I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL I

ANTÔNIO CARLOS DINIZ MURTA

ANA ELIZABETH LAPA WANDERLEY CAVALCANTI

I61

Inteligência Artificial: Desafios da Era Digital I [Recurso eletrônico on-line] organização CONPEDI

Coordenadores: Ana Elizabeth Lapa Wanderley Cavalcanti, Antônio Carlos Diniz Murta. – Florianópolis: CONPEDI, 2025.

Inclui bibliografia
ISBN: 978-65-5274-095-3
Modo de acesso: www.conpedi.org.br em publicações
Tema: Inteligência Artificial e Sustentabilidade na Era Transnacional

1. Direito – Estudo e ensino (Pós-graduação) – Encontros Internacionais. 2. Inteligência Artificial. 3. Desafios da Era Digital. I International Experience Perugia – Itália. (1: 2025 : Perugia, Itália).

CDU: 34

Conselho Nacional de Pesquisa
e Pós-Graduação em Direito Florianópolis
Santa Catarina – Brasil
www.conpedi.org.br

I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA

INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL I

Apresentação

APRESENTAÇÃO DOS ARTIGOS

O Grupo de Trabalho INTELIGÊNCIA ARTIFICIAL: DESAFIOS DA ERA DIGITAL I teve seus trabalhos apresentados nas tardes dos dias 29 e 30 de maio de 2025, durante I INTERNATIONAL EXPERIENCE PERUGIA - ITÁLIA, realizado na cidade de Perugia – Itália, com o tema INTELIGÊNCIA ARTIFICIAL E SUSTENTABILIDADE NA ERA TRANSNACIONAL. Os trabalhos abaixo elencados compuseram o rol das apresentações.

INTELIGÊNCIA ARTIFICIAL: UM NOVO PARADIGMA PARA O PODER JUDICIÁRIO E A REVOLUÇÃO DA JUSTIÇA CONTEMPORÂNEA E DO FUTURO de Eunides Mendes Vieira: Este artigo propõe uma reflexão crítica sobre os impactos da IA no funcionamento da Justiça. Defende que a tecnologia pode reduzir a morosidade e aumentar a previsibilidade das decisões, mas alerta para riscos como viés algorítmico e perda da imparcialidade. Fundamentado em revisão bibliográfica, o texto propõe diretrizes éticas para a adoção da IA no Judiciário, com foco na manutenção dos direitos fundamentais e da equidade no tratamento processual.

A INTELIGÊNCIA ARTIFICIAL NOS TRIBUNAIS: REGULAÇÃO, DESAFIOS E ACCOUNTABILITY de Lais Gomes Bergstein, Douglas da Silva Garcia, Ingrid Kich Severo: O artigo analisa o impacto da inteligência artificial (IA) no Poder Judiciário, destacando sua introdução como mecanismo de automação e celeridade processual. Explora o programa Justiça 4.0 do CNJ, a Plataforma Digital do Poder Judiciário Brasileiro e os marcos regulatórios, como as Resoluções CNJ nº 332 e 335/2020. O texto problematiza a necessidade de governança, transparência e segurança jurídica, especialmente diante da terceirização tecnológica e do uso de dados em nuvem. Conclui-se que o uso da IA deve estar atrelado à ética e à accountability, com observância aos direitos fundamentais.

O USO DA INTELIGÊNCIA ARTIFICIAL NO DIREITO: HARD CASES de Maria de Fátima Dias Santana, Hélcia Macedo de Carvalho Diniz e Silva: O estudo analisa o uso da IA na resolução de hard cases à luz da teoria do Juiz Hércules de Ronald Dworkin. Argumenta que a IA pode contribuir para a celeridade e racionalidade das decisões, mas não substitui a

capacidade de ponderação e interpretação do julgador humano. Traz como exemplo o Projeto VICTOR do STF e propõe que a IA seja usada como instrumento auxiliar, preservando a dimensão humanística da Justiça.

INTELIGÊNCIA ARTIFICIAL E A TRADUÇÃO E GERAÇÃO DE TEXTOS JURÍDICOS de Vanessa Nunes Kaut, Bruno Vinícius Stoppa Carvalho: O texto discute a aplicação de modelos de linguagem (LLMs), como o ChatGPT, na geração e tradução de textos jurídicos. Ressalta o potencial de democratização da escrita jurídica, mas alerta para os riscos à confidencialidade, à autenticidade e à qualidade argumentativa. Aponta que, embora esses sistemas aumentem a produtividade, sua utilização exige regulação adequada, com limites éticos e respeito ao dever de sigilo profissional. O artigo sustenta a importância da supervisão humana e da criação de marcos regulatórios compatíveis com os princípios do Direito.

INTELIGÊNCIA ARTIFICIAL, FISCALIZAÇÃO E CONFORMIDADE TRIBUTÁRIA: DESAFIOS PARA A JUSTIÇA FISCAL de Alexandre Naoki Nishioka, Giulia Ramos Dalmazo: O texto investiga a aplicação da IA na detecção de fraudes fiscais e na conformidade tributária, evidenciando um paradoxo: o mesmo instrumento que fortalece o Fisco também é usado para planejamento tributário abusivo. Analisa a adoção de ferramentas como o SISAM e os desafios éticos e distributivos da automação fiscal. Conclui que é necessário criar estruturas de regulação que conciliem eficiência arrecadatória com justiça fiscal e responsabilidade social.

LIMITES DO CONSENTIMENTO PARENTAL NA PROTEÇÃO DA PRIVACIDADE DOS DADOS PESSOAIS DAS CRIANÇAS NA INTERNET de Gisele Gutierrez De Oliveira Albuquerque: Analisa os desafios jurídicos do consentimento parental no uso de dados de crianças em ambiente digital. Argumenta que a atuação dos pais deve respeitar o princípio do melhor interesse da criança e que o Estado pode e deve impor limites protetivos. Examina normas internacionais e nacionais e conclui pela necessidade de harmonização entre autonomia parental, inovação tecnológica e proteção da infância, principalmente no que tange à coleta e uso de dados pelas plataformas digitais.

PROTEÇÃO DE DADOS PESSOAIS DE CRIANÇAS E ADOLESCENTES E A INTELIGÊNCIA ARTIFICIAL: UM OLHAR SOB A PERSPECTIVA DA LEGISLAÇÃO BRASILEIRA de Ana Elizabeth Lapa Wanderley Cavalcanti, Patrícia Cristina Vasques De Souza Gorisch: Este artigo trata dos desafios específicos enfrentados na proteção de dados pessoais de crianças e adolescentes no contexto da IA e das redes digitais. Analisa a legislação brasileira, como a LGPD, o ECA e a Constituição Federal, destacando a centralidade do princípio do melhor interesse da criança. Argumenta que é necessário rever o

papel do consentimento parental frente à hipervulnerabilidade infantojuvenil e propõe medidas de educação digital, regulação e fiscalização mais efetivas, com foco na proteção integral desse grupo.

QUEM OLHA PELOS SEUS OLHOS? UMA ANÁLISE SOBRE A PROTEÇÃO DE DADOS E A PROVA DE PERSONALIDADE de Edith Maria Barbosa Ramos, Pedro Gonçalo Tavares Trovão do Rosário, Pastora Do Socorro Teixeira Leal: Explora a relação entre a proteção de dados pessoais e a noção de personalidade jurídica, especialmente no contexto da vigilância digital e do uso de IA. Retoma o debate sobre o direito à privacidade a partir de sua construção histórica e reforça que a proteção dos dados é expressão direta da dignidade da pessoa humana. A obra destaca o conceito de "prova de personalidade" como um novo paradigma jurídico, que busca assegurar o controle individual sobre as informações pessoais em tempos de capitalismo de dados.

PRECISAMOS FALAR SOBRE A DISCRIMINAÇÃO ALGORÍTMICA NAS RELAÇÕES DE CONSUMO de Dennis Verbicaro Soares, Loiane da Ponte Souza Prado Verbicaro: O texto aborda como algoritmos utilizados em plataformas digitais e ferramentas de IA têm reproduzido e intensificado práticas discriminatórias contra grupos vulneráveis. Explica que a predição comportamental, quando não supervisionada, pode resultar em decisões automatizadas excludentes, violando o princípio da isonomia. Propõe a criação de um Direito Antidiscriminatório aplicado à tecnologia, bem como a implementação de políticas públicas e marcos regulatórios que evitem a colonização algorítmica do consumidor e assegurem o respeito à dignidade nas relações de consumo.

PERSPECTIVAS E DESAFIOS À GOVERNANÇA TRANSNACIONAL DA INTERNET NA SOCIEDADE DIGITAL de Vanessa De Ramos Keller: O artigo propõe uma reflexão crítica sobre a ausência de uma governança global eficaz da internet. Defende que, em um mundo interconectado, não há mais espaço para ações unilaterais, sendo necessária a criação de um sistema de governança transnacional. Ressalta-se o papel das big techs e a necessidade de coordenação internacional para garantir direitos digitais, proteção de dados, liberdade de expressão e combate à desinformação. A obra argumenta que a sociedade digital demanda novos paradigmas jurídicos e políticos capazes de enfrentar os desafios da era informacional.

OS LIMITES BIOLÓGICOS E COGNITIVOS DA INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE SOBRE A SUSTENTABILIDADE INERENTE AOS IMPACTOS DA IA NA CAPACIDADE SÓCIO-COGNITIVA HUMANA de Aulus Eduardo Teixeira de Souza: Com abordagem interdisciplinar, o artigo discute as barreiras físicas, cognitivas e éticas que limitam a capacidade da inteligência artificial em simular a cognição humana. Contrapõe a

eficiência energética e adaptabilidade do cérebro humano com os altos custos computacionais e a rigidez dos sistemas de IA. Ressalta que a ausência de consciência subjetiva e de empatia torna a IA inadequada para decisões sensíveis. Conclui pela importância de reconhecer os limites biológicos da IA como base para um desenvolvimento tecnológico mais sustentável e responsável.

ORGANIZAÇÕES CRIMINOSAS: A IMPORTÂNCIA DA INTELIGÊNCIA ARTIFICIAL NO ENFRENTAMENTO DO CRIME ORGANIZADO deRoberta Priscila de Araújo Lima, Alice Arlinda Santos Sobral, Raylene Rodrigues De Sena: O estudo destaca o papel da inteligência artificial como aliada estratégica no combate ao crime organizado. Após um panorama da evolução normativa brasileira sobre o tema, especialmente com a Lei 12.850 /2013, o texto evidencia como a IA pode ser utilizada em ações policiais e de inteligência, facilitando a análise de grandes volumes de dados, identificando padrões e prevenindo crimes. A pesquisa conclui que o uso responsável e regulamentado da IA pode fortalecer a segurança pública e otimizar as ações de combate ao crime organizado, respeitando garantias legais e direitos fundamentais.

NEURODIREITOS E INTELIGÊNCIA ARTIFICIAL: MAPEAMENTO PROTETIVO DOS DIREITOS HUMANOS E FUNDAMENTAIS NA SOCIEDADE 4.0 de Simone Gomes Leal, Olivia Oliveira Guimarães: Explora o conceito de neurodireitos como nova categoria de direitos humanos frente à interface entre IA e neurotecnologia. Destaca os riscos à dignidade humana, à identidade e à privacidade mental causados por tecnologias que acessam ou modulam o cérebro. Enfatiza o papel do constitucionalismo digital na proteção desses direitos, propondo sua positivação nas legislações nacionais e internacionais como forma de preservar a integridade do sujeito frente à máquina.

VIESES ALGORÍTMICOS E RECONHECIMENTO FACIAL de Pedro Henrique do Prado Haram Colucci, Sergio Nojiri: Analisa o caso do Projeto Vídeo-Polícia Expansão, implantado na Bahia, e seus efeitos discriminatórios. O artigo mostra como sistemas de reconhecimento facial produzem falsos positivos, especialmente contra pessoas negras, e denuncia a ausência de regulamentação e de auditorias obrigatórias. Propõe modelos internacionais para nortear a regulação brasileira.

IA NA GESTÃO MIGRATÓRIA: INCLUSÃO DIGITAL OU FERRAMENTA DE EXCLUSÃO? de Patricia Cristina Vasques De Souza Gorisch, Ana Elizabeth Lapa Wanderley Cavalcanti: Examina a crescente utilização da IA em políticas migratórias, como triagem de pedidos de refúgio, monitoramento de fronteiras e identificação de migrantes. Denuncia que, embora a tecnologia possa facilitar o acesso a serviços, também é usada para

vigilância e exclusão de grupos vulneráveis. O texto propõe uma regulação ética e baseada nos direitos humanos.

A CIDADANIA ELETRÔNICA DO HOMO DIGITALIS: PERSPECTIVAS JURÍDICAS À LUZ DO REGULAMENTO EU 2024/1689 SOBRE INTELIGÊNCIA ARTIFICIAL de Olivia Oliveira Guimarães, Helen Caroline Cardoso Santos, Lucas Gonçalves da Silva: Trabalha a Inteligência Artificial sob o aspecto da regulação europeia, tendo como base a questão da cidadania digital.

DECISÕES AUTOMATIZADAS E COGNIÇÃO HUMANA: O PAPEL DA INTELIGÊNCIA ARTIFICIAL NO PROCESSO DECISÓRIO JUDICIAL de Sergio Nojiri, Luiz Guilherme da Silva Rangel: Tratando de questões atinentes ao uso da Inteligência Artificial em decisões judiciais.

TRANSAÇÃO NA REFORMA TRIBUTÁRIA COMO MEDIDA DE DESJUDICIALIZAÇÃO de Tammara Drumond Mendes, Antônio Carlos Diniz Murta, Renata Apolinário de Castro Lima.

VEDAÇÃO AO CONFISCO DA PROPRIEDADE ÚNICA QUE ATENDE A FUNÇÃO SOCIAL de Tammara Drumond Mendes, Antônio Carlos Diniz Murta, Renata Apolinário de Castro Lima.

Após duas tardes de intensos debates sobre os temas apresentados, foram encerrados os trabalhos do GT com a elaboração de uma síntese que se chamou de Carta de Perúgia.

Os temas demonstram a abrangência e amplitude do tema que é de grande interesse da ciência jurídica e que permite uma profícua produção acadêmica nacional e internacional. Importante lembrar que os pesquisadores presentes no GT estão vinculados aos mais diversos programas de pós-graduação em Direito, demonstrando a importância de debates como os ocorridos nos dias 28, 29 e 30 de maio de 2025, na cidade de Perúgia – Itália.

Nota-se preocupação de todos quanto à regulação da Inteligência artificial, mormente para que não só, numa visão meramente apocalíptica, se torne um instrumento de maior concentração de poder nas mãos de grandes grupos - big techs - e manipulação comportamental, mas também não possa ser a médio prazo um elemento que possa reduzir a liberdade e autonomia humana no pensar e evoluir seja em questões técnicas seja em questões sociais/filosóficas. Não existem dúvidas que enfrentamos uma nova realidade sem embargo de ser virtual e não materializada que vai exigir da comunidade internacional ou de

cada um de nós adequação para um fenômeno que não pode ser impedido; mas pode ser, a partir de um maior aprofundamento sobre seu poder e efeitos na sociedade, melhor assimilado sem que percamos, sendo otimista, o que nos torna humanos.

Diante da diversidade de temas e das pesquisas de grande qualidade apresentadas neste evento, recomendamos que operadores do direito em todas as suas funções leiam os trabalhos aqui apresentados.

Coordenadores:

Antônio Carlos Diniz Murta

Universidade FUMEC

acmurta@fumec.br

Ana Elizabeth Lapa Wanderley Cavalcanti

Universidade Presbiteriana Mackenzie

ana.cavalcanti@mackenzie.br

VIESES ALGORÍTMICOS, ESTRUTURAS DE DISCRIMINAÇÃO E FALSOS POSITIVOS EM SISTEMAS DE RECONHECIMENTO FACIAL: O CASO DO PROJETO VÍDEO-POLÍCIA EXPANSÃO

ALGORITHMIC BIASES, DISCRIMINATION STRUCTURES AND FALSE POSITIVES IN FACIAL RECOGNITION SYSTEMS: THE CASE OF THE VIDEO-POLICE EXPANSION PROJECT

Pedro Henrique do Prado Haram Colucci [1]
Sergio Nojiri [2]

Resumo

O presente artigo analisa o Projeto Vídeo-Polícia Expansão do Governo do Estado da Bahia como laboratório para a implementação de câmeras de reconhecimento facial com inteligência artificial no Brasil. Trata-se de uma pesquisa qualitativa que emprega estudo de caso para examinar as implicações práticas desta tecnologia. O trabalho busca descontruir o mito da neutralidade tecnológica, demonstrando como algoritmos podem incorporar e amplificar preconceitos estruturais, perpetuando discriminações raciais e de gênero. Evidencia que vieses algorítmicos não são falhas técnicas, mas resultados diretos de dados de treinamento que refletem desigualdades históricas, tornando-se especialmente problemáticos quando aplicados na segurança pública. O estudo destaca a preocupante falta de transparência sobre falsos positivos neste contexto de vigilância em massa, onde cidadãos inocentes são incorretamente identificados como suspeitos de crimes e foragidos. A análise também examina como estas tecnologias podem reforçar práticas discriminatórias já existentes no sistema de justiça criminal. Conclui-se que, para evitar que tecnologias de inteligência artificial intensifiquem a seletividade penal e violações de direitos fundamentais, é essencial estabelecer diretrizes éticas rigorosas e mecanismos efetivos de accountability que reconheçam as profundas implicações sociais destes sistemas tecnológicos.

Palavras-chave: Reconhecimento facial, Inteligência artificial, Viés algorítmico, Segurança pública, Direitos fundamentais

Abstract/Resumen/Résumé

This article analyzes the Video-Police Expansion Project of the Bahia State Government as a comprehensive laboratory for facial recognition cameras with artificial intelligence in Brazil. It is a qualitative research that employs a detailed case study to examine the practical implications of this technology. It deconstructs the myth of technological neutrality,

[1] Mestrando pela Faculdade de Direito de Ribeirão Preto – Universidade de São Paulo (FDRP-USP). Bolsista CAPES. Bacharel em Direito pela Faculdade de Direito de Franca (FDF)

[2] Graduação, Mestrado e Doutorado em Direito pela Pontifícia Universidade Católica de São Paulo. Livre-Docente e Professor pela Faculdade de Direito de Ribeirão Preto (USP). Juiz Federal.

demonstrating how algorithms can incorporate and amplify structural prejudices, perpetuating racial and gender discrimination. It shows that algorithmic biases are not technical failures, but direct results of training data that reflect historical inequalities, becoming especially problematic when applied to public security. The study highlights the concerning lack of transparency about false positives in this mass surveillance context, where innocent citizens are incorrectly identified as suspects. It concludes that, to prevent artificial intelligence technologies from intensifying penal selectivity and rights violations, it is essential to establish rigorous ethical guidelines and accountability mechanisms that recognize the profound social implications of these technological systems.

Keywords/Palabras-claves/Mots-clés: Facial recognition, Artificial intelligence, Algorithmic bias, Public security, Fundamental rights

1 INTRODUÇÃO

O reconhecimento facial a partir da Inteligência Artificial (IA) tem se consolidado como uma tecnologia difundida em diversas aplicações contemporâneas, abrangendo áreas que vão desde a segurança pública até campanhas personalizadas de marketing. Embora a sua popularização represente avanços tecnológicos significativos, ela também suscita questões sensíveis relacionadas à composição dos bancos de dados e sua mobilização, especialmente no que tange aos vieses algorítmicos. Estes vieses, frequentemente subestimados, podem resultar em discriminação racial, de gênero e outras formas de desigualdade, gerando preocupações relativas à confiabilidade e transparência dos sistemas de reconhecimento facial.

No Brasil, a tecnologia está sendo empregada para fins de atividades de policiamento em estados como São Paulo, Rio de Janeiro e Bahia. Este artigo se propõe a realizar um estudo de caso sobre o Projeto Vídeo-Polícia Expansão implementado pelo Governo do Estado da Bahia, iniciado em 2019. O projeto, que integra o reconhecimento facial como uma ferramenta de videomonitoramento, tem como objetivo auxiliar nas atividades de segurança pública do estado.

No entanto, a implementação dessa tecnologia levanta questões críticas sobre suas consequências sociais e éticas. O uso do reconhecimento facial, embora apresente um interesse público aparente, também suscita preocupações relacionadas a liberdades individuais, legalidade e a possibilidade de discriminação, especialmente considerando os riscos de vieses algorítmicos associados à sua aplicação. Este estudo busca analisar esses aspectos críticos, contribuindo para um entendimento sobre a utilização de inteligência artificial no âmbito da segurança pública.

O trabalho apresenta uma análise qualitativa do Projeto Vídeo-Polícia Expansão por meio da metodologia de estudo de caso, visando investigar "[...] um fenômeno contemporâneo dentro de seu contexto da vida real, especialmente quando os limites entre o fenômeno e o contexto não estão claramente definidos [...]" (Yin, 2001, p. 32). Assim, permite-se uma compreensão das suas implicações e desafios no contexto de aplicação de aparatos de inteligência artificial na segurança pública. A revisão bibliográfica realizada fundamenta a discussão, situando o projeto dentro do marco teórico sobre tecnologia, perfilamento racial e vieses algorítmicos.

2 A EUFORIA DA IA E O VÁCUO REGULATÓRIO NA SEGURANÇA PÚBLICA

A inteligência artificial tem gerado uma onda de euforia sem precedentes em diversos setores da sociedade. Empresas, governos e instituições acadêmicas celebram as possibilidades transformadoras que as tecnologias de IA prometem, desde a otimização de processos até a criação de soluções para problemas complexos. Este entusiasmo desenfreado tem impulsionado investimentos bilionários e uma corrida tecnológica global para dominar o campo (Tajra, 2024).

No entanto, quando observamos a aplicação da IA na segurança pública, deparamo-nos com um cenário preocupante. Sistemas de reconhecimento facial e ferramentas de vigilância em massa estão sendo implementados em um ritmo acelerado, muitas vezes sem a devida avaliação de seus impactos sociais. Esta adoção precipitada ocorre em um vácuo regulatório alarmante, onde a ausência de marcos legais claros permite que tecnologias com profundas implicações para direitos fundamentais sejam utilizadas sem supervisão adequada (Nunes, 2022).

A falta de regulamentação específica cria um ambiente onde as decisões sobre como, quando e onde utilizar IA na segurança pública ficam exclusivamente nas mãos de corporações privadas ou de gestores públicos que podem não estar devidamente capacitados para avaliar as consequências éticas e sociais destas tecnologias. Questões sensíveis como vieses algorítmicos, privacidade de dados, transparência e responsabilização permanecem sem respostas claras, enquanto sistemas cada vez mais sofisticados são implementados no policiamento e na vigilância.

Este descompasso entre o avanço tecnológico e o desenvolvimento de estruturas regulatórias adequadas representa um desafio significativo para sociedades democráticas. Sem balizas legais que definam limites claros e estabeleçam mecanismos de controle social, corre-se o risco de que a IA, em vez de contribuir para uma segurança pública mais eficiente, torne-se um instrumento de perpetuação de desigualdades e violações de direitos.

3 DOS ATALHOS COGNITIVOS AOS VIESES ALGORÍTMICOS

Os vieses cognitivos são armadilhas mentais que ocorrem quando processamos e interpretamos informações do mundo ao nosso redor. Presentes em todos os seres humanos,

esses vieses funcionam como atalhos que nos ajudam a tomar decisões rápidas, mas frequentemente nos levam a conclusões imprecisas ou injustas. Eles operam em grande parte como pontos-cegos da cognição, influenciando nossas percepções e decisões sem que tenhamos plena consciência de sua existência (Tversky; Kahneman, 1974).

Nossa percepção de mundo é profundamente moldada por esses eventos mentais. O viés de confirmação, por exemplo, nos leva a dar mais atenção e credibilidade a informações que confirmam nossas crenças preexistentes, enquanto ignoramos ou desvalorizamos dados contraditórios. Já o viés de grupo nos faz favorecer pessoas que identificamos como semelhantes a nós, podendo resultar em preconceitos contra aqueles que percebemos como diferentes. Por meio desses atalhos, os vieses agem como filtros invisíveis que distorcem nossa compreensão da realidade (Wojciechowski; Morais da Rosa, 2021).

Em contextos sociais, os processos de enviesamento podem se manifestar como estereótipos e preconceitos relacionados a características como raça, gênero, idade ou classe social. Tais distorções perceptivas influenciam desde interações cotidianas até decisões institucionais, contribuindo para a perpetuação de desigualdades estruturais. Mesmo pessoas que conscientemente rejeitam preconceitos podem, inadvertidamente, agir com base em associações implícitas formadas ao longo de uma vida de exposição a representações enviesadas na mídia e na cultura.

Com a irrupção da inteligência artificial, especialmente dos sistemas baseados em aprendizado de máquina, ocorreu um fenômeno de transposição: nossas tecnologias mais avançadas começaram a reproduzir fenômenos parecidos com os atalhos cognitivos que os indivíduos experienciam, gerando os vieses algorítmicos. Isso acontece por conta do banco de dados que é abastecido e treinado com dados gerados por humanos, carregando para dentro dos *datasets* distorções e preconceitos presentes em nossa percepção coletiva (Leslie, 2020).

Algoritmos são conjuntos de instruções bem definidas, que seguem uma sequência lógica para resolver problemas ou executar tarefas específicas. No contexto computacional, algoritmos traduzem-se em códigos que orientam o funcionamento de programas e sistemas, determinando como os dados são processados, analisados e transformados (Simões-Gomes; Roberto; Mendonça, 2020). Sua importância cresceu exponencialmente na era digital, sustentando a operação de praticamente todas as tecnologias com as quais interagimos diariamente, desde ferramentas de busca e redes sociais até sistemas bancários.

A taxonomia de vieses algorítmicos proposta por David Danks e Alex John London (2017) oferece uma estrutura concisa para expor e categorizar os diferentes tipos de vieses que podem surgir no desenvolvimento e aplicação de algoritmos em sistemas autônomos. Essa

taxonomia identifica cinco principais fontes de viés: Viés de Dados de Treinamento, Viés de Foco Algorítmico, Viés de Processamento Algorítmico, Viés de Contexto de Transferência e Viés de Interpretação. Cada um desses vieses pode surgir em diferentes etapas do ciclo de vida de um algoritmo, desde a coleta e preparação dos dados até a implementação e interpretação dos resultados.

Tabela 1: Taxonomia de vieses algorítmicos de Danks e London

Tipo de Viés	Ocorrência	Efeitos
Viés de Dados de Treinamento	Dados de treinamento não representativos ou enviesados.	Modelos que refletem desvios dos dados, levando a generalizações incorretas.
Viés de Foco Algorítmico	Uso ou exclusão deliberada de certas informações no algoritmo.	Desvios em relação a padrões morais, legais ou estatísticos.
Viés de Processamento Algorítmico	Escolha de algoritmos ou estimadores estatisticamente enviesados.	Compensação de outros vieses, mas introdução de novos desvios estatísticos.
Viés de Contexto de Transferência	Uso do algoritmo fora do contexto pretendido.	Desempenho enviesado em novos contextos, levando a erros ou injustiças.
Viés de Interpretação	Má interpretação das saídas do algoritmo pelo usuário ou sistema.	Decisões ou ações baseadas em informações incorretas ou mal compreendidas.

Fonte: Danks; London, 2017.

Por sua vez, os algoritmos de reconhecimento facial são sistemas especializados que identificam ou verificam a identidade de pessoas através da análise de suas características faciais. Estes algoritmos operam em etapas distintas: (I) detectam a presença de um rosto em uma imagem ou vídeo; (II) analisam a geometria facial; e (III) mepeiam pontos-chave como a distância entre os olhos, a largura do nariz, o contorno dos lábios e a profundidade das cavidades oculares (Wechsler, 2007).

Estas medidas são convertidas em uma assinatura facial, isto é, uma representação matemática única do rosto de cada indivíduo. Finalmente, esta assinatura é comparada com um banco de dados de rostos previamente registrados para determinar correspondências. A precisão destes sistemas depende fortemente da qualidade das imagens, das condições de iluminação e dos ângulos de captura, além da diversidade e representatividade dos dados de treinamento utilizados para desenvolver o algoritmo.

A inteligência artificial desempenha um papel central neste processo, principalmente através das técnicas de aprendizado profundo (*deep learning*). Diferentemente de algoritmos convencionais programados com regras fixas, os sistemas de IA para reconhecimento facial aprendem a identificar padrões faciais a partir da exposição a milhões de imagens durante seu treinamento. Esta capacidade de aprendizado permite que estes sistemas reconheçam rostos mesmo em condições variáveis ou quando certas características mudam, como o envelhecimento ou presença de acessórios. A IA também possibilita que estes sistemas melhorem continuamente seu desempenho com a exposição a mais dados, refinando sua precisão ao longo do tempo.

No entanto, o papel da IA nesta trama é contraditório. Se por um lado ela potencializa a eficácia dos sistemas de reconhecimento facial, por outro pode amplificar problemas estruturais presentes nos dados com os quais é treinada. Os algoritmos de IA não possuem compreensão contextual ou ética intrínseca, eles simplesmente detectam padrões nos dados que recebem. Quando estes dados refletem disparidades históricas e sociais, como a sub-representação de certos grupos étnicos ou vieses na forma como diferentes populações são fotografadas, a IA incorpora e até amplifica estas distorções.

Longe de serem entidades neutras e objetivas, os sistemas de IA tendem a refletir e reforçar padrões discriminatórios existentes nos dados de treinamento, isto é, essa tecnologia "[...] é um produto social, na medida em que a sociedade interfere nas suas condições de produção e circulação, mas também na constituição dos valores e subjetividades dos agentes que as produzem" (Simões-Gomes; Roberto; Mendonça, 2020, p. 157).

Os vieses algorítmicos referem-se a distorções ou preconceitos que podem surgir em sistemas algorítmicos de inteligência artificial, resultando na perpetuação de preconceitos e discriminações históricas (Pires; Cavagnoli; Cotello; Visani; Gongora, 2021). Esses vieses podem se manifestar de várias maneiras, como a seleção de dados de treinamento que não representam adequadamente a diversidade da população, levando a resultados que favorecem um grupo em detrimento de outros (Schuler; Montardo, 2020).

Segundo Alexandre Morais da Rosa e Paola Wojciechowski (2021, p. 100), "[...] não se trata de um erro do algoritmo e sim do modelo decorrente, em que o data set utilizado para treinar a máquina, já nascia enviesado". Dessa forma, um fator central na perpetuação de vieses algorítmicos é a qualidade dos dados utilizados para treinar os modelos. Se esses dados contêm preconceitos históricos ou refletem desigualdades sociais existentes, o algoritmo pode simplesmente reproduzir e amplificar esses vieses. Isso é particularmente preocupante na área da segurança pública, onde decisões automatizadas podem intensificar significativamente processos de seletividade penal (Monteiro, 2022).

Nos sistemas de reconhecimento facial, esta problemática assume dimensões particularmente preocupantes. Estudos demonstram consistentemente que muitas dessas tecnologias apresentam taxas de erro significativamente mais altas quando analisam rostos de pessoas negras, especialmente mulheres negras, em comparação com homens de pele clara (Buolamwini; Gebru, 2018). Essa disparidade não é acidental, mas resultado direto de conjuntos de treinamento historicamente enviesados que informam os bancos de dados dos sistemas de reconhecimento facial.

Os falsos reconhecimentos resultantes destes processos de enviesamentos tecnológicos podem ter consequências devastadoras quando esses sistemas são implementados no contexto da segurança pública. Casos documentados de pessoas erroneamente identificadas e detidas com base em correspondências incorretas de reconhecimento facial ilustram os perigos reais da implementação desse tipo de tecnologia em um contexto de inexistência regulatória. Um falso positivo pode resultar em acusações indevidas, detenções injustificadas e violações graves de direitos civis, afetando desproporcionalmente grupos já marginalizados (Nunes; Lima; Cruz, 2023).

4 FALSOS POSITIVOS E DÉFICITS DE TRANSPARÊNCIA NA UTILIZAÇÃO DE TECNOLOGIAS DE RECONHECIMENTO FACIAL

Quando um algoritmo é utilizado como instrumento da persecução penal e identifica incorretamente uma pessoa como suspeita, não estamos diante de um simples erro técnico, mas de uma falha que pode dar causa a um erro judiciário, produzindo a condenação de inocentes. Este cenário torna-se ainda mais crítico quando observamos que essas tecnologias estão sendo implementadas em sistemas policiais e de vigilância que já carregam um histórico documentado de seletividade penal (Jefferson, 2020).

Nos Estados Unidos, por exemplo, as práticas de policiamento têm demonstrado padrões consistentes de desigualdade no tratamento de minorias étnicas e raciais. A inserção de ferramentas de reconhecimento facial neste contexto não ocorre em um vácuo social ou institucional, mas em um terreno já marcado por desequilíbrios estruturais que afetam desproporcionalmente pessoas não-brancas (Benjamin, 2019).

Os padrões preexistentes de desigualdade racial no sistema de justiça criminal elevam consideravelmente os riscos associados à adoção dessas novas tecnologias. Para comunidades negras e outras minorias raciais, que já enfrentam maiores taxas de abordagens policiais, buscas, detenções e encarceramento, a implementação de sistemas de reconhecimento facial potencialmente enviesados representa uma nova camada de vulnerabilidade. O fato de que muitos algoritmos demonstram taxas de erro significativamente mais altas quando analisam rostos de pessoas negras amplifica essa preocupação, criando um ciclo de retroalimentação tecnológica para disparidades já estabelecidas (Monteiro, 2022; Silva, 2022).

O que torna esta questão particularmente alarmante é o potencial de institucionalização e amplificação tecnológica de discriminações já existentes. Quando comunidades já sujeitas a maior vigilância policial são também aquelas mais propensas a serem incorretamente identificadas por sistemas automatizados, estabelece-se um ciclo de discriminação retroalimentado e catalisado pela mediação tecnológica. A aura de objetividade científica que envolve a tecnologia de IA e as promessas de eficiência no controle do crime podem mascarar a natureza fundamentalmente social e política das decisões tomadas durante seu desenvolvimento e implementação.

Carvalho Monteiro (2022, p. 95) propõe o conceito de "dupla opacidade", referindo-se à maneira como os discursos dominantes ocultam as dimensões sociais das tecnologias e silenciam debates essenciais sobre questões raciais, mesmo quando suas aplicações evidenciam

claramente essas problemáticas. Dessa forma, as tecnologias de reconhecimento facial devem ser racionalizadas para além do mito da neutralidade tecnológica, isto é, que sugere que algoritmos, por serem baseados em lógicas matemáticas, operam em uma dimensão isenta de questões raciais ou de gênero (Nunes, 2022). Nesse sentido:

> À medida que produtos e sistemas algorítmicos implantados se tornam mais comuns e seus impactos prejudiciais mais visíveis, os esforços para auditá-los tornaram-se cada vez mais convencionais. Uma variedade de indivíduos e organizações agora conduz auditorias algorítmicas de produtos que vão desde mecanismos de recomendação para contratação até modelos de reconhecimento facial, e a auditoria algorítmica emergiu como uma das abordagens mais populares para a responsabilidade algorítmica. Entidades que oferecem serviços de auditoria também proliferaram, mesmo enquanto os processos de auditoria permanecem não padronizados e pouco compreendidos (Constanza-Chock; Harvey; Raji; Czernuszenko; Buolamwini, 2023, p. 1, tradução livre).

Assim, a discussão sobre vieses algorítmicos não se limita apenas à correção técnica, mas também envolve considerações sobre como as tecnologias são desenvolvidas, quem as controla e como são implementadas na sociedade, demandando uma abordagem sobre o uso da inteligência artificial consciencioso sobre seus potenciais de agravamento de desigualdades (Silva, 2022).

4.1 O CASO PROJETO VÍDEO-POLÍCIA EXPANSÃO

Em 2019, a Secretaria de Segurança Pública da Bahia (SSP/BA) lançou o Projeto Vídeo-Polícia Expansão em parceria com a Huawei, visando implementar serviços de monitoramento, pontos de imagem e comunicação crítica com banda larga. Antes da colaboração, existiam mais de 1.900 câmeras sem integração. A nova tecnologia centraliza a gestão de informações, além de adicionar 300 câmeras e licenças de *softwares* de reconhecimento facial, como o *VideoCloud* (Pires; Cavagnoli; Cotello; Visani; Gongora, 2021).

Durante eventos como o Carnaval de Salvador e a Micareta de Feira de Santana, a SSP/BA utilizou extensivamente tecnologias de reconhecimento facial para controle e monitoramento das áreas de grande circulação (Monteiro, 2022). O uso do reconhecimento facial em tamanha escala levanta preocupações sobre falsos positivos que podem levar a violações de direitos.

Os falsos positivos ocorrem quando a tecnologia confunde cidadãos inocentes com criminosos, resultando em prisões indevidas. Apesar de 207 prisões terem sido realizadas com o auxílio desta tecnologia, a SSP/BA não divulga dados sobre quantas foram incorretas, dificultando uma avaliação efetiva do projeto (Pires; Cavagnoli; Cotello; Visani; Gongora; 2021).

Casos como o de uma mãe e seu filho abordados violentamente em setembro de 2019 por conta de um reconhecimento facial equivocado destacam falhas no sistema (Palma; Pacheco, 2020). Além disso, grandes eventos geraram centenas de alertas, mas poucas prisões, levantando questões sobre a confiabilidade da tecnologia (Monteiro, 2022). O alto índice de identificações errôneas exige uma reflexão crítica sobre a real eficácia do reconhecimento facial na segurança pública.

O uso de tecnologias de reconhecimento facial como ferramenta de segurança pública tem se expandido rapidamente pelo Brasil, revelando um cenário preocupante em termos de implementação e consequências. O levantamento junto às secretarias estaduais de Segurança demonstra que quatro estados brasileiros já efetuaram mais de 1.700 prisões utilizando essa tecnologia, sendo a Bahia responsável por aproximadamente 90% desse número. Desde 2019, quando implementou o sistema durante o Carnaval, o estado baiano contabiliza 1.547 detenções realizadas com auxílio do reconhecimento facial, evidenciando a intensidade com que essa ferramenta tem sido empregada nas práticas de policiamento (Tajra, 2024).

O aspecto mais alarmante desse cenário é a ausência de uma legislação específica que regulamente a utilização dessas tecnologias. Os estados brasileiros têm investido volumes significativos de recursos públicos em sistemas que operam num vácuo jurídico, sem parâmetros claros sobre limites, responsabilidades ou mecanismos de prestação de contas. Essa situação cria um ambiente propício para abusos e violações de direitos, onde práticas potencialmente discriminatórias podem se normalizar sob o argumento da eficiência na segurança pública (Nunes, 2023).

Agravando essa problemática está a comprovada tendência desses sistemas a produzirem falsos positivos, especialmente quando analisam a população negra. No contexto brasileiro, essa questão torna-se ainda mais crítica devido à utilização do Banco Nacional de Mandados de Prisão como referência para as identificações. Este banco é majoritariamente composto por pessoas negras, refletindo diretamente as desigualdades e o racismo estrutural presentes no sistema de justiça criminal brasileiro (Nunes; Lima; Cruz, 2023). Cria-se, assim, um ciclo de retroalimentação de produção de dano: um algoritmo já enviesado opera sobre um banco de dados que igualmente carrega vieses de treinamento, potencializando

exponencialmente o risco de falsas identificações e abordagens injustificadas direcionadas a pessoas negras.

A própria Lei Geral de Proteção de Dados (LGPD), que poderia oferecer algum amparo nesse cenário, contém uma exceção significativa em seu Art. 4º, III, a, ao prever sua não aplicação quando se trata do tratamento de dados pessoais exclusivamente para fins de segurança pública[1]. Esta lacuna deixa cidadãos particularmente vulneráveis justamente em interações com instituições que detêm poder coercitivo considerável. Na prática, o Brasil carece de regulamentação específica para o tratamento de dados biométricos, especialmente no que concerne ao uso de tecnologias de reconhecimento facial por forças policiais.

Este cenário demanda atenção urgente por parte de legisladores, juristas e da sociedade civil. A implementação dessas tecnologias sem o devido arcabouço legal e sem mecanismos robustos de supervisão representa não apenas um risco para liberdades individuais, mas também pode agravar desigualdades sociais já profundamente enraizadas. É fundamental que qualquer regulamentação futura aborde diretamente a questão dos vieses algorítmicos, estabeleça protocolos claros para teste e validação desses sistemas antes de sua implementação, e crie mecanismos transparentes de auditoria e responsabilização por identificações incorretas que resultem em abordagens policiais ou detenções injustificadas.

O caso baiano ilustra, portanto, como a introdução de tecnologias de reconhecimento facial orientadas por inteligência artificial no Brasil já nasceu marcada por problemas fundamentais: investimentos vultosos sem regulamentação adequada, implementação acelerada sem mecanismos de controle, e resistência à transparência e ao escrutínio público. O projeto Vídeo-Polícia Expansão, ao funcionar como laboratório nacional dessas tecnologias, não apenas normalizou estas práticas problemáticas, mas estabeleceu precedentes que podem comprometer o desenvolvimento ético e responsável desses sistemas em todo o país.

A relação entre algoritmos, reconhecimento facial e IA constitui, portanto, um território de tensões entre avanço tecnológico e implicações éticas. Os casos de falsos positivos em projetos como o Vídeo-Polícia Expansão da Bahia não representam meras falhas técnicas isoladas, mas sintomas de questões mais profundas sobre representatividade, poder e justiça em uma sociedade cada vez mais mediada por tecnologias automatizadas.

À medida que estas tecnologias se expandem para áreas sensíveis como segurança pública, controle de acesso e vigilância, torna-se imperativo desenvolver não apenas algoritmos mais precisos, mas também estruturas de governança que garantam transparência, prestação de

[1] Art. 4º Esta Lei não se aplica ao tratamento de dados pessoais: [...] III - realizado para fins exclusivos de: a) segurança pública [...] (Brasil, 2018).

contas e proteção contra usos discriminatórios. O verdadeiro desafio não está apenas em aperfeiçoar a tecnologia, mas em assegurar que seu desenvolvimento e implementação ocorram dentro de parâmetros éticos que respeitem a dignidade e os direitos de todos os indivíduos.

5 MODELOS DE REGULAÇÃO E GOVERNANÇA ALGORÍTMICA PARA TECNOLOGIAS DE RECONHECIMENTO FACIAL

Diante dos desafios éticos e técnicos apresentados pela implementação de tecnologias de reconhecimento facial no âmbito da segurança pública, examina-se os diferentes modelos de regulação que estão sendo desenvolvidos globalmente. Estes modelos oferecem perspectivas distintas sobre como equilibrar inovação tecnológica com proteção de direitos fundamentais e podem fornecer diretrizes importantes para o desenvolvimento de um arcabouço regulatório brasileiro.

A União Europeia tem liderado esforços globais nesta área com a proposta do *Artificial Intelligence Act*, que estabelece uma abordagem baseada em risco para a regulação de sistemas de IA. Neste modelo, tecnologias de reconhecimento facial em espaços públicos para fins de identificação são classificadas como de alto risco, exigindo avaliações de impacto, transparência sobre seu funcionamento, supervisão humana significativa e demonstração de conformidade antes da implementação (Veale; Zuiderveen Borgesius, 2021). Esta abordagem equilibra a permissão para o desenvolvimento tecnológico com salvaguardas contra abusos.

Em contraste, cidades como São Francisco e Boston nos Estados Unidos baniram o uso de reconhecimento facial por agências governamentais, refletindo uma abordagem de precaução que prioriza a prevenção de danos potenciais sobre os benefícios alegados. Estas proibições foram motivadas por evidências crescentes sobre disparidades raciais no desempenho desses sistemas e preocupações sobre vigilância em massa (Richardson; Schultz; Crawford, 2019). A abordagem restritiva argumenta que, até que questões fundamentais de equidade e precisão sejam adequadamente resolvidas, o uso governamental dessas tecnologias representa um risco inaceitável para comunidades vulneráveis.

Já o modelo canadense tem enfatizado a governança colaborativa, criando espaços de diálogo entre desenvolvedores, usuários e reguladores para estabelecer padrões éticos e técnicos. A Diretiva sobre Tomada de Decisão Automatizada do Canadá requer que sistemas de IA utilizados pelo governo federal passem por avaliações de impacto algorítmico, estabeleçam

mecanismos claros de recurso e mantenham transparência sobre o funcionamento do sistema (Government of Canada, 2021). Este modelo reconhece que regulação eficaz requer participação ativa de múltiplos *stakeholders* e adaptação contínua às tecnologias emergentes.

No contexto brasileiro, a fragmentação atual da governança de tecnologias de reconhecimento facial — onde cada estado ou município implementa sistemas sem um quadro regulatório comum — cria um cenário de insegurança jurídica e riscos de violações de direitos. Um modelo regulatório eficaz para o Brasil precisaria considerar as particularidades do contexto nacional, especialmente no que se refere às desigualdades raciais estruturais e ao histórico de violência policial, que agravam os riscos associados ao reconhecimento facial (Nunes, 2022).

Uma abordagem promissora poderia combinar elementos do modelo europeu baseado em risco com o foco canadense em governança colaborativa, adaptados à realidade brasileira. Isso incluiria: (I) classificação do reconhecimento facial para segurança pública como tecnologia de alto risco, exigindo avaliações de impacto obrigatórias; (II) estabelecimento de padrões técnicos mínimos de precisão, com requisitos mais rigorosos para igualdade de desempenho entre diferentes grupos demográficos; (III) criação de mecanismos de supervisão independentes com participação da sociedade civil; (IV) transparência obrigatória sobre métodos, limitações e resultados; e (V) responsabilização clara por erros e danos causados pelo uso da tecnologia.

Uma dimensão fundamental frequentemente negligenciada nas discussões regulatórias é a exigência de diversidade nas equipes que desenvolvem essas tecnologias. A homogeneidade nos times de desenvolvedores de IA tem sido identificada como um fator que contribui para a persistência de vieses algorítmicos (West; Whittaker; Crawford, 2019). No caso brasileiro, onde questões de representatividade racial nas áreas de tecnologia são particularmente agudas, políticas que incentivem a diversidade nas equipes responsáveis pelo desenvolvimento e implementação de sistemas de reconhecimento facial poderiam contribuir para mitigar vieses desde as fases iniciais de concepção.

Outra consideração relevante é a necessidade de estabelecer processos de avaliação contínua após a implementação. A eficácia e os impactos sociais dessas tecnologias não podem ser completamente previstos antes de sua aplicação no mundo real. Portanto, um modelo regulatório robusto deve incluir requisitos para monitoramento regular e ajustes baseados em resultados observados. Isto poderia incluir avaliações periódicas independentes sobre taxas de falsos positivos e negativos entre diferentes grupos populacionais e análise de impactos nas práticas policiais.

No entanto, regulação técnica, por si só, pode ser insuficiente para endereçar todas as preocupações éticas e sociais associadas a estas tecnologias. Questões fundamentais sobre os limites apropriados da vigilância estatal em sociedades democráticas, o direito à privacidade em espaços públicos, e o potencial efeito inibidor destas tecnologias sobre liberdades civis, como o direito de protesto, transcendem aspectos puramente técnicos e exigem um debate social amplo (Barocas; Selbst, 2016).

Neste sentido, qualquer moldura regulatória para o reconhecimento facial no Brasil deve ser desenvolvida através de processos deliberativos inclusivos que permitam a participação substantiva de diversos setores da sociedade, incluindo comunidades historicamente marginalizadas que enfrentam os maiores riscos de impactos adversos. A legitimidade e eficácia de tais regulações dependem de sua capacidade de refletir uma variedade de perspectivas e preocupações, especialmente daqueles mais vulneráveis a potenciais abusos.

O complexo panorama de questões éticas, técnicas e sociais levantadas pelas tecnologias de reconhecimento facial no contexto da segurança pública demonstra a urgência de desenvolver um arcabouço regulatório abrangente. A experiência internacional oferece lições e modelos que podem ser adaptados à realidade brasileira através de processos participativos que reflitam as particularidades do contexto nacional e priorizem a proteção de direitos fundamentais.

6 CONSIDERAÇÕES FINAIS

O reconhecimento facial, enquanto tecnologia em ascensão, apresenta desafios éticos e técnicos que não podem ser ignorados. A análise dos vieses algorítmicos e suas consequências sociais deve ser uma prioridade nas discussões sobre a implementação e regulação dessa tecnologia.

Diante deste panorama, torna-se imperativa a implementação de mecanismos de governança que garantam transparência e responsabilização em todas as etapas do processo, desde o design e desenvolvimento até a implementação final dessas tecnologias. Isto é, as práticas de construção e utilização de sistemas de reconhecimento facial precisam ser auditáveis e transparentes como padrão, não como possibilidades acessórias.

Esta exigência de responsabilização aplica-se igualmente aos fornecedores e aos compradores dessas tecnologias. Empresas que desenvolvem algoritmos de reconhecimento facial e entes públicos que os adquirem devem cooperar para satisfazer os requisitos de

transparência e uso responsáveis. Uma cadeia contínua de responsabilidade humana deve ser estabelecida e codificada ao longo de todo o ciclo de vida da produção e uso dessas tecnologias. Este processo deve iniciar na captura de dados e no mapeamento de horizontes tecnológicos, passar pela conceitualização e desenvolvimento, e estender-se até a implementação final no campo. Em cada etapa, deve haver clareza sobre quem são as autoridades responsáveis por decisões e quais são os parâmetros éticos que orientam essas escolhas.

REFERÊNCIAS

BAROCAS, S.; SELBST, A. D. Big Data's Disparate Impact. **California Law Review**, v. 104, p. 671-732, 2016.

BENJAMIN, R. **Race After Technology**: Abolitionist Tools for the New Jim Code. Cambridge: Polity Press, 2019.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD)**. Diário Oficial da União, Brasília, DF, 15 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 13 mar. 2025.

BUOLAMWINI, J.; GEBRU, T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In: **Conference on Fairness, Accountability and Transparency**/ 2018, Proceedings [...]. p. 77-91.

COSTANZA-CHOCK, S.; HARVEY, E.; RAJI, I. D.; et al. Who Audits the Auditors? Recommendations from a field scan of the algorithmic auditing ecosystem. In: **2022 ACM Conference on Fairness, Accountability, and Transparency**. 2022, p. 1571–1583.

DANKS, David; LONDON, Alex John. Algorithmic Bias in Autonomous Systems. *In*: **Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence**. Melbourne, Australia: International Joint Conferences on Artificial Intelligence Organization, 2017, p. 4691–4697. Disponível em: https://www.ijcai.org/proceedings/2017/654. Acesso em: 10 mar. 2025.

GOVERNMENT OF CANADA. **Directive on Automated Decision-Making**. Canadá, 2021. Disponível em: https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592. Acesso em: 10 mar. 2025.

JEFFERSON, B. J. **Digitize and Punish**: Racial Criminalization in the Digital Age. Minneapolis: University of Minnesota Press, 2020.

LESLIE, D. **Understanding bias in facial recognition technologies**: an explainer. Londres: The Alan Turing Institute, 2020.

MONTEIRO, P. D. C. **Reconhecendo faces, enclausurando corpos: terror racial, vigilância racializadora e o uso policial do reconhecimento facial na Bahia**. 2022. Dissertação (Mestrado) – Universidade Federal da Bahia, Faculdade de Direito, Salvador, 2022.

MORAIS DA ROSA, A.; WOJCIECHOWSKI, P. B. **Vieses da Justiça**: como as heurísticas e vieses operam nas decisões penais e a atuação contraintuitiva. Florianópolis: Emais, 2021.

NUNES, P. Mito da Neutralidade Tecnológica e Violência Racial: sobre o Uso de Tecnologias de Reconhecimento Facial para fins de Segurança Pública no Brasil. **Revista direito.UnB**, v. 6, n. 1, p. 105-140, 2022.

NUNES, P.; LIMA, T. G. L.; CRUZ, T. G. **O Sertão vai virar Mar**: Expansão do reconhecimento facial na Bahia. Rio de Janeiro: CESeC, 2023.

PALMA, A.; PACHECO, C. **"O policial já foi com a arma na cabeça dele", diz mãe de rapaz confundido por reconhecimento facial**. Correio, 2020. Disponível em: https://www.metro1.com.br/noticias/cidade/85609,o-policial-ja-foi-com-a-arma-na-cabeca-dele-diz-mae-de-jovem-confundido-por-reconhecimento-facial. Acesso em: 29 set. 2024.

PIRES, A. B. S.; CAVAGNOLI, G.; COTELLO, G.; VISANI, G.; GONGORA, L. Alvos predeterminados: Um estudo de caso sobre a implantação da tecnologia de reconhecimento facial na Bahia. In: **Dados, privacidade e perseguição penal**: Cinco estudos. FGV-SP; Data Privacy Brasil, 2021, p. 18-63.

RICHARDSON, R.; SCHULTZ, J.; CRAWFORD, K. **Dirty Data, Bad Predictions**: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. New York University Law Review Online, v. 94, p. 192-233, 2019.

SCHULER, L. G. B.; MONTARDO, S. P. A política das máquinas: vieses em algoritmos de relevância pública. **Rev. Tecnol. Soc**., Curitiba, v. 16, n. 45, p. 300-312, out./dez., 2020.

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

SIMÕES-GOMES, L.; ROBERTO, E.; MENDONÇA, J. **Viés algorítmico**: um balanço provisório. Estudos de Sociologia, v. 25, n. 48, p.139-166, jan.-jun. 2020.

TAJRA, A. **Ainda sem regulação, estados prendem centenas de pessoas utilizando reconhecimento facial**. Consultor Jurídico, 2024. Disponível em: https://www.conjur.com.br/2024-mai-17/sem-regulacao-estados-prendem-centenas-utilizando-reconhecimento-facial/. Acesso em: 11 mar. 2025.

TVERSKY, A.; KAHNEMAN, D. Judgment under Uncertainty: Heuristics and Biases. **Science**, v. 185, n. 4157, p. 1124-1131, 1974.

VEALE, M.; ZUIDERVEEN BORGESIUS, F. Demystifying the Draft EU Artificial Intelligence Act. **Computer Law Review International**, v. 22, n. 4, p. 97-112, 2021.

WECHSLER, H. **Reliable face recognition methods**: system design, implementation and evaluation. Springer, 2007.

WEST, S. M.; WHITTAKER, M.; CRAWFORD, K. **Discriminating Systems**: Gender, Race and Power in AI. AI Now Institute, 2019.

YIN, R. K. **Estudo de caso**: planejamento e métodos. 2. ed. Porto Alegre: Bookman, 2001.