



# A NEGAÇÃO DO RACISMO E A DISCRIMINAÇÃO RACIAL NAS PLATAFORMAS DIGITAIS: DISCURSOS DE ÓDIO E ALGORITMOS

Mariana Santana Coutinho<sup>1</sup>

Editor Responsável: Silvia Lemos

Revisão textual: Willian Albino

## RESUMO

O artigo analisa como o discurso da neutralidade racial online contribui para a invisibilização e a reprodução do racismo nas plataformas digitais brasileiras. Parte-se da hipótese de que a crença na internet como um espaço racialmente neutro dificulta o reconhecimento das múltiplas formas de discriminação que se manifestam no ambiente digital, tanto por meio de discursos de ódio explícitos e velados quanto pela atuação dos sistemas algorítmicos das plataformas. A pesquisa desenvolve-se mediante revisão bibliográfica e documental, articulando estudos sobre racismo estrutural, discurso de ódio, racismo algorítmico e responsabilidade dos provedores de aplicação de internet, além da análise de dados empíricos produzidos por instituições de pesquisa e organizações dedicadas ao monitoramento da violência digital. Conclui-se que a internet reproduz e potencializa desigualdades raciais já existentes na sociedade, evidenciando que a

---

<sup>1</sup> Pós-Graduada em Direito Digital pela Universidade do Estado do Rio de Janeiro (UERJ), em parceria com o Instituto de Tecnologia e Sociedade do Rio (ITS RIO). Pesquisadora do Lawgorithm, no núcleo "IA e Raça". Membro da Comissão Especial de Inteligência Artificial da OBA/BA. Graduada em Direito pela Universidade Federal da Bahia (UFBA). É advogada, com atuação nas áreas de Direito e Tecnologia. ORCID: <https://orcid.org/0009-0007-2543-7662>. E-mail: [marianasantanacoutinho@gmail.com](mailto:marianasantanacoutinho@gmail.com)

lógica de funcionamento das plataformas e de seus algoritmos desempenha papel relevante na circulação e amplificação de conteúdos discriminatórios. Defende-se, por fim, que o enfrentamento do racismo online exige o abandono do mito da neutralidade racial e o reconhecimento da responsabilidade das plataformas na estruturação e difusão dessas dinâmicas discriminatórias.

**Palavras-chave:** Redes sociais. Discurso de ódio. Plataformas digitais. Racismo estrutural.

## ONLINE RACIAL NEUTRALITY AS A MYTH: HATE SPEECH AND THE ALGORITHMIC COMPLICITY OF DIGITAL PLATFORMS

### ABSTRACT

This article examines how the discourse of online racial neutrality contributes to the invisibility and reproduction of racism on Brazilian digital platforms. It is based on the premise that the belief in the internet as a racially neutral space hinders the recognition of the multiple forms of discrimination that emerge in the digital environment, both through explicit and subtle hate speech and through the operation of platform algorithms. The study is grounded in bibliographic and documentary research, bringing together scholarship on structural racism, hate speech, algorithmic racism, and the liability of internet application providers, as well as empirical data produced by research institutions and organizations dedicated to monitoring digital violence. The findings demonstrate that the internet reproduces and intensifies racial inequalities already present in society, revealing that the operational logic of digital platforms and their algorithms plays a significant role in the circulation and amplification of discriminatory content. Finally, the article argues that addressing online racism requires abandoning the myth of racial neutrality and recognizing the responsibility of digital platforms for structuring and disseminating these discriminatory dynamics.

**Keywords:** Social media. Hate speech. Digital platforms. Structural racism.

## **INTRODUÇÃO**

A internet é uma das maiores aliadas do racismo. Desde sua popularização como rede pública de comunicação, a internet consolidou-se como um verdadeiro habitat para práticas discriminatórias, facilitando o recrutamento de novos adeptos e amplificando a circulação de mensagens violentas contra pessoas racializadas.

Ainda assim, persiste no imaginário coletivo — e na forma como as plataformas digitais são projetadas e operadas — a crença de que a internet seria um espaço neutro, indiferente à raça. Esse imaginário, fundado na falsa ideia de uma suposta neutralidade racial nos discursos dos usuários das redes sociais e em seus algoritmos, não apenas invisibiliza o racismo estrutural que se manifesta online, como também dificulta a formulação de respostas políticas e jurídicas eficazes para combatê-lo.

O presente artigo parte do discurso histórico de negação do racismo, com o objetivo de analisar como essa narrativa influencia a compreensão e o enfrentamento das discriminações raciais no âmbito das plataformas de provedores de aplicação de internet. Sustenta-se que o imaginário da neutralidade racial no ambiente digital não apenas obscurece as formas contemporâneas de manifestação do racismo, como também dificulta a construção de respostas jurídicas adequadas para seu enfrentamento. Para tanto, examina-se criticamente o discurso da neutralidade racial online, evidenciando como essa crença sustenta a reprodução do racismo nas redes sociais brasileiras, seja por meio de manifestações explícitas e veladas, seja por meio da atuação algorítmica das plataformas digitais. De forma específica, busca-se investigar: (i) a origem e os efeitos do imaginário da internet como um espaço racialmente neutro; (ii) as formas de manifestação do discurso de ódio racista nas redes sociais; e (iii) de que modo a estrutura e a lógica algorítmica dessas plataformas contribuem para a reprodução do racismo.

## **1. O MITO DA NEUTRALIDADE RACIAL ONLINE**

A popularização da internet, a partir do final do século XX, inaugurou profundas transformações nas formas de comunicação e interação social. À época, acreditava-se que, impulsionada pela globalização econômica e, sobretudo, pelo avanço das tecnologias da informação, a humanidade experimentaria um novo estágio civilizatório, no qual, embora

ainda marcada por desigualdades e conflitos, formaria uma única comunidade mundial (Lévy, 1999).

Nessa nova era, marcadores sociais historicamente mobilizados para a marginalização de determinados grupos, como a raça, seriam neutralizados pelo simples ato de ingressar no ambiente virtual (Kettrey; Laster, 2014). Assim, seria possível a existência de espaços, como a internet, nos quais a raça não exerceria qualquer influência, partindo-se da premissa de que os indivíduos seriam capazes de ignorar espontaneamente os marcadores raciais e, assim, evitar que estes influenciassem suas falas, pensamentos e comportamentos.

Essa visão adotada na época não se limita a um simples otimismo ou à projeção de um ideal utópico, mas reflete a adesão a uma ideologia racialmente neutra — o chamado *colour-blind racism* (em tradução livre, “racismo daltônico”). Predominante na sociedade norte-americana no período posterior ao movimento em prol dos direitos civis, essa ideologia defende que adotar uma postura racialmente neutra, que ignora a raça, preveniria a ocorrência de preconceitos: se as pessoas ou instituições sequer percebem a raça, então não podem agir de maneira racialmente tendenciosa (Apfelbaum *et al.*, 2012).

Todavia, conforme alerta Moreira (2020), essa concepção parte de uma compreensão individualista do racismo, que o restringe a atitudes intencionais, isoladas e moralmente condenáveis de sujeitos específicos. Trata-se de uma leitura que despolitiza o racismo, interpretando-o apenas como um comportamento desviante, de caráter ético, psicológico ou moral, ou como uma manifestação de preconceito pessoal ou ignorância atribuída a indivíduos específicos (Almeida, 2019).

Desconsidera-se, assim, que o racismo também possui fundamentos históricos e estruturais, e se manifesta como um elemento estruturante da organização econômica, política e cultural da sociedade, influenciando as relações sociais e reproduzindo-se no cotidiano, independente da intenção dos indivíduos ou da ação consciente das instituições (Almeida, 2019). Essa forma estrutural de racismo acarreta desvantagens que mantêm a população negra em situação de desigualdade e a sujeitam a práticas discriminatórias que nem sempre serão explícitas ou manifestadas diretamente por um indivíduo.

Consequentemente, enquanto se mantém a ideia de que adotar uma postura racialmente neutra impede a ocorrência de racismo online, confere-se aos indivíduos — especialmente aos brancos — a prerrogativa de “não ver cor”, como se o simples ato de ignorar marcadores raciais fosse suficiente para eliminar desigualdades historicamente constituídas. Enquanto isso, pessoas brancas seguem usufruindo de privilégios raciais, reproduzindo condutas racistas e, muitas vezes, negando tanto a existência do racismo quanto a sua própria identificação como racistas (Moreira, 2020).

No contexto brasileiro, o discurso da neutralidade racial também se manifesta de forma contundente, ainda que enraizado em uma história particular. Embora o Brasil não tenha instituído um regime nacional de segregação racial nos moldes do modelo norte-americano, sua trajetória jurídica foi marcada pela existência de normas e práticas estatais de caráter discriminatório, que restringiram direitos da população negra e contribuíram para a manutenção das hierarquias raciais. Todavia, os mais de três séculos de escravização de pessoas negras e indígenas, seguidos por uma abolição tardia, movida por interesses econômicos e desprovida de qualquer medida reparatória, somados à vigência de normas jurídicas que, embora aparentemente neutras, atuaram no sentido de marginalizar essas populações e promover uma política sistemática de branqueamento (Silva, F., 2022), consolidaram uma ordem social profundamente marcada pela exclusão, desigualdade e perpetuação do racismo.

Além disso, difundiu-se o discurso de que o país seria uma “democracia racial”, popularizado a partir da primeira metade do século XX com a obra *“Casa Grande & Senzala”*, de Gilberto Freyre. Essa concepção exalta a miscigenação como traço fundante da identidade nacional, sustentando que ela viabilizaria a convivência harmônica entre diferentes grupos e, consequentemente, neutralizaria as tensões étnico-raciais herdadas da escravidão.

A persistência do mito da democracia racial no Brasil não contribuiu para a erradicação do racismo; ao contrário, forneceu sustentação ao discurso de sua negação. Nesse sentido, explica Adilson José Moreira (2020):

A negação de que o racismo seja tema de relevância social no mundo atual é uma das estratégias utilizadas por grupos sociais comprometidos com a preservação das disparidades sociais entre grupos raciais. Esse discurso

assume a mesma forma em diferentes sociedades. Primeiro, ele reduz o racismo a um problema de caráter individual e que pode ser resolvido com a responsabilização específica da pessoa responsável por um ato discriminatório. Segundo, por ter uma natureza individual e comportamental, o racismo não pode ser visto como um fator responsável por disparidades entre pessoas negras e brancas.

Compreender o discurso de negação do racismo e da alegada neutralidade racial é fundamental para reconhecer que a internet jamais constituiu um ambiente racialmente neutro ou imune à reprodução de práticas discriminatórias. Conforme demonstram Kettrey e Laster (2014), com base em pesquisas publicadas entre o final da década de 1990 e o início dos anos 2000, os mundos online e offline estão entrelaçados, de modo que aqueles que acessam o ambiente virtual levam consigo os valores adquiridos por meio de sua socialização no mundo offline.

Assim, em sociedades historicamente marcadas por discriminação racial, o ambiente virtual tende a reproduzir e, em certos casos, a amplificar os preconceitos, mantendo o *status quo* em que os usuários brancos usufruem de maior acesso, legitimidade e poder simbólico em comparação aos usuários racializados. A suposta neutralidade não impede a reprodução das hierarquias raciais, mas frequentemente as encobre.

Diante disso, torna-se imprescindível observar como, apesar do discurso de negação, o racismo opera no Brasil, no âmbito das plataformas de provedores de aplicação de internet, seja por meio das manifestações de discurso de ódio racistas implícitas e explícitas, seja por meio dos próprios sistemas algorítmicos que organizam a experiência digital.

## **2. O DISCURSO DE ÓDIO RACISTA NAS REDES SOCIAIS BRASILEIRAS**

Longe de representar espaços neutros ou meramente comunicacionais, a internet consolidou-se como um ambiente propício à reprodução de práticas discriminatórias. Os provedores de aplicação de internet, aqui entendidos como as pessoas físicas ou jurídicas que oferecem funcionalidades acessadas por meio de terminais conectados à rede, tais como Facebook, Instagram, X e demais redes sociais (Queiroz, 2018), não atuam como meros canais passivos de comunicação, mas viabilizam, por meio de suas arquiteturas

técnicas e lógicas de funcionamento, um ambiente propício à proliferação de discursos discriminatórios.

Essa dinâmica é amparada, sobretudo, pela aparente impunidade proporcionada pelo anonimato e por uma noção distorcida de liberdade de expressão, frequentemente utilizada como escudo para práticas que violam direitos fundamentais e reforçam desigualdades históricas. Assim, indivíduos racistas propagam discursos de ódio, definidos pela Organização das Nações Unidas (ONU, 2019) como qualquer tipo de comunicação oral, escrita ou comportamental que ataque ou utilize termos pejorativos ou discriminatórios direcionados a uma pessoa ou grupo com base em elementos como a etnia e a raça, dentre outros.

Nas plataformas mantidas por esses provedores de aplicação, esses discursos ganham velocidade e alcance inéditos, moldando novas dinâmicas de intolerância. Nesse cenário de proliferação do discurso de ódio racista online, emerge entre pesquisadores da área o *cyber-racism* (em tradução livre, “racismo cibernético”), termo cunhado pelo antropólogo Les Back (Daniels, 2009) para designar uma forma contemporânea de manifestação do racismo presente em movimentos supremacistas brancos da Europa e da América do Norte. Para Back, o racismo cibernético é caracterizado por aspectos como ideias de singularidade racial, nacionalismo e destino comum, bem como por concepções de supremacia racial, superioridade e separação, estruturadas a partir das especificidades do meio digital.

A incidência dessa forma de racismo é expressiva e preocupante. Segundo dados da SaferNet<sup>2</sup>, organização brasileira dedicada ao enfrentamento de crimes e violações de direitos no meio digital, entre os anos de 2006 e 2025 foram registradas 614.383 denúncias anônimas de conteúdos racistas em plataformas de provedores, envolvendo 116.165 páginas (URLs) distintas, das quais 73.719 foram removidas. No mês de julho de 2025, o Brasil ocupava a quinta posição no ranking mundial de países com maior número de denúncias relacionadas a esse tipo de crime.

No contexto dos provedores de aplicação, popularmente conhecidas como redes sociais, que favorecem a conexão, a expressão e a interação entre indivíduos para além de

---

<sup>2</sup> Dados coletados nos indicadores do Safernet em 13 de julho de 2025.

fronteiras geográficas, as manifestações de racismo cibernético têm se intensificado, tanto em volume quanto em diversidade discursiva. Entre os anos de 2010 e 2015, foram registradas 32.866 denúncias de racismo na plataforma Facebook, com crescimento exponencial, passando de apenas 50 casos em 2010 para 11.624 em 2014<sup>3</sup>.

O gráfico a seguir sintetiza os principais dados obtidos pela SaferNet no que se refere à incidência de denúncias de conteúdos racistas nas redes sociais. Verifica-se uma redução progressiva das ocorrências relacionadas ao Orkut, até a cessação completa de registros em 2014, ano em que o Orkut foi desativado. Em sentido oposto, observa-se o crescimento do Facebook entre 2010 e 2014, à medida que a rede social se popularizava e ampliava sua base de usuários.

A partir de 2015, entretanto, o número de denúncias registradas no Facebook passa a diminuir. Essa redução coincide com o surgimento e a crescente popularização de novas plataformas, como Instagram e TikTok, que também passaram a concentrar denúncias de conteúdos racistas. O comportamento dos dados sugere que a diminuição das denúncias em uma plataforma específica não representa, necessariamente, uma redução da incidência do racismo online, mas acompanha a migração dos usuários para novos ambientes digitais.

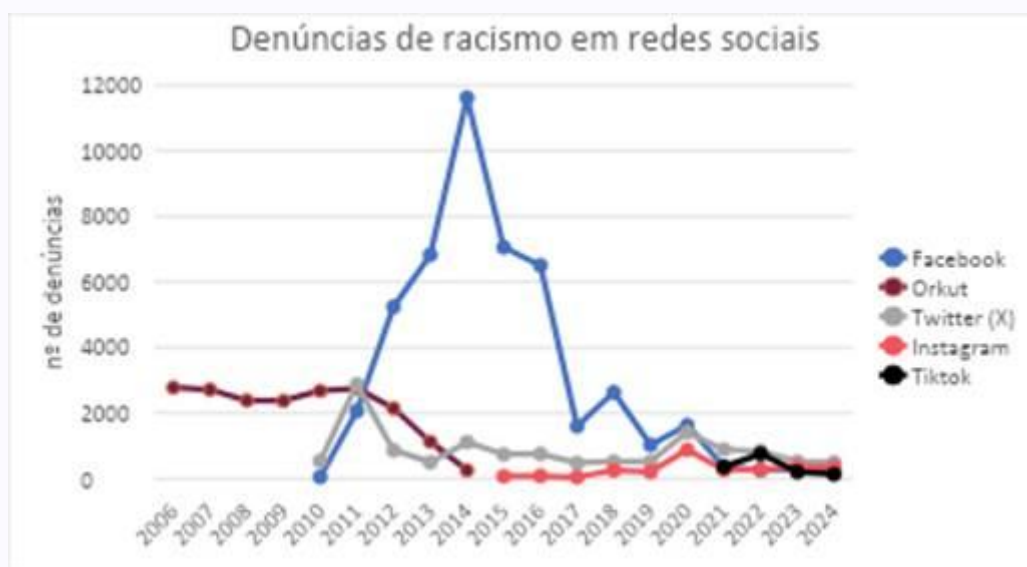


Gráfico criado com base em dados da Safernet. Elaboração própria.

<sup>3</sup> Dados coletados nos indicadores do Safernet em 13 de julho de 2025.

O crescimento no número de denúncias entre 2010 e 2014 reflete não apenas uma possível intensificação das práticas racistas nas redes sociais, mas também um avanço na conscientização social sobre o tema, levando mais usuários a formalizar denúncias. Todavia, a diminuição do número de denúncias a partir de 2014 não significa, necessariamente, uma redução da incidência do racismo nessas plataformas, cujas manifestações continuaram ocorrendo em larga escala, muitas vezes de forma velada, naturalizada ou não denunciada.

No Brasil, um levantamento realizado pela agência Nova/SB, entre abril e junho de 2016, analisou publicações em redes sociais como Facebook, Twitter (atual “X”) e Instagram, utilizando o software de monitoramento Torabit. Foram identificadas ao todo 32.376 menções ao tema do racismo, das quais 97,6% apresentavam conotação negativa, expressando intolerância, preconceito ou discriminação, enquanto apenas 2,1% foram classificadas como positivas, por manifestarem posicionamento contrário ao preconceito racial. As menções neutras, que corresponderam a 0,3% do total, referiam-se a postagens que, segundo a metodologia da pesquisa, “[...] não apoiam nem se opõem a algum tipo de discriminação, geralmente citando ou comentando casos de forma imparcial” (Nova/SB, 2016).

Quanto às formas de manifestação, nota-se que o racismo tem se apresentado tanto de forma implícita, por meio de microagressões e estereótipos racializados, que podem ser intencionais ou não, quanto de forma explícita, por meio de discursos de ódio e da injúria racial.

Para o psicólogo e pesquisador estadunidense Derald Sue (2010), as microagressões são “ofensas, desprezos ou insultos cotidianos, verbais, não verbais ou ambientais, intencionais ou não, que transmitem mensagens hostis, depreciativas ou negativas a pessoas-alvo unicamente com base em seu pertencimento a um grupo marginalizado”. O termo “micro”, conforme alerta Tarcísio Silva (2020), não se refere à gravidade da violência ou ao impacto que ela causa, mas à maneira como essa agressão se manifesta, de forma sutil, cotidiana e muitas vezes disfarçada, como uma piada, um comentário “despretensioso” ou um gesto aparentemente inofensivo

Analisando a obra de Sue, Fernanda Rodrigues Silva (2022) explica as formas como essas microagressões se dividem:

Assim, segundo Sue, elas estariam divididas em três categorias: a) microataques: definidos como crenças ou comportamentos preconceituosos de raça, conscientes, deliberados e sutis ou explícitos, realizados por meio de pistas ambientais, verbalizações ou comportamentos; b) microinsultos: mensagens interpessoais ou ambientais que reproduzem estereótipos e grosseria, além de deturpar a identidade ou herança racial de uma pessoa; e c) microinvalidações: mensagens que excluem, negam ou anulam as experiências pessoais, como sentimentos e pensamentos, de pessoas de cor.

Aplicando o conceito de microataques, conforme definido por Sue, ao ambiente das redes sociais, identifica-se um exemplo concreto quando, diante da divulgação de notícias sobre a implementação de políticas afirmativas, indivíduos associam diretamente a presença negra no ensino superior à criminalidade, à desordem e à degradação moral, compondo um discurso racista que opera por meio da verbalização depreciativa, da construção de estereótipos e da criação de um ambiente simbolicamente hostil.

Como exemplo, observa-se uma postagem realizada pela Agência Brasil, em 28 de junho de 2025, sobre os 20 anos das cotas raciais e seu impacto na transformação do perfil das universidades brasileiras, evidenciando mudanças significativas nesses espaços. Na postagem, ainda disponível durante a elaboração deste artigo, foram encontrados os seguintes comentários<sup>4</sup>:

Os comentários mencionados foram analisados em 17 de julho de 2025. Usuário #1: “É por isso que estamos em 118º lugar, à frente somente da Etiópia? Universidade se tornando antro de drogas? Acorda Brasil”  
Usuário #2: “Mudaram o perfil! Tá tudo incluído! Droga, promiscuidade, vandalismo! São os profissionais que irão atender a população no futuro! Tá ON! KKKKK”  
Usuário #3: “Transformou sim, hoje às universidades Federais estão um caos de militância e depravação. Daqui uns dias vai virar só cota, nenhum estudante de verdade vai poder participar do vestibular.”  
Usuário #4: “Federais virou boca de fumo kkkkkk”  
Usuário #5: “Meu filho não vai por os pés neste antro de maconheiro.”

Para fins de contextualização, a postagem original da Agência Brasil apresentou análises de pesquisadores acerca da democratização do ensino superior por meio das

---

<sup>4</sup> Os comentários mencionados foram analisados em 19 de julho de 2025.

políticas de cotas raciais, sem qualquer referência a drogas ou a conteúdos depreciativos. As manifestações ofensivas e estereotipadas surgiram exclusivamente nos comentários dos usuários, desvinculadas do conteúdo da publicação.

Ainda com base no conceito de Sue, entende-se como exemplo de microinsulto os comentários que minimizam ou ridicularizam práticas culturais tradicionais, como manifestações religiosas afro-brasileiras, pois desconsideram a importância e o valor da herança cultural, reforçando preconceitos e excluindo essas identidades do reconhecimento social legítimo.

Por fim, exemplos de microinvalidações podem ser encontrados com frequência em respostas a discursos antirracistas, muitas vezes travestidas de apelos à neutralidade ou à “união das raças”. Um exemplo emblemático foi observado nas respostas à deputada Talíria Petrone, após sua publicação na rede social X (à época, Twitter), em comemoração à decisão do Superior Tribunal de Justiça que reconheceu a inexistência jurídica do chamado racismo reverso. Nos comentários<sup>5</sup>, observam-se enunciados como

Usuário #6: “deputada Talíria, você tem que pregar a união das raças, não ficar com este vitimismo colocando o negro contra brancos”

Usuário #7: “Se for ver assim os negros de hoje também não foram escravizados. E porque querem direitos e benefícios em cima dos que foram?”

O discurso de ódio racista também se manifesta de forma explícita por meio de ofensas abertas, insultos ou declarações que expõem diretamente o preconceito racial. Nesses casos, a conduta pode configurar o crime de injúria racial, nos termos do art. 2º-A da Lei nº 7.716/1989, com a redação conferida pela Lei nº 14.532/2023, quando é dirigida a um indivíduo com o intuito de ofender sua dignidade ou decoro em razão de raça, cor, etnia ou procedência nacional. Nos casos em que a ofensa se dirige a um grupo ou coletividade, com o objetivo de restringir direitos ou negar o reconhecimento de sua dignidade, caracteriza-se o crime de racismo, previsto na Lei nº 7.716/1989, cuja natureza é imprescritível e inafiançável.

Em mapeamento realizado em 2023 pela Faculdade Baiana de Direito em parceria com a empresa Jusbrasil, foram analisados discursos racistas presentes em ações judiciais

---

<sup>5</sup> Os comentários mencionados foram analisados em 19 de julho de 2025.

de natureza criminal e cível vinculados aos crimes de racismo e injúria racial. Constatou-se que as ofensas mais recorrentes foram a nomenclatura pejorativa (48,60%) e a animalização (39,25%). Já as formas menos frequentes envolveram a imputação de delinquência e a irreligiosidade, ambas com 2,8% dos casos (Nicory; Guanabara; Assumpção, 2023).

A pesquisa também evidenciou a intersecção entre racismo e misoginia. Os dados apontam que 58,54% das vítimas eram mulheres e 18,29% homens, sendo que 23,17% não tiveram o gênero identificado, em razão do caráter coletivo de algumas ofensas. Em relação aos ofensores, 55,56% eram homens. Observou-se, ainda, que as agressões direcionadas a mulheres negras concentraram-se em ataques à sexualidade, higiene e aparência física, enquanto os homens negros foram frequentemente alvos de discursos voltados à inferiorização social e intelectual.

Os dados apresentados corroboram a análise proposta por Luiz Trindade (2020), segundo a qual o racismo, articulado ao machismo, opera para restringir os espaços sociais legitimamente ocupáveis por mulheres negras. As mensagens racistas analisadas, publicadas em resposta a imagens de uma mulher negra em viagem internacional com seu parceiro branco, escancaram um incômodo social com a ascensão e visibilidade dessa mulher em contextos historicamente racializados como “brancos” (Trindade, 2020).

Nesse sentido, o corpo da mulher negra torna-se alvo de múltiplas formas de violência simbólica: não apenas é desumanizado e animalizado, mas também hipersexualizado ou tratado como objeto de escárnio, revelando uma tentativa de reafirmar hierarquias raciais e de gênero. Trindade (2020) exemplifica essa dinâmica com postagens ofensivas extraídas de uma rede social, como: “Ela até que é bonitinha. Pena que os pretos não estão mais à venda (Post #2)”; e “Aguarde até que a neve derreta e então use a água para lavar as louças do hotel (Post #3)”.

Diante do cenário alarmante apresentado, torna-se evidente o profundo impacto que os discursos de ódio nas redes sociais exercem sobre pessoas negras, na medida em que essas pessoas são atingidas cotidianamente por comentários que desumanizam, estereotipam e deslegitimam sua existência, reforçando estruturas históricas de exclusão e inferiorização.

Contudo, esse impacto não se restringe às ações diretas dos usuários. As redes sociais oferecem uma cumplicidade algorítmica, presente nas lógicas de funcionamento das plataformas digitais, que potencializa a circulação desses conteúdos ao favorecer sua disseminação e ampliar o alcance de discursos que geram maior engajamento, mesmo quando carregados de preconceito. Essa dinâmica evidencia que o racismo nas redes não é apenas um reflexo das atitudes individuais, mas também um fenômeno estrutural, sustentado por sistemas automatizados que, sob a aparência de neutralidade, reproduzem e intensificam desigualdades raciais, tema que será aprofundado no tópico a seguir.

### **3. O RACISMO E A CUMPLICIDADE ALGORÍTMICA DAS PLATAFORMAS**

Quando se trata de discurso de ódio racial nas redes sociais, é comum que a responsabilidade pelas manifestações discriminatórias seja atribuída exclusivamente aos usuários, pessoas físicas que, teoricamente, atuam de forma autônoma em relação às plataformas que utilizam, reforçando a ideia de que o racismo constitui um fenômeno meramente individual. Todavia, ainda que, à primeira vista, os provedores de aplicação de internet não sejam autores dessas publicações, desempenham papel ativo na estruturação, circulação e disseminação desses conteúdos por meio das funcionalidades e da arquitetura de suas plataformas.

De início, destaca-se que a própria estrutura dessas plataformas sustenta e amplifica práticas de discurso de ódio. Funcionalidades como o anonimato, a criação de perfis falsos e a utilização de *trolls* ou *bots* contribuem para a proliferação de conteúdos racistas e desinformativos, dificultando a responsabilização direta dos autores das postagens (Rodrigues, 2023). Some-se a isso uma lógica econômica que se beneficia da circulação massiva de conteúdos, mesmo os discriminatórios, em nome do engajamento que gera a monetização: conforme apontam Santos *et al.* (2023), ainda que não sejam autoras da publicação, as plataformas “se apropriam delas para criar um ecossistema narrativo potencialmente comercializável de onde extraem vultosos lucros”.

Para além dessas funcionalidades, as plataformas exercem controle ativo sobre o que é visibilizado ou silenciado por meio de algoritmos de recomendação. Conforme aponta Tarcísio Silva (2022), esses sistemas são programados para filtrar o excesso de

informações, identificar padrões de interesse dos usuários com base em suas interações e entregar conteúdos considerados relevantes buscando maximizar métricas de engajamento, como o tempo de permanência na plataforma e os cliques em anúncios.

Nesse sentido, as “*big techs*” que utilizam esses algoritmos, apesar de alegarem neutralidade na priorização de conteúdos, não explicitam as lógicas e critérios dos algoritmos de hierarquização, visibilização e invisibilização de mensagens, criando uma verdadeira cortina de fumaça. Ainda assim, pesquisas demonstram que esses algoritmos não são ferramentas neutras ou objetivas, mas mecanismos capazes de reproduzir desigualdades estruturais. É nesse contexto que se insere o conceito de racismo algorítmico, entendido como uma dimensão do racismo estrutural na qual a ordenação algorítmica é moldada pela supremacia branca, produzindo efeitos discriminatórios sobre grupos racializados (Silva, T. 2022).

Nesse sentido, a matemática Cathy O’Neil, em sua obra *Weapons of Math Destruction* (2016), afirma que modelos matemáticos — como os algoritmos utilizados pelas grandes plataformas digitais — não são neutros ou objetivos, mas refletem as decisões, prioridades e vieses de seus criadores. Segundo a autora, um modelo é apenas uma representação abstrata de um processo, e os seus pontos cegos revelam os julgamentos e preconceitos embutidos durante sua construção. Quando aplicados ao contexto das redes sociais, esses modelos operam com dados incompletos, descontextualizados e muitas vezes enviesados, reproduzindo, assim, as desigualdades e exclusões que marcam a sociedade fora do mundo digital.

Um exemplo emblemático é trazido por Tarcízio Silva (2022), ao demonstrar como determinadas plataformas digitais lucram com o engajamento gerado por conteúdos violentos e extremistas. Como os algoritmos que operam com base no interesse do usuário acabam amplificando discursos de ódio. Assim, se um indivíduo consome repetidamente conteúdos extremistas, o sistema é calibrado para oferecer conteúdo semelhante.

Situação semelhante ocorre com manifestações racistas. A lógica algorítmica prioriza o tempo de permanência na plataforma e, assim, recomenda conteúdos discriminatórios àqueles que já os consomem, contribuindo para ampliar sua circulação e reforçar comportamentos discriminatórios.

Na prática, nota-se que as próprias redes já reconheceram a ocorrência de racismo em seus algoritmos. Em 2021, o TikTok foi alvo de denúncias de usuários que relataram a supressão de expressões como “*Black Lives Matter*” e “*Black Success*”, diretamente associadas à luta antirracista, ao mesmo tempo em que a plataforma permitia termos abertamente preconceituosos, como “neonazista”, “pró-branco” e “supremacia branca”. Diante da repercussão negativa, a empresa se pronunciou publicamente, pediu desculpas e reconheceu uma falha em seu algoritmo. O episódio evidenciou como sistemas automatizados podem reproduzir, e até amplificar, estruturas de exclusão racial sob a aparência de neutralidade tecnológica (Cruz, 2025).

A discriminação algorítmica revela como o racismo pode ser reproduzido por meio de tecnologias automatizadas. Trata-se de uma forma sofisticada de racismo que opera como uma microagressão, na medida em que se expressa de maneira sutil, persistente e sistemática, impactando a experiência digital de usuários racializados nas plataformas digitais.

Essas plataformas, como já discutido, jamais foram racialmente neutras, a despeito do persistente discurso de negação. Quando se trata de discriminações promovidas por seus usuários, já existem índices de sua ocorrência por conta das denúncias, especialmente as formalizadas criminalmente. No âmbito legal, a Lei nº 12.965/2014, conhecida como o Marco Civil da Internet (MCI), já aborda a responsabilidade civil dos provedores de aplicação de internet por conteúdos publicados por terceiros em suas plataformas.

O tema é, há anos, amplamente debatido nos campos doutrinário, jurisprudencial e legislativo, contando com decisão recente do Supremo Tribunal Federal (STF), no julgamento dos Recursos Extraordinários (REs) n. 1.037.396 e 1.057.258, proferidas em junho de 2025. A decisão alterou o regime de responsabilidade dos provedores de aplicação, visando à proteção suficiente a bens jurídicos constitucionais, inclusive em casos de racismo. Apesar das controvérsias que persistem sobre o tema e das lacunas remanescentes na decisão do Supremo, é inegável que, no tocante ao racismo manifestado por terceiros nas plataformas digitais, observa-se uma mobilização crescente para o enfrentamento do problema. Em contrapartida, o racismo algorítmico — apesar de já

identificado por pesquisadores em diversas áreas, como a ciência da computação e as ciências sociais — segue, em grande medida, invisibilizado.

Essa invisibilidade encontra raízes no discurso de negação da existência do racismo e no mito da neutralidade racial online. A internet e, especialmente, as redes sociais são ambientes onde o racismo é cotidiano, seja na propagação de discursos de ódio racistas explícitos ou velados, seja na cumplicidade algorítmica, o que pode potencializar a entrega desses conteúdos, influenciar padrões de comportamento e propagar ainda mais o racismo nas redes sociais como consequência. A recusa em reconhecer essa realidade contribui para a perpetuação das desigualdades, ao obscurecer as dinâmicas que permitem que o racismo se reproduza sob a aparência de normalidade técnica e imparcialidade algorítmica.

## **CONSIDERAÇÕES FINAIS**

Constata-se que, a despeito do tradicional discurso de negação da relevância da raça no ambiente digital, a internet não é um espaço *racialmente neutro*. Ao contrário, trata-se de um meio no qual questões raciais permanecem socialmente relevantes e onde o racismo se manifesta de maneira recorrente, seja por meio dos discursos de ódio racistas, explícitos ou velados, seja por meio da cumplicidade algorítmica das plataformas, ocasionada pelo racismo presente em seus algoritmos.

A crença na neutralidade racial da internet, bem como a separação fictícia entre os mundos online e offline, não apenas distorcem a realidade como também geram efeitos concretos e nocivos. Como aponta Trindade (2020), ao manter-se o mito de um ambiente virtual “*colour-blind*”, usuários sentem-se autorizados a expressar discursos racistas com maior liberdade no meio digital, sob a falsa premissa de que esses comportamentos não produzem consequências no mundo real. Tal dissociação entre o virtual e o material serve para mascarar as práticas discriminatórias e esvaziar a responsabilidade dos indivíduos e das plataformas.

Além disso, essa negação contribui para a invisibilidade do racismo algorítmico, uma forma sofisticada de discriminação que se manifesta pela reprodução e amplificação de vieses racistas nos sistemas automatizados das plataformas digitais. Esses algoritmos, ao serem programados e alimentados com dados enviesados, promovem a circulação

facilitada de conteúdos discriminatórios, reforçando desigualdades históricas e perpetuando um ciclo de exclusão e marginalização racial que se naturaliza no ambiente virtual.

Negar a existência do racismo online tem implicações graves. Significa deslegitimar as demandas por políticas antirracistas no âmbito das plataformas, dificultar o avanço de marcos regulatórios que enfrentem de forma eficaz os conteúdos discriminatórios e, sobretudo, invisibilizar os obstáculos enfrentados por pessoas racializadas no acesso à justiça. Persistir nessa negação é permitir que as redes sociais sigam sendo um espaço hostil, onde o lucro se sobrepõe ao compromisso com a equidade e os direitos fundamentais.

## REFERÊNCIAS

AGÊNCIA BRASIL. **Cotas raciais nos últimos 20 anos: as ações afirmativas mudaram o perfil da universidade brasileira.** Facebook: @agenciabrasil.ebc. Disponível em: <https://www.facebook.com/agenciabrasil.ebc/posts/cotas-raciais-nos-%C3%BAltimos-20-anos-as-a%C3%A7%C3%B5es-afirmativas-mudaram-o-perfil-da-unive/1143035524527285/>. Acesso em: 17/07/2025.

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

APFELBAUM, Evan P.; NORTON, Michael I.; SOMMERS, Samuel R. **Racial color blindness: emergence, practice, and implications.** *Current Directions in Psychological Science*, [S.l.], v. 21, n. 3, p. 205–209, 2012. Disponível em: <https://doi.org/10.1177/0963721411434980>. Acesso em: 01/07/2025.

CRUZ, Felipe Branco. **TikTok rebate acusação de que seu algoritmo seria racista e 'pró-branco'.** *Veja*, 15 jul. 2021. Disponível em: <https://veja.abril.com.br/cultura/tiktok-rebate-acusacao-de-que-seu-algoritmo-seria-racista-e-pro-branco/>. Acesso em: 22/07/2025

DANIELS, Jessie. **Cloaked websites: propaganda, cyber-racism and epistemology in the digital era.** *New Media & Society*, [S. l.], Summer 2009. *New Media & Society*. Disponível em: [https://academicworks.cuny.edu/hc\\_pubs/271/](https://academicworks.cuny.edu/hc_pubs/271/). Acesso em: 29/06/2025.

DIENSTBACH, Dalby; SANCHES, Danielle; SANTOS, Luiza Carolina dos; MATOS, Eurico; TOMAZ, Renata. **Discurso de ódio on-line Uma análise das políticas das plataformas digitais para moderação de conteúdo.** *Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação*, v. 26, jan–dez, publicação contínua, 2023, p. 1–22. doi.org/10.30962/ec.2709

DOS SANTOS, Luiza Carolina; TOMAZ, Renata.; DIENSTBACH, Dalby; MATOS, Eurico; SANCHES, Danielle. **Discurso de ódio on-line: uma análise das políticas das plataformas digitais para moderação de conteúdo. E-Compós**, [S. l.], v. 26, 2023. DOI: 10.30962/ec.2709. Disponível em: <https://www.e-compos.org.br/e-compos/article/view/2709>. Acesso em: 19/07/2025.

JAKUBOWICZ, Andrew; DUNN, Kevin; MASON, Gail; PARADIES, Yin; BLIUC, Ana-Maria; BAHFEN, Nasya; OBOLER, Andre; ATIE, Rosalie; CONNELLY, Karen. **Cyber racism and community resilience: strategies for combating online race hate**. London: Palgrave macmillian, 2017.

KETTREY, Heather Hensman; LASTER, Whitney Nicole. **Staking territory in the “World White Web”: an exploration of the roles of overt and color-blind racism in maintaining racial boundaries on a popular web site**. *Social Currents*, [S. l.], v. 1, n. 3, p. 257–274, 2014. Disponível em: <https://doi.org/10.1177/2329496514540138>. Acesso em: 21/06/2025.

LÉVY, Pierre. *Ciberespaço*. Tradução de Carlos. Irineu da Costa. São Paulo: Ed. 34, 1999.

MOREIRA, Adilson José. **Tratado de Direito Antidiscriminatório**. São Paulo: Editora Contracorrente, 2020.

NICORY, D.; GUANABARA, D.; ASSUMPÇÃO, V.s (Coord.). *Racismo e injúria racial praticados nas redes sociais: relatório do Observatório das Condenações Judiciais em 2ª Instância até o ano de 2022*. Salvador: Faculdade Baiana de Direito; Jusbrasil, 2023. Disponível em: <https://www.conjur.com.br/dl/pe/pesquisa-racismo-redes-2023.pdf>. Acesso em: 15/06/2025.

NOVA/SB. **Dossiê intolerâncias: visíveis e invisíveis no mundo digital**. São Paulo: Comunica Que Muda, 2016. Disponível em: <https://www.comunicaquemuda.com.br/dossie/intolerancia-nas-redes/>. Acesso em: 16/07/2025.

O'NEIL, Cathy. **Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy**. New York: Crown, 2016.

PETRONE, Taliria. *Racismo reverso não existe!*. X (antes Twitter), publicado em 31 jan. 2020. Disponível em: <https://x.com/taliriapetrone/status/1223264119370997760>. Acesso em: 19 jul. 2025.

PODER360. **76% veem racismo no Brasil, mas só 36% admitem preconceito contra negros. Poder360**. Realizado pelo PoderData, publicado em 26 setembro 2023. Disponível em: <https://www.poder360.com.br/poderdata/76-veem-racismo-no-brasil-mas-so-36-admitem-preconceito-contra-negros/>. Acesso em: 01 jul. 2025.

QUEIROZ, João Quinelato de. **A responsabilidade civil dos provedores de aplicação de internet por danos decorrentes de conteúdo gerado por terceiros: análise na perspectiva civil-constitucional**. 2018. Dissertação (Mestrado em Direito) – Faculdade de Direito, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2018. Disponível em: [https://www.bdt.d.uerj.br:8443/bits-tream/1/9862/1/Joao%20Quinelato%20de%20Queiroz\\_T%20tal.pdf](https://www.bdt.d.uerj.br:8443/bits-tream/1/9862/1/Joao%20Quinelato%20de%20Queiroz_T%20tal.pdf). Acesso em 22/05/2025.

RODRIGUES, Caroline Porfirio. **Identificação de discursos de ódio em Redes Sociais**. Monografia (MBA em Inteligência Artificial e Big Data) Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2023. Disponível em: <https://bdta.abcd.usp.br/directbitstream/4f3e0541-3574-46d0-bba6-4d17ebe36b16/Caroline%20Porfirio%20Rodrigues.pdf>. Acesso em 19/07/2025.

SAFERNET. **Indicadores da Central Nacional de Denúncias de Crimes Cibernéticos**. Safernet Brasil. Disponível em: <https://indicadores.safernet.org.br/index.html>. Acesso em 13/07/2025.

SILVA, Fernanda dos Santos Rodrigues. **Racismo online no Brasil: uma análise a partir da perspectiva de decisões judiciais**. Dissertação (Mestrado em Direito) – Universidade Federal de Santa Maria, Santa Maria, 2022. Disponível em: <https://repositorio.ufsm.br/handle/1/25794>. Acesso em: 07/06/2025.

SILVA, Tarcízio. **Racismo algorítmico em plataformas digitais: microagressões e discriminação em código.** In: SILVA, Tarcízio (org.). Comunidades, algoritmos e ativismos digitais: olhares afrodiaspóricos. São Paulo: LiteraRUA, 2020.

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

SUE, Derald Wing. **Microaggressions and marginality: manifestation, dynamics, and impact.** New Jersey: John Wiley & Sons, Inc., 2010.

TRINDADE, Luiz Valério P. **Mídias sociais e naturalização de discursos racistas no Brasil.** In: SILVA, Tarcízio (org.). Comunidades, algoritmos e ativismos digitais: olhares afrodiaspóricos. São Paulo: LiteraRUA, 2020.

UNITED NATIONS. **United Nations Strategy and Plan of Action on Hate Speech.** New York: United Nations, 2019. Disponível em: [https://www.un.org/en/genocideprevention/documents/advising-andmobilizing/Action\\_plan\\_on\\_hate\\_speech\\_EN.pdf](https://www.un.org/en/genocideprevention/documents/advising-andmobilizing/Action_plan_on_hate_speech_EN.pdf). Acesso em: 17/07/2025.