



PERSPECTIVA RACIAL NA INTERLOCUÇÃO ENTRE A DECLARAÇÃO SOBRE PRINCÍPIOS DE DIREITOS HUMANOS NA ÁREA DE INTELIGÊNCIA ARTIFICIAL E A CONSTITUIÇÃO BRASILEIRA

Rita de Cassia Ladeira¹

Maria do Céu Pinto do Amaral²

Editora Responsável: Gabriela Andrade

Revisão: Tailine Hijaz

RESUMO

Este artigo examina a interlocução entre a Declaração sobre Princípios de Direitos Humanos na Área de Inteligência Artificial (DPDHAIA), do Mercosul, promulgada em 2023 e a Constituição Federal de 1988, com foco na proteção dos direitos fundamentais da população negra frente aos impactos da IA. Investigou-se em que medida os princípios da DPDHAIA encontram respaldo nos dispositivos constitucionais brasileiros, considerando as vulnerabilidades das populações racializadas. O objetivo é identificar correspondências e lacunas normativas na abordagem do racismo estrutural mediado por tecnologias digitais. Trata-se de uma pesquisa qualitativa, de natureza exploratória e documental. O percurso metodológico consistiu na análise dos textos integrais da DPDHAIA e da Constituição de 1988, orientada pela busca de palavras-chave (“raça”, “dignidade”, “igualdade”, “racismo”, “preconceito” e “direitos humanos”) e pela construção de uma matriz de correlação normativa entre os princípios declaratórios e os dispositivos constitucionais. Os resultados evidenciam convergências relevantes quanto à dignidade humana, igualdade e repúdio ao racismo, lacunas semânticas na DPDHAIA e ausência de dispositivos constitucionais sobre

¹ Doutora em Bioética, Ética Aplicada e Saúde Coletiva pela Universidade Federal Fluminense (UFF). Possui pós-graduação em Saúde Coletiva da Família e mestrado acadêmico em Relações Étnico-raciais, ambos pelo CEFET/RJ. ORCID: <https://orcid.org/0009-0005-4591-1355>. E-mail: ritaladeira@msn.com

² Mestra em Odontologia Social pela UFF. Professora assistente da Universidade Federal do Rio de Janeiro. ORCID: <https://orcid.org/0000-0003-2705-4327>. E-mail: pintodoamaral@gmail.com

governança algorítmica. Como contribuição, o estudo oferece um debate sobre o risco da IA reproduzir racismos, reforçar preconceitos e estimular desigualdades. Conclui-se ser urgente formular políticas públicas e marcos legais de regulação algorítmica que incorporem a perspectiva racial, assegurando, assim, que a tecnologia IA seja elaborada, construída, aplicada e utilizada com justiça social e responsabilidade antirracista.

Palavras-chave: Inteligência Artificial; Racismo; Direitos Humanos.

RACIAL PERSPECTIVE IN THE DIALOGUE BETWEEN THE DECLARATION OF PRINCIPLES ON HUMAN RIGHTS IN THE FIELD OF ARTIFICIAL INTELLIGENCE AND THE BRAZILIAN CONSTITUTION

ABSTRACT

This article examines the dialogue between the Declaration on Human Rights Principles in the Field of Artificial Intelligence (DPDHAIA) of Mercosur, promulgated in 2023, and the 1988 Federal Constitution of Brazil, focusing on the protection of the fundamental rights of the Black population in the face of AI-related impacts. The study investigates the extent to which the principles of the DPDHAIA find support in Brazilian constitutional provisions, considering the vulnerabilities of racialized populations. The objective is to identify correspondences and normative gaps in addressing structural racism mediated by digital technologies. This is a qualitative, exploratory, and documentary research. The methodological approach consisted of analyzing the full texts of the DPDHAIA and the 1988 Constitution, guided by the search for keywords (“race,” “dignity,” “equality,” “racism,” “prejudice,” and “human rights”) and by constructing a normative correlation matrix between the declaratory principles and the constitutional provisions. The results reveal significant convergences regarding human dignity, equality, and the repudiation of racism, as well as semantic gaps in the DPDHAIA and the absence of constitutional provisions on algorithmic governance. As a contribution, the study offers a discussion on the risk of AI reproducing racism, reinforcing prejudice, and fostering inequalities. It concludes that it is urgent to formulate public policies and legal frameworks for algorithmic regulation that incorporate a racial perspective, thereby ensuring that AI technologies are designed, developed, applied, and used with social justice and anti-racist responsibility.

Keywords: Artificial Intelligence; Racism; Human Rights.

INTRODUÇÃO

Ao longo de sua história, o Brasil atravessou diferentes períodos de desesperança marcados por regimes autoritários, até alcançar sua consolidação como república democrática. Esse processo teve como marco fundamental a queda da ditadura militar e a posterior elaboração e promulgação da Constituição Federal de 1988 (CF), considerada o ponto de partida da atual ordem democrática (Fischmann, 2009).

Com efeito, a CF de 1988 incorporou os conceitos presentes na Declaração Universal dos Direitos Humanos (DUDH, 1948), estrategicamente distribuídos ao longo de seu texto, adotando a igualdade, a justiça social e a dignidade da pessoa humana como valores fundamentais da república.

Em nível global, grande desesperança foi marcada pelas consequências das atrocidades racistas cometidas pelo regime nazista da Alemanha na segunda guerra mundial, no qual os direitos humanos foram severamente violados pelo Estado. Contudo, desse cenário emergiu um forte anseio no sentido de projetar um futuro com direitos humanos reconstruídos (Piovesan, 2004).

Assim, para reencenar as vozes humanas caladas pelas violências da segunda guerra mundial, foi proclamada, em 1948, a DUDH, sob tutela da Organização das Nações Unidas (ONU).

Ainda que haja críticas relacionadas a efetividade global de seus enunciados, alertando sobre o perigo de tornar-se uma simples retórica inócua, Bobbio (2004, p. 48) reconhece que afirmações sobre o alcance universal da Declaração só podem ser feitas “no âmbito da história profética de que falava Kant”, apontando, assim, para uma visão prospectiva e esperançosa dos direitos humanos, mais próxima de uma profecia histórica³ do que de uma previsão científica, se é que “são possíveis previsões científicas na história humana” conforme questiona Bobbio (2004, p. 48).

Nesse sentido, a Declaração Universal pode ser compreendida como uma espécie de conhecimento histórico profético (Bobbio, 2004) cujo gesto intenciona determinar o

³ O termo “história profética”, em Bobbio (2004), remete à concepção kantiana de antecipação normativa e esperançosa do progresso humano. Nesse sentido, a DUDH expressa um gesto intencional voltado ao “possivelmente universal”, ainda que sua efetivação enfrente contradições históricas.

possivelmente universal, diante da vasta diversidade no número de nações participantes (Fischmann, 2009).

Se os enunciados da DUDH, compreendidos por Bobbio (2004) como um conhecimento histórico profético, buscaram no contexto pós-guerra assegurar garantias universais frente às graves violações da dignidade humana, é nas décadas que se seguiram, marcadas por práticas racistas persistentes como as do Apartheid na África do Sul, e mais recentemente pelo racismo da era digital atual, que novos desafios remodelados à efetivação de tais direitos e valores se impõem, evidenciando como a universalidade desses direitos é constantemente testada.

Se no contexto pós-guerra a DUDH respondeu às violações nazistas, na contemporaneidade emergem novos desafios para a efetivação dos direitos humanos, particularmente no campo tecnológico. No Brasil, tecnologias como o reconhecimento facial exemplificam essa reconfiguração das violações, intensificando padrões históricos de discriminação racial e exigindo a reafirmação dos valores estabelecidos pela DUDH.

Nesse sentido, e para ilustrar a urgência desses novos desafios, merece destaque a audiência pública realizada no 187º Período de Sessões da Comissão Interamericana de Direitos Humanos, fruto da mobilização da Uneafro (União de Núcleos de Educação Popular para Negras/os e Classe Trabalhadora), na qual o uso de tecnologias de reconhecimento facial no Brasil foi debatido como prática que configura-se em “violações sistemáticas dos direitos humanos”, ao intensificar a discriminação, sobretudo contra a população negra (Monteiro, 2024, p. 30).

À luz de discussões como esta, e em paralelo às históricas lutas pela democracia no Brasil que resultaram em documentos fundamentais para a existência e convivência digna, a contemporaneidade, no tocante ao reconhecimento e à proteção dos direitos humanos em ambientes digitais, demanda a elaboração de instrumentos capazes de fortalecer a defesa desses direitos contra violações cada vez mais sofisticadas.

Isso porque os usos de tecnologias “podem trazer impactos negativos ao exercício de direitos pela sociedade, como o direito à privacidade, à proteção contra a discriminação, ao acesso à justiça, e às liberdades de expressão, associação e reunião, dentre outros” (Transparência Brasil, 2020, p. 3).

Isto posto, cabe mencionar a Lei nº 15.123, de 2025, que altera o art. 147-B do Código Penal. Essa Lei visa estabelecer “causa de aumento de pena no crime de violência psicológica contra a mulher quando este for praticado com o uso de inteligência artificial ou de outros recursos tecnológicos que altere imagem ou som da vítima” (Brasil, 2025). Essas novas formas de violação mediadas por IA são capazes de manipular imagens, sons e conteúdos para fins de intimidação, humilhação ou constrangimento ampliando os danos psicológicos e simbólicos e revelando o potencial das tecnologias digitais para reproduzir e intensificar violências de gênero.

Na mesma direção das preocupações com determinados tipos de violações, a Reunião de Altas Autoridades em Direitos Humanos do Mercosul (RAADH) aprovou, em 2023, a Declaração sobre Princípios de Direitos Humanos na Área de Inteligência Artificial (DPDHAIA) pela Reunião de Altas Autoridades em Direitos Humanos do Mercosul (RAADH), com a finalidade de eliminar a “discriminação, a promoção da educação e formação digital, a eliminação de segmentos étnico-raciais na inteligência artificial e a utilização de tecnologias de reconhecimento facial com transparência, privacidade e sem segmentos raciais” (Brasil, 2023).

Destaca-se que embora a literatura existente venha explorando amplamente os impactos gerais da inteligência artificial e a necessidade de sua regulamentação (Silva e Silva, 2024), e que a DPDHAIA representa um avanço importante no mundo digital, permanece inexplorada a análise de como os princípios antirracistas da DPDHAIA dialogam com os marcos constitucionais de seus países membros, particularmente sob a perspectiva das vulnerabilidades raciais.

Até o momento a DPDHAIA não é um documento vinculante. Trata-se de um texto de caráter declaratório, uma reafirmação de intenções e princípios pelos Estados-membros, sem criar obrigações legais compulsórias, similar àquelas em resoluções da ONU sobre inteligência artificial (IA), que servem como guias normativos, mas sem força legal imediata (Brasil, 2023).

O documento possui 17 princípios e estabelece a aplicação dos direitos humanos na era digital a partir da centralidade humana. Seu texto aborda um amplo espectro de garantias, abrangendo desde a igualdade e a não discriminação, privacidade e proteção de

dados, acesso universal à internet, integridade da informação em contextos democráticos, transparência algorítmica e controle humano, até a responsabilização, educação digital, não discriminação algorítmica no mundo do trabalho, combate a assimetrias de gênero e étnico-raciais, regulação ética de tecnologias de reconhecimento facial, preservação de culturas originárias, participação social, colaboração multissetorial, canais de denúncia e governança internacional da IA.

Seu objetivo é construir um consenso para estabelecer formas de governança regional que enfrentem os desafios éticos, sociais e legais dos sistemas de IA. Para tanto, reconhece a necessidade de abordagens cada vez mais inclusivas e responsáveis em um ambiente progressivamente algorítmico e, assim como a CF de 1988 incorporou os direitos humanos presentes na DUDH, a DPDHAIA pode inspirar a regulamentação da IA no Brasil.

1. Aspectos Regulatórios

O impacto inovador da IA no mundo, impulsionado por avanços recentes e, em particular, pela IA generativa, é sem precedentes. Esse caráter inovador tem levado pesquisadores a defenderem que a regulamentação da IA seja abordada sob uma perspectiva global, uma vez que esses sistemas operam sem restrições geográficas e seus efeitos são sentidos mundialmente.

Sobre essas preocupações, Coromina (2024) e Balaguer Callejón (2023) defendem que regulamentações locais podem ser insuficientes, tornando-se essencial o desenvolvimento de padrões éticos e legais comuns para a proteção de direitos em escala global.

Embora haja a tese de que regulamentações locais podem ser insuficientes, tornando-se essencial o desenvolvimento de padrões éticos e legais comuns para a proteção de direitos em escala global na área de IA, é importante considerar as lições de outros campos éticos, como a Bioética.

A Bioética principialista, por exemplo, desenvolvida predominantemente em contextos eurocêntricos e norte-americanos — desde o Norte — inicialmente propunha um conjunto de quatro princípios universais para guiar a conduta em saúde e pesquisa com humanos (Beauchamp e Childress, 2012).

Contudo, a experiência demonstrou que sua aplicação direta e acrítica em diferentes realidades, especialmente na América Latina, revelou-se insuficiente para lidar com desafios éticos específicos em regiões marcadas por profundas desigualdades sociais e raciais em contextos de vulnerabilidades herdadas da colonização (Garrafa, Martorell e Nascimento, 2016).

Essa realidade impulsionou o desenvolvimento de abordagens como a Bioética de Intervenção, que, ao reconhecer a complexidade e as características socioculturais e políticas regionais, propôs uma perspectiva ética mais voltada aos problemas locais.

Traçando um paralelo com a regulação da IA, argumenta-se que a busca por padrões éticos e legais globais, embora importante, não deve negligenciar as reais especificidades regionais e locais, principalmente aquelas relativas às vulnerabilidades das populações racializadas e às manifestações do racismo estrutural. Cumpre informar que na América do Sul processos legislativos e regulatórios nacionais da IA ainda estão em curso e avançam de forma gradual e desigual entre os Estados-membros e associados do Mercosul (Coromina, 2024).

No geral, a América do Sul apresenta algumas iniciativas legislativas e regulatórias sobre IA que refletem as diferentes realidades e prioridades de cada país, sendo o Brasil um exemplo notório de como essas realidades são muitas vezes atravessadas por profundas questões sociais e raciais. Para uma compreensão mais precisa, destacam-se a seguir as iniciativas adotadas pela Argentina e pelo Brasil.

Em 2020, a Argentina lançou o Plano Nacional de Inteligência, denominado ARGENIA, com o objetivo de impulsionar a aplicação e o avanço da tecnologia de IA no país. Este plano adota um modelo colaborativo, envolvendo diversos participantes do campo da IA, e está fundamentado em cinco pilares estratégicos: pesquisa e desenvolvimento, formação de talentos, transformação da produção, governo transparente, ética e direitos humanos. Este último eixo enfatiza a necessidade de incorporar considerações éticas e de direitos fundamentais desde as etapas iniciais do design de sistemas inteligentes.

O ARGENIA se dedica a criar diretrizes robustas e ferramentas de supervisão para assegurar que a IA seja confiável, transparente, inclusiva e responsável, uma abordagem

ética considerada inovadora e distinta no contexto da IA (García Benítez e Ruvalcaba-Gómez, 2021). Entretanto, até o momento, na Argentina não há legislação específica em vigor. As discussões sobre o tema continuam e as poucas iniciativas “podem ser caracterizadas como inerciais (imitam e seguem soluções propostas por outros países/regiões), ‘fragmentadas’ (abordam temas isolados, dispersos ou específicos) e pouco eficazes” (Vercelli, 2024, p. 128).

O Brasil delineou sua estratégia para a inteligência artificial em múltiplas fases, incluindo uma consulta pública em 2019 sobre o arcabouço legal da IA, que contou com a participação de membros da sociedade civil e especialistas. O objetivo central dessa iniciativa foi incentivar a pesquisa ética, a inovação e o progresso da IA estimulando o envolvimento de diversos setores da sociedade.

A estratégia brasileira está organizada em nove eixos temáticos, cada um com ações específicas para seu desenvolvimento. Entre eles, destaca-se a promoção da regulamentação e do uso ético da inteligência artificial. Essa meta visa estabelecer princípios éticos, resguardar a privacidade dos dados, combater a discriminação, incluindo a de caráter racial, assegurar o respeito aos direitos humanos e garantir a responsabilização dos envolvidos.

O plano sublinha a relevância da transparência, diversidade e cooperação no campo da IA. Tais aspectos são vistos como fundamentais para o êxito e a responsabilidade no desenvolvimento e aplicação da IA no Brasil (García Benítez e Ruvalcaba-Gómez, 2021). Convém informar que o Projeto que regulamenta o uso da IA no Brasil (Projeto de Lei 2338/23) teve seu texto aprovado em dezembro de 2024 no Senado e para ser transformado em lei precisa ser aprovado pela Câmara dos Deputados onde ainda se encontra em fase de tramitação.

Cabe ressaltar que o Brasil avançou na questão da proteção de dados ao promulgar em 2018 a Lei Geral de Proteção de Dados (LGPD) e, segundo Lima e Sá (2021, p. 245), a partir do artigo 20 da LGPD, “pode-se visualizar a proteção dos dados pessoais como o início da regulação do uso da inteligência artificial”. A reflexão sobre a experiência brasileira com a LGPD, que serve como um ponto de partida para a regulação da IA no país, ilustra a importância de parâmetros legais que produzam efeitos nas esferas jurídicas.

Nesse sentido, a efetividade da DPDHAIA dependerá de sua capacidade de transposição para os marcos jurídicos e regulatórios internos dos países do Mercosul. Embora represente um avanço normativo no reconhecimento de princípios éticos para o uso da IA, sua implementação prática está condicionada a diversos fatores, como a incorporação de princípios da declaração em políticas públicas, regulamentações específicas e mecanismos de supervisão apropriados e a criação de instâncias regionais de monitoramento, auditoria ou responsabilização. Sem isso, o documento corre o risco de tornar-se uma diretriz simbólica, sem efetiva aplicação jurídica ou institucional.

O compromisso político é importante elemento nesse processo: a ausência de sanções ou metas vinculantes pode fragilizar a potencial eficácia da declaração. Apesar disso, a DPDHAIA pode funcionar como a principal referência normativa para orientar legislações nacionais e fomentar debates sobre aparatos regulatórios específicos para os diversos sistemas tecnológicos de IA no Sul global, tal como tem ocorrido na União Europeia com a proposta de *AI Act*, que institui mecanismos legais obrigatórios para o uso ético da IA (European Parliament, 2024).

2. Aspectos socio-humanos

Assim como a DUDH serviu de inspiração para a Constituição brasileira, a elaboração da DPDHAIA também foi inspirada por marcos internacionais, como a própria DUDH (1948), pela Declaração Americana de Direitos e Deveres do Homem (1948), pela Recomendação da UNESCO sobre Ética na IA (2021) e pelas resoluções recentes da ONU e da União Europeia sobre governança da IA (Mercosul, 2023).

A DPDHAIA está inserida em um contexto socio-histórico global e regional muito específico, marcado por três grandes vetores: o avanço acelerado da inteligência artificial e suas implicações sociais; a pandemia de COVID-19 e a intensificação das iniquidades; e o fortalecimento dos movimentos por justiça digital e pelo antirracismo na região.

Nesse liame, o avanço acelerado da inteligência artificial nos últimos cinco anos, resultou em um crescimento expressivo no desenvolvimento e na aplicação de diferentes

sistemas baseados em IA, tanto em serviços públicos (segurança, saúde, educação) quanto no setor privado (recrutamento, crédito, consumo).

Esse movimento tem sido acompanhado por uma série de violações éticas e sociais que afetam a dignidade e os direitos humanos, pois as diversas tipologias de IA estão sujeitas a reproduzir os conceitos, os preconceitos, as concepções ideológicas, políticas, culturais e religiosas de valores apreendidos e de experiências humanas, ainda que esses sistemas possam conduzir diálogos ou tarefas com ou sem a intervenção humana direta, conforme relatado em diversos estudos que interseccionam os algoritmos às questões sociais (Lima e Sá, 2021; Araújo e Araújo, 2024; Bitencourt e Martins, 2024; Silva, Lima e Brasileiro, 2025).

Tais estudos demonstram que os sistemas algorítmicos perpetuam políticas racistas e a criação de conteúdos discriminatórios, dentro e fora dos meios digitais (Rodrigues e Chai, 2023), sendo esse fenômeno reconhecido como racismo algorítmico. “A concepção política do termo ‘racismo algorítmico’ tem sido levantada por autores que observam como os impactos discriminatórios da IA podem intensificar a opressão racial e reprodução da supremacia branca [...]” (Silva e Silva, 2024, p. 4).

Para Tarcízio Silva (2021), em entrevista ao Centro de Estudos Estratégicos da Fiocruz (CEE-Fiocruz), o racismo algorítmico representa uma atualização do racismo estrutural. Considerando que as tecnologias algorítmicas são atravessadas por relações de poder, é relevante destacar que o racismo constitui um fenômeno social sustentado por essas mesmas relações, responsável por perpetuar desigualdades estruturais relacionadas à raça (Almeida, 2019).

Bitencourt e Martins (2024) ressaltam que a IA pode incorporar preconceitos e vieses presentes nos dados que a alimentam, comprometendo a neutralidade das decisões ao operar com representações estereotipadas que restringem e classificam determinados indivíduos ou grupos sociais, visto ser “inegável que as decisões humanas não são neutras” (Lima e Sá, 2021, p. 228).

Nesse contexto, o uso de decisões automatizadas enviesadas por agentes humanos aprofunda práticas discriminatórias, sobretudo contra populações historicamente marginalizadas.

A menção à pandemia como vetor está no fato de a disseminação de informações sobre a COVID-19 expor a ampliação de desigualdades no acesso às tecnologias digitais e nos sistemas de decisão baseados em algoritmos. Grupos vulneráveis, especialmente as maiorias minorizadas étnico-raciais e de baixa renda, enfrentaram barreiras ainda maiores para acessar e se beneficiar dessas tecnologias, o que aprofundou iniquidades já presentes no mundo físico da saúde, da educação e no acesso à informação (Litchfield et al., 2021).

No campo da e-saúde, embora a telessaúde tenha apresentado muitos aspectos positivos ao longo da pandemia, ela também expôs a persistente exclusão, que mantém populações social e digitalmente marginalizadas com acesso ainda mais limitado a serviços essenciais, tornando determinados grupos populacionais invisíveis no ambiente digital (Faraj et al., 2021).

Ainda em relação ao vetor, Matsiele e Mutula (2024) destacam que estudantes pertencentes a grupos desfavorecidos, notadamente aqueles oriundos de famílias negras e de baixa renda, sofreram significativamente durante a pandemia, não apenas pela dificuldade de acessar a educação online, mas também por enfrentarem diversas formas de exclusão sociodigital, que se sobrepõem e aprofundam as desigualdades estruturais em relação ao acesso à informação.

Apesar de não haver dúvidas sobre a importância da IA em diversos setores, seu avanço tem introduzido dilemas éticos e normativos, como a erosão da privacidade, a discriminação algorítmica, violações de dignidade, comprometimento da autonomia e a opacidade nos processos de tomada de decisão automatizada. Frequentemente, conforme apontado por Bitencourt e Martins (2024), a IA reproduz e amplifica estereótipos, discriminações e preconceitos já existentes na sociedade, em razão dos vieses presentes nos dados de treinamento.

O caso da deputada brasileira Renata Souza⁴, cuja imagem gerada por IA a associou indevidamente à violência em uma favela, é um exemplo de como esta ferramenta pode ser utilizada para manifestar microagressões, como microinsultos e microinvalidações, que desonram a identidade étnico-racial (Araújo e Araújo, 2024).

⁴ Ao pedir a uma IA a imagem de uma mulher negra, com cabelos afro e roupas africanas em uma favela, a IA gerou uma mulher negra armada, associando indevidamente sua identidade e o cenário de favela à violência.

No contexto das garantias de direitos, a implementação de câmeras de reconhecimento facial em estádios de futebol em 2023, sob a justificativa de segurança pública, ampliou a vigilância sobre a população, inclusive sobre crianças e adolescentes. Essa prática tem sido alvo de crescente debate no Brasil e vem levantando sérias preocupações jurídicas e éticas, pois, segundo a LGPD, o tratamento de dados de menores deve ocorrer no melhor interesse da criança e com consentimento expresso dos responsáveis.

O uso dessa tecnologia, conforme alerta o relatório *Esporte, Dados e Direitos* do CESeC, representa um tratamento de alto risco e pode violar não apenas a LGPD, mas também o Estatuto da Criança e do Adolescente e a Constituição. Soma-se a isso o risco de falhas no reconhecimento, que podem resultar em prisões injustas e ferir princípios fundamentais como a dignidade humana e a presunção de inocência. Diante desses paradoxos entre segurança e liberdade, tecnologia e direito, impõe-se a urgente necessidade de refletir sobre os limites éticos do uso dessas ferramentas, buscando um equilíbrio que não viole direitos em nome da proteção.

Mais recentemente, em julho de 2025, a IA Grok, chatbot do X, da startup de Elon Musk, causou espanto no mundo ao emitir uma série de respostas/mensagens consideradas antissemitas e dotadas de informações falsas e enganosas sobre o nazista Adolf Hitler. A empresa esclareceu que essas mensagens foram geradas após uma atualização do código da IA, na qual os programadores treinaram e instruíram a IA a ser franca e sem medo de chocar a sociedade do politicamente correto⁵.

Conforme argumenta Pereira (2024, p. 17), ao afirmar que “a desinformação e o discurso de ódio caminham lado a lado e se retroalimentam”, este incidente constitui uma contundente demonstração de como a desinformação mediada por sistemas de IA pode, ao reproduzir e amplificar conteúdos de ódio, ecoar as atrocidades do racismo nazista que motivaram a própria promulgação da DUDH, reafirmando a relevância de seus enunciados para combater as manifestações emergentes do racismo na era digital.

⁵ Globo. Grok, IA de Musk, pede desculpas por postagens de exaltação a Hitler. G1, 12/7/2025.
<https://g1.globo.com/tecnologia/noticia/2025/07/12/grok-ia-de-musk-pede-desculpas-por-postagens-de-exaltacao-a-hitler.ghml>

No contexto do processo de injustiça digital, encontra-se a exclusão digital, também conhecida como brecha digital. Mais precisamente, refere-se a um problema social complexo que vai além do simples acesso à tecnologia, abrangendo disparidades e desigualdades que podem levar à exclusão e discriminação. É considerada um indicador de pobreza e violação dos direitos humanos. Atinge mulheres e meninas, especialmente as desempregadas ou em situação de vulnerabilidade, com impactos na educação, carreira profissional e empreendedorismo, bem como atinge idosos e populações rurais (Casadei, 2025).

Além disso, algoritmos opacos utilizados em sistemas de contratação automatizada podem discriminar certos grupos raciais sem proporcionar explicação cristalina, perpetuando vieses históricos existentes nos dados de treinamento (Araújo e Araújo, 2024). A falta de transparência sobre a responsabilidade dos algoritmos impacta a capacidade da busca por justiça.

No sentido de reconhecer, apontar e publicizar violações da dignidade dos seres humanos, aspectos da exclusão digital e do racismo algorítmico vêm sendo discutidos tanto por pesquisadores quanto por instâncias governamentais. Entre as soluções propostas, destaca-se a necessidade de implementar políticas públicas e regulamentações que responsabilizem as empresas envolvidas no desenvolvimento e na aplicação de algoritmos com tendência racista e discriminatória.

No entanto, como visto, a regulamentação da inteligência artificial enfrenta o desafio do descompasso entre o ritmo acelerado do desenvolvimento tecnológico e a lentidão dos processos democráticos de criação de leis. Essa disparidade traz implicações diretas para a garantia de direitos, especialmente no que tange à proteção de grupos vulnerabilizados.

Diante da ausência de uma legislação específica em muitos países membros ou associados do Mercosul, a DPDHAIA, surge como instrumento orientador, com princípios voltados à promoção da justiça, da não discriminação e da proteção de direitos fundamentais no contexto da IA.

Nesse sentido, reverbera-se a importância de examinar de que forma esses princípios dialogam com os marcos constitucionais de cada país. No caso brasileiro,

embora a Constituição de 1988 ainda não tenha enfrentado e nem reparado plenamente os legados históricos da violenta escravidão do povo preto, ela representa um marco jurídico importante ao fornecer instrumentos normativos que norteiam o combate às práticas estruturais de racismo que persistem na sociedade brasileira.

Dito isso, a principal questão que o presente estudo buscou responder foi: em que medida os princípios da DPDHAIA encontram interlocuções na Constituição Federal de 1988, especialmente em relação à proteção de direitos fundamentais frente aos impactos da IA considerando as especificidades e vulnerabilidades das populações racializadas?

Com base nisso, este artigo propõe analisar a Declaração sobre Princípios de Direitos Humanos na Área de Inteligência Artificial do Mercosul, buscando identificar suas correlações com os dispositivos constitucionais brasileiros. O enfoque recai sobre os princípios da igualdade racial, da não discriminação e do combate ao racismo, com o objetivo de revelar tanto convergências normativas e aproximações conceituais quanto lacunas que comprometam a efetividade da proteção dos direitos fundamentais diante dos desafios impostos pela inteligência artificial.

3. Método

Trata-se de uma investigação qualitativa — à luz de Minayo (2009, p. 21), para quem esta abordagem "trabalha com o universo dos significados, dos motivos, das aspirações, das crenças, dos valores e das atitudes", compreendendo estes fenômenos como parte integrante da realidade social — e de natureza exploratória no intuito de desenvolver, esclarecer e aprimorar conceitos e ideias com vistas a formular problemas mais delineados ou hipóteses passíveis de investigação em estudos futuros, conforme o entendimento de Gil (2002). Além disso, este trabalho adotou a pesquisa documental, fundamentando-se em Sá-Silva, Almeida e Guindani (2009). Os referidos autores argumentam que os documentos constituem fontes privilegiadas de informação, cuja análise aprofundada não apenas valida

a investigação, mas também expande significativamente a compreensão do objeto em estudo.

A opção por uma análise documental de natureza qualitativa exploratória justifica-se uma vez que os documentos elencados constituem fontes primárias e como os próprios alicerces normativos sobre os quais se ergue a investigação proposta. Além disso, esmiúça e aprofunda conceitos e ideias.

Foram utilizadas como fontes a Declaração sobre Princípios de Direitos Humanos na Área de Inteligência Artificial (DPDHAIA) e a Constituição Federal de 1988. A análise documental foi orientada por busca de palavras-chave — 'raça', 'dignidade', 'racismo', 'preconceito', 'igualdade', 'violação', 'direitos humanos' — no PDF dos documentos, e os resultados foram organizados em uma matriz de correlação normativa.

Contudo, reconhece-se um limite significativo nesse estudo em decorrência da escassez de literatura acadêmica e análises específicas sobre a DPDHAIA (2023), dado o seu caráter recente. Essa carência impossibilitou um diálogo bibliográfico mais denso, bem como uma possível contrastação de interpretações sobre o documento, demandando das pesquisadoras uma análise original do texto declaratório, no que diz respeito às questões raciais.

Ferramentas de IA foram empregadas em etapas específicas deste manuscrito, de forma pontual e auxiliar. O SciSpace foi utilizado como apoio à busca bibliográfica preliminar, facilitando a identificação de fontes relevantes. Já o Google NotebookLM, baseado em modelo de linguagem, contribuiu para a organização inicial das reflexões e o mapeamento temático dos documentos analisados.

Para o SciSpace foi utilizado o seguinte prompt: Busque artigos publicados a partir de 2021 que abordem a relação entre inteligência artificial, reconhecimento facial e discriminação racial, com foco no contexto brasileiro ou latino-americano. Priorize estudos sobre os impactos para a população negra e discussões éticas e regulatórias, incluindo referências à LGPD ou à DPDHAIA, se houver. Apresente título, resumo e link de acesso, se disponível.

Para o Google NotebookLM foi utilizado o seguinte prompt: Gere tópicos temáticos a partir desses documentos e textos selecionados que tratam da relação entre inteligência

artificial, reconhecimento facial e discriminação racial. Considere especialmente os impactos sobre a população negra, os riscos do racismo algorítmico. Crie resumos e mapas mentais organizando as ideias de forma conectadas. Os tópicos devem refletir eixos relevantes para discussão acadêmica.

Ressalta-se, contudo, que todos os processos de pesquisas, resultados, análises, discussão e redação foram integralmente conduzidos pelas autoras.

4. Resultados e discussão

A análise dos princípios da DPDHAIA do Mercosul (2023) evidencia o reconhecimento da questão racial como sendo uma importante dimensão para a governança ética da IA. Na interlocução com a Constituição Federal de 1988, destacam-se três princípios: o Princípio 2 (Igualdade e Não Discriminação), o Princípio 11 (Eliminação de Assimetrias Étnico-raciais na Inteligência Artificial) e o Princípio 12 (Tecnologias de Reconhecimento Facial com Transparência, Privacidade e sem Assimetrias Raciais), especificados no Quadro 1.

Quadro 1 - Matriz de relação entre os princípios antirracistas da DPDHAIA e dispositivos constitucionais brasileiros (CF/1988)

Princípio da Declaração	Dispositivo constitucional	Conteúdo	Tipo de associação	Lacuna
2. Igualdade e Não Discriminação	Art. 3º, IV	promover o bem de todos, sem preconceitos de origem, raça, sexo, cor, idade ou quaisquer formas de discriminação.	Correspondência direta e proteção ampla	Na DPDHAIA: Ausência da palavra “raça” ou “racismo” Lacuna semântica e política que pode comprometer a visibilidade do racismo estrutural.
	Art. 5º I, XLI e XLII	Garante igualdade; proíbe discriminação; considera o racismo crime inafiançável e imprescritível.	Proteção explícita e penalização do racismo	Na CF-1988: Faltam normas que responsabilizem tendências em decisões algorítmicas. Lacuna normativa frente às novas formas de discriminação mediadas por tecnologia.
	Art. 1º, III	Dignidade da pessoa humana como fundamento da República.	Alicerce principiológico	
	Art. 4º, II, VIII	Prevalência dos direitos humanos. O Brasil rege-se nas relações internacionais pelo repúdio ao racismo.	Princípios fundamentais. Alicerce político e internacional	
11. Eliminação de Assimetrias Étnico-raciais na IA	Art. 3º, IV; Art. 5º I e XLII; Art. 1º, III; Art. 4º II, VIII	Idem ao Princípio 2	Idem à associação 2	
	Art. 218 e 219	O desenvolvimento científico e tecnológico deve promover o bem-estar e respeitar os valores nacionais e sociais.	Alicerce ético-científico para desenvolvimento justo	
12. Reconhecimento Facial sem Assimetrias Raciais	Art. 5º I, X, XLII e LXXIX (EC 115/2022- DOU de 11/2/2022)	Idem ao Princípio 2, Proteção de dados pessoais, inviolabilidade de imagem.	Idem à associação 2, sem correspondência direta – respaldo principiológico limitado	Na CF-1988: Ausência de dispositivos específicos sobre governança tecnológica
	Art. 230	Amparo integral a grupos vulneráveis, defesa da dignidade	Proteção implícita	

Fonte: Elaboração própria

O Princípio 2 da DPDHAIA estabelece a igualdade e a não discriminação como fundamentos da utilização da inteligência artificial. Ainda que não mencione explicitamente a palavra raça, a formulação que se refere a “grupos em situação de vulnerabilidade por

suas características, pertencimento ou outras condições” engloba, implicitamente, as comunidades negras, mas não descrever o termo pode comprometer a identificação do racismo estrutural e do discurso de ódio contra essa população.

Conforme Pereira (2024) relata, o estudioso dos discursos de ódio nas redes sociais e da representação de minorias étnicas nos meios de comunicação de massa, o professor Luiz Valério Trindade aponta que as mulheres negras figuram entre os principais alvos dessas práticas, destacando que a disseminação de conteúdos odiosos nas plataformas digitais está vinculada tanto a ideologias de supremacia branca quanto a interesses econômicos das corporações responsáveis pelas redes sociais, que lucram com a propagação do ódio. Essa mesma lógica de reprodução de desigualdades e violência simbólica também pode ser observada, ou mesmo intensificada, nas tecnologias de IA.

A interlocução com a Constituição evidencia respaldo normativo relevante, especialmente no artigo 3º, inciso IV, que estabelece como objetivo fundamental da República a promoção do bem de todos, sem preconceitos de origem, raça, sexo, cor ou idade. A isso se somam os artigos 5º, inciso XLII, que trata da punição de práticas discriminatórias e da tipificação do racismo como crime inafiançável e imprescritível, além do reconhecimento da dignidade da pessoa humana como fundamento da República (art. 1º, III).

Para Lima e Sá (2021, p. 239) “o princípio da dignidade humana é entendido como o reconhecimento do inerente estado humano de ser digno de respeito”. As tecnologias, incluindo os sistemas de IA, devem estar a serviço das pessoas, e não o contrário. Dessa forma, qualquer uso da IA que resulte na violação de direitos fundamentais representa uma afronta direta à Constituição de 1988.

O Princípio 11 é o mais direto quanto à temática racial, ao reconhecer a necessidade de eliminar assimetrias étnico-raciais nos sistemas de IA. Ele propõe que os Estados Partes promovam políticas públicas voltadas ao enfrentamento do racismo algorítmico, incluindo a regulação de tecnologias como o reconhecimento facial. A esse princípio correspondem dispositivos constitucionais que, embora não tratem especificamente de algoritmos, orientam a atuação estatal quanto ao desenvolvimento

tecnológico com responsabilidade social, como os artigos 218 e 219, que vinculam a ciência e a tecnologia à promoção do bem-estar e da justiça social.

Também são relevantes os incisos VIII e II do artigo 4º, que estabelecem o repúdio ao racismo e a prevalência dos direitos humanos como diretrizes que regem as relações internacionais do Brasil, sinalizando a legitimidade da atuação do Estado em fóruns multilaterais sobre a temática.

Já o Princípio 12, que trata das tecnologias de reconhecimento facial com transparência, privacidade e sem assimetrias raciais, apresenta um evidente descompasso com a Constituição de 1988. Embora haja respaldo em dispositivos como o artigo 5º, incisos X e LXXIX (este último introduzido pela Emenda Constitucional nº 115/2022), que tratam da inviolabilidade da intimidade e da proteção de dados pessoais, a ausência de normas constitucionais específicas sobre governança algorítmica evidencia uma lacuna normativa.

Essa lacuna normativa é especialmente grave no Brasil, onde os sistemas de reconhecimento facial vêm sendo associados a erros, falsas identificações e discriminações desproporcionais “contra pessoas negras, em especial mulheres negras” (Monteiro, 2024, p. 28).

A interlocução entre os princípios da DPDHAIA e a Constituição de 1988 revela, portanto, fortes convergências normativas no campo da igualdade racial, do repúdio ao racismo e dos direitos humanos, mas também aponta limitações significativas que comprometem a efetividade do antirracismo na prática, uma vez que o avanço tecnológico sem a devida proteção jurídica abre brechas para a perpetuação e amplificação de iniquidades históricas.

Esse cenário exige ação legislativa, que poderia ser implementada tanto por meio de emenda constitucional quanto pela aprovação de projeto de lei em tramitação no Congresso Nacional (PL nº 2338, de 2023).

A ausência de regulamentação constitucional e infraconstitucional sobre viés algorítmico racial, aliada à falta de parâmetros para tecnologias de vigilância automatizada, favorece a reprodução de iniquidades históricas no ambiente digital. Algoritmos treinados

com dados enviesados tendem a perpetuar injustiças e exclusões contra populações racializadas, ampliando os efeitos do racismo estrutural sob novas roupagens.

Os achados deste trabalho reforçam a urgência de atualizar os marcos normativos e regulatórios, não apenas para acompanhar o ritmo do desenvolvimento tecnológico, mas principalmente para assegurar que a IA seja uma ferramenta de progresso e não um novo vetor de reprodução do racismo e do agravamento das profundas desigualdades, discriminação e opressão que historicamente afetam a população negra brasileira.

Nesse sentido, torna-se prioritária a adoção de medidas que não se limitem a respostas punitivas, mas que também atuem de forma preventiva, buscando evitar a reprodução e o agravamento de preconceitos e discriminações mediadas por tecnologias. Tais medidas devem ser calibradas para considerar as especificidades das diversidades sociorraciais do Brasil, garantindo que as soluções sejam socialmente sensíveis e eficazes.

CONSIDERAÇÕES FINAIS

Este estudo se concentrou na interlocução entre a Declaração sobre Princípios de Direitos Humanos na Área de Inteligência Artificial (DPDHAIA), aprovada em 2023 pelo Mercosul, e a Constituição Federal de 1988. A investigação partiu da indagação sobre em que medida esses princípios encontram respaldo constitucional diante das vulnerabilidades das populações racializadas.

A análise permitiu elencar achados específicos: identificou-se uma sólida convergência normativa quanto aos fundamentos da dignidade humana, igualdade e repúdio ao racismo; constatou-se, porém, uma lacuna normativa crítica na Carta Magna quanto à governança algorítmica e à regulação de tecnologias de IA; e detectou-se uma lacuna semântica na própria DPDHAIA, que ao não empregar explicitamente o termo "raça", pode vir a comprometer o enfrentamento direto ao racismo estrutural.

Ao buscar identificar correspondências e lacunas normativas na abordagem do racismo mediado por tecnologias, a discussão evoluiu ao evidenciar, com base em referências atuais, que o racismo algorítmico representa um risco concreto aos direitos humanos. Os resultados mostram a relevância da CF-1988 como fundamento normativo no

que diz respeito às questões raciais, mas revelam desatualizações em seus marcos normativos, bem como ausências regulatórias em forma de leis específicas em relação à IA.

A relevância dos achados está em sinalizar a urgência de uma regulação que traduza os compromissos constitucionais em salvaguardas técnicas e mecanismos de prestação de contas e responsabilização, tais como auditorias independentes de algoritmos, avaliações prévias de impacto racial e canais efetivos de recurso e reparação, assegurando que a inteligência artificial não seja mais um vetor de reprodução das iniquidades raciais históricas no Brasil.

Conclui-se, portanto, ser urgente formular políticas públicas e marcos legais de regulação algorítmica que incorporem a perspectiva racial, assegurando que a IA seja uma ferramenta de justiça e não de perpetuação de desigualdades. A superação desses desafios será decisiva para a efetividade dos direitos humanos. Futuras pesquisas poderão, à medida que a literatura sobre a DPDHAIA se expanda, aprofundar análises com contribuições de diversas áreas do conhecimento, superando assim uma das limitações do presente estudo e consolidando o vínculo entre regulação algorítmica e justiça racial.

REFERÊNCIAS

ALMEIDA, S. L. Racismo estrutural (Feminismos Plurais). São Paulo: Pólen, 2019.

ARAÚJO, Júlio; ARAÚJO, Júlio. Racismo algorítmico e inteligência artificial: uma análise crítica multimodal. **Revista Linguagem em Foco**, v.16, n.2, p. 89-109, 2024. DOI: <https://doi.org/10.46230/lef.v16i2.13108>. Disponível em: <https://revistas.uece.br/index.php/linguagememfoco/article/view/131008.10.46230/lef.16i2.13108>. Acesso em: 15 jun. 2025.

BALAGUER C. F. A Constituição do Algoritmo. Rio de Janeiro: Forense, 2023.

BEAUCHAMP, T.L.; CHILDRESS, J.F. Principles of Biomedical Ethics. 7. ed. New York: Oxford University Press, 2012.

BITENCOURT, C. M.; MARTINS, L. H. N. O estado da arte da utilização de inteligência artificial nos órgãos de controle de contas da administração pública brasileira. **Revista Jovens Pesquisadores**, Santa Cruz do Sul, v. 14, n. 2, p. 152-168, dez 2024.

DOI: <https://doi.org/10.17058/rjp.v14i2.19230>. Disponível em: <https://online.unisc.br/seer/index.php/jovenspesquisadores/article/view/19230>. Acesso em: 8 jun. 2025.

BOBBIO, N. A era dos direitos. Tradução de Carlos Nelson Coutinho. Apresentação de Celso Lafer. Nova edição. Rio de Janeiro: Elsevier, 2004. 7ª reimpressão.

BRASIL. Constituição da República Federativa do Brasil. Brasília, DF: Presidência da República, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/ConstituicaoCompilado.htm. Acesso em: 14 abr. 2025.

BRASIL. Lei nº 15.123, 24 de abril de 2025. Altera o art. 147-B do Decreto-Lei nº 2.848, de 7 de dezembro de 1940 (Código Penal), para estabelecer causa de aumento de pena no crime de violência psicológica contra a mulher quando praticado com o uso de inteligência artificial ou de qualquer outro recurso tecnológico que altere imagem ou som da vítima. Diário Oficial da União, Brasília, DF, 24 abr. 2025. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2023-2026/2025/Lei/L15123.htm. Acesso em: 12 set. 2025.

BRASIL. Ministério dos Direitos Humanos e da Cidadania. Em plenária final, RAADH aprova Declaração de Princípios de Direitos Humanos no âmbito da Inteligência Artificial no Mercosul. Disponível em: <https://www.gov.br/mdh/pt-br/assuntos/noticias/2023/novembro/em-plenaria-final-raadh-aprova-declaracao-de-principios-de-direitos-humanos-no-ambito-da-inteligencia-artificial-no-mercosul>. Acesso em: 12 abr. 2025.

CASADEI, T. Brechas digitais: el reto de las nuevas tecnologías para los derechos humanos. **Revista de la Facultad de Derecho de México**, v. 74, n. 290, p. 149-178 13 dez. 2024. DOI: 10.22201/fder.24488933e.2024.290.90069. Disponível em: <https://www.revistas.unam.mx/index.php/rfdm/article/view/90069>. Acesso em: 11 abr. 2025.

COROMINA, M. P. U. Análisis comparado de los distintos enfoques regulatorios de la inteligencia artificial en la Unión Europea, EE. UU., China e Iberoamérica. **Anuario Iberoamericano de Justicia Constitucional**, v. 28, n.1, p. 129-156, jul. 2024. DOI: <https://doi.org/10.18042/cepc/aijc.28.05>. Disponível em: <https://recyt.fecyt.es/index.php/AIJC/article/view/108247>. Acesso em: 10 abr 2025

EUROPEAN PARLIAMENT. Artificial Intelligence Act: MEPs adopt negotiating position. 2024. Disponível em: <https://www.europarl.europa.eu/news/en/pressroom/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law>.

FARAJ, S.; RENNO, W.; BHARDWAJ, A. Unto the breach: What the COVID-19 pandemic exposes about digitalization. **Information and Organization**, v.31, p. 1-7, mar. 2021. Doi: <https://doi.org/10.1016/j.infoandorg.2021.100337>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1471772721000038>. Acesso em: 12 jul. 2025.

FISCHMANN, R. Constituição brasileira, direitos humanos e educação. **Revista**

Brasileira de Educação, v. 14, n. 40, p. 156–167, jan. 2009. DOI: <https://doi.org/10.1590/S1413-24782009000100013>. Disponível em: <https://www.scielo.br/j/rbedu/a/QPz7bgW7FmF3K4tbVRHVNmt/?lang=pt>. Acesso em: 5 abr. 2025.

GARCÍA BENÍTEZ, V. H.; RUVALCABA-GÓMEZ, E. Análisis de las estrategias nacionales de inteligencia artificial en América Latina: estudio de los enfoques

de ética y de derechos humanos. **Revista de Gestión Pública**, Guadalajara, v. 10, n. 1, p. 5-32, jun. 2021. DOI: <https://doi.org/10.22370/rgp.2021.10.1.3151>. Disponível em: <https://revistas.uv.cl/index.php/rgp/article/view/3151>. Acesso em: 14 jun. 2025.

GARRAFA, V.; MARTORELL, L. B. e NASCIMENTO, W. F. Críticas ao principialismo em bioética: perspectivas desde o norte e desde o sul. **Revista Saúde e Sociedade**, São Paulo, v. 25, n. 2, p. 442-451, abr 2016. DOI: <https://doi.org/10.1590/S0104-12902016150801>. Disponível em: <https://www.scielo.br/j/sausoc/a/v46P8wRBDz6588xLJYLjcRh/>. Acesso em: 12 jun. 2025.

GIL, Antônio Carlos. 1946 - **Como elaborar projetos de pesquisa**. - 4. ed. - São Paulo: Atlas, 2002.

LIMA, Taisa Maria Macena de; SÁ, Maria de Fátima Freire de. Inteligência artificial e Lei Geral de Proteção de Dados Pessoais: o direito à explicação nas decisões automatizadas. **Revista Brasileira de Direito Civil**, v. 26, n. 04, p. 227-246, fev. 2021. DOI: <https://doi.org/10.33242/rbdc.2020.04.011>. Disponível em: <https://rbdcivil.ibdcivil.org.br/rbdc/article/view/584>. Acesso em: 14 mai. 2025.

LITCHFIELD, I.; SHUKLA, D.; GREENFIELD, S.. Impact of COVID-19 on the digital divide: a rapid review. **BMJ Open**, v. 11, n. 10, p. 19, out. 2021. DOI: <https://doi.org/10.1136/bmjopen-2021053440>. Disponível em: <https://bmjopen.bmj.com/content/11/10/e053440>. Acesso em: 10 abr. 2025.

MATSIELE, M.; MUTULA, S. COVID-19 and digital transformation in higher education institutions: towards inclusive and equitable access to quality education. **Education Sciences**, Basel, v. 14, n. 8, p. 1-19, jul. 2024. DOI: <https://doi.org/10.3390/educsci14080819>. Disponível em: <https://www.mdpi.com/2227-7102/14/8/819>. Acesso em: 12 jun. 2025.

MERCOSUR. Declaração sobre Princípios de Direitos Humanos na Área de Inteligência Artificial. Disponível em: https://documentos.mercosur.int/simfiles/declaraciones/98118_ATTXNW0C.docx. Acesso em: 16 abr. 2025.

MINAYO, M. C. S. O desafio da pesquisa social. In: MINAYO, M. C. S. (Org.). Pesquisa social: teoria, método e criatividade. Rio de Janeiro, RJ: Vozes, 2009. p. 9-29.

MONTEIRO, Pedro Diogo. A ascensão do reconhecimento facial no Brasil, a fase digital da aplicação da lei e a ameaça às comunidades negras (e esforços para detê-la). In:

Inteligência Artificial e Discriminação Racial no Brasil: questões principais e recomendações. Belo Horizonte: Instituto de Referência em Internet e Sociedade, p. 28- 31, 2024

Disponível em: <https://irisbh.com.br/publicacoes/inteligencia-artificial-e-discriminacao-racial-no-brasil/>. Acesso em: 16 abr. 2025.

PEREIRA, Gabriela de Almeida. Desinformação e discurso de ódio como forma de legitimar a brutalidade e o assassinato de pessoas negras e pobres no Brasil. In: Inteligência Artificial e Discriminação Racial no Brasil: questões principais e recomendações. Belo Horizonte: Instituto de Referência em Internet e Sociedade, p. 15-18, 2024. Disponível em: <https://irisbh.com.br/publicacoes/inteligencia-artificial-e-discriminacao-racial-no-brasil/>. Acesso em: 16 abr. 2025.

PIOVESAN, Flavia. Direitos sociais, econômicos e culturais e direitos civis e políticos. **Revista internacional de direitos humanos**, São Paulo, v.1, n. 1, 2004, p. 20-47. DOI:

<https://doi.org/10.1590/S1806-64452004000100003>. Disponível em: <https://www.scielo.br/j/sur/a/vv3p3pQXYPv5dhH3sCLN46F/abstract/?lang=pt>. Acesso em: 01 mai. 2025.

RODRIGUES, Júlia Carvalho; CHAI, Cássius Guimarães. Inteligência artificial e algoritmo antirracismo: análise da neutralidade de dois algoritmos frente a episódios de violação de direitos em meus dispositivos digitais. **Revista Eletrônica [do] Tribunal Regional do Trabalho da 9ª Região**, Curitiba, v. 12, n. 118, p. 92-103, mar. 2023. Disponível em: <https://juslaboris.tst.jus.br/handle/20.500.12178/215797>. Acesso em: 8 jul. 2025.

SÁ-SILVA, Jackson Ronie; ALMEIDA, Cristóvão Domingos de; GUINDANI, Joel Felipe. Pesquisa documental: pistas teóricas e metodológicas. **Revista Brasileira de História e Ciências Sociais**, São Leopoldo, RS, v. 1, n. 1, p. 1-14, jul. 2009. Disponível em: <https://periodicos.furg.br/rbhcs/article/view/10351>. Acesso em: 5 abr. 2025

SILVA, Fernanda dos Santos Rodrigues; SILVA, Tarcízio (orgs). Inteligência Artificial e Discriminação Racial no Brasil: questões principais e recomendações. Belo Horizonte: Instituto de Referência em Internet e Sociedade, 2024. Disponível em: <https://irisbh.com.br/publicacoes/inteligencia-artificial-e-discriminacao-racial-no-brasil/>. Acesso em: 16 abr. 2025.

SILVA, Tarcízio. O racismo algorítmico é uma espécie de atualização do racismo estrutural. Centro de Estudos Estratégicos da Fiocruz - CEE-Fiocruz, 16 jun. 2021. Disponível em: <https://cee.fiocruz.br/?q=Tarcizio-Silva-O-racismo-algoritmico-e-uma-especie-de-atualizacao-do-racismo-estrutural>. Acesso em: 12 abr. 2025.

SILVA, Vanessa; LIMA, Thaise; BRASILEIRO, Fellipe. Algoritmos da branquitude: vieses e representações racistas em sistemas de inteligência artificial. **Lumina**, Juiz de Fora, v. 19, n. 1, p. 50-69, 30 abr. 2025. Disponível em: <https://periodicos.ufjf.br/index.php/lumina/article/view/46384/28857>. Acesso em: 20 jun. 2025.

SOUSA, Raquel et al. Esporte, Dados e Direitos: o uso de reconhecimento facial nos estádios brasileiros. Rio de Janeiro: Centro de Estudos de Segurança e Cidadania (CESeC); Observatório Social do Futebol/UERJ, 1 jan. 2024. Disponível em: <https://observatoriosocialfutebol.org/2024/01/01/esporte-dados-e-direitos-o-uso-de-reconhecimento-facial-nos-estadios-brasileiros/>. Acesso em: 16 abr. 2025.

TRANSPARÊNCIA BRASIL. Recomendações de Governança – Uso de Inteligência Artificial pelo Poder Público. São Paulo: Transparência Brasil, fev. 2020. Disponível em: https://www.transparencia.org.br/downloads/publicacoes/Recomendacoes_Governanca_Uso_IA_PoderPublico.pdf. Acesso em: 14 jul. 2025.

VERCELLI, Ariel. Regulaciones e inteligencias artificiales en Argentina. **In Mediaciones de la Comunicación**, Montevideu, v. 19, n. 1, p. 107-135, 2024. DOI: <https://doi.org/10.18861/ic.2024.19.1.3549>. Disponível em: <http://www.scielo.edu.uy/scielo.php?script=sciarttext&pid=S168886262024000100052>. Acesso em: 10 jul. 2025.