

Jean Carlos de Oliveira, Priscila Almeida Cunha Arantes *

Alexa, você é racista? Racismo algorítmico, vieses e intencionalidade

*

Jean Carlos de Oliveira é graduado em Tecnologia e Mídias Digitais, com ênfase em Arte e Tecnologia, pela PU-CSP (2007), possui especialização em Sócio Psicologia pela FESP (2011) e em Docência no Ensino Superior pelo Centro Universitário Senac (2020). Foi aluno bolsista no Programa de Mestrado em Design na UAM (2022 a 2024) e integrou o grupo de pesquisa “Design, arte e memória: perspectivas contemporâneas”, orientado pela Profa. Dra. Priscila Arantes, na UAM. Atualmente, é docente nos cursos de Design na Universidade São Judas.

<jean.artec@gmail.com>

ORCID 0000-0002-3578-0137

Priscila Almeida Cunha Arantes é pesquisadora, crítica e curadora no campo da arte e estética contemporânea. Possui graduação em Filosofia pela USP, com pós-doutorado pela UNICAMP e pela Penn State University (EUA). É professora do Departamento de Artes da PUC/SP e diretora adjunta da Facul-

Resumo Este artigo busca examinar de forma crítica os algoritmos e suas implicações, levantando questões sobre equidade, justiça social e responsabilidade ética. Além disso, propor medidas para mitigar os efeitos negativos do viés algorítmico, como o desenvolvimento de práticas de design mais inclusivas, revisões sistemáticas de algoritmos existentes, maior transparência e prestação de contas por parte das empresas e organizações que os utilizam. Através dessas análises e propostas, o artigo pretende contribuir para um debate mais amplo e informado sobre o viés algorítmico, incentivando ações e soluções que promovam a equidade e a justiça nos sistemas impulsionados por algoritmos.

Palavras-chave Racismo algorítmico, Design e ativismo, Algoritmo, Vieses racistas.

dade de Filosofia, Comunicação, Letras e Artes. Além disso, leciona no PPG em Design da Universidade Anhembi Morumbi. Entre 2007 e 2020, foi diretora e curadora do Paço das Artes, instituição da Secretaria de Estado da Cultura de SP, onde desenvolveu diversos projetos curatoriais, seminários nacionais e internacionais, eventos e publicações. É autora de “Arte @ Mídia: perspectivas da estética digital” (FAPESP/SENAC), finalista do 48º Prêmio Jabuti, “Reescrituras da Arte Contemporânea: história, arquivo e mídia” (Ed. Sulina), “Urgências na Arte”, entre outros trabalhos. Também atua como vice-diretora da ABCA (Associação Brasileira de Críticos de Arte) e integra, como convidada, o grupo de pesquisadores do ID+, instituto de pesquisa da Universidade de Porto e Aveiro (Portugal).

<priscila.a.c.arantes@gmail.com>
ORCID 0000-0002-0500-0849

Alexa, are you racist? Algorithmic racism, biases and intentionality

Abstract *This article seeks to critically examine algorithms and their implications, raising questions about equity, social justice, and ethical responsibility. Furthermore, it proposes measures to mitigate the negative effects of algorithmic bias, such as the development of more inclusive design practices, systematic reviews of existing algorithms, increased transparency, and accountability from the companies and organizations that use them. Through these analyses and proposals, the article aims to contribute to a broader and more informed debate about algorithmic bias, encouraging actions and solutions that promote equity and justice in algorithm-driven systems.*

Keywords *Algorithmic racism, Design, and activism, Algorithm, Racist biases.*

Alexa, ¿eres racista? Racismo algorítmico, sesgos e intencionalidad

Resumen *Este artículo busca examinar de forma crítica los algoritmos y sus implicaciones, planteando cuestiones sobre equidad, justicia social y responsabilidad ética. Además, propone medidas para mitigar los efectos negativos del sesgo algorítmico, como el desarrollo de prácticas de diseño más inclusivas, revisiones sistemáticas de algoritmos existentes, mayor transparencia y rendición de cuentas por parte de las empresas y organizaciones que los utilizan. A través de estos análisis y propuestas, el artículo pretende contribuir a un debate más amplio e informado sobre el sesgo algorítmico, incentivando acciones y soluciones que promuevan la equidad y la justicia en los sistemas impulsados por algoritmos.*

Palabras clave *Racismo algorítmico, Diseño y activismo, Algoritmo, Sesgos racistas.*

Introdução

“Um dos maiores problemas com a nossa confiança cega nos algoritmos é que podemos propagar padrões discriminatórios sem reconhecer qualquer tipo de intenção.” - Cathy O’Neil

A democratização da informação proporcionada pelo avanço tecnológico, especialmente no que diz respeito à internet, representa um avanço significativo em diversos aspectos da vida em sociedade, destacando-se pelo impacto positivo no acesso e abrangência a conteúdos culturais, acadêmicos e artísticos. Essa transformação tem ampliado a visibilidade e o reconhecimento de expressões anteriormente marginalizadas pelas mídias tradicionais (revista, jornal, rádio e TV)¹ e circuitos ideológicos (exposições, festivais, galerias, mostras e museus)².

Com a disseminação global da internet, artistas e designers têm tido a sua disposição uma série de recursos – muitos deles gratuitos – acessíveis e uma seara de sites e plataformas específicos para compartilhar suas criações e produções diretamente para audiências amplas e diversificadas. As plataformas de mídia social, blogs, portfólios online, bancos de imagens e outras formas de exposição digital têm proporcionado um ambiente propício para criadores independentes, incluindo aqueles pertencentes a grupos historicamente excluídos, para divulgarem e promoverem seu trabalho. A facilidade na troca instantânea de informações e inspirações tem possibilitado a ampliação de oportunidades de visibilidade e reconhecimento para diversos talentos.

Apesar desses avanços, ainda persistem inúmeros desafios na garantia do alcance e acesso a essas informações e produções, sejam eles pela existência de barreiras estruturais, como a falta de acesso equitativo à tecnologia e a desigualdade no acesso à educação de parte expressiva da população, ou então por meio de algoritmos enviesados que reproduzem discriminações raciais e limitam o alcance de determinados conteúdos, principalmente aqueles produzidos por pessoas negras. A esse fenômeno, que impõe dificuldades adicionais à população negra na internet, conveniou-se chamar de racismo algorítmico.

O termo “algoritmo” refere-se a um conjunto de instruções, ou regras bem definidas e ordenadas, que devem ser seguidas para resolver um problema ou executar uma tarefa específica. Em termos simples, um algoritmo é como uma receita de bolo: ele descreve passo a passo o que deve ser feito para alcançar um determinado resultado. No caso da computação os algoritmos são usados para realizar cálculos, processar dados e executar diversas operações. (Cf. Hora, 2021).

Já racismo algorítmico é um fenômeno recente na nossa sociedade e que tem chamado a atenção de artistas, designers, pesquisadores e ativistas antirracistas em todo o mundo, em especial, em países atravessados pela colonização e diáspora africana. Entre eles, destacam-se pesquisadores

como a americana Safiya Umoja Noble³ (2018), que aborda questões de gênero e raça a partir dos mecanismos de busca comerciais como a Google; a canadense Joy Buolamwini⁴, que investiga os efeitos discriminatórios de sistemas de reconhecimento facial em relação a pessoas de pele mais escura; e aqui no Brasil, Tarcízio Silva⁵, que possui trabalhos dedicados à pesquisa e incidência sobre transparência, responsabilidade e antirracismo na inteligência artificial.

À medida que a tecnologia avança, as decisões automatizadas baseadas em algoritmos se tornam cada vez mais presentes em diversos aspectos da vida contemporânea. Isso inclui desde a triagem e análise de perfis em processos de recrutamento e seleção a aprovação de crédito em transações financeiras. Também são objetos de automação algorítmica a operação de sistemas de reconhecimento facial, sistemas de recomendação e moderação artificial de conteúdos em plataformas online e mecanismos de busca, e, mais recentemente, a condução de veículos autônomos.

Para os designers, a incorporação da inteligência artificial (IA)⁶, por meio do algoritmo robustecido, tem gerado impactos significativos, afetando diferentes aspectos do processo criativo e da prática profissional. Essas transformações abrangem desde a automação de tarefas repetitivas, possibilitando que o profissional foque no processo criativo, até a geração de ideias e prototipagem com base em parâmetros específicos. A assistência criativa inclui também sugestões para paletas de cores, tipografias e tendências de design, além da geração e produção de elementos visuais, sonoros, textuais, entre outros.

O emprego de algoritmos matemáticos para a manipulação de dados na computação não é uma novidade, a diferença reside no fato de que, agora, eles são potencializados por IA. Segundo as empresas desenvolvedoras, as chamadas Big Techs, esses algoritmos são fundamentados nos inputs (treinamentos) fornecidos pelos programadores e enriquecidos pela interação contínua com os usuários.

Em síntese, essa interação é o que possibilita a adaptação do algoritmo a diversas situações, permitindo que ele aprenda com novos dados e aprimore suas respostas ao longo do tempo. A eficácia do algoritmo seria, portanto, influenciada pela qualidade dos dados de treinamento e pela diversidade e quantidade de interações com usuários reais, além da utilização de Modelos de Linguagem Grande (LLMs), contribuindo para sua capacidade de tomar decisões e fornecer respostas contextualmente relevantes. Segundo a Google, o LLMs consiste em

um modelo de linguagem estatística, treinado com uma quantidade enorme de dados, que pode ser usado para gerar e traduzir texto e outros conteúdos, além de executar outras tarefas de processamento de linguagem natural (PLN). Normalmente, os LLMs são baseados em arquiteturas de aprendizado profundo, como o Transformer desenvolvido pelo Google em 2017, e podem ser treinados com bilhões de texto e outros conteúdos. (Google Cloud, s/d)

Essa abordagem adotada pelas empresas destaca a importância da colaboração mútua entre programadores e usuários para a evolução e o refinamento contínuos dos sistemas de AI. Diante dessa interação social, emergem questões cruciais para os pesquisadores dedicados ao tema, incluindo a possibilidade de os sistemas de inteligência artificial, alimentados por algoritmos treinados com base na interação entre desenvolvedores e usuários, manifestarem comportamentos presentes na sociedade, como racismo, sexismo, misoginia, entre outros.

Os problemas relativos à suposta neutralidade do algoritmo não são novos e estão presentes no uso da internet há algum tempo. Sobre os mecanismos de busca e pesquisa, por exemplo, a pesquisadora Safiya Umoja Noble (2013) considera que a implementação desses algoritmos nem sempre é neutra, pois eles determinam, a partir de diversos fatores não muito claros aos usuários, quais assuntos e conteúdos terão maior ou menor relevância, aderência e potencial de engajamento do público. Desse modo, é possível afirmar que muitas vezes os algoritmos incorrem no risco de perpetuar e amplificar, mesmo que de maneira sutil, as desigualdades sociais e raciais existentes.

Tarcízio Silva desenvolve crítica análoga ao destacar a falta de transparência nos processos adotados por empresas e desenvolvedores, bem como a dificuldade de se identificar e denunciar possíveis violações nesse tipo de tecnologia. Para ele,

A compreensão de muitas manifestações do racismo na forma de micro-agressões em práticas cotidianas é especialmente difícil quando se trata de tecnologias algorítmicas de comunicação, que trazem novas camadas de opacidade em seu funcionamento. Se a tecnologia é erroneamente enquadrada e percebida como neutra, a tal equívoco se soma a negação do racismo como fundante de relações e hierarquias sociais em países como o Brasil. (Silva, T., 2022a, p. 29)

Na internet, abundam relatos e exemplos de situações em que, ao realizar pesquisas específicas nos mecanismos de busca sobre profissões ou atividades, os resultados destacaram predominantemente imagens ou informações associadas a pessoas brancas. Além disso, em bancos de imagens, observa-se viés racial na categorização ou sugestão de imagens, ao passo em que sistemas de reconhecimento facial apresentam taxas de erro mais elevadas para indivíduos de pele negra.

Adicionalmente, há denúncias frequentes de que algoritmos empregados em processos de recrutamento e seleção possam incorporar viés racial ao analisar currículos ou perfis online. Da mesma forma, algoritmos de classificação de conteúdo em redes sociais ou plataformas de streaming também são acusados de apresentar esse viés, influenciando quais tipos de conteúdo são promovidos ou restringidos com base em características racializadas.

Como parte de sua pesquisa, Tarcízio Silva desenvolveu o projeto “Desvelar – Justiça Racial na Inteligência Artificial e TICs”, um site dedicado

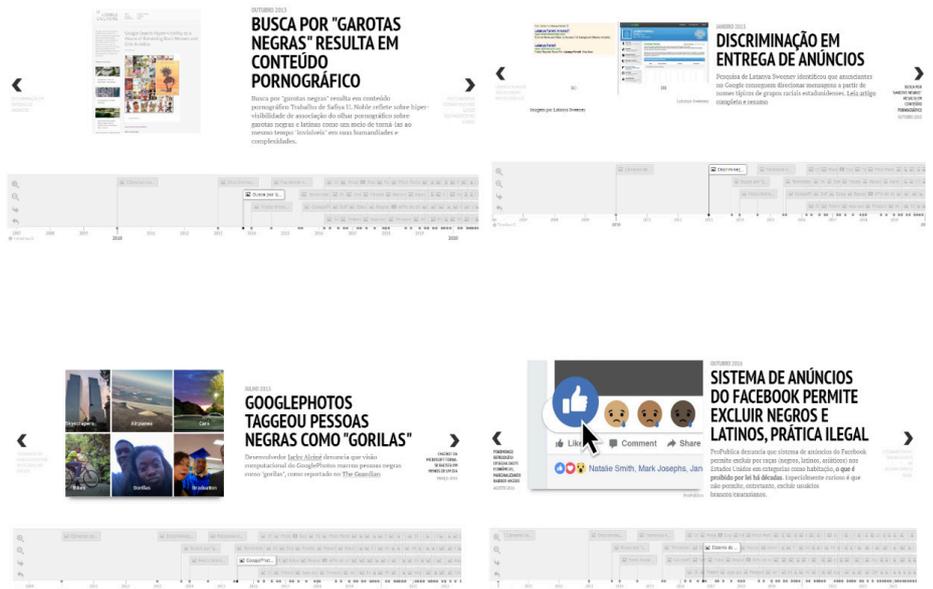
a contribuir para os desenvolvimentos e debates sobre tecnologias digitais, com foco especial em políticas públicas e regulamentações relacionadas à IA no Brasil. O projeto destaca a atenção à replicação de práticas discriminatórias existentes no país, como o racismo estrutural, especialmente contra populações negras, por sistemas de IA e algoritmos (Figura 1).

Figura 1. Print da página inicial do Projeto Develar
 Fonte: @Develar - <https://desvelar.org>



O site conta ainda com a seção “Danos e discriminação algorítmica” que traz um mapeamento de casos, reportagens e reações sobre discriminação algorítmica ao longo dos últimos anos e que podem ser visualizados através de uma timeline interativa (Figura 2).

Figura 2. Linha do tempo sobre racismo algorítmico
 Fonte: @Develar - <https://desvelar.org>



Como vimos, atualmente existe uma infinidade de tipos de algoritmos, cada um projetado para automatizar, sistematizar e solucionar uma ampla gama de problemas. Esses algoritmos podem lidar com diversas combinações de variáveis e dados, adaptando-se às necessidades específicas de cada situação para, assim, fornecerem soluções estruturadas e, pretensamente, eficientes para uma variedade de problemas em diferentes campos de estudo e setores da indústria.

Com essas observações em mente, nos interessa analisar o “algoritmo de relevância pública”, termo adotado pelo pesquisador Tarleton Gillespie (2018b). Ele se refere às rotinas de programação desenvolvidas para selecionar e apresentar conteúdos relevantes para uma determinada audiência, levando supostamente em consideração os interesses específicos do usuário. Esses algoritmos são projetados com o objetivo de fornecer informações e conteúdos personalizados de acordo com a experiência do usuário, suas preferências e suas necessidades individuais previamente mapeadas e catalogadas na base de dados.

Os algoritmos de relevância pública têm uma ampla aplicação, sendo extensivamente utilizados por plataformas de busca como Google, Bing e Yahoo, redes sociais como Facebook, Instagram, TikTok e X (antigo Twitter). Eles são responsáveis por filtrar e exibir conteúdo considerado mais adequado para cada usuário, levando em consideração fatores como histórico de interações, preferências declaradas e padrões de comportamento. Segundo Amanda Chevtchouk Jurno e Silvia Dalben:

O trabalho desses agentes é escolher o que o usuário X gostaria de ver dentre a miríade de conteúdo disponível em uma determinada plataforma ou na web de forma geral, seja com base nas suas escolhas pessoais de X, seja com base em critérios como relevância e importância. (JURNO; DALBEN, 2018; p. 20)

A falta de transparência nos critérios utilizados no treinamento desses algoritmos tem suscitado crescentes preocupações quanto à possível formação de “bolhas” de informação, sobre as quais Gillespie (2011 apud JURNO; DALBEN, 2018, p. 22) alerta para “uma importante tensão que emerge entre o que nós esperamos que esses algoritmos sejam, e o que eles são de fato”. Diante disso, surgem questionamentos pertinentes: quais critérios são aplicados na triagem de conteúdo? Quais interesses são priorizados pelos algoritmos – os dos usuários, mantenedores da plataforma ou dos anunciantes e patrocinadores? Com base em quais dados essas decisões são tomadas?

Essas indagações destacam, portanto, a necessidade de maior transparência no funcionamento desses sistemas para mitigar preocupações relacionadas a vieses, manipulação e falta de imparcialidade nos resultados apresentados aos usuários. Nos termos propostos pelo jornalista Nicholas Diakopoulos:

A ética e transparência algorítmica devem, portanto, considerar os algoritmos como objetos criados por humanos e levar em conta as intenções, de grupos ou processos institucionais, que influenciaram o design, assim como a agência de atores humanos na interpretação do output em processos superiores de tomada de decisões (DIAKOPOULOS, 2014, p. 5 apud JURNO; DALBEN, 2018; p. 24).

Essencial, portanto, é a busca por maior transparência e responsabilidade por parte dessas empresas, incentivando o diálogo público e a reflexão sobre essas questões. A conscientização e o debate em torno da ética e da governança dos algoritmos são fundamentais para garantir uma maior equidade e compreensão dos impactos dessas tecnologias em nossas vidas.

As origens do racismo algorítmico

O termo “racismo algorítmico” não possui uma origem única ou criador específico. Ele emergiu da necessidade de descrever e analisar os casos em que algoritmos e sistemas de inteligência artificial demonstram viés racial ou perpetuam práticas discriminatórias uma vez que, de acordo com Amanda Jurno (2020), estabelecem uma lógica distinta na administração dos fluxos de informações e reorganizam os conteúdos com o objetivo de atender a interesses individuais, criando um sistema específico de visibilidade e invisibilidade.

Um dos principais fatores que contribuem para o racismo algorítmico é a qualidade dos dados utilizados no treinamento desses algoritmos, ou seja, se os dados históricos usados para treinar os modelos contêm preconceitos raciais, esses preconceitos podem ser aprendidos e replicados pelos algoritmos. Por exemplo, se os dados de treinamento de um sistema de reconhecimento facial contêm principalmente rostos de indivíduos de determinadas etnias, o algoritmo pode ter dificuldade em reconhecer com precisão rostos de outras etnias, resultando em discriminação.

Além disso, os próprios desenvolvedores dos algoritmos podem introduzir seus próprios vieses inconscientes durante o processo de criação, assim como os profissionais responsáveis e habilitados para a moderação, mediação e filtragem de conteúdo disponíveis nas plataformas e sistemas de dados. Assim, as decisões tomadas durante a seleção de recursos, definição de métricas de desempenho e ajuste de parâmetros podem refletir preconceitos raciais e sociais. Outro fator importante é a própria falta de diversidade na indústria de tecnologia, revelada pela predominância de desenvolvedores de software de determinados grupos étnicos podendo levar a uma visão limitada dos problemas e uma falta de sensibilidade para questões raciais.

Sobre este ponto em específico da qualificação e moderação dos conteúdos disponibilizados, Tarcízio Silva (2022) aponta que

O próprio trabalho árduo de moderação de bilhões de unidades de conteúdo é distribuído geográfica e racialmente no mundo. Enquanto profissionais nos centros de poder da tecnologia, como Vale do Silício, definem as regras de moderação objetivando a lucratividade e a fuga de implicações legais, são profissionais precarizados de países do Sul global que efetivamente veem os conteúdos violentos e perturbadores. (SILVA, T., 2022, p.41)

Segue:

Alegando melhorar índices como esses, mas também motivadas pela chance de enxugar as equipes de moderação, as plataformas têm experimentado a remoção automatizada de conteúdo baseada em processamento de linguagem natural e visão computacional. A adoção de regras desenvolvidas por especialistas nas temáticas como discurso de ódio, conflitos e violência sexual se soma ao aprendizado de máquina sobre as milhões de decisões anteriores consideradas adequadas. (SILVA, T., 2022, p. 42)

Outro recurso muito utilizado pelas empresas responsáveis pelas plataformas e redes sociais digitais para redução de custos com pessoal e a pretexto de dar maior agilidade aos serviços de qualificação, moderação e exclusão de conteúdos e materiais de cunho potencialmente ofensivos e – muitas vezes – criminosos, de acordo com a legislação vigente em cada país, é relegar parte dessa atividade aos próprios usuários. Desse modo, a partir de canais de denúncia disponibilizados, ou até mesmo, pela alta rejeição (dislikes) a determinados conteúdos, a comunidade de usuários se co-responsabiliza na tentativa de ocultar ou remover tais conteúdos que, segundo critérios próprios, violariam as Diretrizes da Comunidade.

Essas práticas de atribuir aos algoritmos ou aos próprios usuários a responsabilidade de apontar os assuntos e conteúdos considerados nocivos à “comunidade”, a partir da quantidade de interações e denúncias negativas, possuem inúmeros efeitos colaterais, sendo o mais imediato o impacto enviesado e empobrecido do debate público. Isso ocorre, pois assuntos considerados polêmicos movimentam as redes e provocam o engajamento de grupos heterogênicos na tentativa de pautar as discussões, ou ainda, rechaçar e invalidar determinados assuntos, mesmo que por vezes relevantes e legítimos, mas que contrariam certas ideologias, inclinações políticas, religiosas e morais. Nestes casos, grupos de pessoas com perfis reais ou fakes e, por vezes, impulsionados por “robôs” ou bots (programas de auto execução com finalidade de executar ações repetitivas) se mobilizam com o objetivo de “derrubar” os conteúdos que consideram indesejáveis.

Para inibir este fenômeno, algumas plataformas como Instagram, Facebook, Youtube entre outras, tem apostado na ocultação numérica da quantidade de interações negativas sobre os assuntos e conteúdos posta-

dos, incidindo novamente em um processo sem transparência e de critérios duvidosos. Para o pesquisador Tarleton Gillespie,

A opacidade esconde não apenas o fato da seleção, mas também os valores que motivam tal seleção, incluindo a palatabilidade do conteúdo para uma audiência imaginada e o potencial para sua comercialização e viralidade, de um lado, e a probabilidade de causar ofensas e danos à marca, de outro. (GILLESPIE, 2018^a, p.160 apud SILVA, T., 2022^a, p.44)

Assim, ao lidar com algoritmos, é fundamental ter em mente que eles são o produto de complexas redes sociotécnicas, compostas por diversos atores humanos e não humanos, que influenciam e participam ativamente da construção de significados para os usuários das plataformas em que operam (JURNO; DALBEN, 2018, p. 25).

Vieses racistas em algoritmos

Como vimos, os algoritmos são criados por seres humanos e treinados com conjuntos de dados que contêm informações históricas. Se esses dados refletirem desigualdades sociais, discriminação ou estereótipos raciais, os algoritmos podem aprender e perpetuar esses padrões. Essas inclinações, denominadas vieses, podem surgir de várias maneiras, como nos dados de treinamento utilizados para desenvolver os algoritmos, nas suposições e decisões tomadas pelos desenvolvedores, bem como nas interações com os usuários.

Ao abordar essa questão, no episódio #1 - Eu Existo? – O Racismo Algoritmo nas redes, para o podcast Comunicando em Preto, o jornalista Giovane Gabriel Ramos André, relembra um caso envolvendo a plataforma Facebook que ficou emblemático nos noticiários a época. Ele cita a manchete “IA do Facebook rotulou vídeo de homens negros como ‘primatas’; empresa pede desculpas”, publicada pelo portal Olhar Digital, em 2021. Segundo a matéria, “usuários que assistiram um vídeo com homens negros, postado no dia 27 de junho pelo Daily Mail do Reino Unido no Facebook, receberam uma mensagem automática da plataforma perguntando se eles queriam “continuar assistindo vídeos de primatas?”.

Em resposta, um porta-voz do Facebook, classificou o fato como um “erro claramente inaceitável” e seguiu: “Como já dissemos, enquanto fizemos melhorias em nossa IA, sabemos que ela não é perfeita e que há mais progressos a serem feitos. Nos desculpamos com qualquer pessoa que possa ter visto essas recomendações ofensivas.” (SCHNOOR, 2021)

O fato é que, para além de mero aborrecimento quanto a “recomendações ofensivas”, “erros claramente inaceitáveis” ou a insuficiência na melhoria da IA, a existência de vieses racistas em algoritmos torna-se ainda mais preocupante, na medida em que esses sistemas estão cada vez mais presentes em várias esferas da vida, incluindo emprego, saúde, justiça

criminal e tomada de decisões governamentais. Quando esses algoritmos são usados para tomar decisões importantes que afetam a vida das pessoas, é crucial que eles sejam justos, imparciais e transparentes. Silva (2023) reforça que o

Racismo algorítmico não é uma questão de programação ou engenharia. Mais importante que as linhas de código é saber quais são as relações de poder e quais decisões são habilitadas pela implementação de alguma tecnologia. (SILVA T., 2023)

Sendo assim, para mitigar os vieses racistas em algoritmos, é necessário um esforço conjunto de desenvolvedores, pesquisadores e reguladores. Isso inclui uma maior diversidade na equipe de desenvolvimento, a seleção cuidadosa e a limpeza dos dados de treinamento para evitar desequilíbrios e a realização de auditorias e testes regulares para identificar e corrigir vieses. Além disso, é fundamental estabelecer diretrizes e regulamentos para garantir que os algoritmos sejam justos e transparentes, além da responsabilização quando da identificação desses vieses.

Por fim, a conscientização e o debate público também desempenham um papel importante na abordagem dos vieses racistas em algoritmos. É necessário educar e envolver as pessoas sobre os riscos desses vieses e a importância de um uso ético da inteligência artificial. Somente através de uma abordagem coletiva e multidisciplinar podemos trabalhar para minimizar os vieses racistas em algoritmos e promover uma sociedade mais justa e igualitária.

Considerações finais

Como podemos observar, a inteligência artificial, especialmente o aprendizado de máquina com base em dados, pode alimentar sistemas algorítmicos que perpetuam preconceitos e promovem a discriminação. Portanto, ao explorar os algoritmos por trás dessas tecnologias e os desafios associados à redução do viés, torna-se evidente que a responsabilidade recai tanto sobre as empresas desenvolvedoras quanto sobre os usuários, haja vista que somente por meio de uma abordagem intencional e consciente do tema, de investimentos na diversidade na equipe de desenvolvimento e ações coletivas é que poderemos buscar a equidade em nossas interações com esses dispositivos tecnológicos e promover uma sociedade mais justa e inclusiva para todos.

O combate ao racismo algorítmico requer uma abordagem multidisciplinar que envolva tecnólogos, pesquisadores, legisladores e a sociedade como um todo. Somente por meio de esforços conjuntos podemos garantir que a tecnologia seja usada para promover a igualdade e combater as desigualdades raciais, em vez de perpetuá-las. E isso envolve a utilização de dados de treinamento mais diversificados, a implementação de revisões de

viés e testes abrangentes durante o desenvolvimento de algoritmos e a promoção de uma maior diversidade na indústria tecnológica, tanto em termos de gênero quanto de raça.

Além disso, é fundamental adotar uma abordagem crítica e reflexiva em relação à aplicação da tecnologia, o que implica em questionar os pressupostos e impactos dos algoritmos, envolver a sociedade civil nas discussões sobre ética e tecnologia, e garantir a transparência e a prestação de contas na implementação de sistemas automatizados.

Notas

1. As mídias tradicionais são aquelas que existiam antes da popularização da internet e das redes sociais. São geralmente meios de comunicação de massa unidirecionais, como televisão, rádio, jornais e revistas, que alcançam um grande público de forma padronizada e linear. Essas mídias tradicionais têm sido historicamente importantes para o processo de disseminação de informações, entretenimento e publicidade. Para saber mais, ver o trabalho de Kasele Kazadi (2012).
2. Circuitos ideológicos referem-se a redes de instituições, práticas, discursos e valores que contribuem para a reprodução e legitimação de determinada ideologia dominante em uma sociedade. Esses circuitos podem incluir mídia, educação, religião, cultura popular, entre outros elementos que influenciam a forma como as pessoas pensam e percebem o mundo ao seu redor (Cf. Althusser, 1970).
3. Safiya Umoja Noble é uma professora catedrática da Universidade da Califórnia em Los Angeles e cofundadora e codiretora do UCLA Center for Critical Internet Inquiry.
4. Joy Buolamwini é uma cientista da computação e ativista digital ganense-americana-canadense que trabalha no MIT Media Lab.
5. Tarcízio Silva atualmente é Senior Tech Policy Fellow pela Fundação Mozilla. Também é curador na Desvelar, colaborador na Tecla / Ação Educativa e Instituto Sumaúma. É de sua autoria o livro “Racismo Algorítmico: inteligência artificial e discriminação nas redes digitais”, lançado em 2021.
6. A Inteligência artificial (IA) refere-se “à ciência e engenharia de criar máquinas inteligentes, especialmente programas de computador inteligentes. Está relacionada à tarefa semelhante de usar computadores para entender a inteligência humana, mas a IA não precisa se limitar a métodos que sejam biologicamente observáveis.” (McCarthy, 2007, p.2 – tradução nossa).

Referências

GILLESPIE, Tarleton. **Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media**. New Haven: Yale University Press, 2018, p.160.

JURNO, Amanda Chevtchouk. **(In)visibilidade algorítmica no “feed de notícias” do Facebook**. Contemporânea: Comunicação e Cultura, Salvador, v. 15, n. 2, p. 463-484, maio/ago. 2017. DOI: <https://doi.org/10.9771/contemporanea.v15i2.17796>. Disponível em:

<https://periodicos.ufba.br/index.php/contemporaneaposcom/article/view/17796> - Acesso em 07/06/2023.

JURNO, Amanda Chevtchouk. **Questões e apontamentos para o estudo de algoritmos.** Parágrafo, São Paulo, v.6, n. 1, p. 17-29, jan./abr. 2018. Disponível em: <https://revistaseletronicas.fiamfaam.br/index.php/recicofi/article/view/709> - Acesso em 07/06/2023.

NADAL, K. L. (2014). **A guide to responding to microaggressions.** In CUNY Forum (Vol. 2, No. 1, pp. 71-76).

NOBLE, Safiya Umoja. **Algorithms of oppression: how search engines reinforce racism.** New York: New York University Press, 2018.

NOBLE, Sofiya Umoja. **Google Search: hiper-visibility as a means of rendering black woman and girls invisible.** InVisible Culture: an Electronic Journal for Visual Culture, Rochester, v. 19, 29 out. 2013. Disponível em: <http://ivc.lib.rochester.edu/google-search-hyper-visibility-as-a-means-of-rendering-black-women-and-girls-invisible/> - Acesso em 07/06/2023.

O'NEIL, Cathy. **Weapons of math destruction: how big data increases inequality and threatens democracy.** New York: Broadway Books, 2016.

OPENAI. **“ChatGPT é uma inteligência artificial de linguagem natural desenvolvida pela OpenAI, que usa uma arquitetura de rede neural para gerar respostas a perguntas feitas por usuários.”** Disponível em: <https://chat.openai.com/share/45f91530-60f9-49e6-a96b-f1e23e0fa76f> - Acesso em 13/06/2023.

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

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

SILVA, Tarcízio. **#1 - Eu Existo? O Racismo Algorítmico na rede.** Entrevistador: Giovane Ramos. Comunicando em Preto - Spotify, 29/06/2022. Podcast. Disponível em: <https://open.spotify.com/episode/2HKqIEHX1yFU6dyMfUz6Gg?si=lsZbA9AJQmOAtwFRTnHP4A&nd=1> . Acesso em 13/06/2023.

SCHNOOR, Marina. **IA do Facebook rotulou vídeo de homens negros como ‘primatas’; empresa pede desculpas.** Olhar Digital. 2021. Disponível em: <https://olhardigital.com.br/2021/09/04/internet-e-redes-sociais/ai-facebook-rotulou-video-homens-negros-primatas/> - Acesso em 07/06/2023.

VÍÉS. In: HOUAISS, Antônio. **Dicionário Houaiss da Língua Portuguesa.** Rio de Janeiro, Ed. Objetiva, 2023 [online]. Disponível em: https://houaiss.uol.com.br/corporativo/apps/uol_www/v6-1/html/index.php#0 - Acesso em 07/06/2023.

Recebido: 06 de setembro de 2023

Aprovado: 21 de fevereiro de 2024