

Pelo enfrentamento do racismo no ambiente digital:

Contribuições do IRIS à Consulta da SECOM e MIR sobre combate ao racismo nas plataformas digitais

Autoria

Fernanda dos Santos Rodrigues Silva

Paulo Rená da Silva Santarém

Revisão

Ana Bárbara Gomes Pereira

Como referenciar em ABNT:

SILVA, Fernanda dos Santos Rodrigues; SANTARÉM, Paulo Rená da Silva. **Pelo enfrentamento do racismo no ambiente digital:** Contribuições do IRIS à Consulta da SECOM e MIR sobre combate ao racismo nas plataformas digitais. Belo Horizonte: Instituto de Referência em Internet e Sociedade, 2026. Disponível em: xxxxxxxx. Acesso em: dd mmm aaaa.



Pelo enfrentamento do racismo no ambiente digital

Contribuições do IRIS à Consulta da SECOM e MIR sobre combate ao racismo nas plataformas digitais

Eixo 1 - Formas de violência racial nas plataformas digitais

Descrição: identificação e classificação das manifestações racistas, discriminatórias e preconceituosas, desde o discurso de ódio até o racismo algorítmico; abordagens sobre o silenciamento de vozes negras, indígenas e de outros grupos discriminados.

As diferentes formas com que o racismo atua *em* e *através* de novas tecnologias exige a adoção de uma terminologia adequada a cada contexto. Nesse sentido, gostaríamos de contribuir para a consolidação de dois conceitos que já vêm sendo utilizados pela literatura.

1.1 Racismo e discriminação racial online

O primeiro conceito é de racismo online ou discriminação racial online. Este termo busca dar conta das ofensas racistas intencionais propagadas na internet, que podem ocorrer em diferentes formatos (não somente em texto), seja contra um indivíduo ou um grupo de pessoas. Nas palavras de Tynes et al, discriminação racial online constitui-se de “ataques intencionais verbais ou visuais pessoalmente dirigidos a um indivíduo baseado na raça, exclusão de sites ou conversas comunicadas por texto, vídeo, imagens e símbolos”. A este tipo de violência, os autores destacam que se soma a discriminação racial online “vicária”, que consistiria em “agressões verbais ou visuais intencionais, incluindo piadas sobre o grupo étnico de uma pessoa, testemunhadas por meio de texto, vídeos, imagens e símbolos”, isto é, sem ser destinada a alguém de forma específica.

Em ambos os casos, é possível perceber que a intenção é um dos requisitos da discriminação racial online, demonstrando que se trata de uma classificação que guarda correspondência com os tipos penais de racismo e injúria racial (Silva, 2021). Nestes, de acordo com o ordenamento jurídico e jurisprudência brasileiras, é necessária a comprovação do dolo do agressor, ou seja, a comprovação da intenção de discriminar.

No Brasil, o prof. Dr. Luiz Valério P. Trindade possui alguns dos estudos mais aprofundados sobre o tema, incluindo a análise de casos de racismo na internet, conforme se verifica da imagem abaixo:

Imagem 1 - Casos de discriminação racial online analisados pelo prof. Luiz Valério Trindade.

Post #1: Lugar de mulher preta é no campo colhendo algodão, e não viajando pela Europa.

Post #2: Ela até que é bonitinha. Pena que os pretos não estão mais à venda.

Post #3: Por acaso você obteve permissão do IBAMA para andar por aí com uma macaca?

Post #4: No caso de um *blackout*, a única coisa que se vê são os dentes dela.

Post #5: Quanto custa para um preto viajar de navio?

Post #6: Já é tempo de pôr um fim a este racismo. É chegado o momento de vivermos em paz (os humanos e os pretos).

Fonte: Trindade, 2020, p. 34-35.

Com efeito, verifica-se que são postagens que se utilizam de ofensas explicitamente racistas para, intencionalmente, atacar pessoas negras. Mesmo comentários que, em outro momento, poderiam tentar se valer de uma suposta natureza humorística, são hoje enquadrados como racismo, diante da tipificação do racismo recreativo como crime pela Lei 14.532/2023, que também passou a equiparar a racismo os crimes de injúria racial.

Assim, como forma de melhor compreender o fenômeno do discurso de ódio racista em ambiente digital sugere-se que a terminologia a ser utilizada em casos como os supramencionados seja a de discriminação racial online ou racismo online.

1.2 Racismo algorítmico

Com o avanço de novas tecnologias digitais, em especial de sistemas de inteligência artificial, as formas de violência racista também puderam se atualizar. Ao lado dos casos de discriminação direta, em que a violência está explícita em imagens, símbolos ou palavras, surgem casos em que o racismo é perpetuado pelos próprios resultados e efeitos da tecnologia, ainda que ela não tenha sido programada para tanto.

Sistemas de reconhecimento facial que falham mais sobre o rosto de pessoas negras (Buolamwini; Gebru, 2018), gerando de detenções indevidas a restrições de trabalho (Cintra, Souza e Silva, 2025); moderação de conteúdo automatizada que erra e penaliza mais pessoas negras e trans; sistemas de geração de imagens que reforçam estereótipos degradantes e negativos (Okolo, 2023); e outros inúmeros casos são exemplos de um fenômeno que não pode ser tratado como algo isolado. Ao contrário de um mero problema de *datasets* enviesados, eles representam a forma com que o racismo, historicamente arraigado na sociedade em diferentes âmbitos (desde a família até as relações sociais na escola, universidade, trabalho e afins), impacta profundamente a concepção, construção e desenvolvimento de sistemas de IA.

Para o Dr. Tarcízio Silva, autor do livro *Racismo Algorítmico* (2022), este pode ser definido como “o modo pelo qual a disposição de tecnologias e imaginários sociotécnicos em um mundo moldado pela supremacia branca realiza a ordenação algorítmica racializada de classificação social, recursos e

violência em detrimento de grupos minorizados". A porta de entrada para essa forma de discriminação estrutural em sistemas algorítmicos não está, portanto, somente no uso de dados que contenham padrões racistas, mas também na ausência ou baixa presença de pessoas negras em cursos de tecnologia e empresas de desenvolvimento (Pretalab, 2022), assim como na interpretação humana sobre os resultados de uma IA, que os considera objetivos e inquestionáveis (Ruback, Carvalho e Ávila, 2022), por exemplo. Seja como for, a importância do reconhecimento do conceito de racismo algorítmico está em evidenciar que, mesmo adotando critérios aparentemente neutros, inclusive a não utilização da categoria de raça como uma variável, é plenamente possível que um sistema algorítmico perpetue ideais racistas e excludentes.

Para a prof^a. Dr^a. Bianca Kremer (2023), o racismo algorítmico seria concretizado, na prática, por meio de vieses algorítmicos discriminatórios, que mobilizam as categorias de preconceito e discriminação racial. Nesse sentido, ganha relevância a definição de discriminação indireta, segundo a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, ratificada com status de emenda constitucional por meio do Decreto 10.932/2022. De acordo com a norma, essa forma de discriminação ocorre “quando um dispositivo, prática ou critério aparentemente neutro tem a capacidade de acarretar uma desvantagem particular para pessoas pertencentes a um grupo específico, (...) ou as coloca em desvantagem, a menos que esse dispositivo, prática ou critério tenha um objetivo ou justificativa razoável e legítima (...)”. Ou seja, já consta em nosso ordenamento jurídico uma definição com peso constitucional para os casos de discriminação racial indireta, de modo que em casos de racismo algorítmico, em que não se verifique a intenção do agente, seja possível a sua responsabilização, pelo menos cível, no caso concreto.

Na seara penal, embora a responsabilização se dê somente com a comprovação do dolo do agressor, recomenda-se a revisão da doutrina e jurisprudência sobre o tema, diante do avanço de teorias que abordam a estruturalidade do racismo. Falar sobre a necessidade de intenção em um contexto onde a racismo é regra, e não exceção, pode significar fragilizar a proteção jurídica daqueles que são vítimas, ao mesmo tempo em que representa um desincentivo para que a sociedade busque medidas ativas para evitar práticas racistas não intencionais, isto é, para que adote medidas *antirracistas*.

1.3 Referências bibliográficas

BUOLAMWINI, Joy; GEBRU, Timnit. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. **Proceedings of Machine Learning Research**, Vol. 81, pp. 1–15, 2018.

CINTRA, Juliane; SOUZA, Gustavo; SILVA, Tarcízio. iFood e o racismo algorítmico... dos outros. **Le Monde Diplomatique Brasil**, 2025. Disponível em: <https://diplomatie.org.br/ifood-e-o-racismo-algoritmico-dos-outros/>

Kremer, Bianca. **Racismo algorítmico**. Rio de Janeiro : CESeC, 2023.

OKOLO, Chinasa T. The promise and perils of generative AI: case studies in an African context. **AfriCHI '23**, November 27 – December 01, 2023, East London. South Africa.

Pretalab. **Report 2022**. Disponível em: <https://www.pretalab.com/report-2022>

Ruback, L., Carvalho, D. & Avila, S. (2022). Mitigating Bias in Machine Learning: A Socio-technical Analysis (Mitigando Vieses no Aprendizado de Máquina: Uma Análise Sociotécnica). **iSys: Revista Brasileira de Sistemas de Informação**, 15(1), 23:1-23:31. DOI: 10.5753/isys.2022.2396

Silva, Fernanda dos Santos Rodrigues. Racismo online no Brasil: uma análise a partir da perspectiva de decisões judiciais. 2021. Disponível em: <https://repositorio.ufsm.br/handle/1/25794>

SILVA, Tarcízio. **Racismo algorítmico**. São Paulo: Sesc Edições, 2022.

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

Eixo 4 - Mecanismos de denúncia e atendimento

Descrição: avaliação da acessibilidade, transparência e efetividade dos canais de denúncia atualmente disponíveis tanto nas plataformas quanto no sistema de Justiça.

Pessoas negras que sofrem racismo em ambientes digitais enfrentam um percurso marcado por obstáculos sistêmicos que configuram verdadeira denegação de acesso à Justiça. Primeiro, deparam-se com plataformas que oferecem canais de denúncia confusos, dispersos ou inacessíveis, frequentemente sem opções adequadas para categorizar racismo como crime. Segundo, ao buscar o sistema de Justiça formal, encontram delegacias despreparadas para registrar ocorrências de crimes digitais, ausência de protocolos específicos para racismo online e resistência institucional em reconhecer a gravidade dessas violações. Terceiro, experimentam a chamada "zona de impunidade": denúncias são feitas, mas vítimas raramente recebem retorno sobre encaminhamentos, investigações ou responsabilizações.

A realidade evidencia uma lacuna crítica: inexistem estatísticas públicas consolidadas sobre quantas denúncias de racismo digital são realizadas, qual percentual é efetivamente investigado, quantos casos resultam em responsabilização e quais os tempos médios de resposta (Silva et al, 2024). Essa opacidade informacional impede diagnósticos precisos, formulação de políticas baseadas em evidências e accountability institucional. Ademais, plataformas aplicam protocolos globais de moderação que ignoram a especificidade constitucional brasileira, onde racismo é crime imprescritível e inafiançável, e desconsideram o racismo estrutural que permeia relações sociais e instituições.

A Lei nº 15.211/2025 (ECA Digital) representa avanço normativo significativo ao estabelecer mecanismos de denúncia acessíveis, obrigações de remoção de conteúdos violadores e, fundamentalmente, salvaguardas contra abusos nos instrumentos de denúncia (arts. 28-33). Essa

arquitetura jurídica pode inspirar modelo ampliado de enfrentamento ao racismo digital, incorporando transparência, celeridade e garantias processuais. Internacionalmente, a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância e as Recomendações da ONU sobre Coleta de Dados Antidiscriminatórios fornecem parâmetros para sistemas integrados de denúncia baseados em direitos humanos.

Assim, propomos a criação de uma Rede Nacional de Atendimento e Justiça para Denúncias de Racismo em e nas novas tecnologias, com coordenação compartilhada entre Ministério da Igualdade Racial (MIR) e Secretaria de Políticas Digitais (SPDigi). Em primeiro lugar, a sua estrutura envolveria a criação de um **Observatório Nacional de Racismo Digital**, com o objetivo de criar uma base pública de dados desagregados sobre denúncias de racismo digital, com recortes por raça, gênero, faixa etária, região, tipo de plataforma e natureza da violação. Esse observatório deve publicar relatórios periódicos com análises de série histórica, taxas de resposta institucional, tempos médios de processamento e desfechos, seguindo padrões internacionais de coleta de dados antidiscriminatórios. A transparência de dados é condição para controle social, formulação de políticas públicas baseadas em evidências e responsabilização de plataformas e instituições, sendo que a condução de pesquisas sobre racismo é um dever do Estado brasileiro, conforme a Convenção Interamericana contra o Racismo, a Discriminação Racial e Formas Correlatas de Intolerância, ratificada com peso de emenda constitucional pelo Decreto 10932/2022.

Como forma de contribuir para um melhor registro e tramitação dos casos de racismo, tanto na seara administrativa como judicial, é importante que a Rede proposta contenha **Núcleos Especializados Interinstitucionais**, que poderão envolver Ministérios Públicos estaduais e federal, Defensorias Públicas e Polícias Cíveis, por meio de equipes multidisciplinares capacitadas em direitos digitais, criminologia racial e atendimento humanizado. Esses núcleos devem oferecer acolhimento especializado, realizar triagem qualificada e assegurar encaminhamentos céleres, com atenção especial à vítima. Junto a eles, é importante que o Estado implemente **programas contínuos de capacitação em letramento racial** para operadores do Direito e agentes de segurança pública sobre racismo estrutural, interseccionalidade, trauma racial e acolhimento não-revitimizante, alinhados à jurisprudência da Corte Interamericana de Direitos Humanos sobre discriminação racial e acesso à Justiça.

No que tange às plataformas digitais, é importante assegurar a **cooperação entre plataformas-Estado**, como por meio do desenvolvimento de um protocolo nacional de comunicação direta entre ambas as partes, permitindo fluxo automatizado e seguro de denúncias de racismo. Inspirado no art. 28 do ECA Digital, esse sistema deve garantir que denúncias graves sejam imediatamente encaminhadas às autoridades competentes, preservando evidências digitais e metadados essenciais para investigação, sem depender exclusivamente de iniciativa das vítimas. Na mesma direção, é fundamental assegurar que todos **os canais de denúncia (plataformas e órgãos públicos) adotem linguagem simples e não-técnica**, design acessível conforme diretrizes WCAG, múltiplos formatos de comunicação (texto, áudio, vídeo, Libras) e tradução automática para idiomas indígenas e de comunidades imigrantes. Conforme art. 33 do ECA Digital, acessibilidade não é concessão, mas direito que remove barreiras à participação e ao acesso à Justiça.

Por fim, destaca-se também o papel das plataformas em adotar medidas de **devido processo e garantias procedimentais**, no sentido de aplicar integralmente os princípios de notificação clara, fundamentação, identificação de processos automatizados, direito de recurso e prazos definidos previstos no ECA Digital a todas as denúncias de racismo, protegendo tanto vítimas quanto denunciados de arbitrariedades. Estabelecer priorização legal de denúncias de racismo, reconhecendo sua natureza de crime imprescritível, com metas de tempo máximo de resposta inicial (48h) e conclusão de análise preliminar (15 dias).

Eixo 5 - Protocolos de resposta das plataformas

Descrição: análise das ações (ou omissões) após denúncias de racismo e a existência de medidas preventivas, reparatorias ou educativas.

1. Reconhecimento do racismo como crime e incorporação de perspectivas afrocentradas

No Brasil, o racismo é crime imprescritível e inafiançável (art. 5º, XLII, da Constituição Federal). Essa premissa exige das plataformas digitais uma postura proativa no enfrentamento do racismo em conteúdos online. Contudo, a maioria das plataformas adota protocolos globais de moderação "neutros" ou "race-blind", desenvolvidos em países onde o racismo é enquadrado apenas como discurso de ódio, e não como crime. Essa abordagem universalista ignora o racismo estrutural e suas manifestações específicas em contextos como o brasileiro, reproduzindo assimetrias de poder globais que subalternizam experiências do Sul Global.

Sistemas de IA e algoritmos de moderação aplicam padrões mais rigorosos a postagens de pessoas negras, mesmo quando versam sobre justiça racial, enquanto conteúdos racistas permanecem online. Esse fenômeno configura o que pode ser denominado racismo algorítmico: a reprodução e amplificação de discriminações raciais por meio de sistemas tecnológicos pretensamente neutros. A neutralidade racial nas políticas de moderação não é neutra em seus efeitos: ela perpetua desigualdades e silencia vozes historicamente marginalizadas.

Para corrigir esse desequilíbrio estrutural, os protocolos de resposta das plataformas devem incorporar perspectivas afrocentradas e antirracistas que reconheçam o racismo como crime, como estrutura social e como experiência vivida. Isso requer: (i) formação e capacitação de equipes de revisão por especialistas negros e negras, com ênfase em letramento racial e compreensão contextualizada das manifestações do racismo no Brasil; (ii) diversificação racial e epistêmica das equipes de governança e moderação, assegurando representatividade substantiva e com poder de decisão, não apenas simbólica; (iii) desenvolvimento de critérios de moderação contextualizados, que considerem o impacto desproporcional do racismo sobre comunidades negras e a historicidade e natureza estrutural das violências raciais; (iv) implementação de ações educativas e reparatorias junto às vítimas e comunidades atingidas, incluindo suporte psicossocial, orientação jurídica e medidas de restauração da dignidade; (v) publicação de relatórios periódicos de transparência com dados desagregados por tipo de discriminação racial, resposta adotada, tempo de processamento e taxa de remoção, permitindo accountability e controle social.

Falta diversidade racial na governança das plataformas e é necessário incorporar, ainda, epistemologias do Sul Global e da afrocentricidade para mitigar o racismo algorítmico. A afrocentricidade não significa exclusão de outras perspectivas, mas o recentramento das experiências e saberes negros como fundamentais para a construção de sistemas justos.

2. Garantia do devido processo na moderação de conteúdo

O combate ao racismo deve harmonizar-se com a proteção de direitos fundamentais, especialmente liberdade de expressão e ampla defesa frente a decisões privadas com impacto público. O devido processo na moderação de conteúdo é instrumento indispensável para assegurar transparência, previsibilidade e justiça procedimental, evitando que medidas antirracistas sejam deslegitimadas por arbitrariedades ou que vítimas sofram dupla vitimização.

O ordenamento jurídico brasileiro avançou significativamente nessa direção. O Supremo Tribunal Federal, ao julgar parcialmente inconstitucional o artigo 19 do Marco Civil da Internet, determinou que plataformas criem regras próprias de moderação incluindo devido processo, com direito a compreender fundamentos das decisões e recorrer. A Lei nº 15.211/2025 (ECA Digital) consolidou esse direito, estabelecendo no artigo 30 cinco garantias essenciais: notificação, explicação de motivos, indicação de ação humana ou automatizada, meios para recurso e prazos definidos. Trata-se de marco regulatório que democratiza a liberdade de expressão e fortalece proteção contra decisões discriminatórias.

Esses parâmetros devem ser aplicados aos protocolos de resposta a denúncias de racismo, evitando que vítimas sejam silenciadas por falhas na moderação. Recomenda-se que as plataformas: (i) apliquem os cinco elementos do devido processo do ECA Digital a todos os casos de racismo, independentemente da idade do usuário; (ii) assegurem notificação nítida, fundamentada e em linguagem acessível tanto à vítima quanto ao denunciado, explicitando a natureza do racismo como crime; (iii) garantam revisão humana qualificada por equipes diversas racialmente e capacitadas em questões raciais, evitando decisões puramente automatizadas em casos sensíveis; (iv) mantenham canais acessíveis de recurso com interface inclusiva, considerando barreiras digitais e socioeconômicas; (v) estabeleçam priorização temporal de denúncias de racismo, reconhecendo sua gravidade constitucional; (vi) publiquem relatórios detalhados descrevendo tempos de resposta, taxas de procedência, medidas adotadas e impactos mensuráveis.

O devido processo não é entrave à proteção antirracista: é o instrumento que a torna legítima, equitativa e efetiva. Ele constitui pilar normativo que alinha combate à discriminação com defesa de direitos humanos, fortalecendo confiança social nas medidas de moderação e permitindo controle democrático sobre o poder corporativo das plataformas digitais.

Eixo 6 - Inteligência artificial, moderação algorítmica e impactos desiguais no acesso, distribuição ou circulação de conteúdos

Descrição: discussão sobre como a inteligência artificial e os sistemas automatizados podem perpetuar desigualdades raciais no ambiente digital; temas relacionados ao debate sobre o ecossistema informacional das plataformas digitais, seus modelos de negócios e suas interfaces com aspectos étnico-raciais.

Criadores de conteúdo negros e negras têm apontado impactos desproporcionais sobre os seus perfis de rede social nos últimos anos. Em 2020, influenciadores acusaram o TikTok de reduzir o alcance de suas publicações sobre o movimento #BlackLivesMatter (McCluskey, 2020); no mesmo ano, *creators* brasileiros realizaram experimentos em seus perfis, postando conteúdo com pessoas brancas no Instagram e relataram perceber um aumento exponencial no seu número de visualizações (Morais, 2020); em 2022, entrevistas feitas com influenciadores negros e negras no Brasil reforçou a percepção de que criadores de conteúdo brancos, do mesmo segmento e tamanho que os entrevistados, tendem a ter mais alcance e desempenho nas plataformas (Thamara, 2022).

A ausência de informações adequadas sobre como funcionam os mecanismos de moderação de conteúdo em plataformas digitais, principalmente aqueles de recomendação, é uma situação que enfraquece a identificação e resolução desse tipo de viés discriminatório. No lugar de buscar compreender as razões pelas quais inúmeras pessoas negras relatam ter de enfrentar o mesmo tipo de problema, as plataformas frequentemente se limitam a afirmar que a raça do usuário não é um critério para a determinação do alcance do seu conteúdo. No entanto, diante de uma opacidade quase total sobre como operam os sistemas de recomendação, tais experiências devem ser reconhecidas enquanto formas autênticas de auditoria algorítmica independente, pois a repetição dos casos tende a indicar um padrão.

Em um contexto de discriminação racial estrutural, é necessária a rejeição de abordagens *colour-blind* (“cega às raças”), com o objetivo de compreender adequadamente como mecanismos de recomendação têm atuado, comparativamente, em relação às contas de usuários brancos e negros. Nesse sentido, recomenda-se que:

1. Plataformas digitais que contam com mecanismos de recomendação de conteúdo devem informar adequadamente seus usuários sobre as razões pelas quais determinados conteúdos são recomendados para os seus perfis;
2. Plataformas digitais que contam com mecanismos de recomendação de conteúdo devem realizar avaliação de impacto algorítmico, de modo a identificar possíveis vieses discriminatórios, especialmente raciais, e adotar as medidas de mitigação necessárias;
3. Plataformas digitais que contam com mecanismos de recomendação de conteúdo devem elaborar relatórios de moderação de conteúdo, que contenham informações mínimas sobre o sistema de recomendação.

Referências

McCluskey, Megan. These TikTok Creators Say They're Still Being Suppressed for Posting Black Lives Matter Content. Time, 2020. Disponível em: <https://time.com/5863350/tiktok-black-creators/>

Eixo 7 - Responsabilização das plataformas

Descrição: construção de caminhos normativos e institucionais para garantir que plataformas digitais respondam juridicamente por omissões e falhas.

A transformação da Agência Nacional de Proteção de Dados em agência reguladora de natureza especial, por meio da Medida Provisória nº 1.347/2025, representa inflexão institucional decisiva para a governança digital brasileira. A reestruturação da ANPD consolida modelo regulatório independente, capaz de equilibrar inovação tecnológica com proteção de direitos fundamentais, superando a subordinação ministerial que limitava sua atuação anterior.

Paralelamente, a decisão do Supremo Tribunal Federal sobre o artigo 19 do Marco Civil da Internet estabeleceu novo regime de corresponsabilidade das plataformas digitais, afirmando que empresas devem implementar sistemas próprios de moderação com garantias de devido processo, transparência e direito de recurso. Esse precedente vinculante reconhece que plataformas não são meras intermediárias neutras, mas agentes econômicos com poder de curadoria que exercem governança privada sobre o espaço público digital, devendo responder por omissões sistêmicas na proteção de direitos.

O ECA Digital (Lei nº 15.211/2025) complementa essa arquitetura ao estabelecer deveres de cuidado específicos, incluindo remoção expedita de conteúdos ilícitos, canais de denúncia acessíveis, relatórios de transparência e auditorias de sistemas algorítmicos. Embora focado na proteção de crianças e adolescentes, seus princípios estruturantes (devido processo, prestação de contas algorítmica, e prevenção de danos) são extensíveis ao enfrentamento do racismo no contexto digital.

A partir desse cenário, propomos a consolidação de um regime jurídico de responsabilização civil das plataformas por omissão sistêmica no combate ao racismo online, articulado nos seguintes pilares:

a) Responsabilidade civil por falha no dever de cuidado: assim como ficou estabelecido no julgamento do art. 19, do Marco Civil da Internet, pelo Supremo Tribunal Federal, nos casos de conteúdos racistas, as plataformas digitais devem ser civilmente responsabilizadas por omissão sistêmica, caracterizada pela ausência reiterada de políticas, procedimentos e recursos adequados para prevenir, identificar e remover conteúdos racistas. A responsabilidade deriva do descumprimento do dever de cuidado exigível de empresas que exercem curadoria algorítmica sobre bilhões de interações, em conformidade com o Código de Defesa do Consumidor, a Lei Geral de Proteção de Dados e a jurisprudência do STF. Diferentemente de responsabilidade objetiva por publicação, trata-se de responsabilidade subjetiva por negligência estrutural na governança de conteúdo.

b) Regulação de plataformas digitais com o estabelecimento de obrigações de governança interna e transparência algorítmica: uma vez que o Marco Civil da Internet não aborda especificamente como as plataformas devem conduzir o procedimento de moderação de conteúdo, bem como diante

do estabelecimento de parâmetros mínimos pelo STF no julgamento do art. 19, do MCI, recomendamos a retomada imediata da discussão legislativa em torno da uma regulação de plataformas digitais. Dentre as previsões, como forma de combater o racismo online e algorítmico, é necessário que o texto preveja, pelo menos, a necessidade (i) canais de denúncia multilíngues, acessíveis e seguros, com prazos máximos de resposta ; (ii) sistemas de registro auditável de denúncias, decisões de moderação, recursos e desfechos, com preservação de evidências digitais e anonimização pelo prazo mínimo de 5 anos, respeitadas as leis aplicáveis; (iii) auditorias internas periódicas sobre vieses algorítmicos, taxas de erro de moderação e impactos desproporcionais sobre grupos racializados; (iv) relatórios públicos semestrais de transparência desagregados por tipo de conteúdo, resposta adotada, tempo de processamento e taxa de procedência, incluindo seção específica sobre discriminação racial; (v) avaliações de impacto sobre direitos humanos antes da implementação de novos algoritmos de moderação de conteúdo, incluindo aqueles de recomendação.

c) Protocolo de cooperação entre plataformas digitais e o sistema de justiça brasileiro em casos de racismo: é importante que o Estado estabelece medidas legais que obriguem as plataformas a colaborar ativamente com autoridades policiais e judiciais na investigação de crimes de racismo, fornecendo dados de tráfego e registros de acesso conforme LGPD e Marco Civil, sem que isso implique responsabilidade penal direta. Essa cooperação deve ser condição para exercício legítimo de atividade econômica no território nacional, não mera liberalidade empresarial.