em casos recentes que ganharam repercussão pública. Em novembro de 2019, uma reportagem da emissora TV Itapoan, na Bahia⁴, trouxe à tona a história de um adolescente que foi erroneamente identificado como traficante em decorrência de falhas na análise de imagens captadas por câmeras de segurança. O jovem foi abordado no metrô de Salvador e conduzido à delegacia, vivenciando uma situação de extrema exposição e constrangimento. As consequências emocionais foram significativas, levando-o a desenvolver medo de frequentar a escola e de utilizar meios de transporte público, como ônibus e metrô.

No mesmo contexto geográfico e temporal, outro episódio ilustra de forma ainda mais contundente os efeitos dessas falhas. Um homem negro, portador de deficiência mental, foi confundido com um suspeito de assalto e abordado de maneira violenta por policiais enquanto se dirigia a uma consulta médica acompanhado de sua mãe. A situação evidencia não apenas a falibilidade da tecnologia, mas também a forma como ela se articula com práticas policiais marcadas por violência e desproporcionalidade. O relato da mãe, transcrito a seguir, revela a dimensão humana e cotidiana desse tipo de ocorrência:

Aconteceu no mês de setembro. Hoje, ele está tranquilo. Ele é especial, tem uma deficiência mental, mas é assistido por médicos especialistas, ele tem acompanhamento psiquiátrico. No momento da abordagem, ele ficou parado, não esboçou reação nenhuma, nem a que os policiais exigiam que ele tivesse - colocar as mãos na cabeça. Ele ficou parado só me olhando. Um policial ficou com a arma na cabeça dele e ele só me olhava. [...] A gente estava indo para uma consulta médica no Santa Izabel. Eu entrei numa padaria para tomar café porque era muito cedo. Seguiram a gente do metrô. A padaria estava muito cheia na hora, era horário de pico, eu tinha pedido um lanche pra ele, não deu tempo nem de a gente pegar o lanche, já foi com arma na cabeça dele, outra nas costas. [...] Lá mesmo, no ato da abordagem, eu me controlei e depois eu comecei a perguntar, falei que ele era especial: 'Eu sou acompanhante dele, ele tem 25 anos, não moro aqui, moro em Lauro de Freitas'. E foi aí que o policial do lado de fora entrou com uma imagem no celular que eu não vi, que eu tive a atitude de pegar a identidade do meu filho no bolso dele, apresentei e perguntei o que é que estava acontecendo. Aí foi que ele constatou que meu filho não era quem ele estava procurando, pediu desculpas ali no momento,

⁴ Trecho da matéria incluída em: *Tarcízio Silva, "Reconhecimento facial na Bahia: mais erros policiais contra negros e pobres", 21 nov. 2019, disponível em: <<https://tarciziosilva.com.br/blog/reconhecimento-facial-na-bahia-mais-erros-policiais-contr-negros-e-pobres/>>, acesso em: out. 2021.*

falaram que estavam procurando duas pessoas por assalto e que meu filho foi reconhecido nas câmeras.⁵

A recorrência de episódios como esses indica que o problema não se limita a falhas pontuais, mas se insere em uma dinâmica mais ampla, que envolve fatores tecnológicos, institucionais e políticos. O barateamento dessas tecnologias, aliado à ausência de regulamentação rigorosa, contribui para sua rápida disseminação. Empresas do setor passam a ofertar soluções com limitações técnicas significativas, encontrando receptividade em gestores públicos que priorizam respostas imediatistas à criminalidade.

Nesse contexto, consolida-se um ambiente em que sistemas automatizados passam a ser percebidos como instrumentos de autoridade incontestável, muitas vezes recebendo mais credibilidade do que a própria análise humana. Tal inversão de confiança tende a reforçar práticas policiais baseadas em abordagens agressivas e na presunção de culpa, ampliando o risco de violações de direitos fundamentais. Ainda que decisões equivocadas sejam posteriormente reconhecidas, os danos causados — especialmente no plano psicológico e social — frequentemente são irreversíveis.

Assim, torna-se imprescindível questionar as prioridades que orientam as políticas públicas de investimento. A alocação de recursos significativos em tecnologias controversas e potencialmente discriminatórias contrasta com a negligência de áreas essenciais ao desenvolvimento humano, como educação, saúde e inclusão social. Essa escolha revela não apenas uma aposta em soluções tecnológicas como resposta à complexidade da segurança pública, mas também uma opção política que merece ser analisada de forma crítica, sobretudo à luz de seus impactos sociais e de sua compatibilidade com os princípios democráticos e os direitos fundamentais.

b) Estereotipação de Corpos Negros nas Tecnologias

Fato ocorrido em meados de outubro de 2023, envolvendo a deputada estadual Renata Souza, do PSOL-RJ e então líder da CPI de Reconhecimento Fotográfico nas Delegacias da ALERJ, igualmente levantou em pauta preocupações relevantes relacionadas ao racismo algorítmico. O episódio ganhou notoriedade em razão de uma experiência relatada pela própria parlamentar ao utilizar ferramentas de geração de imagens por inteligência artificial, que à época se popularizavam amplamente nas redes sociais e passaram a integrar, de forma acelerada, o cotidiano digital de diversos usuários.

Ao aderir à tendência de criação de imagens inspiradas no estilo de animação do estúdio Pixar — reconhecido mundialmente por produções como Toy Story e Ratatouille —, a deputada forneceu um comando (prompt) para gerar uma representação de si mesma. Na descrição, contextualizou sua identidade e trajetória: uma mulher negra, vestindo um blazer de inspiração africana e inserida no

⁵ ALMA, Amanda. *O policial já foi com a arma na cabeça dele', diz mãe de rapaz confundido por reconhecimento facial*. *Correio*. n.p., 2019. Disponível em: <https://www.correio24horas.com.br/entre/o-policial-ja-foi-com-a-arma-na-cabeca-dele-diz-mae-de-rapaz-confundido-por-reconhecimento-facial-0120>. Acesso em: 26 ago. 2024.

cenário de uma favela, tudo conforme a estética visual característica dos pôsteres de animações. Ressalte-se que, em sua solicitação, não houve qualquer menção à presença de armas de fogo ou a elementos que remetessem à violência, o que torna ainda mais relevante a análise do resultado produzido.

Entretanto, ao final do processamento, a imagem gerada apresentou a parlamentar empunhando uma pistola. O resultado evidencia, de forma contundente, os limites da alegada neutralidade dos sistemas algorítmicos, demonstrando que tais tecnologias podem não apenas reproduzir, mas também intensificar vieses estruturais presentes nos dados que as alimentam e orientam seu funcionamento.

Não se trata de uma falha meramente técnica ou de uma interpretação equivocada isolada. O episódio revela a reprodução de estereótipos profundamente enraizados no imaginário social. Por um lado, reforça a associação reducionista das favelas brasileiras à criminalidade e à violência, ignorando fatores estruturais como a histórica omissão estatal e a insuficiência de políticas públicas voltadas ao desenvolvimento dessas regiões. Por outro, evidencia a recorrente vinculação entre pessoas negras e a ideia de periculosidade, contribuindo para a manutenção e naturalização de padrões discriminatórios que já se encontram amplamente disseminados no tecido social.

A relevância do caso se amplia em razão da posição institucional ocupada pela deputada, que, à época, presidia uma comissão parlamentar voltada justamente à análise crítica dessas tecnologias e de seus impactos no sistema de justiça e segurança pública. Nesse sentido, o episódio não apenas reforça as preocupações já existentes, como também ilustra, de maneira concreta, os riscos associados à adoção acrítica dessas ferramentas.

Diante disso, sua reação não se limitou ao âmbito pessoal ou a uma insatisfação pontual com o resultado obtido. Ao contrário, houve uma mobilização no sentido de dar visibilidade ao ocorrido, utilizando suas plataformas públicas para denunciar o problema, fomentar o debate e ampliar a conscientização social acerca dos limites e perigos dessas tecnologias.

A manifestação realizada por meio da rede social X (antigo Twitter) insere-se nesse contexto de conscientização, ao buscar alertar a sociedade sobre os riscos associados ao uso indiscriminado de inteligências artificiais, especialmente quando empregadas sem critérios técnicos rigorosos ou sem a devida regulação.

Segue abaixo a manifestação realizada pela deputada:

“Racismo algorítmico!

Ao criar uma arte inspirada nos pôsteres da Disney, me deparei com uma imagem gerada a partir de Inteligência Artificial que me retratava como uma mulher negra com uma arma na mão. A descrição pedida era de uma mulher negra, de cabelos afro, com roupas de estampa africana num cenário de favela. E essa foi a imagem gerada. Não pode uma mulher negra, cria da favela, estar num espaço que não da violência? O que leva essa “desinteligência artificial” a associar o meu corpo, a minha

identidade, com uma arma?

Em um mundo forjado pela branquitude, que tem o domínio sobre as tecnologias, a reprodução da estrutura racista se dá em todos os espaços e ferramentas. Não há isenção ou imparcialidade, é claramente racismo algorítmico. Enquanto presidenta da CPI do Reconhecimento Fotográfico observo o quanto pessoas negras podem ser vitimadas por mecanismos, como o reconhecimento facial, mesmo sendo inocentes.

O imaginário de violência nas favelas e de criminalização de corpos e territórios negros também está presente em tecnologias, o que coloca em risco a segurança de pessoas negras. É preciso rever essas tecnologias e procedimentos.”

O episódio, assim, reforça a necessidade de uma abordagem mais crítica, responsável e cautelosa no desenvolvimento e na aplicação dessas ferramentas, sobretudo quando seus impactos atingem diretamente direitos fundamentais, a representação social e a igualdade racial.

c) Embranquecimento como Capital Simbólico e Apagamento da Influência Negra

No ano de 2024, a Academia Brasileira de Letras (ABL) se viu no centro de intensos debates e polêmicas após o desenvolvimento de um projeto que envolveu o uso de inteligência artificial, em parceria com a instituição Euvatar Storytelling. O objetivo principal do projeto era recriar um dos maiores expoentes da literatura brasileira, Machado de Assis, por meio de um avatar humanizado que permitisse ao público interagir de forma direta e dinâmica, com respostas em tempo real sobre a vida e a obra do autor. A iniciativa buscava, em termos técnicos, aproximar o público contemporâneo do escritor, permitindo um tipo de interação inédita com um ícone cultural.

Conforme relatado em matéria do Jornal de Letras, “foram necessários mais de três meses para treinar o banco de dados” que serviria de base para a criação do avatar. Apesar do investimento de tempo e recursos, a recepção do público foi marcada por críticas, especialmente no que diz respeito à aparência física do avatar. O personagem foi apresentado com pele clara e cabelos grisalhos, o que gerou acusações de “embranquecimento” do escritor. Registros históricos indicam que Machado de Assis era um homem de pele negra, nascido e criado em um contexto de pobreza no Rio de Janeiro do século XIX, descendente de Francisco José de Assis, um escravo alforriado, e de Maria Leopoldina Machado da Câmara, lavadeira portuguesa oriunda dos Açores.

Essa representação, além de evidenciar padrões eurocêntricos de estética e beleza branca, acarreta efeitos simbólicos significativos: ao alterar características essenciais do escritor, o projeto contribui para o apagamento da presença e da influência negra na literatura e na cultura brasileira. Mais do que uma questão estética, trata-se de um deslocamento simbólico que reduz a visibilidade histórica da população negra, minimizando seu protagonismo e importância na formação cultural e literária do país.

O caso ressalta que, mesmo com os avanços tecnológicos e a sofisticação das ferramentas de inteligência artificial, ainda existem grandes desafios no uso ético e responsável dessas tecnologias. A necessidade de vigilância constante, ajustes permanentes e debates críticos sobre representações culturais se torna evidente, especialmente quando se trata de figuras históricas cuja identidade está ligada a trajetórias de resistência, inclusão e significado social. Projetos como este demonstram que, sem critérios sensíveis e conscientes, a tecnologia pode reproduzir e intensificar preconceitos já existentes, mesmo em contextos que buscam inovação ou experimentação artística.

Casos como o do avatar de Machado de Assis ilustram, portanto, a importância de se manter uma reflexão crítica sobre os impactos culturais, sociais e simbólicos da tecnologia, destacando a necessidade de proteger a memória e a identidade de indivíduos e grupos historicamente marginalizados. Eles evidenciam que a sofisticação técnica por si só não garante justiça ou representatividade, e reforçam a urgência de incorporar critérios éticos e de diversidade na aplicação de ferramentas digitais voltadas à educação, cultura e patrimônio histórico.

CONSIDERAÇÕES FINAIS

Conforme demonstrado ao longo do presente trabalho, conclui-se que os sistemas algorítmicos estão longe de representar o mito da neutralidade. Pelo contrário, eles se mostram constantemente influenciados pelas construções sociais, históricas e culturais nas quais são desenvolvidos, assim como pelos dados que alimentam seu funcionamento. Esses elementos moldam decisões automatizadas de forma que refletem, muitas vezes de maneira imperceptível, preconceitos e desigualdades presentes na sociedade.

Nesse contexto, o racismo algorítmico emerge como uma das formas mais sofisticadas e preocupantes de perpetuação da discriminação racial. Sua gravidade se deve justamente à aparência de imparcialidade que esses sistemas aparentam possuir. Como apontado ao longo do estudo, essa falsa neutralidade tem contribuído para a realização de práticas discriminatórias em diversas esferas, afetando direitos fundamentais e reforçando desigualdades estruturais, especialmente entre grupos historicamente marginalizados.

Adicionalmente, a opacidade inerente a grande parte desses sistemas — frequentemente desenvolvidos como verdadeiras “caixas-pretas” sob o argumento de proteção à propriedade intelectual — dificulta a compreensão dos critérios utilizados nas decisões automatizadas. Tal característica impede que os indivíduos afetados contestem ou questionem essas decisões de forma efetiva, fragilizando direitos fundamentais e colocando em evidência a necessidade urgente de mecanismos que promovam maior explicabilidade, transparência e responsabilidade.

Diante desse cenário, torna-se imprescindível a participação ativa de grupos historicamente marginalizados em todas as etapas do desenvolvimento e aplicação dessas tecnologias. Essa participação é essencial não apenas para a

identificação e correção de vieses existentes, mas também para garantir que os sistemas algorítmicos produzam respostas mais precisas e, ao mesmo tempo, respeitem a diversidade, a pluralidade e os princípios de justiça social. A inclusão desses grupos fortalece o caráter democrático da tecnologia, promovendo o uso responsável de ferramentas que, quando aplicadas de maneira crítica e consciente, podem contribuir para reduzir desigualdades e ampliar o acesso a direitos de forma equitativa.

REFERÊNCIAS

ALMEIDA, José Antonio Caldeira de; SOUZA, Patrícia Verônica Nunes Carvalho Sobral de. **Tecnologias de reconhecimento facial: o racismo algorítmico como instrumento de política de segurança pública**. *Direito e Política*, v. 20, n. 1, 2025.

ALVES, Marco Antônio Sousa; ANDRADE, Otávio Morato de. **Da “caixa-preta” à “caixa de vidro”: o uso da Explainable Artificial Intelligence (XAI) para reduzir a opacidade e enfrentar o enviesamento em modelos algorítmicos**. *RDP*, Brasília, v. 18, n. 100, p. 349-373, out./dez. 2021.

COSTA, Ingrid Coelho da; ARAÚJO, Thamiris Lima de; MOURÃO, Rosalia Maria Carvalho. Racismo algorítmico: o reconhecimento facial em desconformidade aos direitos fundamentais. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 11, n. 5, maio 2025. DOI: 10.51891/reaase.v11i5.19333.

CRESTANE, Dérique Soares; LEAL, Mônia Clarissa Hennig. **Discriminação algorítmica e discriminação estrutural: standards protetivos da Corte Interamericana de Direitos Humanos e do Supremo Tribunal Federal**. 1. ed. São Paulo: Tirant lo Blanch, 2024.

FERRARI, Manoela. A inteligência artificial ressuscita Machado de Assis. **Jornal de Letras**, Rio de Janeiro, n. 301, p. 10-11, jul. 2024.

GUIMARÃES, Luíza Resende. **Inteligência artificial e enviesamento algorítmico: novas formas de discriminação contra pessoas com deficiência**. *civilistica.com*, a. 13, n. 2, 2024. DOI: 10.5281/zenodo.18616376.

LIMA, Thaise Marques de; SILVA, Vanessa Maria Gomes da; BRASILEIRO, Fellipe Sá. Algoritmos da branquitude: vieses e representações racistas em sistemas de inteligência artificial. **Revista do Programa de Pós-Graduação em Comunicação**, Juiz de Fora, v. 19, n. 1, p. 50-69, jan./abr. 2025.

MATOS, Gilson Ely Chaves de; LEMOS, Marcelo Augusto Rodrigues de; SILVA, Clara Virgínia de Oliveira. Racismo e o emprego de novas tecnologias pelo Estado no âmbito da segurança pública. **Revista Direito em Debate**, ano XXXIII, n. 62, p. 1-11, jul./dez. 2024. DOI: 10.21527/2176-6622.2024.62.15865.

MELLO, Ana Carolina Paes de. Discriminação codificada: o impacto do racismo algorítmico na sociedade. **Revista Rua**, Campinas, v. 31, n. 1, p. 329-342, jun. 2025. DOI: 10.20396/rua.v31i1.8680812.

MELO, Jair Alcindo Lobo de. **Racismo algorítmico à luz da análise do biopoder de Michel Foucault**. Tese (Doutorado em Letras) — Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, 2025.

NINJA. Deputada Renata Souza denuncia racismo em plataformas de IA. Mídia NINJA, [s. l.], 17 nov. 2023. Disponível em: <https://midianinja.org/deputada-renata-souza-denuncia-racismo-em-plataformas-de-ia/>. Acesso em: 5 abr. 2026.

OLIVEIRA, Jean Carlos de; ARANTES, Priscila Almeida Cunha. Alexa, você é racista? Racismo algorítmico, vieses e intencionalidade. **DATJournal**, v. 9, n. 1, 2024.

RIBEIRO, Adriana; ELOI, Dárcio Luiz Costa; RODRIGUES, Renata Vicentini Monteiro. Transparência e ética no uso de IA generativa. **Revista Base Científica (RBC)**, v. 3, e202501, p. 1-17, 2025. DOI: 10.52832/rbc542.

SANTOS, Milena Barros Marques dos. **Racismo algorítmico e microagressões raciais no mundo digital**. In: IASULAITIS, S.; SILVEIRA, S. A. (org.). Estudos sociopolíticos da inteligência artificial. Campina Grande: EDUEPB, 2025. p. 163-174. DOI: 10.7476/9786552210234.0007.

SILVA, Isabela Maria Soares; BARBOSA, Letícia Mendes. Inov(ação): discriminação algorítmica racial e as inteligências artificiais no Brasil. **Revista do Centro Acadêmico Afonso Pena**, Belo Horizonte, v. 28, n. 2, jul./dez. 2023.

SILVA, Tarcízio da. Visão computacional e racismo algorítmico: branquitude e opacidade no aprendizado de máquina. **Revista da ABPN**, v. 12, n. 31, p. 428-448, dez. 2019-fev. 2020. DOI: 10.31418/2177-2770.2020.v12.n.31.p428-448.

SOARES, Cláudio. **Machado de Assis “embranquecido” em avatar levanta debate sobre viés racial da IA**. Brasil de Fato, Rio de Janeiro, 13 mar. 2024. Disponível em: <https://www.brasildefato.com.br/2024/03/13/artigo-machado-de-assis-embranquecido-em-avatar-levanta-debate-sobre-vieis-racial-da-ia/>. Acesso em: 5 abr. 2026.

TEIXEIRA, Lucas de Barros. **Transparência algorítmica em soluções utilizadas por governos mundo afora, e o contexto Brasil**. Interface Tecnológica, v. 18, n. 1, 2021. DOI: 10.31510/infa.v18i1.1083.