

Racismo Algorítmico: uma análise da inteligência artificial a partir de Djamila Ribeiro

Algorithmic Racism: an analysis of artificial Intelligence based on the work of Djamila Ribeiro

Racismo algorítmico: un análisis de la inteligencia artificial basado en la obra de Djamila Ribeiro

Alex Ricardo Weber

Mestrando em Educação, Universidade Comunitária da Região de Chapecó (UNO)
<https://orcid.org/0009-0009-6562-909x>; E-mail: alex.weber@ifsc.edu.br

Andrei Zandoná Bollis

Doutorando em Educação, Universidade Comunitária da Região de Chapecó (UNO)
<https://orcid.org/0009-0003-7745-7754>; E-mail: andrei.bolis@gmail.com

Leonel Piovezana

Pós-Doutor em Educação, Universidade Comunitária da Região de Chapecó (UNO)
<https://orcid.org/0000-0001-8577-319x>; E-mail: leonel@unochapeco.edu.br

RESUMO

Este artigo apresenta uma análise do racismo algorítmico em inteligências artificiais (IA) no Brasil, tendo como eixo analítico as categorias de Ribeiro (estruturalidade, neutralidade algorítmica e práticas antirracistas). O objetivo central é examinar como os estudos contemplam essas três dimensões. Trata-se de uma pesquisa qualitativa, de caráter bibliográfico. A busca na BDTD, na CTD/CAPES e no Portal de Periódicos CAPES recuperou 38 trabalhos; após a aplicação dos critérios de inclusão, cinco (uma tese, uma dissertação e três artigos) compuseram o corpus. O material foi submetido à análise temática de conteúdo com base nas categorias de Ribeiro (2019). O recorte revela a incipiência, mas também a crescente atenção acadêmica à interface entre IA e o racismo.

Palavras-chave: inteligência artificial; racismo algorítmico; Ribeiro; educação antirracista.

ABSTRACT

This article presents an analysis of algorithmic racism in artificial intelligence (AI) in Brazil, using Ribeiro's categories (structurality, algorithmic neutrality, and anti-racist practices) as an analytical axis. The central objective is to examine how studies address these three dimensions. This is qualitative, bibliographical research. The search in the BDTD, CTD/CAPES, and CAPES Periodicals Portal retrieved 38 works; after applying the inclusion criteria, five (one thesis, one dissertation, and three articles) comprised the corpus. The material was subjected to thematic content analysis based on Ribeiro's categories (2019). The selection reveals the incipient, but also the growing academic attention to the interface between AI and racism.

Keywords: artificial intelligence; algorithmic racism; Ribeiro; anti-racist education.

RESUMEN

Este artículo presenta un análisis del racismo algorítmico en la inteligencia artificial (IA) en Brasil, utilizando las categorías de Ribeiro (estructuralidad, neutralidad algorítmica y prácticas antirracistas) como eje analítico. El objetivo central es examinar cómo los estudios abordan estas tres dimensiones. Se trata de una investigación cualitativa y bibliográfica. La búsqueda en el BDTD, CTD/CAPES y el Portal de Publicaciones Periódicas de CAPES arrojó 38 trabajos; tras aplicar los criterios de inclusión, se seleccionaron cinco (una tesis, una disertación y tres artículos) para conformar el corpus. El material fue sometido a un análisis temático de contenido basado en las categorías de Ribeiro (2019). La selección revela una atención académica incipiente, pero también creciente, hacia la relación entre la IA y el racismo.

Palabras clave: inteligência artificial; racismo algorítmico; Ribeiro; educação antirracista.

INTRODUÇÃO

A irrupção dos grandes modelos de linguagem, como GPT-3, deu origem ao ChatGPT, Gemini, DeepSeek, Manus e outros sistemas, como os geradores de imagens, por exemplo, MidJourney, DALL.E 3 e ImageFX, ou os geradores de vídeo como o Gemini VEO 3, OpenAI Sora, ainda o Llama 3. Desde o lançamento do ChatGPT 3, pela OpenAI, em novembro de 2022, o ecossistema sociotécnico brasileiro reconfigurou-se de forma acelerada. A 35ª Pesquisa Anual do Uso de TI, do FGVcia, indica que 56% das médias e grandes empresas nacionais já integram pelo menos uma IAG aos seus fluxos de trabalho (Meirelles, 2024); tendência corroborada por um levantamento empírico que identifica expansão semelhante entre micro e pequenas organizações de cinco setores produtivos (Bonifácio; Schapachnik; Porto, 2025). Paralelamente, multiplicam-se evidências de racismo algorítmico: como o caso da imagem ofensiva criada contra a deputada Renata Souza, que demonstra a reprodução de estereótipos racistas em sistemas de IAG (Araújo; Araújo, 2024). Também “uma análise de mais de 5.000 imagens criadas com a IA Stable Diffusion, feita pela Bloomberg, apontou que o modelo generativo é extremamente enviesado” (Samarão, 2023, p. 64), bem como outros episódios correlatos reportados pela imprensa nacional que pautaram o debate público sobre a discriminação.

A opressão racista também é uma realidade no universo digital. O racismo online tem como objetivo manter o poder político, cultural e econômico dos brancos no espaço digital e, para isso, alimenta um sistema de opressão contra grupos historicamente minorizados (Tynes *et al.*, 2019). Contudo, no ambiente digital, o racismo ocorre em meio a tecnologias que podem potencializar ou inibir essas práticas, dependendo de um conjunto de fatores. Silva (2022) explica que, nos ambientes digitais, evidencia-se um desafio mais profundo relacionado à compreensão de como o racismo se entrelaça às tecnologias por meio de processos “invisíveis”, presentes em recursos automatizados e/ou definidos pelas plataformas, como a recomendação de conteúdo, a moderação, o reconhecimento facial e o processamento de imagens. Por conseguinte, torna-se imperativo compreender as manifestações discriminatórias que se estruturam e se expressam na própria infraestrutura das plataformas, nos algoritmos ou nas interfaces — materializadas em símbolos, imagens, vozes, textos e representações gráficas.

O racismo algorítmico é uma forma moderna e atualizada de discriminação sistêmica por raça, que oprime um grupo em favor de outro e se manifesta por meio da tecnologia (Almeida, 2019; Araújo e Araújo, 2024). Apesar de existirem relatórios e notícias, são poucos os trabalhos e as revisões da literatura sobre IAG e racismo no Brasil.

O objetivo do presente estudo é analisar os artigos selecionados e verificar se, e como, contemplam as categorias: estruturalidade, mito da neutralidade e práticas antirracistas com base no referencial teórico de Ribeiro (2019). Os objetivos específicos visam identificar como cada estudo apresenta a relação entre a tecnologia, a IAG e o racismo em suas respectivas categorias: a) estruturalidade: analisar se o artigo trata o racismo algorítmico como uma falha pontual ou como expressão do racismo estrutural; b) neutralidade: descrever se o artigo considera a tecnologia e a IAG neutras, bem como os argumentos utilizados; c) práticas antirracistas: relacionar as práticas antirracistas apresentadas para promover transformações.

Considerando que a IAG tornou-se popularmente conhecida em 2022, sabe-se que há menos artigos publicados sobre IAG do que sobre IA. Logo, além de analisar os estudos existentes sobre IAGs em conformidade com as categorias de Ribeiro, busca-se também explorar os apontamentos sobre racismo nesses sistemas, delimitando-os como o foco empírico e recente dentro de um debate maior sobre a IA.

Esta pesquisa caracteriza-se como qualitativa, bibliográfica e do tipo Estado do Conhecimento (EC) (Morosini; Kohls-Santos; Bittencourt, 2021), que analisa os resultados a partir de três categorias extraídas de Ribeiro (2019): estruturalidade, neutralidade algorítmica e práticas antirracistas.

As categorias de análise utilizadas neste estudo foram construídas a partir da leitura interpretativa da obra *Pequeno Manual Antirracista*, de Ribeiro (2019). A categoria estruturalidade fundamenta-se na compreensão do racismo como elemento constitutivo das relações sociais e institucionais, ultrapassando o âmbito das ações individuais. A categoria de neutralidade algorítmica foi elaborada a partir da problematização dos discursos que negam ou invisibilizam as relações de poder e as desigualdades raciais, aproximando-se da crítica ao mito da democracia racial discutida pela autora. Já a categoria práticas antirracistas deriva das reflexões de Ribeiro (2019) sobre a necessidade de ações conscientes de enfrentamento ao racismo, que envolvem posicionamentos éticos,

educacionais e políticos. Essas categorias orientaram a leitura e a análise dos trabalhos selecionados, permitindo identificar de que forma os estudos compreendem e discutem o racismo algorítmico nas inteligências artificiais generativas.

O presente estudo tem relevância social, uma vez que o debate sobre a regulamentação de IAs está em andamento no Congresso Nacional. O projeto de lei (PL) 2338/2023, “cria regras para os sistemas de IA disponibilizados no Brasil e estabelece os direitos das pessoas afetadas pelo seu funcionamento” (Samarão, 2023, p. 90). A relevância acadêmica também é notável, considerando a “invasão” da IAG em diferentes áreas, inclusive na educação, onde a educação antirracista permanece relevante por meio de práticas pedagógicas. Além disso, o avanço da ética no uso de IAG, especialmente em contextos educacionais, carece de discussão. Assim, recomendações derivadas do eixo de práticas antirracistas podem subsidiar legislações e currículos antirracistas.

A seguir, na seção 2, descrevemos uma base conceitual antirracista para a geração das três categorias de análise, considerando, em especial, a obra *Pequeno Manual Antirracista* de Ribeiro (2019). Na seção 3, focamos no Estado do Conhecimento sobre o racismo algorítmico a partir de IAGs no Brasil. Na seção 4, apresentamos a análise dos artigos selecionados no EC, com base nas três categorias extraídas da obra de Ribeiro (2019).

O RACISMO NO BRASIL: UM RECORTE A PARTIR DA OBRA PEQUENO MANUAL ANTIRRACISTA

Silva (2022) argumenta que o racismo é um sistema de opressão de pessoas ou grupos de pessoas que nega direitos. Ele possui caráter estrutural e pode parecer paralisante diante da magnitude de uma estrutura social historicamente racista; contudo, também incentiva a luta por meio de práticas antirracistas, visto que esta é a única forma de mitigar os efeitos do racismo estrutural. Este termo pode ser definido como o que descreve Almeida (2019) em seu livro *Racismo estrutural*, que parte do entendimento de que o racismo integra a estrutura social e, portanto, pode se manifestar independentemente de intenção individual. Ainda que o silêncio diante do racismo não implique, necessariamente, culpa moral ou responsabilidade jurídica, ele acarreta implicações éticas e políticas ao contribuir para a manutenção do racismo. Desse modo, a

transformação social não se efetiva apenas por meio de denúncias ou da reprovação moral do racismo, mas exige, sobretudo, a adoção de posicionamentos e práticas antirracistas.

Silva (2022) concorda com Almeida (2019) quanto ao fato de o racismo não ser algo individual, e sim estrutural, profundo e perpetuado ao longo da história, e que deve ser combatido por cada um por meio de práticas antirracistas, especialmente em tempos contemporâneos, com as tecnologias digitais.

O racismo também está nos meios digitais e repercute na continuidade ou não do sistema opressor. Neste sentido, busca-se identificar, nos artigos, a menção ao racismo estrutural, perpetuado por meio da tecnologia de IAG, para posterior análise de como isso é percebido e do que se pode fazer para mitigar essa situação.

Silva (2022), ao discutir o mito da neutralidade algorítmica, apresenta a experiência relatada pela pesquisadora Buolamwini com sistemas robóticos atravessados por vieses algorítmicos. Ao recorrer a um relato pessoal, Buolamwini evidencia um processo de desumanização vivenciado ainda durante sua formação acadêmica, quando sistemas robóticos ignoraram seu rosto e interagiram com todos os demais colegas, sendo ela a única pessoa negra do grupo. Diante da hipótese de que a distinção residia na cor da pele e em traços fenotípicos, realizou um experimento simples, porém significativo: ao utilizar uma máscara branca genérica, sem feições, percebeu que até mesmo essa representação foi reconhecida pelos sistemas, evocando, de modo involuntário, a metáfora presente em *Pele negra, máscaras brancas*, de Frantz Fanon (Silva, 2022).

Essa passagem ajuda na afirmação de que a tecnologia não é neutra, ou que os algoritmos não são neutros (Silva, 2022; Buolamwini; Gebru, 2018; Noble, 2018). Isto porque são feitos por pessoas e, no caso da IAG, são treinados com dados produzidos por pessoas que, em sua maioria, vivem em um sistema que perpetua a discriminação racial. A não neutralidade dos algoritmos pode ser comparada ao mito da democracia racial, apontado por Ribeiro (2019). Silva, autor do livro intitulado *Racismo algorítmico: inteligência artificial e discriminação nas redes sociais*, destaca que a neutralidade e a democracia racial são conceitos distantes, mas que comungam no objetivo de ocultar relações de poder que naturalizam e aprofundam explorações e desigualdades (Silva, 2022). Neste contexto, é comum o argumento de que o algoritmo e os computadores não são racistas nem propagam o racismo, pois seriam mais “objetivos”, sem vieses, por terem suas saídas — as

respostas dos programas — resultantes de cálculos, matemática e/ou estatística. Neste sentido, os defensores da neutralidade podem estar tomados pelo conceito de tecnochauvinismo de Meredith Broussard, que consiste na confiança de que a tecnologia será sempre a solução e de que a tecnologia ou os algoritmos são puramente matemáticos ou estatísticos (Silva, 2022).

Pode-se ainda pensar que a afirmação de que a tecnologia é, ou pode vir a ser, um mito, derivado do anterior, cujo objetivo é o mesmo: argumentar pela manutenção das desigualdades. Contudo, num contexto diferente, no meio digital, nas relações estabelecidas pela internet. Seria uma atualização contemporânea das forças do sistema racista já apontadas por Ribeiro (2019) e exemplificadas pela diferenciação do racismo no Brasil em relação a outros, como o regime nazista, o apartheid sul-africano ou o racismo contra a população negra nos Estados Unidos. Nestes casos, o racismo era explícito e institucionalizado por meio de leis e práticas oficiais (Ribeiro, 2019).

Já no Brasil, o mito da democracia racial é um dos principais operadores que enraíza, de forma implícita, o preconceito nas relações sociais. Conforme destaca Ribeiro (2019), formulado e difundido por sociólogos vinculados à elite econômica na metade do século XX, esse mito sustenta a ideia de que, no Brasil, os conflitos raciais teriam sido superados por uma suposta harmonia entre negros e brancos, expressa na miscigenação e na inexistência de leis segregacionistas.

Essa formulação ideológica sustenta a premissa de que as tensões raciais e as assimetrias históricas teriam sido superadas pela miscigenação, bem como pela ausência de aparatos legislativos segregacionistas oficiais no país. Segundo a autora, esse mito consolidou, no imaginário social brasileiro, uma postura discursiva ambígua, caracterizada pela negação generalizada do próprio preconceito por indivíduos que, paradoxalmente, reconhecem a existência do racismo no país. Para ilustrar esse ponto de vista, a autora usa um exemplo de pesquisa do Datafolha de 1995, na qual se verificou que 89% dos brasileiros reconheciam a existência de preconceito de cor no país, enquanto 90% não se identificavam como racistas. À época, o estudo foi considerado o mais abrangente sobre o tema, tendo entrevistado 5.081 pessoas com mais de 16 anos, em 121 cidades de todas as unidades da federação.

Nos últimos anos, a percepção do preconceito permanece elevada na sociedade

brasileira. Em uma pesquisa do Datafolha de novembro de 2024, foram entrevistadas 2.004 pessoas; dentre elas, 59% disseram que a maioria da população brasileira é racista. Ainda que 73% dos pretos e pardos já tenham se sentido discriminados por conta da cor da pele (Alves, 2024).

Fica evidente que algoritmos aprendem a partir de bases de dados permeadas por essa estrutura racista; portanto, os vieses não são “falhas”, mas uma continuidade histórica. Na análise a ser realizada neste estudo, pretende-se evidenciar que os artigos encontrados corroboram essa visão.

Quanto ao racismo algorítmico, faz-se necessário conceituar as microagressões. Elas se configuram como ofensas verbais, comportamentais ou ambientais recorrentes, intencionais ou não, que expressam desrespeito por meio de insultos hostis, depreciativos ou negativos. Essas manifestações recaem sobre grupos socialmente vulnerabilizados, como pessoas racializadas, mulheres e migrantes, bem como sobre as interseções entre essas condições (Silva, 2022; Sue, 2010).

O contexto das microagressões é complexo e cheio de nuances, uma boa explicação é oferecida por Silva (2022, p. 28): “A reprodução das culturas hegemônicas na produção de artefatos midiáticos ficcionais, bem como na midiaticização da sociedade, disfarça-se de” e é completada por Viana e Belmiro (2019, p. 199) “liberdade de expressão, de padrões de beleza, de gostos pessoais, de humor e de uma infinidade de microagressões que subjagam, ofendem, diminuem, maltratam, excluem e submetem à violência a população negra”.

As microagressões são classificadas em deseducação, microinvalidações, microinsultos e desinformação. Araújo (2024), com base na teoria de Silva (2022), apresenta as seguintes definições das categorias de microagressões. Entre elas, destacam-se os microinsultos, caracterizados por comentários aparentemente inofensivos, muitas vezes disfarçados de humor ou de críticas indiretas, mas que carregam sentidos pejorativos em relação à identidade racial.

Também se evidenciam as microinvalidações, que operam ao invalidar ou ao minimizar as experiências raciais dos indivíduos, deslegitimando seus sentimentos e vivências. Soma-se a isso a deseducação, entendida como a disseminação de informações equivocadas ou de estereótipos que reforçam concepções prejudiciais sobre determinados

grupos raciais. Por fim, a desinformação manifesta-se na propagação de conteúdos falsos ou distorcidos que contribuem diretamente para a manutenção de estereótipos e da discriminação racial no ambiente virtual. Conforme destaca Silva (2022), tais processos podem ocorrer de forma intencional ou não, mas, em ambos os casos, resultam em impactos significativos ao perpetuarem representações estigmatizantes de pessoas pretas.

Com a compreensão de microagressões aplicada, torna-se perceptível, na prática, como o racismo algorítmico se manifesta, e pode-se defini-lo conceitualmente. Além do termo racismo algorítmico, encontra-se na literatura o termo preconceito algorítmico e, de forma mais ampla, abrangendo outros preconceitos, um termo moderno é algoritmos enviesados.

A importância da prática antirracista: em meio a essa discussão, Ribeiro (2019, p. 9) afirma que “mais grave é não reconhecer e não combater a opressão”. Em seguida, ela pergunta: “O que cada um de nós tem feito e pode fazer pela luta antirracista?” (Ribeiro, 2019, p. 9).

A partir desse questionamento fundamental, a autora inaugura uma reflexão profunda acerca das posturas e ações a serem adotadas por sujeitos comprometidos com o enfrentamento do racismo. O primeiro passo pode ser: “o autoquestionamento — fazer perguntas, entender seu lugar e duvidar do que parece natural — é a primeira medida para evitar reproduzir esse tipo de violência, que privilegia uns e oprime outros” (Ribeiro, 2019, p. 9).

A autora ressalta que o enfrentamento da questão racial passa também pela superação de sua invisibilização, o que implica reconhecer explicitamente as diferenças raciais. Nesse sentido, problematiza-se a ideia de “não ver cor”, uma vez que o cerne da questão não reside na cor em si, mas no modo como ela é historicamente utilizada para justificar práticas de segregação e opressão. Assim, torna-se fundamental reconhecer a diversidade racial e promover o debate tanto sobre a negritude quanto sobre a branquitude, considerando que as relações sociais são atravessadas por dimensões raciais (Ribeiro, 2019).

Sobre negritude e branquitude, a autora reflete que os brancos naturalizaram a ausência de negros nos espaços de poder. “Para desnaturalizar isso, todos devem

questionar a ausência de pessoas negras em posições de gerência, de autores negros em antologias, de pensadores negros na bibliografia de cursos universitários, de protagonistas negros no audiovisual” (Ribeiro, 2019, p. 13). Em seguida, enfatiza-se que o sujeito branco deve compreender os privilégios estruturais atrelados à branquitude, visto que, embora esse grupo não constitua a maioria da população, ocupa desproporcionalmente os espaços de poder. Assim, “Isso é importante para que privilégios não sejam naturalizados ou considerados apenas esforço próprio” (Ribeiro, 2019, p. 13).

A autora afirma que “a prática antirracista é urgente e se dá nas atitudes mais cotidianas” (Ribeiro, 2019, p. 6). Neste sentido, busca-se identificar e discutir de que modo os artigos encontrados sugerem práticas antirracistas no contexto da IAG.

Além de estimular o reconhecimento do racismo estrutural, sustentado por mitos como o da democracia racial, agora reformulado na ideia de neutralidade algorítmica, e de promover a prática antirracista cotidiana, a autora destaca a necessidade de problematizar a branquitude, valorizar as produções intelectuais negras e assegurar a inclusão efetiva. Essas categorias têm relevância equivalente no processo antirracista, mas foram omitidas neste artigo para permitir uma análise mais aprofundada das categorias escolhidas, por se considerar que elas apareceram com maior destaque nos trabalhos examinados. As categorias estruturalidade, neutralidade e práticas antirracistas serão utilizadas na análise para identificar como os artigos selecionados abordam essas temáticas.

ESTADO DO CONHECIMENTO SOBRE O RACISMO ALGORÍTMICO EM INTELIGÊNCIAS ARTIFICIAIS GENERATIVAS NO BRASIL

O Estado do Conhecimento (EC) compreende a identificação, o registro, a categorização, a reflexão e a síntese das publicações científicas de uma área ao longo de um período (Morosini; Fernandes, 2014). O EC visa “fornecer um mapeamento das ideias já existentes, dando-nos segurança quanto às fontes de estudo e apontando subtemas passíveis de maior exploração” (Morosini; Fernandes, 2014, p. 158). Para isso, faz-se necessário desenvolver o trabalho com rigor ao longo do percurso metodológico. Neste sentido, definem-se, em Morosini, Kohls-Santos e Bittencourt (2021), os seguintes passos: definição do objetivo, estabelecimento das bases de dados a serem utilizadas, definição dos descritores ou palavras-chave para a busca, realização da bibliografia anotada, da

bibliografia sistematizada, da bibliografia categorizada e, por fim, da bibliografia propositiva. A seguir, apresentam-se os procedimentos metodológicos e os principais resultados obtidos no EC.

Com o propósito de delimitar o que se quer conhecer sobre determinada área de estudo ou tema (Morosini; Kohls-Santos; Bittencourt, 2021), define-se o objetivo desta busca pelo Estado do Conhecimento como: Conhecer e analisar as publicações nacionais, realizadas a partir de 2023, sobre racismo algorítmico no contexto de inteligências artificiais generativas, com um olhar pelas lentes da obra de Ribeiro (2019). O recorte temporal é inferior a 5 anos, uma vez que esta tecnologia se tornou conhecida mundialmente em novembro de 2022.

O passo seguinte foi definir os repositórios a serem utilizados. Em consonância com o objetivo do estudo, realizaram-se buscas em bases de dados de acesso livre ou disponibilizadas por meio institucional. As buscas foram realizadas nas seguintes bases de dados: Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), Catálogo de Teses e Dissertações (CTD) da Capes e o Portal de Periódicos da CAPES (PPC). Os critérios de seleção de trabalhos para o EC são: artigos de periódicos, teses ou dissertações de nível *stricto sensu* (mestrado, doutorado ou pós-doutorado); textos em português, espanhol ou inglês; e com foco em IAG.

Como critérios de exclusão deste EC, adotaram-se os seguintes: indisponibilidade de acesso gratuito ao trabalho completo; estudos internacionais; ausência de abordagem com IAG; estudos já selecionados em outra fonte; trabalhos de conclusão de curso em nível de graduação.

Os descritores ou palavras-chave utilizadas para essa busca foram: “racismo algorítmico”. As strings de busca de cada base de dados são detalhadas no Quadro 1. Não foi usado o termo inteligência artificial generativa nessa busca, o que mais tarde se revelou importante, pois alguns trabalhos encontrados usam o termo IA, mesmo quando a referência era a uma IAG, como o ChatGPT.

Além das palavras-chave de busca, o quadro 1 apresenta um resumo dos caminhos percorridos para selecionar os estudos em cada base de dados. Descreve-se a palavra-chave utilizada e, caso necessário, os filtros aplicados para restringir melhor os resultados, em conformidade com o objetivo apresentado, o tipo de material (tese, dissertação ou

artigo), a quantidade de trabalhos encontrados, de trabalhos selecionados e de trabalhos excluídos. As buscas dos estudos foram realizadas entre 12 e 16 de junho de 2025.

Quadro 1: Mapeamento de trabalhos a partir de cada base pesquisada

Palavras-chave	Base de dados	Tipo de Material	Total de trabalhos encontrados	Trabalhos selecionados	Total de excluídos
“racismo algorítmico”	BDTD	Teses e Dissertações	12	0	12
“racismo algorítmico”	CTD	Teses e Dissertações	13	2	11
“racismo algorítmico”	PPC	Teses, Dissertações e Artigos	13	3	10
TOTAL			38	5	33

Fonte: Elaborado pelos autores (2026).

Foram 38 resultados no total, dos quais 5 foram selecionados. Dentre os trabalhos selecionados, a maioria (3) foi identificada por meio de uma busca no Portal de Periódicos da CAPES. Um trabalho selecionável foi encontrado na BDTD; contudo, ele já havia sido selecionado na CTD, na qual dois trabalhos foram selecionados. Os trabalhos foram escolhidos após a leitura do título, do resumo e, quando necessário, de trechos do trabalho completo. A seguir, apresenta-se um quadro com informações importantes dos trabalhos selecionados, bem como a descrição da sua caracterização.

O Quadro 2 apresenta o detalhamento de cada trabalho selecionado. Os autores, o ano de publicação, o título do artigo e o tipo: tese (T), dissertação (D) ou artigo (A). Observou-se que um autor produziu dois trabalhos selecionados nesta revisão. Em um desses estudos, o autor conta com a colaboração de outro com o mesmo nome e sobrenome. Os trabalhos analisam o mesmo fenômeno, mas sob perspectivas e metodologias distintas.

Todos os trabalhos são nacionais, conforme o objetivo da busca que é conhecer/analisar as publicações nacionais, realizadas a partir de 2023, sobre racismo algorítmico no contexto de inteligências artificiais generativas.

Quanto ao ano de publicação, a tese e a dissertação são de 2023, assim como um dos 3 artigos encontrados. Os outros artigos são de 2024. Supõe-se que trabalhos de mestrado e doutorado possam ter sido concluídos até meados de 2025 e ainda não terem

o trabalho final disponibilizado online, assim como artigos que possam estar em avaliação.

Quadro 2: Dados sobre teses, dissertações e artigos sobre racismo algorítmico considerando a inteligência artificial generativa

ANO	AUTORES	TÍTULO	T/D/A
2023	Samarão, Lilianny Alves	IA e algoritmos enviesados: as dimensões do viver na sociedade 5.0	T
2023	Nepomuceno, Thiago de Godoy	Reflexões sobre a Colonialidade Digital: Explorando a dialética entre o ensino de arte e a análise crítica sociológica dos estudantes enquanto sujeitos inseridos na cibercultura.	D
2024	Araújo, Júlio	Racismo algorítmico e microagressões nas redes sociais	A
2024	Araújo, Júlio; Araújo, Júlio	Racismo algorítmico e inteligência artificial: uma análise crítica multimodal	A
2023	Damasceno, Adriana; Peruzzo, Cicilia Krohling	O tecnicismo na comunicação contemporânea	A

Fonte: elaborado pelos autores (2026).

Quanto aos detalhes metodológicos dos estudos selecionados, apresentam-se as abordagens de pesquisa, os sujeitos do estudo, o *locus*, ou local onde a pesquisa foi realizada e um resumo dos objetivos.

Todas as cinco pesquisas selecionadas adotam uma abordagem qualitativa. Quanto ao delineamento, o *corpus* divide-se em: um estudo empírico (Nepomuceno, 2023), conduzido com duas turmas do Ensino Médio; dois estudos de base bibliográfica ou documental (Damasceno; Peruzzo, 2023; Samarão, 2023); e dois estudos de caso (Araújo, 2024; Araújo; Araújo, 2024).

Os trabalhos apresentam objetivos complementares sobre o racismo algorítmico. Ao reconhecer que os algoritmos carregam marcas históricas de poder e, portanto, “não são neutros”, Samarão (2023) propõe identificar como esses vieses são gerados e como se espalham, discutir suas consequências tecnológicas, sociais e éticas e, assim, sublinhar a urgência de uma regulação democrática da IA. Essa premissa, de que algoritmos não são neutros, orienta a literatura recente a explicitar como distorções tecnológicas se materializam em diferentes esferas socioculturais: da colonialidade digital que permeia a

cibercultura e o ensino de artes, onde a arte é acionada como mediação crítica (Nepomuceno, 2023), às microagressões direcionadas a figuras públicas negras, como a deputada Renata Souza, após a circulação de uma imagem racista gerada por inteligência artificial generativa (Araújo, 2024).

Ao detalhar as estratégias multimodais que constituem essa imagem e seus efeitos ideológicos, Araújo e Araújo (2024) reforçam a necessidade de desvendar as interações entre modos de representação que sustentam o racismo algorítmico. Nesse mesmo horizonte crítico, Damasceno e Peruzzo (2023) mostram como a introdução de redatores-robôs no jornalismo aprofunda o tecnochauvinismo, apontando para transformações socioculturais que demandam vigilância ética e regulamentação participativa. Em conjunto, esses estudos articulam a urgência de compreender, denunciar e regular as dinâmicas de poder inscritas na IA, seja na sala de aula, na esfera midiática ou na produção automatizada de conteúdo, a fim de mitigar seus impactos discriminatórios e promover justiça social. As demais informações dos trabalhos, pertinentes ao escopo deste estudo, são apresentadas e analisadas a seguir.

ANÁLISE DOS RESULTADOS

Os textos serão elaborados e desenvolvidos por meio da análise de conteúdo de Bardin (2016), cujo objetivo é superar as incertezas e enriquecer a leitura dos dados. A análise de conteúdo é um procedimento cujo ponto de partida é a mensagem. A mensagem pode ser “verbal (oral ou escrita), gestual, silenciosa, figurativa, documental ou diretamente provocada” (Franco, 2005, p. 13).

Após a leitura dos trabalhos selecionados, trechos desses textos foram codificados nas três categorias criadas a partir dos temas extraídos do texto de Ribeiro (2019). Como os textos dos artigos selecionados tratavam de racismo algorítmico, pressupôs-se que eles trariam as categorias obtidas conforme descrito na obra de Ribeiro (2019) e que, por consequência, fazem parte da base teórica de vários trabalhos. O objetivo da análise foi evidenciar as publicações nacionais específicas sobre racismo algorítmico no contexto das inteligências artificiais generativas, à luz da obra de Ribeiro (2019). Abaixo, trechos extraídos dos trabalhos selecionados para cada uma das três categorias.

A estruturalidade do racismo é a compreensão de que ele tem caráter estrutural:

não é algo individual; é profundo e se perpetua ao longo da história. É um sistema de opressão de pessoas ou grupos de pessoas que retiram direitos. O racismo também está nos meios digitais e repercute na continuidade ou não do sistema opressor. O entendimento desta estruturalidade e de sua profundidade, inclusive nos meios digitais, é significativo para a transformação e para iniciar o movimento em direção às práticas antirracistas. Nos trabalhos selecionados, ele é descrito de diferentes formas em todos os textos.

Embora Samarão (2023) não recorra explicitamente ao termo racismo estrutural, a análise de seus argumentos evidencia o reconhecimento do caráter estrutural e sistêmico do racismo, bem como de sua relação com as desigualdades históricas que atravessam tanto os dados de treinamento quanto a atuação dos desenvolvedores de inteligência artificial. Dessa forma, a autora aponta as dificuldades de treinar sistemas de IA livres de preconceitos em contextos sociais ainda conduzidos hegemonicamente por pessoas brancas, destacando que a ausência de cuidados no processo de treinamento e de testes com grupos diversos tende a produzir e reforçar vieses discriminatórios.

Fica clara a relação entre os argumentos da autora e as características da estruturalidade apontadas por Ribeiro (2019) e outros. Samarão (2023, p. 49) ainda afirma que “[...] muitos dos problemas têm muito mais relação com as informações históricas do que, necessariamente, com os algoritmos”, enfatizando o peso do fator estrutural do racismo sobre o aspecto técnico das IAGs.

Do texto de Nepomuceno (2023) emerge a premissa de que a propagação do racismo por meio dos algoritmos é um problema estrutural, e não uma falha técnica ou humana de programação. Os vieses raciais presentes nesses sistemas refletem a cultura e os valores das pessoas envolvidas em seu desenvolvimento, desde acionistas, proprietários e gestores até programadores e demais profissionais. Desse modo, o racismo algorítmico constitui uma manifestação do racismo estrutural que permeia a sociedade e se reproduz nas tecnologias digitais. Em vista disso, o autor argumenta que os artefatos algorítmicos superam a condição de mera falha técnica ou humana de programação, configurando-se como sistemas complexos de reprodução estrutural. Nesses arranjos, elementos como a estética e a classe articulam-se a diferentes matrizes simbólicas. Essa dinâmica torna evidente a inter-relação entre colonialidade, capitalismo e racismo, aspectos que se

refletem diretamente no ambiente digital.

Araújo (2024) analisa os comentários das pessoas sobre o caso da deputada estadual do RJ, cuja imagem correspondente — gerada por um sistema de IA — apresentou teor preconceituoso, ao mostrar uma mulher negra na favela com uma arma. Com relação à IAG que gerou a imagem com viés, ele aponta a perpetuação das desigualdades em função de que a “influência dos algoritmos vai muito além de instrumentos inócuos de processamento de dados. Eles desempenham um papel ativo na intensificação das disparidades sociais já existentes, criando e perpetuando desigualdades” (Araújo, 2024, p. 6).

Em outro texto do mesmo autor, fica explícita a perspectiva de que o racismo algorítmico é uma expressão do racismo estrutural, não uma falha pontual. Ele conecta a lógica dos dados de treino, a desigualdade de acesso e as histórias coloniais à forma como a IA perpetua o racismo.

O racismo algorítmico é uma forma moderna dessa discriminação, perenizando preconceitos raciais através da tecnologia, com as mulheres negras sendo particularmente impactadas. Esse fenômeno reforça preconceitos raciais e de gênero, limitando seu acesso a oportunidades econômicas, sociais e políticas, e mantendo estruturas de poder históricas (Araújo; Araújo, 2024, p. 90).

Damasceno e Peruzzo (2023, p. 6) afirmam que “apesar da positividade dessa tecnologia, a estrutura técnico-algorítmica tende a facilitar e reproduzir expressões de racismo, uma vez que vivemos em uma sociedade historicamente e estruturalmente racista, [...]”. Ao dialogarem com Noble (2018), as autoras reforçam que os resultados produzidos pelos algoritmos de IAG não apenas decorrem desse contexto, mas também o alimentam, na medida em que podem difundir informações prejudiciais sobre indivíduos, contribuir para invisibilizações sistêmicas e promover formas de demarcação digital que intensificam e legitimam situações de opressão.

Portanto, percebeu-se que todos os estudos destacam a necessidade de abordar o racismo a ponto de compreender sua engrenagem estrutural, intimamente imbricada à cultura. Posteriormente, é possível adotar ações antirracistas. Especialmente no contexto tecnológico relacionado aos algoritmos, que perpetuam essa estruturalidade no meio digital.

Ao dialogar com Ribeiro (2019), percebe-se que o racismo algorítmico presente nas inteligências artificiais generativas não pode ser compreendido como resultado de falhas técnicas isoladas, mas como expressão de um sistema historicamente estruturado pelas desigualdades raciais produzidas desde o período escravocrata. Os dados utilizados no treinamento das IAGs operam como repositórios de uma memória social marcada pela exclusão da população negra do acesso a direitos, educação, produção intelectual e representatividade social. A sub-representação de sujeitos negros em bases de dados e a associação recorrente entre negritude, violência e marginalização reproduzem, em linguagem algorítmica, a lógica histórica de demarcação de posições sociais inferiores. Assim, os vieses identificados nos estudos analisados evidenciam que as IAGs não apenas refletem desigualdades estruturais, mas também podem contribuir para a continuidade e a legitimação dessas desigualdades no ambiente digital.

A neutralidade algorítmica é um argumento de defesa da IAG, da IA e de outras tecnologias, utilizado por autores geralmente descomprometidos com a educação antirracista. Em contrapartida, as obras que integram o corpus desta pesquisa alinham-se explicitamente à perspectiva crítica, dedicando-se a descortinar os mecanismos sociotécnicos que perpetuam as assimetrias raciais. Estes autores estão imbuídos do esclarecimento das artimanhas estruturais que o racismo emprega para se perpetuar. O argumento de neutralidade algorítmica, de que basta usar mais dados no treinamento da IAG ou de que os algoritmos são puramente cálculos matemáticos e, por isso não seriam capazes, ou responsáveis por reproduzir o racismo, já foi questionado por diversos autores (Silva, 2022; Buolamwini; Gebru, 2018; Noble, 2018).

Nesta categoria, observa-se a defesa da não neutralidade dos algoritmos em todos os textos selecionados. Samarão (2023) indica que “os dados não são neutros”. O texto utiliza a obra de Wachter-Boettcher para argumentar contra a neutralidade. Samarão (2023) argumenta que os sistemas automatizados não têm a capacidade de reconhecer, por si mesmos, a inadequação de suas conclusões e, até o momento, dependem de feedback externo de agentes humanos responsáveis por sua programação. Dessa forma, quando um algoritmo apresenta vieses, não se trata de uma falha do computador em si, mas de limitações inerentes à própria estrutura algorítmica, que refletem as escolhas e direcionamentos de quem o desenvolveu. Tal compreensão conduz a uma reflexão mais

ampla sobre a suposta neutralidade dos algoritmos, evidenciando a presença de vieses e preconceitos em seu funcionamento.

Já Nepomuceno (2023) realiza uma pesquisa empírica e descreve suas práticas pedagógicas para discutir a não neutralidade dos algoritmos. O autor apresenta aos estudantes algumas obras antirracistas e, depois, realiza práticas em sala de aula para discutir e evidenciar a não neutralidade dos algoritmos. Ele descreve as práticas em que ele e seus estudantes fizeram consultas para que o ChatGPT criasse personagens para uma história literária. Para tanto, solicitaram que descrevessem as características físicas, como traços do rosto e a cor da pele, da empregada doméstica e do empresário. O trabalho também apresenta buscas no Google, nas quais as respostas com imagens de profissões e outras denunciam respostas racistas. Dentre outros referenciais, o autor adota como fundamento central a obra *Algoritmos da Opressão*, de Noble (2018), que destaca a necessidade de problematizar os sujeitos responsáveis pela criação dos códigos e a forma como esses sistemas são desenvolvidos. A autora questiona a concepção de que os algoritmos seriam inerentemente neutros e objetivos por se basearem em cálculos matemáticos, argumentando, ao contrário, que tais tecnologias refletem os valores, crenças e visões de mundo de seus criadores, os quais podem incluir perspectivas racistas, sexistas ou fundamentadas em noções equivocadas de meritocracia. Sob o prisma de Ribeiro (2019), o viés algorítmico denuncia o “lugar de fala” corporativo que desenha a tecnologia: um espaço majoritariamente masculino, branco e ocidental, cujas vivências e privilégios são codificados como se fossem o padrão universal neutro da humanidade.

Ao considerar os algoritmos como uma forma de linguagem, segundo Araújo (2024, p. 5), “é possível perceber que, por trás deles, estão representações que moldam e direcionam a comunicação e a interação entre indivíduos”. Ele ainda considera preocupante o fato de que os algoritmos fomentam o discurso racista nas redes sociais, ao contribuírem para a disseminação de mensagens prejudiciais e discriminatórias e para a reprodução de preconceitos (Araújo, 2024).

O segundo trabalho do autor analisa, mais especificamente, a imagem gerada pela IAG no caso da deputada estadual do Rio de Janeiro, Renata Souza. Ela pediu à IAG que criasse uma imagem no estilo Pixar de uma mulher negra, com cabelos afro e roupas com estampa africana, em um cenário de favela. A repercussão desse caso foi grande, pois a

imagem produzida pela IAG apresentou os elementos visuais solicitados, incluindo uma arma nas mãos da personagem. Este artigo mostra empiricamente, por meio de análise multimodal de imagem, como a IA produziu uma saída sintética enviesada, reforçando o argumento de que os dados não são neutros. Conforme apontam Araújo e Araújo (2024), a articulação de determinados elementos visuais pode resultar na construção de mensagens de cunho discriminatório, como, por exemplo, a associação de uma mulher negra à violência em um contexto de favela. Tal representação evidencia uma forma de racismo algorítmico, na medida em que reproduz e reforça estereótipos raciais e preconceitos já presentes na estrutura social, os quais também se manifestam nas imagens geradas por sistemas de inteligência artificial.

Damasceno e Peruzzo (2023) citam trechos de O'neil (2020) para explicar que “apesar da reputação de imparcialidade atribuída aos algoritmos, eles refletem objetivos e ideologias”. Também cita Noble (2018) para destacar que os algoritmos não são benignos, neutros nem objetivos. As autoras ainda usam o conceito de Tecnochauvinismo citando Silva (2022). “É o tecnochauvinismo que faz parecer que notícias racistas, que podem vir a ser geradas por esses programas, seriam apenas um espelho da realidade social. Dessa forma, textos problemáticos não apenas pareceriam normais, mas também inescapáveis” (Damasceno; Peruzzo, 2023, p. 8).

A partir das reflexões de Ribeiro (2019), é possível compreender que o discurso da neutralidade algorítmica opera de modo semelhante ao mito da democracia racial no Brasil, ao ocultar relações históricas de poder sob a aparência de imparcialidade técnica. Assim como a ideia de harmonia racial contribuiu para invisibilizar práticas racistas estruturais, a crença de que algoritmos e inteligências artificiais produzem resultados objetivos por meio de cálculos matemáticos tende a mascarar os interesses, valores e escolhas humanas envolvidos em seu desenvolvimento. O tecnochauvinismo atua como mecanismo discursivo que naturaliza desigualdades produzidas pelas tecnologias digitais, atribuindo legitimidade a resultados discriminatórios sob o argumento da racionalidade estatística. Dessa forma, os vieses algorítmicos deixam de ser percebidos como construções sociais e passam a ser tratados como consequências inevitáveis do funcionamento técnico das plataformas.

Compreender a estruturalidade do racismo e sua reconfiguração no ambiente

digital, especialmente por meio da perpetuação de práticas discriminatórias em sistemas algorítmicos e plataformas online, possibilita refletir sobre estratégias capazes de enfrentar ou minimizar os efeitos racistas presentes nas inteligências artificiais generativas. Dessa forma, apresenta-se a categoria práticas antirracistas, voltada às ações antirracistas sugeridas nos estudos analisados.

Resgatando um pouco da fundamentação teórica da categoria de práticas antirracistas, destaca-se a afirmação de Ribeiro (2019) de que a inação contra a opressão e a falta de reconhecimento do racismo são atitudes mais graves do que o uso de palavras como branco, preto, negro, racismo e racista. Em seguida, ela pergunta o “que cada um de nós tem feito e pode fazer pela luta antirracista?”. Esta pergunta é o que impulsiona a terceira categoria de análise. A própria autora inicia a resposta, que seguirá sendo respondida considerando os apontamentos dos trabalhos selecionados. Ribeiro (2019, p. 9) sugere “o autoquestionamento - fazer perguntas, entender seu lugar e duvidar do que parece ‘natural’ [...]” para perceber o sistema de opressão e de privilégios em funcionamento. Ela também sugere que se atue para retirar o preconceito racial da invisibilidade, nomeando as opressões.

Nos textos selecionados, várias sugestões são apresentadas para mitigar o racismo no uso de IAs. De forma resumida, a categoria práticas antirracistas identificou cinco ações sugeridas nos artigos selecionados. A educação antirracista foi mencionada em todos os trabalhos. Três trabalhos apontaram para ações de regulamentação das IAs (Samarão, 2023; Nepomuceno, 2023; Araújo; Araújo, 2024). A mesma quantidade de artigos apontou para a importância da diversidade e da representatividade das equipes de desenvolvimento de IAs (Araújo; Araújo, 2024; Damasceno; Peruzzo, 2023; Samarão, 2023). Dois trabalhos defendem a transparência nos processos algorítmicos com auditorias (Araújo; Araújo, 2024; Damasceno; Peruzzo, 2023). E um descreveu iniciativas afrocentradas (Damasceno; Peruzzo, 2023).

As práticas antirracistas identificadas nos estudos analisados dialogam diretamente com a compreensão de Ribeiro (2019) de que não basta reconhecer a existência do racismo sem assumir uma postura ativa de enfrentamento. A inação diante das opressões contribui para a manutenção das desigualdades estruturais, exigindo ações éticas, políticas e educacionais contínuas. As propostas encontradas na literatura, como educação

antirracista, regulamentação das inteligências artificiais, transparência algorítmica, ampliação da diversidade nas equipes de desenvolvimento e fortalecimento de iniciativas afrocentradas, representam estratégias voltadas à retirada do racismo algorítmico da invisibilidade técnica e social. Além disso, essas ações contribuem para questionar a centralidade da branquitude nos processos de produção tecnológica e para ampliar a presença de sujeitos, conhecimentos e bases de dados historicamente marginalizados no treinamento e desenvolvimento das IAGs.

Em busca de um detalhamento qualitativo das ações sugeridas na categoria de práticas antirracistas, apresentam-se as ações consideradas mais relevantes. No que tange à regulamentação, o estudo mais abrangente é o da Samarão (2023). Devido à natureza acadêmica e ao fôlego de uma tese, sua extensão e seu escopo analítico apresentam-se estruturalmente mais amplos do que os observados em artigos. A autora relaciona algumas entidades que atuam na pesquisa sobre a regulamentação da IA e também aponta a situação da legislação em diversos países. Sobre a necessidade e a urgência de uma regulação a nível global, a autora destaca que grande parte das tecnologias de inteligência artificial ainda se concentra no domínio das grandes empresas de tecnologia. Com o avanço da IA generativa, essa preocupação se intensifica, uma vez que esses sistemas passam a produzir conteúdos em larga escala sem garantias quanto à ausência de vieses em seus processos. Nesse contexto, torna-se fundamental compreender de que modo essas tecnologias são treinadas, como os dados são tratados e como as saídas sintéticas são disponibilizadas aos indivíduos.

A educação para o uso da IA também foi abordada em diversos textos. Um deles (Nepomuceno, 2023) apresentou e detalhou as intervenções pedagógicas realizadas em seu estudo. Várias delas, por meio da arte, abordam temas contemporâneos, como as fake news. Araújo e Araújo (2024, p. 107) também defendem a educação sobre racismo algorítmico “capacitando usuários a questionar preconceitos em algoritmos”. Araújo (2024, p. 31) afirma que “é importante promover a educação e conscientização sobre como nossas palavras e ações podem afetar os outros”.

Sobre as iniciativas afrocentradas citadas por Damasceno e Peruzzo (2023), elas descrevem o Nappy, um banco de imagens com fotos gratuitas em alta resolução de pessoas negras em situações do cotidiano. Também tem o Young, Gifted & Black, que é

outro banco de imagens produzido por mulheres negras. Estes têm como objetivo combater representações negativas, oferecendo diversas imagens de pessoas negras sem discriminação. Outros exemplos, inclusive nacionais, de empoderamento são apresentados no trabalho das autoras. Contudo, não estão relacionadas com IAG.

Assim, analisamos, a partir da perspectiva de Bardin (2016), que se evidencia um movimento interpretativo que articula recorte, categorização e inferência a partir das mensagens presentes nos textos (Franco, 2005), permitindo observar como os significados sobre racismo algorítmico são construídos no campo. Ao organizar os excertos nas categorias derivadas de Ribeiro (2019), nota-se que essas categorias operam não apenas como instrumentos de classificação, mas também como eixos que tornam visíveis determinadas regularidades discursivas, especialmente no que se refere à compreensão do racismo como fenômeno estrutural que se reconfigura no ambiente digital.

Nesse processo, os estudos analisados, ainda que com enfoques distintos, como a ênfase nas dimensões históricas em Samarão (2023), a abordagem empírica e pedagógica em Nepomuceno (2023), a análise discursiva e multimodal em Araújo (2024), em Araújo e Araújo (2024) e a problematização sociotécnica em Damasceno e Peruzzo (2023), mobilizam argumentos que se relacionam e se reforçam mutuamente. Tal recorrência também se manifesta na problematização da neutralidade algorítmica, tensionada por autores como Silva (2022), Buolamwini e Gebru (2018) e Noble (2018), evidenciando que os algoritmos incorporam valores e estruturas sociais. Ao mesmo tempo, a categoria práticas antirracistas revela um deslocamento do plano analítico para o propositivo, no qual as sugestões identificadas — educação, regulação, diversidade e transparência — emergem como desdobramentos das próprias interpretações realizadas nos estudos.

Dessa forma, evidenciam-se as formas pelas quais os estudos analisados vêm produzindo, reiterando e articulando sentidos sobre o racismo algorítmico em inteligências artificiais generativas, ao mesmo tempo em que se delineiam compreensões recorrentes, tensões e lacunas que indicam a necessidade de aprofundamento e diversificação das abordagens no debate.

CONSIDERAÇÕES FINAIS

Este estudo partiu do objetivo de mapear e analisar a produção nacional, a partir de

2023, sobre racismo algorítmico no contexto das inteligências artificiais generativas (IAG), à luz das categorias de Ribeiro (2019): estruturalidade, neutralidade algorítmica e práticas antirracistas. O *corpus*, formado por uma tese, uma dissertação e três artigos, permitiu identificar o estado incipiente, porém em rápida expansão, desse campo no Brasil.

Os cinco trabalhos analisados convergem nos três pontos centrais a partir de Ribeiro (2019): a) reconhecem o racismo algorítmico como expressão do racismo estrutural, evidenciando a continuidade das desigualdades históricas no ambiente digital; b) refutam a neutralidade dos algoritmos, mostrando que o preconceito surge das escolhas humanas inscritas nos dados e nos códigos; e c) propõem frentes de ação antirracista que incluem educação crítica, regulamentação democrática, transparência e auditoria, diversidade nas equipes de desenvolvimento e bases de dados afrocentradas. Esses achados reforçam a urgência de políticas públicas alinhadas ao debate legislativo em curso sobre IA no país (PL 2338/2023).

Este ensaio oferece uma síntese criteriosa da literatura recente, preenchendo uma lacuna sobre IAG e racismo no Brasil e fornecendo um arcabouço analítico replicável para futuras pesquisas. Do ponto de vista social, sustenta a formulação de currículos antirracistas e orienta a atuação de educadores, desenvolvedores e legisladores na mitigação de discriminações algorítmicas.

A revisão concentrou-se em publicações nacionais disponíveis até junho de 2025; a amostra reduzida (cinco estudos) e o curto intervalo temporal refletem tanto a novidade do tema quanto a escassez de pesquisas empíricas com dados primários. Além disso, não foram explorados trabalhos que abordam outras dimensões críticas, como branquitude ou produção intelectual negra, mencionadas na fundamentação teórica, mas ausentes nos estudos analisados. Também não foram analisados artigos que conceituam o racismo algorítmico por meio de outros termos.

Recomendações para pesquisas futuras incluem ampliar o escopo empírico, especialmente por meio de investigações de campo em escolas de ensino médio, a fim de compreender as percepções dos estudantes sobre IAG e racismo. Ampliar as estratégias de busca, incorporando descritores como “viés”, “preconceito algorítmico” e “algoritmos enviesados” para captar estudos que empregam terminologias distintas. Explorar recortes interseccionais (gênero, classe, território) para ampliar a compreensão dos efeitos da IAG.

Ao demonstrar que o racismo algorítmico se perpetua na IAG e que enfrentá-lo exige ações complexas e em várias frentes, esta revisão reafirma que a tecnologia não é neutra e, portanto, deve ser regulada, auditada e educada em chave antirracista. Espera-se que as possibilidades aqui sistematizadas sirvam de fundamento para investigações futuras e para a construção de ecossistemas sociotécnicos mais justos e inclusivos.

REFERÊNCIAS

ARAÚJO, Júlio. Racismo algorítmico e microagressões nas redes sociais. **Domínios de Lingu@gem**, Uberlândia, v. 18, p. e1849, 2024. Disponível em: <https://seer.ufu.br/index.php/dominiosdelinguagem/article/view/74207>. Acesso em: 13 jul. 2025.

ARAÚJO, Júlio; ARAÚJO, Júlio de. Racismo algorítmico e inteligência artificial: uma análise crítica multimodal. **Revista Linguagem em Foco**, Fortaleza, v. 16, n. 2, p. 89-109, 2024. Disponível em: <https://revistas.uece.br/index.php/linguagememfoco/article/view/13108>. Acesso em: 13 jul. 2025.

ALMEIDA, Silvio. **Racismo estrutural**. São Paulo: Pólen, 2019.

ALVES, Beatriz. Datafolha: 59% dos brasileiros acreditam que a maioria da população é racista. **CNN Brasil**, São Paulo, 20 nov. 2024. Disponível em: <https://www.cnnbrasil.com.br/politica/datafolha-59-dos-brasileiros-acreditam-que-a-maioria-da-populacao-e-racista/>. Acesso em: 11 jul. 2025.

BARDIN, Laurence. **Análise de conteúdo**. São Paulo: Edições 70, 2016.

BONIFÁCIO, Carmen; SCHAPACHNIK, Fernando; PORTO, Fabio. **Generative AI and the transformation of Work in Latin America – Brazil**. arXiv preprint, 2025. Disponível em: <https://arxiv.org/abs/2505.13490>. Acesso em: 17 jul. 2025.

BUOLAMWINI, Joy; GEBRU, Timnit. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. **Proceedings of Machine Learning Research**, v. 81, 2018, p. 1-15. Disponível em: https://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline&ref=akusion-ci-shi-dai-bizinesumedeia. Acesso em: 11 jul. 2025.

DAMASCENO, Adriana; PERUZZO, Cicilia Krohling. O tecnicismo na comunicação contemporânea: controvérsias sobre a relação entre redatores-robôs e racismo. **Tríade: Comunicação, Cultura e Mídia**, Sorocaba, SP, v. 11, n. 24, p. e023014, 2023. DOI: 10.22484/2318-5694.2023v11id5177. Disponível em: <https://periodicos.uniso.br/triade/article/view/5177>. Acesso em: 13 jul. 2025.

FRANCO, Maria Laura P.B. **Análise de Conteúdo**. 2. ed. Brasília: Líber Livro Editora, 2005.

MEIRELLES, Fernando S. **Pesquisa do Uso da TI nas Empresas** – 35. ed. São Paulo: FGV EAESP, 2024. Disponível em: https://eaesp.fgv.br/sites/default/files/legacy/u68/pesti-fgvcia-2024_0.pdf. Acesso em: 17 jul. 2025.

MOROSINI, Marília Costa; FERNANDES, Cleoni Maria Barboza. Estado do Conhecimento: conceitos, finalidades e interlocuções. **Educação Por Escrito**, Porto Alegre, v. 5, n. 2, p. 154–164, 2014. Disponível em: <https://revistaseletronicas.pucrs.br/poescrito/article/view/18875>. Acesso em: 10 dez. 2024.

MOROSINI, Marília Costa; KOHLS-SANTOS, Priscila; BITTENCOURT, Zoraia Aguiar. **Estado do Conhecimento: teoria e prática**. 1. ed. Curitiba: CRV, 2021.

NEPOMUCENO, Thiago de Godoy. **Reflexões sobre a colonialidade digital**: explorando a dialética entre o ensino da arte e a análise crítica sociológica dos estudantes enquanto sujeitos inseridos na cibercultura. Dissertação (Mestrado). Universidade Federal de Pelotas, Programa de Pós-Graduação em Artes, Centro de Artes. Pelotas, 2023. Disponível em: https://sucupiralegado.capes.gov.br/sucupira/public/consultas/coleta/trabalhoConclusao/viewTrabalhoConclusao.jsf?popup=true&id_trabalho=14954039. Acesso em: 14 jul. 2025.

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

O'NEIL, Cathy. **Algoritmos de destruição em massa**: como o big data aumenta a desigualdade e ameaça a democracia. Santo André: Rua do Sabão, 2020.

RIBEIRO, Djamila. **Pequeno manual antirracista**. São Paulo: Companhia das Letras, 2019.

SAMARÃO, Lilianny Alves. **IA e algoritmos enviesados**: as dimensões do viver na sociedade 5.0. Tese (Doutorado). Pontifícia Universidade Católica de São Paulo, Programa de Estudos Pós-Graduados em Comunicação e Semiótica. São Paulo: [s.n.], 2023. 108 p. Disponível em: <https://ariel.pucsp.br/handle/handle/40758>. Acesso em: 13 jul. 2025.

SILVA, Tarcízio. **Racismo algorítmico**: inteligência artificial e discriminação nas redes digitais. São Paulo: Edições Sesc São Paulo, 2022. Disponível em: <https://assets.pubpub.org/7bmo6qr3/61661883837114.pdf>. Acesso em: 11 jul. 2025.

SUE, Derald Wing. **Microaggressions in Everyday Life**: Race, Gender, and Sexual Orientation, Hoboken, NJ: John Wiley & Sons, 2010.

TYNES, Brendesha M. et al. From Racial Microaggressions to Hate Crimes: A Model of Online Racism Based on the Lived Experiences of Adolescents of Color. In: TORINO, Gina C. et al. (org.). **Microaggression Theory**: Influence and Implications, Hoboken, NJ: John Wiley & Sons, 2019.

VIANA, Pablo Moreno F.; BELMIRO, Dalila Maria M. Midiatização do racismo brasileiro: Todecachinho, consumo, cidadania no caso da Youtuber Ana Clara Barbosa. **Revista**

Ícone, Caruaru v. 17, n. 2, p. 189–201. Disponível em:
<https://periodicos.ufpe.br/revistas/icone/article/view/240854>. Acesso em: 23 mar. 2026.

NOTA SOBRE A AUTORIA

Autor 1 foi responsável pela realização da pesquisa, levantamento bibliográfico, análise dos dados e redação do manuscrito.

Autor 2 contribuiu para o aprofundamento teórico e a organização do estudo.

Autor 3 atuou na orientação e no direcionamento do artigo.

REVISÃO DO ARTIGO

Ana Valéria Razera. Graduação em Letras – Português e Inglês e respectivas literaturas.

NOTA SOBRE USO DE INTELIGÊNCIA ARTIFICIAL

Durante o processo de produção deste artigo, especificamente nas etapas de estruturação e de revisão textual, foram utilizadas as IAGs ChatGPT (modelo o3) e Grammarly como ferramentas de apoio ao aprimoramento da redação e da fluidez do trabalho. Após a utilização das IAGs, os autores revisaram e aprimoraram o texto com base nas sugestões geradas, assumindo inteira responsabilidade pelo conteúdo científico e intelectual aqui apresentado.

Recebido em: 07/04/2026

Parecer em: 04/06/2026

Aprovado em: 03/07/2026